International Journal of Mental Health and Addiction
https://doi.org/10.1007/s11469-021-00587-y

ORIGINAL ARTICLE

Mental Health During the COVID-19 Pandemic in Japan: Applying Topic Modeling in Daily Life Descriptions

Yuta Chishima[1] · I-Ting Huai-Ching Liu[2]

Accepted: 31 May 2021

Abstract

The novel coronavirus disease pandemic is threatening not only physical but also mental health. Although some recent quantitative studies have been conducted and revealed the influence of the pandemic on mental health and its relevant factors, it is impossible to obtain and explore all possible variables strongly related to mental health. Therefore, we attempted to adopt a bottom-up approach using text mining of participants' narratives. We examined how participants' descriptions of daily life during the pandemic were categorized into various topics, and which topics were related to their mental health in a sample of 776 Japanese citizens in the general population over 18 years old. Results of a topic modeling with 2,594 unique words provided nine topics (mask, physical symptoms, children, infection anxiety, disinfection items, economic influence, remote work, going out, and change of lifestyle). Those who wrote about economic influence, physical symptoms, and disinfection items experienced lower life satisfaction and higher depression and negative affect, whereas those who mentioned their children were likely to have higher life satisfaction. This study highlighted that monitoring the mental health of individuals with economic impacts and physical symptoms may reduce the damage of COVID-19.

Keyword COVID-19; Depression; Life satisfaction; Mental health; Text mining; Japan

A novel coronavirus (COVID-19) has spread worldwide, infecting over twenty million people and killing more than seven hundred thousand as of August 10, 2020. The pandemic has drastically impacted ways of life to avoid being infected, infecting loved ones, or spreading the virus in communities. These experiences have been severely damaging to not only physical but also mental health. According to a survey conducted between April 24 and May 4 in the USA, the stress level for American adults this year (5.4 out of 10) was significantly higher than the average stress level reported in the 2019 survey (4.9), marking the first significant increase in average reported stress since the survey began in 2007 (American Psychological Association, 2020). Moreover, Luo et al. (2020) conducted meta-analyses using 62 studies with 162,639 participants from 17 countries and demonstrated

✉ Yuta Chishima
chishima.yuta.fw@u.tsukuba.ac.jp

1 University of Tsukuba, Tsukuba, Ibaraki 305-8572, Japan

2 Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan

that the prevalence of anxiety and depression among the general population was 32% and 27%, respectively. Also, the prevalence of those symptoms was very high among patients with pre-existing clinical conditions (56%) and COVID-19 infection (55%).

Quantitative Surveys About Factors of Mental Health

Researchers need to monitor and report the rates of mental health and then determine the mechanisms that explain it (Holmes et al., 2020). Indeed, many researchers conducted quantitative surveys to explore the mechanism of poor mental health within general populations. Luo et al. (2020) reviewed the existing literature and summarized that common risk factors included being a woman or a nurse, having low socioeconomic status or a high risk of COVID-19 infection, social isolation, and watching COVID-19-related news over long periods. Protective factors included sufficient medical resources, up-to-date and accurate information, and precautionary measures.

Since the publication of the review (Luo et al., 2020), other studies have replicated and extended the results. Poor mental health outcomes were associated with females (Smith et al., 2020; Varshney et al., 2020; Verma & Mishra, 2020), younger ages (de Bruin, 2020; Smith et al., 2020; Varshney et al., 2020), stay-at-home orders (Tull et al., 2020), the perceived negative impact of the pandemic on livelihood (Guo et al., 2020), worry about infection (Choi et al., 2020; Satici et al., 2020), not having enough surgical masks (Choi et al., 2020), the inability to work from home (Choi et al., 2020), low annual income (Smith et al., 2020), unemployment (Verma & Mishra, 2020), smoking (Smith et al., 2020), binge drinking (Verma & Mishra, 2020), chronic physical disease (Gorrochategi et al., 2020; Smith et al., 2020; Varshney et al., 2020), and intense exposure through the media (Chao et al., 2020; Guo et al., 2020).

