Article title: Self-Perceived Loneliness and Depression During the COVID-19 Pandemic: a Two-Wave Replication Study

Authors: Alessandro Carollo[1], Andrea Bizzego[2], Giulio Gabrieli[3], Keri Ka-Yee Wong[4], Adrian Raine[5], Gianluca Esposito[6]

Affiliations: Department of Psychology and Cognitive Science, University of Trento, Italy[1], School of Social Sciences, Nanyang Technological University, Singapore[2], Department of Psychology and Human Development, University College London, London, UK[3], Departments of Criminology, Psychiatry, and Psychology, University of Pennsylvania[4]

Orcid ids: 0000-0002-2737-0218[1], 0000-0002-1586-8350[2], 0000-0002-9846-5767[3], 0000-0002-2962-8438[4], 0000-0002-3756-4307[5], 0000-0002-9442-0254[6]

Contact e-mail: gesposito79@gmail.com

Preprint statement: This article is a preprint and has not been peer-reviewed, under consideration and submitted to UCL Open: Environment Preprint for open peer review.

Links to data: www.doi.org/10.5522/04/16583861

Funder: UCL Global Engagement Fund

DOI: 10.14324/111.444/000095.v2

Preprint first posted online: 30 June 2022

Keywords: COVID-19, depression, lockdown, loneliness, global study, machine learning, SARS-CoV-2, Health

Self-Perceived Loneliness and Depression During the COVID-19 Pandemic: a Two-Wave Replication Study

Alessandro Carollo[a], Andrea Bizzego[a], Giulio Gabrieli[b], Keri Ka-Yee Wong[c], Adrian Raine[d], Gianluca Esposito*[a]

[a] *Department of Psychology and Cognitive Science, University of Trento, Italy*
[b] *School of Social Sciences, Nanyang Technological University, Singapore*
[c] *Department of Psychology and Human Development, University College London, London, UK*
[d] *Departments of Criminology, Psychiatry, and Psychology, University of Pennsylvania*

Abstract

Background: The global COVID-19 pandemic has forced countries to impose strict lockdown restrictions and mandatory stay-at-home orders with varying impacts on individual's health. Combining a data-driven machine learning paradigm and a statistical approach, our previous paper documented a U-shaped pattern in levels of self-perceived loneliness in both the UK and Greek populations during the first lockdown (17 April to 17 July 2020). The current paper aimed to test the robustness of these results by focusing on data from the first and second lockdown waves in the UK. **Methods:** We tested *a*) the impact of the chosen model on the identification of the most time-sensitive variable in the period spent in lockdown. Two new machine learning models - namely, support vector regressor (SVR) and multiple linear regressor (MLR) were adopted to identify the most time-sensitive variable in the UK dataset from wave 1 ($n = 435$). In the second part of the study, we tested *b*) whether the pattern of self-perceived loneliness found in the first UK national lockdown was generalizable to the second wave of

UK lockdown (17 October 2020 to 31 January 2021). To do so, data from wave 2 of the UK lockdown (n = 263) was used to conduct a graphical inspection of the week-by-week distribution of self-perceived loneliness scores. **Results:** In both SVR and MLR models, depressive symptoms resulted to be the most time-sensitive variable during the lockdown period. Statistical analysis of depressive symptoms by week of lockdown resulted in a U-shaped pattern between week 3 to 7 of wave 1 of the UK national lockdown. Furthermore, despite the sample size by week in wave 2 was too small for having a meaningful statistical insight, a graphical U-shaped distribution between week 3 and 9 of lockdown was observed. **Conclusions:** Consistent with past studies, these preliminary results suggest that self-perceived loneliness and depressive symptoms may be two of the most relevant symptoms to address when imposing lockdown restrictions.

Keywords: COVID-19; depression; lockdown; loneliness; global study; machine learning; SARS-CoV-2

Correspondence: 31 Corso Angelo Bettini, Rovereto, 38068, Italy. gianluca.esposito@unitn.it

1. Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a novel and highly pathogenic coronavirus that originated in bats and hosted by pangolins before the spillover to humans [1, 2, 3, 4]. SARS-CoV-2 disease was first documented in the Hubei province of China in December 2019 and has since rapidly spread throughout the world with the World Health Organization declaring it a pandemic on 11 March 2020 [5]. As of September 2021, over 224 million people have been infected by COVID-19 and more than 4.6 millions of deaths have been reported globally [6].

With no available vaccine to prevent COVID-19, many countries were initially forced to adopt lockdown restrictions, which greatly impacted the environments in which people were legally allowed to work in, play in, and socialise in - all in the efforts to slow down the spread of the invisible virus. Across countries, restrictions varied in period, length, and strictness - but all mandates resulted in reduced physical contact between humans in environments that we are used to. In particular, the UK's first lockdown imposed on 23rd March 2020 encountered a 'must-stay-home' order [7], forcing many individuals to renegotiate the home environment as simultaneously also a place of play, learning, rest, and socialising. Leaving the house was allowed only once a day and for essentials only like shopping, exercising, medical needs, caring duties, and essential travel for work [8]. These restrictions were accompanied by physical distancing measures, which were aimed at reducing the person-to-person transmission of the virus by encouraging the population to stay at least 2 meters away from others [9]. Though these policies were effective at reducing the number of new cases and the spread of the airborne virus, individuals had to endure long periods of social isolation, reduced activity in confined indoor spaces, skepticism towards others, and little to no contact with others (e.g., friends, parents, siblings, partners), which may have had short and longer-term impacts on their health.

