Title

Machine-learning approaches to identify determining factors of happiness during the COVID-19 pandemic: retrospective cohort study

Journal

BMJ Open, 12(12)

ISSN

2044-6055

Authors

Osawa, Itsuki
Goto, Tadahiro
Tabuchi, Takahiro
et al.

Publication Date

2022-12-01

DOI

10.1136/bmjopen-2021-054862

Peer reviewed

BMJ Open

Machine-learning approaches to identify determining factors of happiness during the COVID-19 pandemic: retrospective cohort study

Itsuki Osawa ,[1] Tadahiro Goto ,[2,3] Takahiro Tabuchi ,[4] Hayami K Koga ,[5] Yusuke Tsugawa [6,7]

To cite: Osawa I, Goto T, Tabuchi T, *et al*. Machine-learning approaches to identify determining factors of happiness during the COVID-19 pandemic: retrospective cohort study. *BMJ Open* 2022;**12**:e054862. doi:10.1136/bmjopen-2021-054862

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-054862).

Received 02 July 2021
Accepted 29 November 2022

For numbered affiliations see end of article.

Correspondence to
Dr Itsuki Osawa;
ioosawa-tky@umin.ac.jp

ABSTRACT

Objective To investigate determining factors of happiness during the COVID-19 pandemic.
Design Observational study.
Setting Large online surveys in Japan before and during the COVID-19 pandemic.
Participants A random sample of 25 482 individuals who are representatives of the Japanese population.
Main outcome measure Self-reported happiness measured using a 10-point Likert scale, where higher scores indicated higher levels of happiness. We defined participants with ≥8 on the scale as having high levels of happiness.
Results Among the 25 482 respondents, the median score of self-reported happiness was 7 (IQR 6–8), with 11 418 (45%) reporting high levels of happiness during the pandemic. The multivariable logistic regression model showed that meaning in life, having a spouse, trust in neighbours and female gender were positively associated with happiness (eg, adjusted OR (aOR) for meaning in life 4.17; 95% CI 3.92 to 4.43; p<0.001). Conversely, self-reported poor health, anxiety about future household income, psychiatric diseases except depression and feeling isolated were negatively associated with happiness (eg, aOR for self-reported poor health 0.44; 95% CI 0.39 to 0.48; p<0.001). Using machine-learning methods, we found that meaning in life and social capital (eg, having a spouse and trust in communities) were the strongest positive determinants of happiness, whereas poor health, anxiety about future household income and feeling isolated were important negative determinants of happiness. Among 6965 subjects who responded to questionnaires both before and during the COVID-19 pandemic, there was no systemic difference in the patterns as to determinants of declined happiness during the pandemic.
Conclusion Using machine-learning methods on data from large online surveys in Japan, we found that interventions that have a positive impact on social capital as well as successful pandemic control and economic stimuli may effectively improve the population-level psychological well-being during the COVID-19 pandemic.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This is the first study that investigated the determinants of happiness during the COVID-19 pandemic using data on more than 25 000 individuals in Japan.
⇒ We used machine-learning methods to comprehensively investigate various factors that may affect people's self-reported happiness during the COVID-19 pandemic.
⇒ Our surveys using a single question based on a 10-point Likert scale could capture only some aspects of happiness, which is a multifaceted concept.
⇒ Given the exploratory design of our research, we could not elucidate causal relationships between the candidate factors and individuals' levels of happiness during the COVID-19 pandemic.
⇒ Our finding may not be generalisable to the populations in other countries or among people who are not experiencing the pandemic.

INTRODUCTION

The COVID-19 pandemic has impacted the physical, psychological and social aspects of our lives. Recent studies have reported the negative impacts of the pandemic on mental health across the world.[1–3] For example, Zacher and Rudolph[4] found that self-reported well-being of the population in Germany has decreased during the COVID-19 pandemic, and Wang and Tang[5] reported 44.8% of study participants were depressed under the initial COVID-19 pandemic in China. The deterioration of mental health during the pandemic may be attributable to multiple factors including physical and social isolation from lockdowns imposed to control the outbreaks, and economic downturns due to the pandemic.[6 7] Additionally, the psychological stress itself was likely to cause excessive preventive behaviour against the COVID-19, which could exacerbate population psychological state.[8] Given that the worsening of population mental health could result in a rise in suicide rates,[9] it is critically important for policymakers to design interventions that could effectively buffer the negative impact of the COVID-19 pandemic on the psychological well-being of the population.

However, designing interventions to maintain the well-being of the population during the pandemic is not straightforward because the general foundation of well-being is highly complicated.[10 11] According to GENIAL framework, well-being is defined as positive psychological experience, which is impacted by sociocontextual factors beyond the control of each individual,[11] and the influence of sociocontextual factors on well-being could be mediated by three major pathways: behavioural, psychological and physiological mechanisms.[12] Therefore, in this study, we focused on determinant factors of happiness (ie, psychological well-being) for designing interventions to ameliorate major societal challenges during the dramatic changes in our lifestyle due to the pandemic.

Studies conducted before the COVID-19 pandemic found that sociodemographic factors (eg, academic attainment and income) are among the strongest predictors of happiness.[13 14] However, with the dramatic alterations in our lifestyle caused by the COVID-19 outbreak, the perceptions and priorities that determine our happiness also may have changed. Although recent research has identified several factors regarding COVID-19 that could act as psychological stressors—such as longer duration of quarantines, financial loss and stigma[1 15–17]—few studies have investigated the determinants of happiness during the COVID-19 pandemic.[4 16]

To address this important knowledge gap, we investigated the factors determining happiness during the COVID-19 pandemic using machine-learning methods on data from large online surveys in Japan.

METHODS

Data source and study participants

We retrospectively analysed the data of two sequential, large, cross-sectional, self-reported questionnaire online surveys: (1) the *Japan 'COVID-19 and Society' Internet Survey* (JACSIS) study and (2) the *Japan 'Society and New Tobacco' Internet Survey* (JASTIS) study. Both surveys, conducted by a large internet research agency (Rakuten Insight, Tokyo, Japan),[18] included approximately 2.2 million panellists who are representatives of the Japanese population from the perspectives of age, gender and socioeconomic status.[19] Using the unique individual number assigned to each participant, we were able to identify the subjects who participated in both the JACSIS and JASTIS studies, providing us data from before and during the pandemic for these participants.