In Japan, a state of emergency was declared on April 7, 2020 due to the spread of COVID-19. People were requested to stay at home as much as possible and limit social contact. Shigemura and Kurosawa (2020) reported that Japanese citizens experienced increased anxiety; faced serious shortages of personal protective equipment, surgical masks, and other disinfection items; and were skeptical about the effectiveness of the response of the national government to the economy and health care. Indeed, some surveys in Japan have shown that the pandemic has caused depression and anxiety. A weekly online survey (Macromill, 2020) on April 15, 2020 showed that 48% of the respondents felt depressed, showing a 15% increase compared to 1 month ago, and a 29% increase compared to 1 year ago. A longitudinal survey from February 25–27, 2020, to April 1–6, 2020 (Kikuchi et al., 2020) also demonstrated a significant increase of distress within 1 month. Additionally, according to a survey ($N = 11{,}333$), comparing people's distress in May 2020 with past surveys in 2010, 2013, 2016, and 2019 (Yamamoto et al., 2020), the proportion of highly distressed people in 2020 was much higher than before, with a more than 20% increase compared to previous years (e.g., 48.1% in 2020 vs. 26.9% in 2019). In terms of factors influencing mental health, similar factors as in other countries were confirmed: younger ages (Ueda et al., 2020; Yamamoto et al., 2020), decreased income (Kikuchi et al., 2020; Sugawara et al., 2021; Ueda et al., 2020; Yamamoto et al., 2020), being unemployed (Ueda et al., 2020), medical histories of mental illness (Yamamoto et al., 2020), COVID-19-related anxiety (Sugawara et al., 2021; Yamamoto et al., 2020), longer time spending outside (Sugawara et al., 2021), and poor interpersonal relationships (Yamamoto et al., 2020).

Qualitative Approach to Obtain Social Reactions to the Pandemic

Although some quantitative surveys have been conducted, it is impossible to design surveys with all possible variables related to mental health. Teti et al. (2020) suggest that qualitative inquiries are the best method for capturing social responses to the pandemic, since participants' natural responses from open-ended questions are well suited to identifying or focusing on vulnerable populations. Analyzing qualitative data could also be useful to reveal unexpected factors of mental health. Indeed, the text mining method has been adopted in some research to explore social responses to the COVID-19 pandemic (Abd-Alrazaq et al., 2020; Han et al., 2020; Kleinberg et al., 2020; Stokes et al., 2020). For example, Abd-Alrazaq et al. (2020) analyzed 167,073 tweets in English about COVID-19 between February and March 2020 and identified 12 topics. The topic with the highest likes for the tweets was economic loss. Furthermore, Kleinberg et al. (2020) extracted 20 topics about participants' worry about COVID-19 using 2,500 texts and showed that "following the rules related to the lockdown" and "worries about employment and the economy" were prevalent. Although these texts were used to approximate the corresponding emotional responses (e.g., participants' emotional responses to "worry" were partially predicted by their texts about "worry"), how general texts about COVID-19 are related to mental health has not yet been explored. Therefore, this research explored the relation between general descriptions from an open-ended question and mental health status during the COVID-19 pandemic, with the hope of detecting potential factors leading to mental illness.

Current Study

To overcome the limitation of qualitative and quantitative methodology and extend the existing results, we adopted a mixed approach (Creswell & Clark, 2017) using qualitative (description texts) and quantitative data (Likert-type scales of mental health) in the current study. This study aimed to examine how individuals' descriptions of daily life during the pandemic were categorized and to clarify which topics they wrote were related to mental health. The mental health variables included life satisfaction, depression, and some aspects of negative affect. Although some quantitative surveys about the relationship between COVID-19 and mental health have been conducted (see Luo et al., 2020), it is still difficult to generate any specific hypotheses based on text data because the topics that are extracted rely completely on what participants write about their daily life. Therefore, instead of making specific hypotheses, we consider the current study as an exploratory bottom-up research to comprehend phenomena and generate hypotheses for future studies.