Considering the impact of social isolation on people's physical and mental health [10, 11, 12, 13], we hypothesized that lockdown measures, specifically lockdown duration (in days), may impact several important aspects of an individual's daily lives. Globally, studies have documented links between restrictions and poorer mental health, such as more post-traumatic stress symptoms, anxiety, depression, insomnia, and trust in others [14, 15, 16, 17, 18]. Similarly, in a previous data-driven study, we identified that, by using a machine learning model, self-perceived loneliness was most impacted by the time in lockdown, over and above other mental health indicators [19]. Further statistical analyses were conducted to assess the variations in participants' levels of self-perceived loneliness as a function of time spent in lockdown (in weeks). Specifically, participants from the UK who took part in the study during week 6 of national lockdown reported significantly lower levels of self-perceived loneliness compared to their counterparts who completed the survey during week 3 of lockdown. Likewise, lower levels of self-perceived loneliness were observed for participants who completed the survey in weeks 4 and 6 of the Greek national lockdown. This pattern of results together with a graphical inspection suggested the existence of a U-shaped distribution in self-perceived loneliness levels by weeks in lockdown in both the UK and Greece. An effect of restrictions on an individual's perceived loneliness during the first lockdown period was replicated and substantiated by other

COVID studies in the literature [20, 21, 22, 23].

Building on previous findings, the current study aims to replicate and extend on the previous results. In particular, the current study consists of two parts. In the first part, the work aims to test whether the identification of the most time-sensitive variable by Carollo et al. [19] depended on the chosen machine learning model. To do so, we applied two new machine learning models on the same set of UK data from the first lockdown period to identify the most time-sensitive variable. In this way, we wanted to verify if, when changing the predictive model, new variables with different patterns of time-sensitivity could be identified and studied under a statistical approach. This would provide insight into other time-sensitive variables that might have been overlooked by the previously adopted model - namely, the RandomForest model. In the second part, the study aims to test whether the documented distribution of self-perceived loneliness levels by week in lockdown depended on the specific wave of lockdown. To do so, we graphically analyzed self-perceived loneliness distribution by week on data from the second UK national lockdown, with data collected from the UCL-Penn Global COVID Study between 17 October 2020 and 31 January 2021 [24]. The current study provides the opportunity to uncover other aspects that may be significantly influenced by the lockdown restrictions in both the first and second waves of lockdown.

2. Methods

2.1. Questionnaire

The current study is based on survey data from the UCL-Penn Global COVID Study, a 12-month study of COVID-19's impact on mental health in adults conducted between 17 April 2020 and 31 July 2021 [24]. Specifically, this study will use data from wave 1 collected between 17 April 2020 and 10 July 2020, and data from wave 2 collected between 17 October 2020 and 31 January 2021. Briefly, the survey was available in 8 languages and anyone 18 years and above with access to the survey link through several social media channels (website - www.GlobalCOVIDStudy.com -, email, LinkedIn, Whatsapp, Instagram, Facebook, and Reddit) was able to take part in the study. Participants received a randomized presentation of 13 standardized questionnaires assessing mental health including self-perceived loneliness, anxiety, depression, aggression, physical health, social relationships (empathy), living conditions, and background variables. For this study, 12 indices derived from the previous questionnaires were included in the analytic sample (see Table 1). As an index of internal reliability, Cronbach's alpha was computed over the scores based on multiple items. This study received ethical approval from the University College London Institute of Education Research Ethics Committee (REC 1331; April 2020).

2.2. Participants

Participants from the first wave of lockdown

During the first period of lockdown, a total of 2,276 adults from 66 different countries participated in the study. We excluded participants who: i)

Score	Description	Reference	Domain	Cronbach's Alpha (C.I. 95%)	Observed Range
Mild Activity Difference	Difference between days of mild physical activity post- and pre- COVID-19 lockdown.	*International Physical Activity Questionnaire – Short Form* (IPAQ-SF, 6-items) [25]	Physical Activity	Not applicable	[-7, 6]
Mild Activity Time Difference	Difference between minutes of mild physical activity post- and pre- COVID-19 lockdown.	*International Physical Activity Questionnaire – Short Form* (IPAQ-SF, 6-items) [25]	Physical Activity	Not applicable	[-480, 510]
Moderate Activity Difference	Difference between days of moderate physical activity post- and pre- COVID-19 lockdown.	*International Physical Activity Questionnaire – Short Form* (IPAQ-SF, 6-items) [25]	Physical Activity	Not applicable	[-6, 7]
Sleep Quality	Self-reported sleep quality and quantity, where higher scores reflect better sleep quality.	*Pittsburgh Sleep Quality Index* (2-items) [26], *Epworth Sleepiness Scale* [27], *Subjective and Objective Sleepiness Scale* [28]	Sleep Quality	0.73 (0.70-0.77)	[7, 23]
Empathy	Self-reported affective, cognitive, and somatic empathy, where higher scores reflect higher empathy.	*Cognitive, Affective, Somatic Empathy Scale* (CASES, 30-items) [29]	Empathy	0.87 (0.85-0.88)	[29, 60]
Anxiety	Higher scores reflect higher anxiety.	*General Anxiety Disorder-7* (GAD-7) [30]	Anxiety	0.89 (0.88-0.91)	[0, 20]
Depression	Higher scores reflect higher depression.	*Patient Health Questionnaire-9* (PHQ-9, 9-items) [31]	Depression	0.87 (0.86-0.89)	[0, 22]
Perceived Loneliness	Higher scores reflect higher perceived loneliness.	*Loneliness Questionnaire* (LQ, 20-items) [32]	Perceived Loneliness	0.94 (0.93-0.95)	[23, 71]
Living Conditions/Environment	Higher scores reflect more chaotic home environments.	*Chaos, Hubbub, and Order Scale and Health Risk Behaviors* (CHAOS, 6-items) [33]	Demographic Information	0.66 (0.62-0.67)	[6, 24]
Beliefs	Perceived effectiveness of government guidelines on social distancing, schools closing, face masks and gloves as protection. Higher scores reflect stronger beliefs.	Summed 9-items on COVID-19 beliefs	Worries and Beliefs	0.81 (0.78-0.83)	[19, 45]
Schizotypal Traits	Higher scores reflect more schizotypal traits.	Schizotypal Personality Questionnaire–Brief [34]	Social Suspicions and Schizotypal Traits	0.73 (0.70-0.77)	[0, 19]
Reactive-Proactive Aggression	Higher score reflects more aggression.	Reactive-Proactive Aggression Questionnaire [35]	Aggression	0.86 (0.84-0.87)	[0, 21]