In the JACSIS study, the questionnaires were randomly distributed to 224 389 panellists selected by each category of gender, age and prefecture (covering all 47 prefectures in Japan) among 2.2 million panellists on 25 August 2020. The target numbers of respondents for each gender, age and prefecture category were met on 30 September 2020 (target numbers for each gender, age and prefecture category had been determined in advance based on the population distribution in 2019; 28 000 respondents; and response rate of 12.5%). We finally identified 25 482 respondents—defined as 'total respondents'—after excluding 2518 respondents with unnatural or inconsistent responses from 28 000 respondents using the original algorithm we developed. We used this cross-sectional data set, including 25 482 respondents, to investigate determinants of happiness during the COVID-19 pandemic as our primary analysis.

Similarly, 11 000 people responded to the questionnaires in the JASTIS study, which was conducted prior to the JACSIS study between 9 February and 2 March 2020 (before the COVID-19 pandemic), and we identified 7434 respondents who replied to both the JASTIS and JACSIS studies using the unique individual number assigned to each participant. Among the 7434 subjects, we defined 6965 individuals with ≥3 out of the 10-point self-reported happiness scale in the JASTIS study as 'longitudinal respondents' to investigate the determinants of declined happiness (a decrease by ≥2 on the 10-point self-reported happiness scale) during the COVID-19 pandemic. Given that the state of emergency over COVID-19 was first declared in Japan on 7 April 2020, we defined the COVID-19 pandemic period as beginning after April 2020 in this study. We used this longitudinal data set, including 6965 panellists, to investigate determinants of declined happiness during the COVID-19 pandemic.

Candidate determinants of happiness

We selected the candidate determinants of happiness from the JACSIS study data based on prior studies looking at the factors associated with well-being—a superordinate concept of happiness.[10–12] Specifically, we used respondents' demographics (ie, age,[20 21] gender,[13 22] body mass index), sociodemographic characteristics (eg, having a spouse,[13] academic attainment,[13] type of occupation,[23] reception of public assistance,[24] annual household income more than average in Japan, ≥5 000 000 yen (≥US$48 000 based on an exchange rate of 105 yen per US dollar as of August 2020)[13]), COVID-19-related sociodemographic characteristics (eg, anxiety about future household income, one-time receipt of the 100 000 yen (US$950) support allowance for the COVID-19 pandemic,[25] trust in communities, online interactions, feeling isolated during the COVID-19 pandemic[26]) and health-related or psychological characteristics (chronic conditions other than psychiatric diseases (ie, hypertension, diabetes, asthma, coronary disease, stroke, chronic obstructive pulmonary disease (COPD), cancer),[27] depression,[28] psychiatric diseases other than depression (eg, schizophrenia),[29] chronic pain,[30] meaning in life,[31 32] self-reported poor health,[33] worsening self-reported health during the pandemic). The respondents with activities of daily living (ADL) intact were defined as those who had no problems walking, washing or dressing by themselves; nor with the usual activities in the EuroQoL 5 Dimensions 5-Level questionnaire used in the JACSIS study.[34] All the candidate determinants are shown in table 1.

 Osawa I, *et al. BMJ Open* 2022;**12**:e054862. doi:10.1136/bmjopen-2021-054862

Table 1 Sociodemographic and health-related characteristics of respondents during the COVID-19 pandemic

Variables	Total respondents (n=25 482)	Longitudinal respondents (n=6965)
Age (year), median (IQR)	49 (35–64)	49 (32–61)
Female gender	12 809 (50)	2926 (42)
BMI (kg/m^2), median (IQR)	21.9 (19.8–24.2)	21.9 (19.7–24.2)
Sociodemographic characteristics		
Living alone	4997 (20)	1380 (20)
Having a spouse	15 230 (60)	3880 (56)
Having any children	8584 (34)	2179 (31)
Family caregiver	1923 (8)	513 (7)
Academic attainment, college or higher	13 195 (52)	3862 (55)
Type of occupation: office work	7498 (29)	2400 (34)
Type of occupation: sales work	3793 (15)	1005 (14)
Type of occupation: manual labour	4163 (16)	1114 (16)
Receiving public assistance	151 (0.6)	37 (0.5)
Household annual income, more than average (≥5 000 000 yen [US$48 000])	9979 (39)	2975 (43)
COVID-19-related sociodemographic characteristics		
Living in the area most affected by COVID-19	11 586 (45)	3724 (53)
The rate of changes in household income during the pandemic (%), median (IQR)	0 (–1 to 0)	0 (0–0)
Anxiety about household income in the future	5324 (21)	1114 (16)
Receipt of the support allowance (100 000 yen [US$950]) for COVID-19	23 047 (90)	6223 (89)
Interaction with neighbours	7382 (29)	1746 (25)
Trust in neighbours and communities	16 502 (65)	4543 (65)
Online interaction with family members or friends	4707 (18)	1142 (16)
Face-to-face interaction with family members or friends not living together	6468 (25)	1514 (22)
Feeling isolated during the pandemic	3662 (14)	832 (12)
Health-related or psychological characteristics		
ADL intact	3020 (12)	643 (9)
Chronic conditions other than psychiatric diseases	6360 (25)	1642 (24)
Depression	969 (4)	249 (4)
Psychiatric diseases other than depression	950 (4)	2207 (32)
Chronic pain	8926 (35)	2207 (32)
Meaning in life	15 176 (60)	4037 (58)
Self-reported poor health	3561 (14)	889 (13)
Worsening self-reported health during the pandemic	3918 (15)	1004 (14)
Self-reported happiness scale (1–10, 10 being the best), median (IQR)	7 (6–8)	7 (6–8)
Self-reported happiness scale ≥8	11 418 (45)	2849 (41)
Decrease by ≥2 of the self-reported happiness scale during the COVID-19 pandemic	–	781 (11)

Longitudinal respondents were defined as subjects with ≥3 out of the 10-point self-reported happiness scale in the JASTIS study conducted before the COVID-19 pandemic (February to March 2020) who also participated in the JACSIS study conducted during the COVID-19 pandemic (August to September 2020). Values represent n (%), unless otherwise indicated.
ADL, activities of daily living; BMI, body mass index; JACSIS, Japan 'COVID-19 and Society' Internet Survey; JASTIS, Japan 'Society and New Tobacco' Internet Survey.