Method

Participants

Participants were recruited using two Japanese crowdsourcing platforms. All participants live in Japan. We excluded five participants who did not follow the instructions. The final sample consisted of 776 participants.

Procedures

This study was approved by the research ethics committee of Kyoto University (Ref: 30-P-24). Parts of the data came from previous research samples (Chishima et al., in press). The participants consented to participation in the study before responding to the questionnaires.

Participants were recruited by Japanese cloud sourcing platforms "Lancers (https://www.lancers.jp)" and "CloudWorks (https://crowdworks.jp)." The workers in the platforms can find our survey when we post the recruitment on the websites. When the workers finished responding to the survey, we checked their responses and paid 300 JPY. The online anonymous data collection was conducted from April 13 to 15, 2020, approximately 1 week after the emergency declaration in Japan. Participants were first asked to answer demographic and COVID-19-related questions.

Daily Life Descriptions

Then, they were asked to write about their current daily life during the pandemic. The instruction was, "how does the spreading of the novel coronavirus disease influence your daily life? Please write about your current daily life in detail." To make it easier for them to start writing, we provided some example topics as follows: "For example, your health condition, your daily work/commuting, your anxiety/concern for COVID-19, your preventive behaviors for COVID-19, your family/partner/friends, your neighborhood/residential area, and the daily necessities (e.g., food, toilet paper, or mask)." The average length of the answers was 184.53 Japanese characters ($SD = 126.54$, range = 9–1019). Lastly, participants completed questions about life satisfaction, depression, and negative affect.

Measures

Life Satisfaction

Life satisfaction was assessed using a scale developed by Prenda and Lachman (2001). The Japanese version was used in the national survey called Midlife in Japan. Participants were asked, "How would you rate your daily life overall during the past week?" They responded to the question using using an 11-point Likert scale ranging from 0 ("worst") to 10 ("best").

Depression

Depression was assessed using the Kessler Psychological Distress Scale (K6; Kessler et al., 2003). The scale asks participants how frequently they have experienced each of the six symptoms of major depression and generalized anxiety disorder during the past month. We changed this time interval to the past "week" to focus on the situation after the emergency declaration in Japan. Items include "How often did you feel hopeless?" and "How often did you feel worthless?" Items were answered on a 5-point Likert scale ranging from 1 ("none of the time") to 5 ("all of the time"). Evidence in support of the reliability and validity of the K6 was found in an American sample (Kessler et al., 2003) and a Japanese sample (Furukawa et al., 2008). The alpha coefficient in the current study was 0.90.

Negative Affect

We assessed negative affect using the Emotion and Arousal Checklist (EACL) developed by Oda et al. (2015). We used four subscales to assess negative affect (fear, anger, sadness, and disgust), with four items each. They were rated on a 5-point Likert scale, ranging from 0 ("strongly disagree") to 3 ("strongly agree"). Participants were asked to what extent they currently felt each feeling. Items for fear were "fearful," "frightened," "scared," and "anxious." Items for anger were "angry," "irritated," "grumpy," and "edgy." Items for sadness were "sad," "gloomy," "depressed," and "overwhelmed." Items for disgust were "disgusting," "uncomfortable," "unpleasant," and "bored." Evidence in support of the reliability and validity of the EACL was found in Japanese samples (Oda et al., 2015). The internal consistency reliabilities of the negative affect in the current study were $\alpha = 0.92$ for fear, $\alpha = 0.93$ for anger, $\alpha = 0.92$ for sadness, and $\alpha = 0.92$ for disgust. Evidence in support of the reliability and validity of the EACL was found in Japanese samples (Oda et al., 2015).