Table 1: Variables that are computed to quantify participants' mental and physical health and living environment during lockdown. Cronbach's Alpha was computed on multiple-item scores and it refers to the scores collected during the first wave of lockdown.

dissented to take part (n = 32), had incomplete (n = 712) or missing data (n = 165); ii) did not complete the survey within two days from the start date (n = 76); iii) filled in the survey from a country that was different from their original country of residence (n = 132). Criterion ii) was applied to exclude possible confounds in the amount of time passed from the start to the end of survey completion. This was a particularly key point in the data processing procedure since we were interested in the effects that the amount of time in lockdown had on people's mental and physical health. Similarly, criterion iii) was applied to exclude confounds of different types of lockdown restrictions that were adopted by the various countries of the world. All of these participants were excluded from the final analysis.

In contrast to Carollo et al. [19], the current study examined UK participants only. After also excluding the participants who completed the survey after week 9 of lockdown (n = 40), the analytic sample (N = 435) had the following demographic features: female = 345 (79.31%), male = 81 (18.62%), non-binary = 4 (0.92%), prefer not to say = 2 (0.46%), self-identified = 3 (0.69%); age: Range = 18-88 years, *Mean* = 37.62, *SD* = 13.83 (missing = 1).

Participants from the second wave of lockdown

With regard to the second wave of lockdown, 2,280 participants completed the survey. The same exclusion criteria described in the section above were applied to wave 2 data. Thus, 1,341 and 140 participants were excluded because they had incomplete and missing data respectively. Other 206 were excluded because they did not complete the survey within two days. Finally, 43 did not filled in the survey from their original country of residence and, therefore, were excluded from the analysis.

To be consistent with the sample used in our previous study, the statistical analysis applied to uncover the pattern of self-perceived loneliness in wave 2 was conducted uniquely on the UK participants (n = 263). The sample had the following demographic features: female = 216 (82.13%), male = 39 (14.83%), non-binary = 5 (1.90%), prefer not to say = 2 (0.76%), self-identified = 1 (0.38%); age: Range = 18-89 years, *Mean* = 38.28, *SD* = 13.74 (missing = 2).

2.3. Data Analysis

All the scripts for the data analysis are available at the following link: https://doi.org/10.5522/04/20183858. Prior to data analysis, we computed the variable "Weeks in lockdown" for each participant in both wave 1 and wave 2 of the UK national lockdown. The variable "Weeks in lockdown" corresponds to the difference between the date in which the UK adopted lockdown preventive measures (either the beginning of the first or the second lockdown wave) and the survey completion date. This new numerical variable referred to the week of lockdown into which the single participant completed the survey. Table 2 reports the number of participants by week across the first and second waves of the UK national lockdown.

Wave of lockdown	**Before Week 3**	**Week 3**	**Week 4**	**Week 5**	**Week 6**	**Week 7**	**Week 8**	**Week 9**	**After Week 9**	**TOT**
Wave 1	0	42	100	80	76	110	23	4	0	435
Wave 2	244	5	2	3	1	0	0	4	4	263

Table 2: Number of participants from the UK by week during the first and second period of lockdown.

Using data from waves 1 and 2 of the UCL-Penn Global COVID Study and the same health variables across both time-points, we conducted two sets of analyses to answer our research questions. To test whether the identification of the most time-sensitive variable in Carollo et al. [19] depended on the chosen machine learning model, we used wave 1 data and we adopted a data-driven machine learning approach. As compared to the RandomForest model adopted in Carollo et al. [19], in the current work we used two different machine learning models to identify the most time-sensitive variable (out of the 12 indices included). The distribution of scores by week of the identified most time-sensitive variable was then examined through a statistical approach with significance tests corrected for multiple comparisons.

To test whether the U-shaped pattern of self-perceived loneliness found in Carollo et al. [19] was unique to wave 1 of lockdown, we used wave 2 data to conduct a graphical inspection of the distribution of scores by week in lockdown.

Data-driven and statistical replication of the results in wave 1

The current paper first adopted a machine learning approach to test whether the identification of the most time-sensitive variable in Carollo et al. [19] was specific to the RandomForest model or whether we would replicate the result using new models - namely, Support Vector Regressor (SVR) [36] and Multiple Linear Regression (MLR). While RandomForest's predictions are based on the creation of an ensemble of decision trees from the input variables, SVR is rooted on the derivation of a best-fit hyperplane and the MLR on linear relations between variables. Data from 12 variables of interest (outlined in Table 1) were included in the models to predict the independent variable "Weeks in lockdown". The assumption behind this approach was that the independent variable "Weeks in lockdown" would modulate, to a different extent, the scores of the dependent variables included in the dataset. Particularly, the most time-sensitive variable would be strongly modulated by time in lockdown and its scores would systematically co-vary with the variable "Weeks in lockdown". Therefore, the most time-sensitive variable would also be the most informative and important for the model when trying to predict "Weeks in lockdown". Under these assupmtions, first, we applied a standardized 10x5fold cross-validation scheme to train the SVR and the MLR on 75% of the data. Once the models were established, we then applied them to the remaining 25% of data, the 'testing set' data. The cross-validation and the train-test split procedures are common practice in machine learning as they help to control the model's overfitting by evaluating the model's performances on unseen data [37]. Overall, the models' accuracy was assessed by comparing real and predicted values. In particular, the models' performances were evaluated by Mean Squared Error (MSE), which consists of the average squared difference between predicted and real values. Thus, a lower MSE value corresponds to a higher overlap between the real and predicted data. For every training iteration, the variables were ranked by their absolute coefficient value to reflect their influence on the model's built. On all the training' importance rankings, we computed a Borda count to determine the most important and informative variable for the model's prediction of the Weeks in lockdown. Borda count is a method to derive a single list summarizing the information coming from a set of lists [38]. For the SVR model, by comparing the several training-evaluation iterations, we derived the optimal hyper-parameter C. In SVR, the parameter C is a cost regularization parameter which determines the trade-off cost between minimizing the training error and minimizing model complexity [39]. The resulting optimized C parameter was equal to the value of 0.01, and it was implemented in the final model. The final models (i.e., SVR with C parameter set at 0.01 and the MLR) were then trained by using all the data from the training set and their performances were evaluated on the testing set data.