COVID-19 pandemic in Japan and related policies

We defined the areas most affected by the COVID-19 pandemic as the seven prefectures (Tokyo, Kanagawa, Saitama, Chiba, Osaka, Hyogo and Fukuoka) where the state of emergency over COVID-19 was first declared on 7 April 2020. The Japanese government established the support allowance system (100 000 yen per person) under the COVID-19 outbreak in April 2020.[25]

Outcome measures (happiness)

Our primary outcome was self-reported happiness. Participants were asked to rate their levels of happiness using

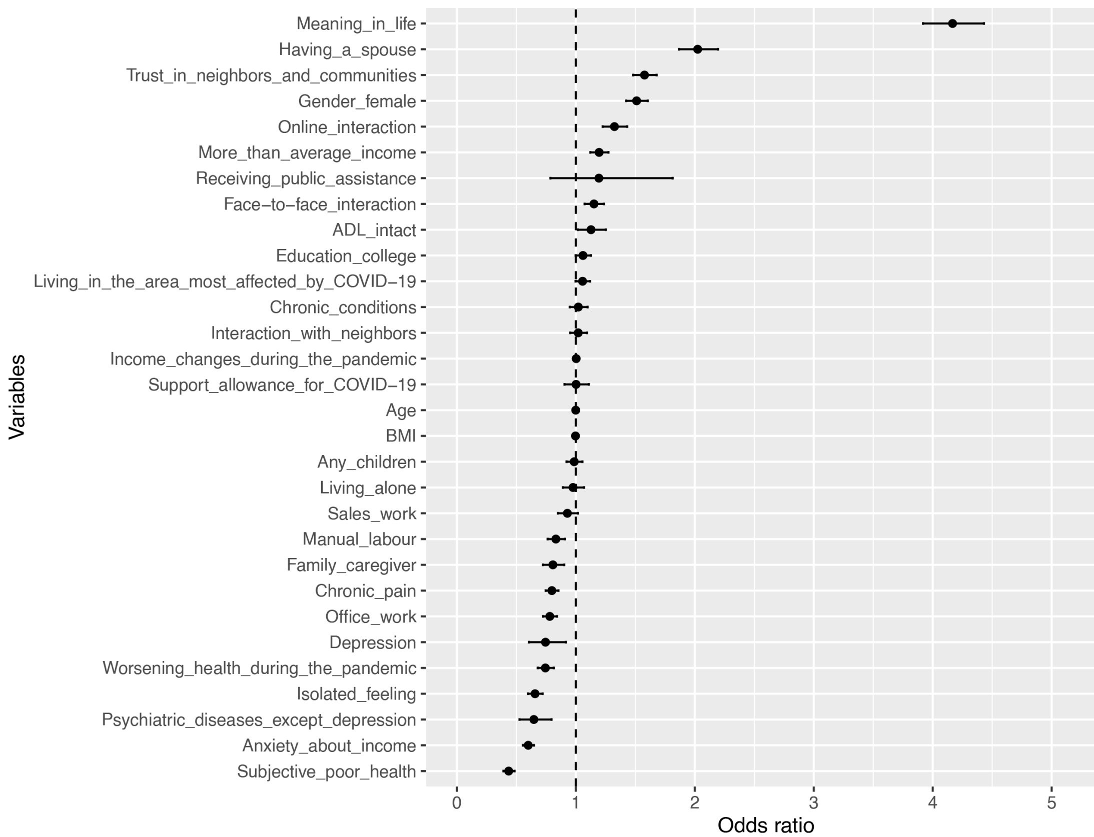

Figure 1 Association between the respondents' characteristics and happiness. We examined the association between the respondents' characteristics and happiness, defined as ≥8 on the 10-point self-reported happiness scale during the COVID-19 pandemic. Points and error bars indicate the OR and 95% CI of variables, respectively. The corresponding values are presented in online supplemental table 1. ADL, activities of daily living; BMI, body mass index.

a 10-point Likert scale, where higher scores indicated higher levels of happiness.

Statistical analysis

We first delineated the sociodemographic and health-related characteristics of respondents and showed the distribution of the self-reported happiness scale during the COVID-19 pandemic. We then explored the determinants of happiness as a binary outcome using the multivariable logistic regression model. For the logistic regression, the outcome variable was defined as high levels of happiness, defined as ≥8 on the 10-point self-reported happiness scale, with 10 being the best. To address the potential collinearity of the variables included in each model, we calculated the variance inflation factor (VIF) using the *car* package.[35 36]

Next, to explore determinants of happiness as a 10-point discrete outcome without model misspecification, we developed two machine-learning-based models (random forest and gradient-boosted decision tree) as a non-parametric approach that can accommodate non-linearities and interactions without prespecification of a particular parametric model. For machine-learning-based models, the outcome measure was defined as the 10-point happiness scale of each responder. Random forest is an ensemble of decision trees using bootstrap aggregation and random feature selection.[37 38] Gradient-boosted decision tree is an additive model of decision trees estimated by gradient descent.[38 39] To identify the best tuning hyperparameters, we used a grid search strategy using *ranger, xgboost* and *caret* packages for the random forest and gradient-boosted decision tree models.[40–42] Both 10-fold cross-validation and out-of-bag estimation techniques were used to minimise potential overfitting and improve the internal validity of our machine-learning-based models. We estimated the contribution of each candidate determinant to machine-learning-based models based on the variable importance computed by the random forest and the gradient-boosted decision tree. The variable importance is a scaled measure to have a maximum value of 100. Additionally, we calculated Shapley Additive Explanations (SHAP) values to assess the detailed contribution of each determinant factor on the regression output by using the *SHAPforxgboost* package.[43 44] Higher SHAP values represent higher possibilities that the given determinant factor contributes to higher level of happiness. If a variable is binary, the high feature value of the variable means that respondents answered 'Yes' to the question.