Analyses

Word Tokenization

We used an open-source text segmentation library for the Japanese language called "MeCab" (Kudo, 2006), which has been commonly used for the tokenization of the Japanese language. We selected only nouns so we could get understandable topics when running topic modeling. In addition, some stop words including special symbols and numbers were removed. Also, words about COVID-19 such as "COVID," "novel," and "coronavirus" were removed because most participants reported them, which could cause interference with the results in categorizing the words into some topics. As a final step, we reviewed all words and deleted or modified some problematic or uncertain words. The final number of words was 2,594, which appeared 15,920 times in total.

Association Between Words and Dependent Variables

Out of the 2594 words, we used only words that have appeared at least 10 times to explore the overall relationships between the obtained words and dependent variables. This left us with 287 words. We calculated the mean score of the dependent variables of participants who used those 287 words, regardless of the frequency they used each word.

Topic Modeling

To explore how many topics the descriptions can be classified into, and how those topics relate to mental health, we first adopted structural topic modeling (Roberts et al., 2016) using the R package "stm" (Roberts et al., 2019), which is a method for topic modeling with document-level covariate information (mental health in the current study). To determine the optimal number of topics, we estimated the model fittings using the *searchK*-function in the "stm" package. The function provided some statistical values (e.g., held-out likelihood and semantic coherence) for choosing among the different number of topics (Taddy, 2013; Wallach et al., 2009). The held-out likelihood is the probability of words appearing within a document when these words have been excluded in the estimation step. Some proportion of the data is held out for estimation, and it is used for validation later. Semantic coherence is a criterion that is maximized when the most probable words in a given topic frequently co-occur together (Mimno et al., 2011). Thus, the higher both values are, the better the model is.

Correlation Analyses

Next, we ran correlation analyses with demographic and COVID-19 related questions to provide evidence proving that the classification and naming of topics were reasonable. Each of the 9 topics has a different probability to appear in each participant's description (776 in total). This is called "topic probability." The topic probability indicates how likely each topic is included in a participant's description. Because all nouns in participants' descriptions fall under the 9 topics, the total of the 9 topic probabilities in each participant's description is equal to 1 (=100%). The mean of topic probability indicates the mean score of each topic probability among 776 participants.

Regression Analyses

We conducted a series of multiple regression analyses using a stepwise method. The topic probabilities were included as independent variables, while life satisfaction, depression, and negative affect (fear, anger, sadness, and disgust) were included as dependent variables.

Results

Descriptive Statistics

The mean age of participants was 38.92 years ($SD = 10.42$, range = 18–76), and 54.3% were women. Participants' socioeconomic status was assessed by asking what they think their standard of living is compared to other people in Japan, rating from 1 ("poorest") to 10 ("richest"). The mean was 4.43 ($SD = 1.52$). Participants' employment status was as follows: "student" = 4.6%, "full-time employee" = 38.7%, "part-time employee" = 14.4%, "housewife/househusband" = 17.0%, "unemployed" = 12.4%, and "other" = 12.9%. Participants whose working style has partially/completely changed into remote working from home were 22.6%. Those who have one or more children were 29.9%. Only three participants had been tested for COVID-19, and the reported results were negative for all. Those who had any physical symptoms were 17.8%. The descriptive statistics and correlations among the variables are presented in Supplemental Table 1.

Association Between Words and Dependent Variables

The words with the highest scores for each dependent variable are displayed in Table 1. The frequency refers to how many times the word is mentioned in participants' descriptions. The mean indicates the average score of each dependent variable (e.g., life satisfaction) among those who included the word in their description. For example, the word "married couple" was included in the description 11 times in total, and the average score