Next, focusing on the most time-sensitive variable identified with the SVR and MLR models, we applied a multipair Kruskal-Wallis test to assess whether the variable scores changed over the lockdown period. Scores from participants belonging to weeks 3 (since at the beginning of the data collection, the UK lockdown was already started) to 7 were compared. As the study had a cross-sectional design across waves of lockdown, participants were grouped by the "Week in lockdown" variable. "Week in lockdown" groups were compared in terms of scores reported for the identified most time-sensitive variable. In this way, a significant result in the multipair Kruskal-Wallis test would indicate that levels of the identified variable significantly differed by "Weeks in lockdown" for at least two groups of weeks. If the multipair Kruskal-Wallis test suggested the existence of significant weekly variations, we conducted multiple pairwise Kruskal-Wallis tests with Bonferroni correction to compare week 7 scores to other weeks. Eta-squared was computed to estimate the magnitude of significant results [40, 41].

Graphical replication of the results in wave 2

To test whether the distribution of weekly self-perceived loneliness levels were unique to wave 1 of lockdown, a graphical qualitative inspection was conducted on wave 2 data. Again, participant's self-perceived loneliness scores were clustered by week of lockdown and the distribution of scores from week 3 to 9 was inspected with boxplots. It is worth noting that, considering the limited sample size that was available for wave 2 from week 3 to 9, no statistically meaningful insight could be derived from the comparisons of groups, so the second part of the study can only have a qualitative and descriptive significance, and must be considered as a preliminary approach.

3. Results

3.1. Replication of the results in wave 1

MSEs for the SVR performances were 2.04 and 2.29 for the training and test data, respectively. For the MLR, MSEs were 1.97 and 2.39 for the training and test data, respectively. While both models' performances on the training set are slightly worse than in Carollo et al. [19], the performances on the test are in line with the previous paper. Furthermore, depression scores were found to be the most informative for both the SVR and MLR's training, above and beyond the other variables in the models (see Figure 1).

A closer look at boxplots representing depressive symptoms divided by week in lockdown suggests that, from week 3 to 7, the median score decreased in the first period (week 3 to week 4) and then increased again (from week 4 to week 7; see Figure 2). A decrease followed by an increase in scores suggests a U-shaped pattern for depressive symptoms in the first wave of UK lockdown.

A Kruskal-Wallis test confirmed that at least one week (in the period from the 3rd to the 7th week of lockdown) differed significantly from the others in terms of depressive symptoms (H=22.03, $p < 0.001$, $\eta^2 = 0.042$).

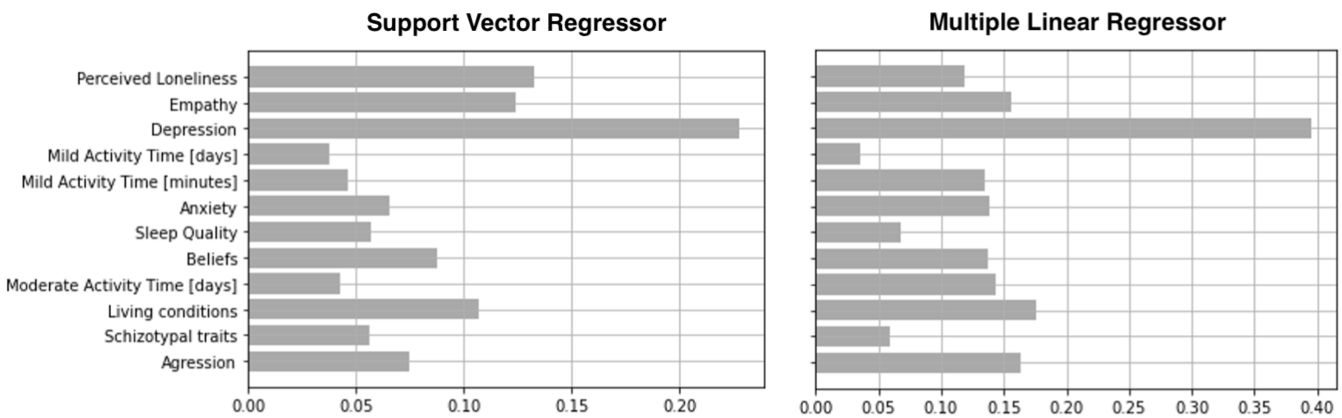

Figure 1: Normalized average importance of the selected variables when training a Support Vector Regressor model (on the left) and a Multiple Linear Regressor (on the right) on data from the first lockdown period. The importance of the variables was derived from the trained predictive models as the absolute value of the variables' weights or coefficients for the SVR and MLR, respectively.