Sensitivity analyses

We conducted several sensitivity analyses. First, we used different thresholds for high levels of happiness for multivariable logistic regression: (1) ≥7 and (2) ≥9 on the

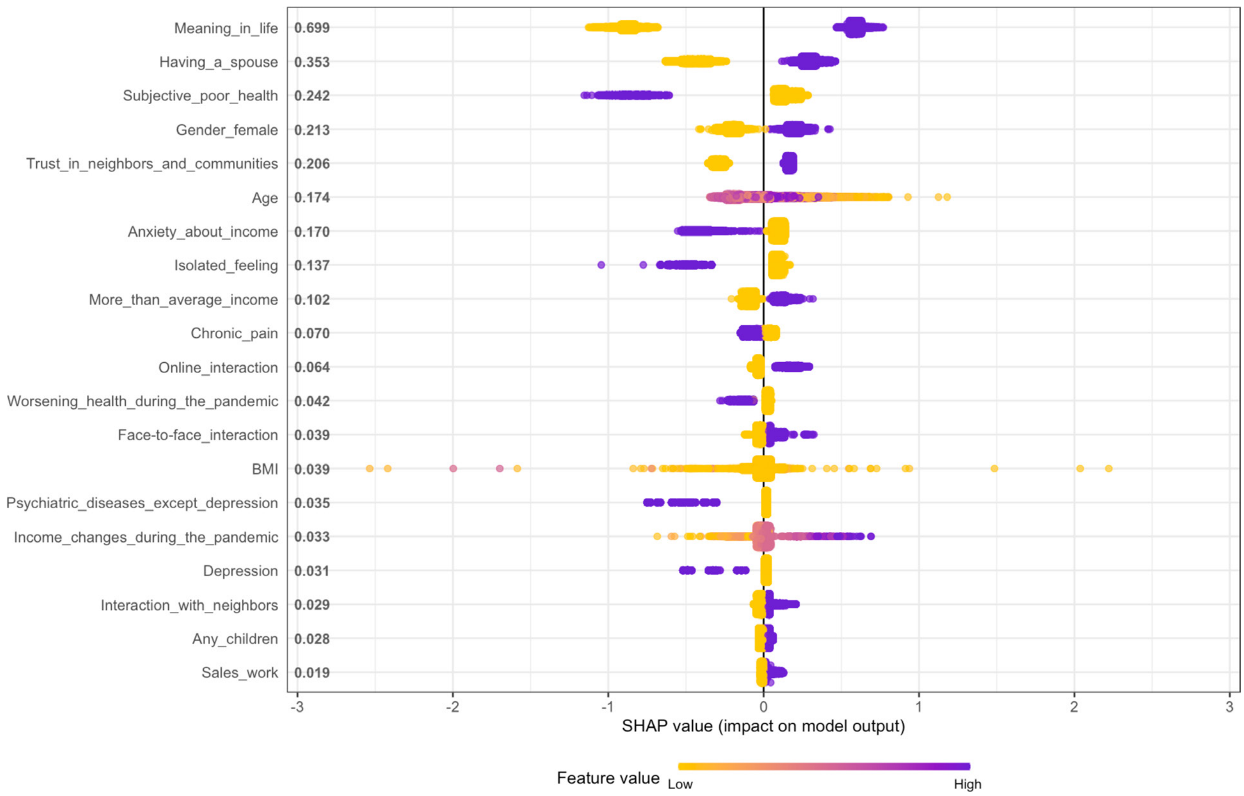

Figure 2 Shapley Additive Explanations (SHAP) summary plot of top 20 variables of the gradient-boosted decision tree model for happiness. The higher the SHAP value, the higher the possibility of the self-reported happiness scale becomes. A dot indicates each attribution of each variable at a feature value. The colour of a dot shows the absolute value of each variable (eg, purple dots represent higher feature values, but yellow ones represent lower feature values). If a variable is binary, the high feature value of the variable means that respondents answered 'Yes' to the question. For example, for self-reported poor health, we interpret that (1) while respondents with self-reported poor health were likely to have lower levels of happiness scale, those without self-reported poor health were likely to have higher levels of happiness scale, and (2) the level of the attribution of self-reported poor health to the happiness scale was larger than the one of no self-reported poor health to the happiness scale in the gradient-boosted decision tree prediction model. BMI, body mass index.

10-point scale. Second, to minimise an issue of sample selection bias (eg, younger people may be more likely to enrol in an online study), we developed a weighted multivariable logistic regression model (ie, the multivariable logistic regression model in our primary analysis with inverse probability weighting added) (see online supplemental appendix 1 for details). Third, to address the possibility that our perceptions and determinant factors of happiness may vary depending on age,[13 20 45] we described the average of the happiness scale by age and conducted multivariable logistic regression analyses stratified by age: young (15–39 years), middle-aged (40–59 years) and elderly (60–79 years) respondents. Fourth, to investigate whether the determinants of happiness differ by other characteristics of the respondents,[22 28 29] we also stratified analyses by gender and history of psychiatric diseases, respectively. Fifth, we investigated the determinants of declined happiness during the COVID-19 pandemic using the multivariable logistic regression model for longitudinal responders. A decline in happiness was defined as a decrease by ≥2 on the 10-point happiness scale during the COVID-19 pandemic. Sixth, to evaluate the association between age and declines in happiness levels, we described the levels of decline in happiness by age and stratified analyses by age.

A p value <0.05 was considered statistically significant. All analyses were performed using R version 4.0.2. (The R Foundation for Statistical Computing).[46] We created the SHAP summary plot using the *SHAPforxgboost* package,[44] other figures using the *ggplot2* package.[47]

Patient and public involvement

No patients were involved in setting the research question or the outcome measures, nor were they involved in developing plans for the design or implementation of the study. No patients were asked to advise on interpretation or writing up of results. There are no plans to disseminate the results of the research to study participants or the relevant patient community. Patient consent was not required for the study.