Table 1 Words with the highest means in each dependent variable

Life Satisfaction (0-10)			Depression (1-5)			Fear (0-3)		
Top 20 words	Frequency	*M*	Top 20 words	Frequency	*M*	Top 20 words	Frequency	*M*
Married couple	11	6.44	Depression	12	3.39	Depression	12	2.32
Stock	13	5.85	Country	19	3.25	Throat	16	2.09
Prevention	11	5.82	This month	10	3.18	This month	10	2.05
Everything	11	5.73	Disease	12	2.92	Salary	12	2.05
Traffic	13	5.58	Test	11	2.90	Cough	36	2.03
School closure	34	5.57	Economy	32	2.86	Abnormality	11	2.00
Housewife	12	5.50	Others	12	2.85	Family	10	1.97
Custom	12	5.50	Cough	36	2.84	Mother	16	1.95
Full-time	10	5.50	Abnormality	11	2.83	Industry	21	1.95
Children	142	5.44	Out of stock	23	2.82	Country	19	1.94
Online	18	5.40	Check	14	2.79	Mood	37	1.91
Tissue	21	5.40	Place	20	2.79	Test	11	1.91
School	43	5.37	Industry	21	2.79	Disease	12	1.88
Cancellation	11	5.36	Contract	11	2.79	Symptoms	35	1.88
Phone	16	5.31	Mood	37	2.79	Part-time job	14	1.85
Sheet	10	5.30	Large amount	10	2.78	Each time	10	1.85
Public	15	5.29	Nervousness	15	2.78	Immunity	18	1.85
Playing	15	5.29	Severe symptoms	14	2.77	Cup noodles	11	1.84
Chance	49	5.28	Sneeze	12	2.76	Out of stock	23	1.82
Everyone	11	5.27	Mentality	40	2.76	Cold	15	1.82
Anger (0–3)			Sadness (0–3)			Disgust (0–3)		
Top 20 Words	Frequency	*M*	Top 20 Words	Frequency	*M*	Top 20 Words	Frequency	*M*
Depression	12	1.93	Depression	12	2.27	Depression	12	2.05
Country	19	1.85	Country	19	2.02	Spray	10	1.86
Sneeze	12	1.65	Test	11	1.84	Sneeze	12	1.83
Spray	10	1.64	Mood	37	1.79	Cough	36	1.81

Table 1 (continued)

Mood	37	1.61	Economy	32	1.76	Country	19	1.79
Cough	36	1.60	This month	10	1.75	Economy	32	1.67
Cold	15	1.55	Staying the same	26	1.72	Throat	16	1.64
Paper	13	1.50	Salary	12	1.70	Immunity	18	1.63
Abnormality	11	1.48	End	10	1.69	Salary	12	1.61
Damage	10	1.47	Human	13	1.69	City	10	1.58
Economy	32	1.47	Nervousness	15	1.68	Cancellation	21	1.58
Cancellation	21	1.45	Cough	36	1.65	Part-time job	14	1.56
Broadcasting	26	1.44	Throat	16	1.64	This month	10	1.55
Immunity	18	1.40	Getting worse	11	1.64	Instability	12	1.55
Throat	16	1.39	Industry	21	1.64	Mood	37	1.54
Government	21	1.39	Postponement	18	1.63	Broadcasting	26	1.54
Crisis	19	1.38	Unemployment	12	1.63	Paper	13	1.52
Test	11	1.38	Mentality	40	1.61	End	10	1.50
Coping	35	1.34	Cup noodles	11	1.59	Industry	21	1.50
Severe symptoms	14	1.34	Months	31	1.59	Others	12	1.50

Only words which emerged over 10 times were selected. All words are single-word nouns in Japanese of life satisfaction among those who used the word "married couple" was 6.44. For life satisfaction, family-related (e.g., married couple, children, and housewife) and school-related words (e.g., school and school closure) were highly ranked. As for depression and negative affect, words about mental aspects (e.g., depression, nerve, and mentality), economy (e.g., economy, industry, and salary), and symptoms (e.g., disease, cough, and sneeze) were in the top 20.