Specifically, symptoms between week 4 and week 7 (H=22.52, $p < 0.001$, $\eta^2 = 0.050$), and between week 5 and week 7 (H=9.69, p=0.002, $\eta^2 = 0.020$) were statistically different. Conversely, the comparisons between week 3 to week 7 (H=4.64, p=0.031), and week 6 to week 7 (H=4.02, p=0.045) were not significant after applying the Bonferroni bias-correction.

3.2. Qualitative replication of the results in wave 2

A graphical inspection of boxplots with self-perceived loneliness scores divided by week suggests that, between week 3 to 9 of wave 2 UK national lockdown, another U-shaped pattern could be reported. Specifically, participants who took part at the study during the 4th and 5th week of lockdown reported lower levels of self-perceived loneliness than did participants in the survey during week 3. Although there were not enough participants for week

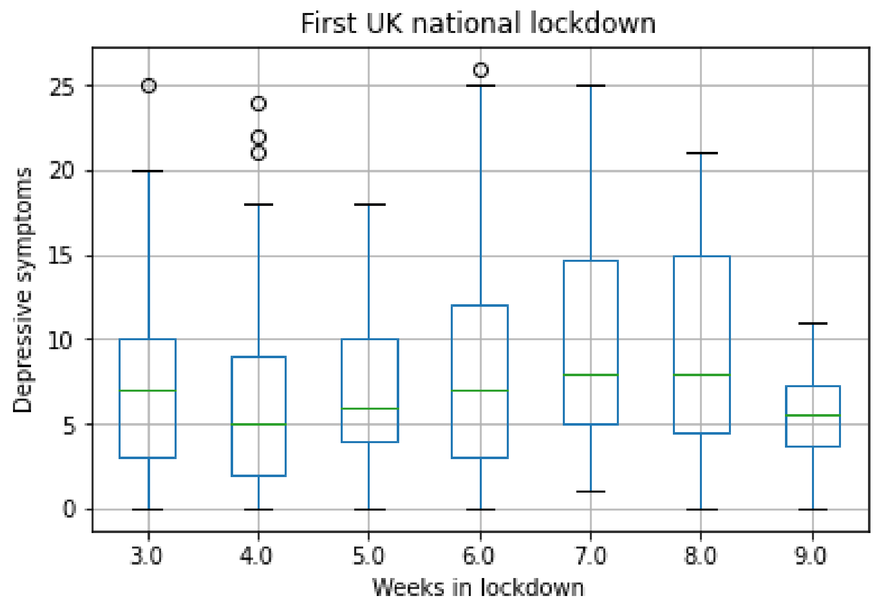

Figure 2: Symptoms of Depression reported by week during the first UK national lockdown.

6, 7, and 8, self-perceived loneliness scores during week 9 were reportedly higher again (see Figure 3).

4. Discussion

This study applying a machine learning approach alongside a statistical approach to data from waves 1 (17 April to 31 July 2020) and 2 (17 October 2020 to 31 January 2021) of the UCL-Penn Global COVID Study [24] identifies the mental health variable(s) most influential in predicting UK lockdown duration, and how the variable varies by week. This gives an indication of how people were fairing when confined in the limited, often shared, space in which they have to work, learn, play, and rest in. With the aim of replicat-

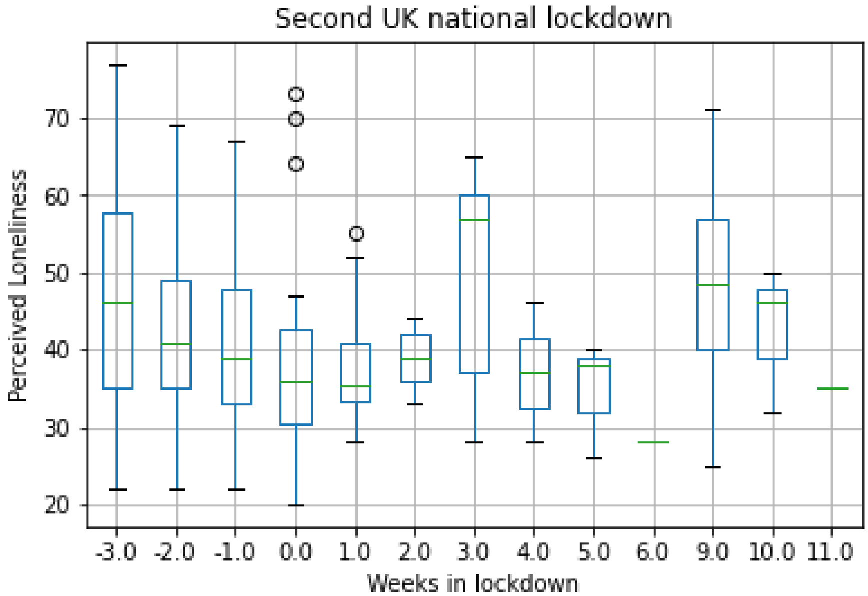

Figure 3: Reports of Perceived Loneliness by week during the second UK national lockdown.

ing and extending the results from our previous paper, Carollo et al. [19], we applied a Support Vector Regressor (SVR) model and a Multiple Linear Regression (MLR) model instead of a RandomForest model to predict participant's weeks in lockdown. Based on the variables importance ranking, depressive symptoms, over and above the other 11 health indices, were the most important variable for both the SVR and MLR models when determining the model best-fit to the data and were the best at predicting lockdown duration in weeks. Depressive symptoms were therefore identified by both the SVR and MLR models as the most time-sentivie variable in the dataset. Since the focus of the study was not to assess the variables' predictive capability *per se*, it is worth noting that the low model performance did not affect the reliability of the variable importance ranking and, therefore, the identification of the most time-sensitive variable in the dataset [19]. Specifically, depressive symptoms reported across the 9 lockdown weeks resulted in a U-shaped pattern where symptoms were lowest during weeks 4 and 5 compared to week 7.