RESULTS

Characteristics of study participants

Among 25 482 respondents (total respondents), 12 809 (50%) were female, and the median age was 49 years. The

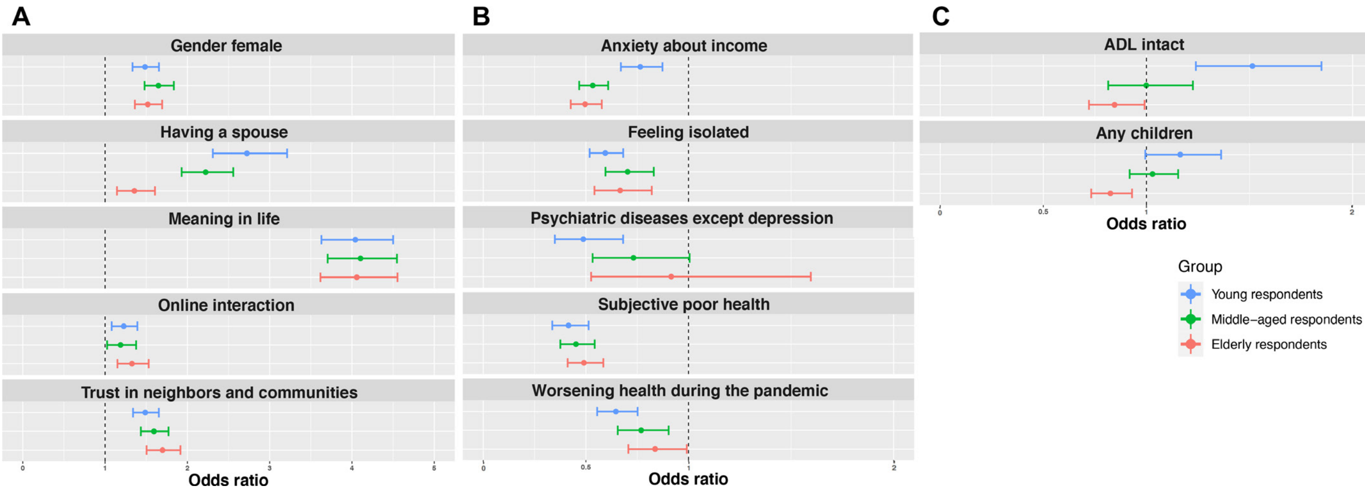

Figure 3 Differences in the ORs of the respondents' characteristics for happiness by age. (A) Top five characteristics with a high OR for happiness among total respondents. (B) Top five characteristics with a low OR for happiness among total respondents. (C) Characteristics with a different OR for happiness depending on age. We showed differences in the ORs of the respondents' characteristics for happiness by age (ie, young, middle-aged and elderly respondents). The corresponding values are presented in online supplemental table 4. ADL, activities of daily living.

median happiness score during the pandemic was 7 (IQR 6–8), and 11 418 (45%) were defined as having high levels of happiness (≥8 on the 10-point Likert scale) (table 1). Among 6965 respondents (longitudinal respondents), 781 (11%) experienced a decline in happiness levels during the pandemic. The distribution of the happiness scale during the pandemic and changes in the score from the prepandemic assessment are shown in online supplemental figure 1.

Determinants of happiness

In the multivariable logistic regression, we found that meaning in life (adjusted OR (aOR) 4.17; 95% CI 3.92 to 4.43; p<0.001), having a spouse (aOR 2.22; 95% CI 1.93 to 2.56; p<0.001), trust in neighbours and communities (aOR 1.59; 95% CI 1.43 to 1.77; p=0.02) and female gender (aOR 1.51; 95% CI 1.42 to 1.60; p<0.001) were associated with higher odds of having high levels of happiness. On the other hand, self-reported poor health (aOR 0.44; 95% CI 0.39 to 0.48; p<0.001), anxiety about household income in the future (aOR 0.60; 95% CI 0.55 to 0.65; p<0.001), psychiatric disorder diseases other than depression (aOR 0.65; 95% CI 0.53 to 0.80; p<0.001) and feeling isolated (aOR 0.66; 95% CI 0.60 to 0.72; p<0.001) were associated with lower odds of having high levels of happiness. We did not observe a statistically significant association between age and happiness (aOR 1.00; 95% CI 1.00 to 1.00; p=0.24) (figure 1 and online supplemental table 1). We found no variables with high VIF (>10) among parameters used in the multivariable logistic regression model, indicating that collinearity among the candidate determinants is not an issue for our analyses (online supplemental table 2).

In our machine-learning models (random forest and gradient-boosted decision tree), the variable importance, indicating the contribution of each candidate determinant, was quantified (online supplemental figure 2). Consistent with the results using logistic regression, meaning in life, self-reported poor health and having a spouse were the common and important determinants of happiness during the pandemic in both models. To identify the detailed contribution of each feature to the model, we depicted the SHAP summary plot of the top 20 variables of the gradient-boosted decision tree model (figure 2). This plot suggested that meaning in life, having a spouse and female gender had positive contributions to the prediction of happiness. In contrast, self-reported poor health, anxiety about household income in the future and feeling isolated had a negative impact on the prediction of happiness.

Sensitivity analyses

When using different thresholds for logistic regression to define high levels of happiness, we found similar variables as determinants of happiness (online supplemental table 1). The weighted multivariable regression model showed results that were overall consistent with our main analysis (online supplemental table 3). With regard to age as the effect modifier of happiness, we found a U-shaped relationship between the average of the self-reported happiness scale and age (online supplemental figure 3) and identified similar determinant factors of happiness among each age group (ie, young, middle-aged and elderly respondents) (figure 3, online supplemental table 4 and figure 4). However, some factors (ie, ADL intact and having any children) may have different impacts on happiness across age groups (figure 3). The results of stratified analyses indicated that observed relationships do not vary substantially by gender and history of psychiatric diseases (online supplemental table 5 and 6 and figure 5 and 6). We found poor health and feeling isolated were the important determinants of a decline in happiness during the COVID-19 pandemic, whereas meaning in life and trust in communities may calm deteriorating happiness during the pandemic (online supplemental table 7 and figure 7). Although the decline in happiness during the COVID-19 pandemic varied with age and seemed to be less serious among elderly respondents than young ones, as shown in online supplemental figure 3, there was no systemic difference in the patterns as to determinants of declined happiness during the pandemic after the stratification by age (online supplemental table 8 and figure 8).