Topic Modeling

First, we roughly set the number of topics between 5 and 50, which is recommended as a suitable starting place for a few hundred to a few thousand words (Roberts et al., 2019). Since the results reported that more than 20 topics did not increase the model fittings, we ran the same analysis after changing the range to 5–20 topics (Fig. 1). According to the diagnostic values and interpretability of the topics, we selected 9 topics. Each of the 2594 words has a different probability to appear in each of the 9 topics. This is called "word probability." Word probability indicates how likely each word is included in a topic. The words with the highest-ranked word probability for each topic are shown in Fig. 2. According to each topic's highest-ranked words, we named the 9 topics as follows: Mask, Physical symptoms, Children, Infection anxiety, Disinfection items, Economic influence, Remote work, Going out, and Change of lifestyle. Examples of descriptions and topic probabilities for each topic are shown in Table 2.

Correlation Analyses

As shown in Table 3, the high topic probability of topic 2 (Physical symptoms) was positively correlated with having actual symptoms ($r = 0.12$). The high topic probability of topic 3 (Children) was positively correlated with being female ($r = 0.21$), having children ($r = 0.33$), and high socioeconomic status ($r = 0.22$). The high topic probability of topic 6 (Economic influence) was negatively correlated to socioeconomic status ($r = -0.10$), and the high topic probability of topic 7 (Remote work) was positively correlated to working remotely ($r = 0.15$). Lastly, the probability of topic 9 (Change of lifestyle) was positively correlated to working remotely ($r = 0.12$).