Variation in the population's depressive symptoms during lockdown has been reported by past studies as depressive symptoms have been a key mental health issue during the COVID-19 pandemic [42, 43, 44, 45]. Specifically, Ammar et al. [46] compared the scores pre- and post-lockdown in symptoms of depression and found higher depressive symptoms as a result of home confinement. Notably, this study relied on self-report ratings of depression from participants internationally (e.g., Asia, Europe, and Africa), thus further substantiating the reliability of our finding. This is not surprising, given that social isolation is a common precursor of poorer mental and physical health [47], with increased risk for depression [48, 49, 50]. In another study by Delmastro and Zamariola [51] of lockdown in Italy, people living alone, or not being allowed to leave the house to go to work, tended to have higher depressive symptoms. Like self-perceived loneliness, symptoms of depression have varied during the first UK lockdown. Self-report data from the US during their first three-months of lockdown also showed that self-perceived loneliness was positively correlated with depression and suicide ideation at various time-points [52]. In fact, during the COVID-19 pandemic, self-perceived loneliness - a discrepancy between desired and perceived social connection - seemed to be one of the most important risk-factors for depression (and anxiety) [53], and social trust [18]. Specifically, higher perceived social support during lockdown - in other words, lower self-perceived loneliness - was associated with lower depressive symptoms [54]. After such periods, instead, self-perceived loneliness appeared to act as a moderator between stress and depression [55].

While the limited sample size by week in wave 2 data did not allow to use the statistical approach adopted in [19], a graphical U-shaped pattern of self-perceived levels of loneliness seems to emerge again across the lockdown weeks. Again, qualitatively, the self-perceived levels of loneliness were low during weeks 4 and 5, and highest during the third and ninth weeks of the lockdown period. These results have to be considered only as a qualitative and preliminary insight, since the sample size collected for the weeks of interest did not allow to make any meaningful statistical inference. In fact, graphical disparities among scores might be mere random variation and they might not reflect real differences. Nonetheless, our study findings suggest that local and nation-wide initiatives to help reduce self-perceived loneliness and increase solidarity and community cohesion may be helpful at improving people's mental health during lockdowns.

In conclusion, both self-perceived loneliness and depressive symptoms appear to follow U-shaped curves across periods of lockdown (although no statistical test was computed over scores of self-perceived loneliness by week in the second wave of the UK lockdown). Knowing the unfolding of these trajectories might be helpful for conveying the adequate support to the population in lockdown with the right timing. People might also be made aware of the possible fluctuations in self-perceived loneliness and depressive symptoms throughout the lockdown period. Overall, this knowledge can help manage expectations in populations and support systems to ensure that resources are allocated effectively, especially in future lockdown environments. Of course, "why" both perceived levels of loneliness and depression follow U-shaped patterns will necessarily involve the examination of individual-level characteristics (e.g., age, gender), or other variables, that were not assessed and explored in the current study. For the same aim, a longitudinal investigation - opposed to the cross-sectional design of the current study - could also result useful. Furthermore, to fully pursue the replication aims of the current study, it would be useful to apply the same machine learning and statistical approach across different data sources. As we did not find any dataset similar enough to the one we adopted, the results from the current paper can only be considered as preliminary. Although these limitations, the present study has also some clear strengths. First of all, a wide range of mental and physical variables could be studied in a data-driven fashion thanks to the adopted machine learning approach. In this way, we were able to identify and, in a second phase, statistically characterize the index that varied the most accordingly to the time spent in lockdown. Moreover, given the differences across lockdown restrictions, cross-cultural comparisons of the impacts of COVID-19 on populations are challenging. Thus, a strength of the current study is to focus just on the UK. Generally, the study highlighted the importance of considering the potential weekly variation in mental health across a wide range of variables and the variation that may exists across individuals and countries with different lockdown restrictions.

Author contribution

Conceptualization: A.B., G.G., K.K.W., G.E.; Data curation: A.C., A.B., G.G., K.K.W.; Data analysis, Data interpretation, Writing: A.C., A.B.; Revision: A.C., A.B., G.G., K.K.W., A.R., G.E.; Supervision: G.E. All authors read and agreed to the published version of the manuscript.

Conflicts of interest

The authors declare no conflict of interest.

Ethics

This study was pre-registered (https://osf.io/4nj3g/) on 17 April 2021 and ethical approval for the COVID-19 Social Study was granted by the University College London Institute of Education Ethics and Review Committee in April 2020 (REC 1331; [24]). The study is GDPR compliant.

References

[1] F. Wu, S. Zhao, B. Yu, Y.-M. Chen, W. Wang, Z.-G. Song, Y. Hu, Z.-W. Tao, J.-H. Tian, Y.-Y. Pei, et al., A new coronavirus associated with human respiratory disease in china, Nature 579 (2020) 265–269.

[2] T. Zhang, Q. Wu, Z. Zhang, Probable pangolin origin of sars-cov-2 associated with the covid-19 outbreak, Current biology 30 (2020) 1346–1351.

[3] P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li, C.-L. Huang, et al., A pneumonia outbreak associated

with a new coronavirus of probable bat origin, nature 579 (2020) 270–273.

[4] T. T. Nguyen, P. N. Pathirana, T. Nguyen, Q. V. H. Nguyen, A. Bhatti, D. C. Nguyen, D. T. Nguyen, N. D. Nguyen, D. Creighton, M. Abdelrazek, Genomic mutations and changes in protein secondary structure and solvent accessibility of sars-cov-2 (covid-19 virus), Scientific Reports 11 (2021) 1–16.

[5] W. H. Organization, Coronavirus Disease 2019 - Situation Report – 51, https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200311-sitrep-51-covid-19.pdf, 2020.