DISCUSSION

Using the machine-learning methods and data collected through large-scale, nationwide surveys conducted before and during the COVID-19 pandemic, we found that meaning in life and social capital (eg, having a spouse and trust in communities) are the strongest positive determinants of happiness during the pandemic. On the other hand, poor health, anxiety about future household income and feeling isolated are the key factors negatively associated with happiness. These associations were consistent across different groups of age, gender and history of psychiatric diseases. Although poor health, anxiety about future household income and isolation were determinants of the decline in happiness during the pandemic, meaning in life and social capital may be able to alleviate deteriorating happiness during the pandemic. These findings should be informative for policymakers to design policies that can minimise the negative impact of the COVID-19 pandemic on the psychological well-being of the population.

Our findings indicated that uncertainty about the future due to the COVID-19 pandemic could be a key cause of low levels of happiness. A recent study found that elderly respondents—who are likely to have greater mindfulness and present-moment perspectives—seemed to have experienced less negative impact on psychological well-being during the pandemic compared with younger people,[21] indicating that focusing on the present moment rather than excessive consideration of the future may retain psychological well-being in this unprecedented crisis. Paying attention to the present moment is an essential part of mindfulness, which has the potential to help people live happier as well as reduce mental health problems.[48] Accordingly, to cope with future uncertainty from the psychological perspective, meaning in life could be a key factor of happiness during the time when uncertainties about the future are high. Indeed, the positive association between meaning in life and positive psychological well-being has long been known and has encouraged psychiatrists to take salutogenic approaches for their patients' mental health—focusing on factors related to happiness, instead of factors directly causing diseases.[31] As one of the salutogenic approaches, interventions that give people an opportunity to reflect on their life's meaning, such as life-crafting interventions, could be a potential solution for improving happiness in such an unusual crisis with future uncertainties.[49] On top of that, providing counselling on the basis of second wave positive psychology—confronting, accommodating or transforming negative emotions and finding another deep joy in such an unprecedented nightmare based on the reality that life itself is full of evil and suffering—was recently proposed to achieve sustainable well-being for individuals from psychological perspectives.[50]

Social capital was also identified as a strong determinant of happiness during the COVID-19 pandemic in our study. Given that physical isolation and lockdowns imposed to control the COVID-19 outbreak in many regions of the world could decrease social interactions and increase psychological isolations, we need urgent measures to reinforce social interactions without increasing the risk of infections (eg, social interactions through online networks, social interactions in outdoor settings while maintaining social distancing and wearing masks). A study conducted before the pandemic reported that improving components of individual social capital (eg, trust in communities and social interaction) could enhance both individual well-being and community social capital, which would subsequently create a virtuous circle that improves individual social capital.[51] Moreover, although poor health status was negatively associated with happiness during the pandemic in our analyses, prior studies have found that higher levels of social capital are associated with improved health,[52] and potentially lead to a higher uptake of preventive measures for infections (eg, wearing masks and washing hands) and social distancing that could reduce the transmission of the virus during the COVID-19 pandemic.[53] Indeed, a study by Ye *et al* reported people who shared their activities with others or interacted with friends were more likely to take preventive measures for the COVID-19.[54] However, as misinformation regarding the COVID-19 by non-experts through online networks can be disseminated and sometimes affect negatively the population mental health,[55] the reliance only on social networking service for social interactions may be a double-edged sword in such a special time.

Therefore, while it is obviously important to introduce psychological counselling (eg, second wave positive psychology) for each individual, interventions for improving social capital as well as infection control and economic stimulus policies during the pandemic may effectively have positive effects on happiness. Of note, to improve population-level psychological well-being, it is important for policymakers to promote people's voluntary and collective activities, ideally by developing asset-based communities and schemes, rather than by directly enforcing social capital increases.[56]

Limitations

Our study has several limitations. First, we used the machine-learning methods to comprehensively investigate various factors that may affect people's self-reported well-being during the COVID-19 pandemic, and our study was not designed to evaluate causal relationships. Therefore, it remains unclear whether improving important determining factors identified in this study could actually lead to better well-being. Second, it is possible that our surveys using a single question based on a 10-point Likert scale captured only some aspects of happiness, psychological well-being which is a multifaceted concept.[10–12] However, as we considered various factors, including indicators of material well-being, eudaimonia and ill-being as our candidate determinants of well-being during the COVID-19 pandemic,[11] we believed the strength of our study was exhaustive analyses of determinants of happiness based on multiple domains. Third, given that we used data collected through internet surveys conducted in Japan, it is possible that respondents of the surveys may differ from the general population in meaningful ways (there is a risk of sample selection bias). However, the results of the sensitivity analysis using weighted analyses accounting for the probability of responding to the surveys were not qualitatively different from our main analyses (unweighted analyses), supporting the robustness of our analyses. Fourth, our study was conducted in Japan, and therefore our findings may not be generalised to populations in other countries. Lastly, our study was not pre-registered in an authorised clinical trial registry (eg, UMIN-CTR), but our study protocol was approved by an institutional review board, which did not require us to pre-register our study in such an authorised clinical trial registry beforehand.

CONCLUSIONS

Using nationwide online surveys before and during the COVID-19 pandemic in Japan, we identified that meaning in life and social capital components were the strongest positive determinants of happiness. On the other hand, poor health, anxiety about future household income and feeling isolated were key factors that negatively impacted happiness. These findings should be informative for policymakers in designing policies and interventions that could effectively improve the psychological well-being of the population during the COVID-19 pandemic.