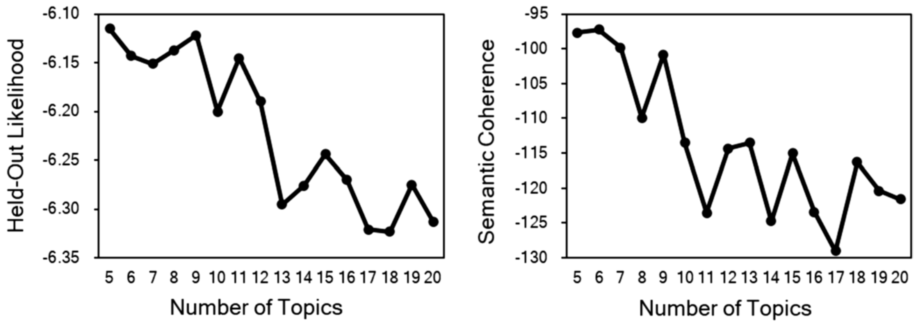

Fig. 1 Diagnostic values by number of topics

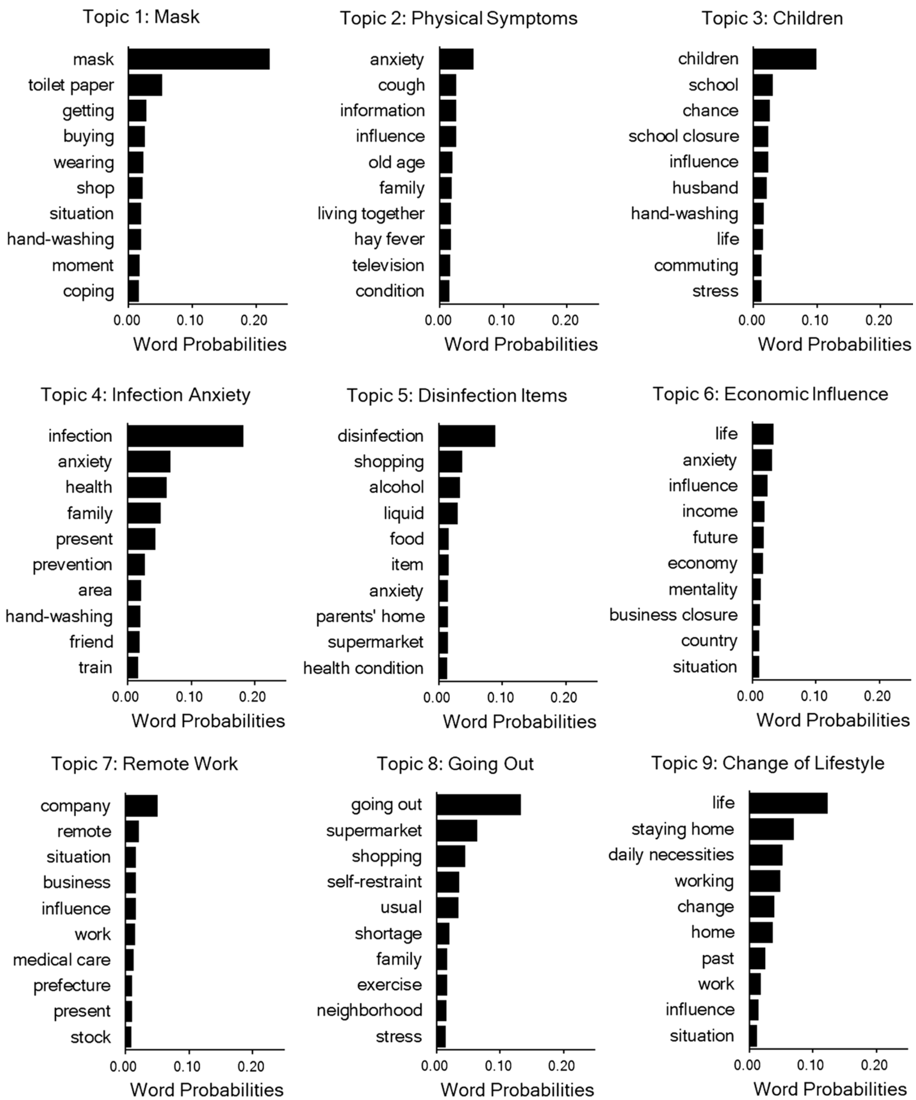

Fig. 2 Highest word probabilities for each topic

Regression Analyses

As shown in Table 4, writing about economic influence (topic 6) predicted lower life satisfaction and higher depression and negative affect. Also, writing about physical symptoms (topic 2) and disinfection items (topic 5) had a similar effect, although the coefficients were relatively small. Conversely, those who wrote about children (topic 3) and remote work (topic 7) experienced higher life satisfaction and lower depression and negative affect.

Table 2 Description examples and topic probabilities in each topic

Topic names	Description examples	TP1	TP2	TP3	TP4	TP5	TP6	TP7	TP8	TP9
Topic 1: Mask	I think it can't be helped that life is no longer the same as usual. However, many daily items such as toilet paper, masks, mouthwash, and food, are in short supply, so I have to spend a lot of time and costs to go to shops far away, or buy them at higher prices than usual. Sometimes I'm scared when I am sick and having a slight fever. I reuse masks while washing disposable ones.	**0.77**	0.01	0.03	0.06	0.04	0.00	0.01	0.05	0.03
Topic 2: Physical symptoms	Whenever I have a slight sore throat, cough, or difficulty breathing, I suspect that I'm infected with COVID-19. I'm even more anxious about infecting my parents who live with me because they have high-risk diseases. Even when I go to a supermarket, I can't help keeping distance from people who are coughing. I'm nervous and tired.	0.04	**0.73**	0.00	0.10	0.02	0.02	0.01	0.05	0.04
Topic 3: Children	My children's elementary school, kindergarten, and lessons are all closed, so we live together 24 h a day. My children are stressed out because they can't go out to play. It's also hard to find good ways to spend each day with them. Food costs are also higher than usual.	0.03	0.01	**0.82**	0.03	0.01	0.00	0.02	0.05	0.03
Topic 4: Infection anxiety	I'm in good health right now, but I'm scared that I don't know when I'm going to get infected with the coronavirus. I wash my hands and gargle thoroughly to prevent infection, but I'm anxious because I don't know if these are effective. My family is not infected now, but I'm anxious every day that they might get infected because they are working outside.	0.06	0.03	0.00	**0.73**	0.01	0.05	0.02	0.05	0.04
Topic 5: Disinfection items	I always take medication to prevent my asthma from getting worse. I used to hate wearing a mask because it was uncomfortable, but now I wear a cloth mask to prevent the spread of infection. I'm worried that my favorite restaurant will go deserted and close its doors. I have no problem getting food at supermarkets and vegetable stores. Paper and sanitary items have returned to the shelves, but I can't find alcohol disinfectants without many tours to the drug stores.	0.22	0.01	0.01	0.06	**0.65**	0.00	0.01	0.02	0.02
Topic 6: Economic influence	I am most anxious about the increase in the number of infected people, but I am most deeply concerned about the economic collapse caused by COVID-19, including the future of my own work. In spite of high GDP, Japan's economic measures have lagged far behind those of other countries, and the country has taken a series of foolish measures. To be honest, I can't stand to see bankruptcies and a drastic increase in the unemployment rate in the near future. After all, as long as we live in the same society, the instability of one's own work and life is linked to the instability of the world.	0.01	0.01	0.00	0.03	0.00	**0.91**	0.01	0.01	0.01