[6] W. H. Organization, Weekly epidemiological update on covid-19 - 14 september 2021, https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---14-september-2021, 2021.

[7] G. Iacobucci, Covid-19: Uk lockdown is "crucial" to saving lives, say doctors and scientists, 2020.

[8] R. S. Ogden, The passage of time during the uk covid-19 lockdown, Plos one 15 (2020) e0235871.

[9] M. Qian, J. Jiang, Covid-19 and social distancing, Journal of Public Health (2020) 1–3.

[10] J. S. Kuiper, M. Zuidersma, R. C. O. Voshaar, S. U. Zuidema, E. R. van den Heuvel, R. P. Stolk, N. Smidt, Social relationships and risk of

dementia: A systematic review and meta-analysis of longitudinal cohort studies, Ageing research reviews 22 (2015) 39–57.

[11] J. Holt-Lunstad, T. B. Smith, M. Baker, T. Harris, D. Stephenson, Loneliness and social isolation as risk factors for mortality: a meta-analytic review, Perspectives on psychological science 10 (2015) 227–237.

[12] N. K. Valtorta, M. Kanaan, S. Gilbody, S. Ronzi, B. Hanratty, Loneliness and social isolation as risk factors for coronary heart disease and stroke: systematic review and meta-analysis of longitudinal observational studies, Heart 102 (2016) 1009–1016.

[13] N. Cotterell, T. Buffel, C. Phillipson, Preventing social isolation in older people, Maturitas 113 (2018) 80–84.

[14] T. Elmer, K. Mepham, C. Stadtfeld, Students under lockdown: Comparisons of students' social networks and mental health before and during the covid-19 crisis in switzerland, Plos one 15 (2020) e0236337.

[15] A. J. Idrissi, A. Lamkaddem, A. Benouajjit, M. B. El Bouaazzaoui, F. El Houari, M. Alami, S. Labyad, A. Chahidi, M. Benjelloun, S. Rabhi, et al., Sleep quality and mental health in the context of covid-19 pandemic and lockdown in morocco, Sleep medicine 74 (2020) 248–253.

[16] R. Rossi, V. Socci, D. Talevi, S. Mensi, C. Niolu, F. Pacitti, A. Di Marco, A. Rossi, A. Siracusano, G. Di Lorenzo, Covid-19 pandemic and lockdown measures impact on mental health among the general population in italy, Frontiers in psychiatry 11 (2020) 790.

[17] C. Pieh, S. Budimir, J. Delgadillo, M. Barkham, J. R. Fontaine, T. Probst, Mental health during covid-19 lockdown in the united kingdom, Psychosomatic medicine 83 (2021) 328–337.

[18] K. K.-Y. Wong, Y. Wang, G. Esposito, A. Raine, A three-wave network analysis of covid-19's impact on schizotypal traits, paranoia and mental health through loneliness., UCL Open: Environment Preprint (2021).

[19] A. Carollo, A. Bizzego, G. Gabrieli, K. K.-Y. Wong, A. Raine, G. Esposito, I'm alone but not lonely. u-shaped pattern of self-perceived loneliness during the covid-19 pandemic in the uk and greece, Public Health in Practice 2 (2021) 100219.

[20] J. M. Groarke, E. Berry, L. Graham-Wisener, P. E. McKenna-Plumley, E. McGlinchey, C. Armour, Loneliness in the uk during the covid-19 pandemic: Cross-sectional results from the covid-19 psychological wellbeing study, PloS one 15 (2020) e0239698.

[21] W. D. Killgore, S. A. Cloonan, E. C. Taylor, N. S. Dailey, Loneliness: A signature mental health concern in the era of covid-19, Psychiatry research 290 (2020) 113117.

[22] S. G. S. Shah, D. Nogueras, H. C. van Woerden, V. Kiparoglou, The covid-19 pandemic: A pandemic of lockdown loneliness and the role of digital technology, Journal of Medical Internet Research 22 (2020) e22287.

[23] M. T. Tull, K. A. Edmonds, K. M. Scamaldo, J. R. Richmond, J. P. Rose, K. L. Gratz, Psychological outcomes associated with stay-at-home

orders and the perceived impact of covid-19 on daily life, Psychiatry research 289 (2020) 113098.

[24] K. K. Wong, A. Raine, Covid-19: Global social trust and mental health study, https://doi.org/10.17605/OSF.IO/FE8Q7, 2020.

[25] P. H. Lee, D. J. Macfarlane, T. H. Lam, S. M. Stewart, Validity of the international physical activity questionnaire short form (ipaq-sf): A systematic review, International Journal of Behavioral Nutrition and Physical Activity 8 (2011) 115.

[26] D. J. Buysse, C. F. Reynolds III, T. H. Monk, S. R. Berman, D. J. Kupfer, The pittsburgh sleep quality index: a new instrument for psychiatric practice and research, Psychiatry research 28 (1989) 193–213.

[27] M. W. Johns, A new method for measuring daytime sleepiness: the epworth sleepiness scale, sleep 14 (1991) 540–545.

[28] T. Åkerstedt, M. Gillberg, Subjective and objective sleepiness in the active individual, International Journal of Neuroscience 52 (1990) 29–37.

[29] A. Raine, F. R. Chen, The cognitive, affective, and somatic empathy scales (cases) for children, Journal of Clinical Child & Adolescent Psychology 47 (2018) 24–37.

[30] R. L. Spitzer, K. Kroenke, J. B. Williams, B. Löwe, A brief measure for assessing generalized anxiety disorder: the gad-7, Archives of internal medicine 166 (2006) 1092–1097.

[31] K. Kroenke, R. L. Spitzer, J. B. Williams, The phq-9: validity of a brief depression severity measure, Journal of general internal medicine 16 (2001) 606–613.