Author affiliations
[1]Department of Emergency and Critical Care Medicine, The University of Tokyo Hospital, Tokyo, Japan, Bunkyo, Tokyo, Japan
[2]Department of Clinical Epidemiology and Health Economics, School of Public Health, The University of Tokyo, Bunkyo, Tokyo, Japan
[3]TXP Medical Co. Ltd, Chiyoda, Tokyo, Japan
[4]Cancer Control Center, Osaka International Cancer Institute, Osaka, Osaka, Japan
[5]Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA
[6]Division of General Internal Medicine and Health Service Research, David Geffen School of Medicine at UCLA, Los Angeles, California, USA
[7]Department of Health Policy and Management, UCLA Fielding School of Public Health, Los Angeles, California, USA

Twitter Itsuki Osawa @Osawa_Itsuki, Tadahiro Goto @GtoDr, Takahiro Tabuchi @TakahiroTabuchi, Hayami K Koga @KogaHayami and Yusuke Tsugawa @Yusuke_Tsugawa

Contributors IO and TT had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. IO, TG, TT and YT conceived and designed the study. IO, TG, HKK, TT and YT interpreted the data, critically revised the manuscript for important intellectual content and approved the final manuscript. IO, TG and YT drafted the initial manuscript. YT supervised the study. IO, TG and TT are the guarantors of the study.

Funding This study was funded by the Japan Society for the Promotion of Science (JSPS) KAKENHI Grants (grant numbers: 17H03589, 19K10671, 19K10446, 18H03107, 18H03062), the JSPS Grant-in-Aid for Young Scientists (grant number: 19K19439), Research Support Program to Apply the Wisdom of the University to Tackle COVID-19 Related Emergency Problems, University of Tsukuba, and Health Labour Sciences Research Grant (grant numbers: 19FA1005, 19FG2001). YT was supported by the National Institutes of Health (NIH)/NIMHD Grant (R01MD013913) and NIH/NIA Grant (R01AG068633) for other work not related to this study.

Disclaimer The findings and conclusions of this article are the sole responsibility of the authors and do not represent the official views of the funders/sponsors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval The study protocol was approved by the Institutional Review Board of the Osaka International Cancer Institute (No 20084).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available. No additional data are available.

ORCID iDs
Itsuki Osawa http://orcid.org/0000-0001-7790-9130
Tadahiro Goto http://orcid.org/0000-0002-5880-2968
Takahiro Tabuchi http://orcid.org/0000-0002-1050-3125
Hayami K Koga http://orcid.org/0000-0002-9716-2502
Yusuke Tsugawa http://orcid.org/0000-0002-1937-4833

REFERENCES

1 Pierce M, Hope H, Ford T, *et al*. Mental health before and during the COVID-19 pandemic: a longitudinal probability sample survey of the UK population. *Lancet Psychiatry* 2020;7:883–92.
2 Ettman CK, Abdalla SM, Cohen GH, *et al*. Prevalence of depression symptoms in US adults before and during the COVID-19 pandemic. *JAMA Netw Open* 2020;3:e2019686.
3 Kikuchi H, Machida M, Nakamura I, *et al*. Changes in psychological distress during the COVID-19 pandemic in Japan: a longitudinal study. *J Epidemiol* 2020;30:522–8.
4 Zacher H, Rudolph CW. Individual differences and changes in subjective wellbeing during the early stages of the COVID-19 pandemic. *Am Psychol* 2021;76:50–62.
5 Wang G-Y, Tang S-F. Perceived psychosocial health and its sociodemographic correlates in times of the COVID-19 pandemic: a community-based online study in China. *Infect Dis Poverty* 2020;9:148.
6 Singh S, Roy D, Sinha K, *et al*. Impact of COVID-19 and lockdown on mental health of children and adolescents: a narrative review with recommendations. *Psychiatry Res* 2020;293:113429.
7 Fiorillo A, Sampogna G, Giallonardo V, *et al*. Effects of the lockdown on the mental health of the general population during the COVID-19 pandemic in Italy: results from the comet collaborative network. *Eur Psychiatry* 2020;63:e87.