[32] D. W. Russell, Ucla loneliness scale (version 3): Reliability, validity, and factor structure, Journal of personality assessment 66 (1996) 20–40.

[33] A. P. Matheny Jr, T. D. Wachs, J. L. Ludwig, K. Phillips, Bringing order out of chaos: Psychometric characteristics of the confusion, hubbub, and order scale, Journal of applied developmental psychology 16 (1995) 429–444.

[34] A. Raine, D. Benishay, The spq-b: A brief screening instrument for schizotypal personality disorder, Journal of Personality Disorders 9 (1995) 346–355.

[35] A. Raine, K. Dodge, R. Loeber, L. Gatzke-Kopp, D. Lynam, C. Reynolds, M. Stouthamer-Loeber, J. Liu, The reactive–proactive aggression questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys, Aggressive Behavior: Official Journal of the International Society for Research on Aggression 32 (2006) 159–171.

[36] V. Vapnik, The nature of statistical learning theory, Springer science & business media, 2013.

[37] A. Bizzego, G. Gabrieli, M. H. Bornstein, K. Deater-Deckard, J. E. Lansford, R. H. Bradley, M. Costa, G. Esposito, Predictors of contemporary under-5 child mortality in low-and middle-income countries: a machine

learning approach, International journal of environmental research and public health 18 (2021) 1315.

[38] G. Jurman, S. Riccadonna, R. Visintainer, C. Furlanello, Algebraic comparison of partial lists in bioinformatics, PloS one 7 (2012) e36540.

[39] C.-H. Wu, G.-H. Tzeng, R.-H. Lin, A novel hybrid genetic algorithm for kernel function and parameter optimization in support vector regression, Expert Systems with Applications 36 (2009) 4725–4735.

[40] M. Tomczak, E. Tomczak, The need to report effect size estimates revisited. an overview of some recommended measures of effect size, Trends in sport sciences 1 (2014) 19–25.

[41] A. Carollo, W. Chai, E. Halstead, D. Dimitriou, G. Esposito, An exploratory analysis of the effect of demographic features on sleeping patterns and academic stress in adolescents in china, International Journal of Environmental Research and Public Health 19 (2022) 7032.

[42] C. Pieh, S. Budimir, T. Probst, The effect of age, gender, income, work, and physical activity on mental health during coronavirus disease (covid-19) lockdown in austria, Journal of psychosomatic research 136 (2020) 110186.

[43] D. A. Antiporta, Y. L. Cutipé, M. Mendoza, D. D. Celentano, E. A. Stuart, A. Bruni, Depressive symptoms among peruvian adult residents amidst a national lockdown during the covid-19 pandemic, BMC psychiatry 21 (2021) 1–12.

[44] J. A. Cecchini, A. Carriedo, J. Fernández-Río, A. Méndez-Giménez, C. González, B. Sánchez-Martínez, P. Rodríguez-González, A longitudinal study on depressive symptoms and physical activity during the spanish lockdown, International Journal of Clinical and Health Psychology 21 (2021) 100200.

[45] M. Daly, A. R. Sutin, E. Robinson, Depression reported by us adults in 2017–2018 and march and april 2020, Journal of affective disorders 278 (2021) 131–135.

[46] A. Ammar, P. Mueller, K. Trabelsi, H. Chtourou, O. Boukhris, L. Masmoudi, B. Bouaziz, M. Brach, M. Schmicker, E. Bentlage, et al., Psychological consequences of covid-19 home confinement: The eclb-covid19 multicenter study, PloS one 15 (2020) e0240204.

[47] N. Leigh-Hunt, D. Bagguley, K. Bash, V. Turner, S. Turnbull, N. Valtorta, W. Caan, An overview of systematic reviews on the public health consequences of social isolation and loneliness, Public health 152 (2017) 157–171.

[48] X. Wang, L. Cai, J. Qian, J. Peng, Social support moderates stress effects on depression, International journal of mental health systems 8 (2014) 1–5.

[49] R. T. Han, Y.-B. Kim, E.-H. Park, J. Y. Kim, C. Ryu, H. Y. Kim, J. Lee, K. Pahk, C. Shanyu, H. Kim, et al., Long-term isolation elicits depression and anxiety-related behaviors by reducing oxytocin-induced

gabaergic transmission in central amygdala, Frontiers in molecular neuroscience 11 (2018) 246.

[50] L. Pancani, M. Marinucci, N. Aureli, P. Riva, Forced social isolation and mental health: A study on 1,006 italians under covid-19 lockdown, Frontiers in Psychology 12 (2021) 1540.

[51] M. Delmastro, G. Zamariola, Depressive symptoms in response to covid-19 and lockdown: a cross-sectional study on the italian population, Scientific reports 10 (2020) 1–10.

[52] W. D. Killgore, S. A. Cloonan, E. C. Taylor, M. A. Miller, N. S. Dailey, Three months of loneliness during the covid-19 lockdown, Psychiatry Research 293 (2020) 113392.

[53] Y. Palgi, A. Shrira, L. Ring, E. Bodner, S. Avidor, Y. Bergman, S. Cohen-Fridel, S. Keisari, Y. Hoffman, The loneliness pandemic: Loneliness and other concomitants of depression, anxiety and their comorbidity during the covid-19 outbreak, Journal of affective disorders 275 (2020) 109.

[54] A. Sommerlad, L. Marston, J. Huntley, G. Livingston, G. Lewis, A. Steptoe, D. Fancourt, Social relationships and depression during the covid-19 lockdown: longitudinal analysis of the covid-19 social study, Psychological medicine (2021) 1–10.

[55] T. Probst, S. Budimir, C. Pieh, Depression in and after covid-19 lockdown in austria and the role of stress and loneliness in lockdown: a longitudinal study, Journal of Affective Disorders 277 (2020) 962.