8 Ye Y, Wang R, Feng D, *et al*. The recommended and excessive preventive behaviors during the COVID-19 pandemic: a community-based online survey in China. *Int J Environ Res Public Health* 2020;17:6953.
9 Tanaka T, Okamoto S. Increase in suicide following an initial decline during the COVID-19 pandemic in Japan. *Nat Hum Behav* 2021;5:229–38.
10 Huppert FA. Psychological well-being: evidence regarding its causes and consequences. *Applied Psychology: Health and Well-Being* 2009;1:137–64.
11 Mead J, Fisher Z, Kemp AH. Moving beyond disciplinary silos towards a transdisciplinary model of wellbeing: an invited review. *Front Psychol* 2021;12:642093.
12 Kemp AH, Arias JA, Fisher Z. Social Ties, Health and Wellbeing: A Literature Review and Model. In: *Neuroscience and social science*, 2017: 397–427.
13 Oskrochi G, Bani-Mustafa A, Oskrochi Y. Factors affecting psychological well-being: evidence from two nationally representative surveys. *PLoS One* 2018;13:e0198638.
14 Zhang N, Liu C, Chen Z, *et al*. Prediction of adolescent subjective well-being: a machine learning approach. *Gen Psychiatr* 2019;32:e100096.
15 Brooks SK, Webster RK, Smith LE, *et al*. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *Lancet* 2020;395:912–20.
16 Wang C, Pan R, Wan X, *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* 2020;17. doi:10.3390/ijerph17051729. [Epub ahead of print: 06 Mar 2020].
17 Jia R, Ayling K, Chalder T, *et al*. Mental health in the UK during the COVID-19 pandemic: cross-sectional analyses from a community cohort study. *BMJ Open* 2020;10:e040620.
18 Rakuten Insight, Inc. About us: Rakuten insight, 2020. Available: https://insight.rakuten.co.jp/en/aboutus.html
19 Tabuchi T, Gallus S, Shinozaki T, *et al*. Heat-not-burn tobacco product use in Japan: its prevalence, predictors and perceived symptoms from exposure to secondhand heat-not-burn tobacco aerosol. *Tob Control* 2018;27:e25–33.
20 Ramsey MA, Gentzler AL. Age differences in subjective well-being across adulthood: the roles of savoring and future time perspective. *Int J Aging Hum Dev* 2014;78:3–22.
21 Wilson JM, Lee J, Shook NJ. COVID-19 worries and mental health: the moderating effect of age. *Aging Ment Health* 2021;25:1–8.
22 Batz C, Tay L. Gender differences in self-reported well-being. In: Diener E, Oishi S, Tay L, eds. *Handbook of well-being*. Salt Lake City: DEF Publishers, 2018. nobascholar.com
23 Hofmann J, Gander F, Ruch W. Exploring differences in well-being across occupation type and skill. *Transl Issues Psychol Sci* 2018;4:290–303.
24 Ministry of Health, Labor and welfare. Public assistance system, 2011. Available: https://www.mhlw.go.jp/english/wp/wp-hw5/dl/23010809e.pdf
25 Ministry of International Affairs and Communications. Guide to special cash payments, 2020. Available: https://www.soumu.go.jp/main_content/000715668.pdf
26 Lim LL, Kua E-H, alone L. Living alone, loneliness, and psychological well-being of older persons in Singapore. *Curr Gerontol Geriatr Res* 2011;2011:1–9.
27 Megari K. Quality of life in chronic disease patients. *Health Psychol Res* 2013;1:27.
28 Malone C, Wachholtz A. The relationship of anxiety and depression to subjective well-being in a mainland Chinese sample. *J Relig Health* 2018;57:266–78.
29 Stanga V, Turrina C, Valsecchi P, *et al*. Well-being in patients with schizophrenia, mood and personality disorders attending psychiatric services in the community. A controlled study. *Compr Psychiatry* 2019;91:1–5.
30 Gureje O, Von Korff M, Simon GE, *et al*. Persistent pain and well-being: a world Health organization study in primary care. *JAMA* 1998;280:147–51.
31 Zika S, Chamberlain K. On the relation between meaning in life and psychological well-being. *Br J Psychol* 1992;83:133–45.
32 Dezutter J, Casalin S, Wachholtz A, *et al*. Meaning in life: an important factor for the psychological well-being of chronically ill patients? *Rehabil Psychol* 2013;58:334–41.
33 Ngamaba KH, Panagioti M, Armitage CJ. How strongly related are health status and subjective well-being? systematic review and meta-analysis. *Eur J Public Health* 2017;27:879–85.
34 Shiroiwa T, Fukuda T, Ikeda S, *et al*. Japanese population norms for preference-based measures: EQ-5D-3L, EQ-5D-5L, and SF-6D. *Qual Life Res* 2016;25:707–19.
35 Austin PC, Steyerberg EW, variable Eper. Events per variable (EPV) and the relative performance of different strategies for estimating the out-of-sample validity of logistic regression models. *Stat Methods Med Res* 2017;26:796–808.
36 Car: companion to applied regression, 2019. Available: https://cran.r-project.org/web/packages/car/index.html
37 Breiman L. Random forests. *Mach Learn* 2001;45:5–32.
38 Badillo S, Banfai B, Birzele F, *et al*. An introduction to machine learning. *Clin Pharmacol Ther* 2020;107:871–85.
39 Natekin A, Knoll A. Gradient boosting machines, a tutorial. *Front Neurorobot* 2013;7:21.
40 Ranger: a fast implementation of random forests, 2020. Available: https://cran.r-project.org/web/packages/ranger/index.html
41 xgboost: extreme gradient boosting, 2021. Available: https://cran.r-project.org/web/packages/xgboost/index.html
42 Caret: classification and regression training, 2020. Available: https://cran.r-project.org/web/packages/caret/index.html
43 Lundberg S, Lee S-I. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst*, 2017:4768–77.
44 SHAPforxgboost: SHAP Plots for 'XGBoost', 2021. Available: https://cran.r-project.org/web/packages/SHAPforxgboost/index.html
45 Karasawa M, Curhan KB, Markus HR, *et al*. Cultural perspectives on aging and well-being: a comparison of Japan and the United States. *Int J Aging Hum Dev* 2011;73:73–98.
46 R Core Team. R: a language and environment for statistical computing, 2021. Available: https://www.r-project.org
47 ggplot2: create elegant data Visualisations using the grammar of graphics, 2021. Available: https://cran.r-project.org/web/packages/ggplot2/index.html
48 Keng S-L, Smoski MJ, Robins CJ. Effects of mindfulness on psychological health: a review of empirical studies. *Clin Psychol Rev* 2011;31:1041–56.
49 de Jong EM, Ziegler N, Schippers MC. From shattered goals to meaning in life: life crafting in times of the COVID-19 pandemic. *Front Psychol* 2020;11:577708.
50 Wong PTP. Second wave positive psychology's (PP 2.0) contribution to counselling psychology. *Couns Psychol Q* 2019;32:275–84.
51 Rocco L, Suhrcke M. Is social capital good for health? a European perspective. Copenhagen, who regional office for Europe; 2012.
52 Villalonga-Olives E, Wind TR, Kawachi I. Social capital interventions in public health: a systematic review. *Soc Sci Med* 2018;212:203–18.
53 Makridis CA, Wu C. How social capital helps communities weather the COVID-19 pandemic. *PLoS One* 2021;16:e0245135.
54 Ye Y, Wu R, Ge Y, *et al*. Preventive behaviours and family inequalities during the COVID-19 pandemic: a cross-sectional study in China. *Infect Dis Poverty* 2021;10:100.
55 Zhou J, Ghose B, Wang R, *et al*. Health perceptions and misconceptions regarding COVID-19 in China: online survey study. *J Med Internet Res* 2020;22:e21099.
56 Fabian M, Pykett J. Be happy: navigating normative issues in behavioral and well-being public policy. *Perspect Psychol Sci* 2022;17:169–82.