Table 2 (continued)

Topic names	Description examples	TP1	TP2	TP3	TP4	TP5	TP6	TP7	TP8	TP9
Topic 7: Remote work	I have been working remotely for two weeks. My workload is less than before the coronavirus outbreak. I work for an advertising design company, and most of the events have been postponed or cancelled. I try to stay at home as much as possible, including weekends. However, I feel that my physical strength is weakening due to staying at home all the time, so I take a walk or go jogging at least once a day. My wife continues to commute to work because her job does not allow her to work remotely. Fortunately, her workplace is within walking distance, so she can avoid the crowd.	0.01	0.00	0.02	0.03	0.01	0.00	**0.91**	0.01	0.02
Topic 8: Going out	I don't go out anymore and I keep my shopping to a minimum. I try to use the Internet to avoid going out as much as possible. I also try to encourage my family not to go out. However, not going out causes stress and lack of exercise, which makes me feel uncomfortable and unwell.	0.04	0.02	0.01	0.04	0.01	0.01	0.00	**0.83**	0.04
Topic 9: Change of lifestyle	Even before the COVID-19 pandemic, I was working at home most of the time, so my lifestyle remained the same. I'm not too concerned about infection because I have limited contact with other people, but it has become difficult to obtain some daily necessities due to false rumor and excessive media coverage.	0.03	0.01	0.01	0.07	0.01	0.01	0.02	0.08	**0.75**

TP, topic probability

Table 3 Correlations between topic probabilities and demographic variables

	Age	Gender(male = 1, female = 2)	Employment (unemployed=1, employed=2)	Having symptoms(no = 1, yes = 2)	Having children(no = 1, yes = 2)	Socioeconomic Status(poorest = 1, richest = 10)	Doing remote work(no = 1, yes = 2)
TP1: Mask	.04	− .05	.02	.05	− .06	− .12**	− .11**
TP2: Physical symptoms	− .03	.03	.04	.12***	− .15***	− .07*	− .05
TP3: Children	.00	.21***	− .05	− .10**	.33***	.22***	− .06
TP4: Infection anxiety	.02	− .09*	.01	− .04	− .04	− .06	.04
TP5: Disinfection items	− .02	.06	− .02	.05	− .03	− .02	− .03
TP6: Economic influence	.06	− .04	− .03	.08*	− .04	− .10**	.00
TP7: Remote work	− .01	− .06	− .06	− .02	− .01	.10**	.15***
TP8: Going out	− .02	− .01	.11**	− .06	− .02	.00	− .07*
TP9: Change of lifestyle	− .05	− .07*	− .01	− .06	− .05	.01	.12**

$^*p < .05$. $^{**}p < .01$. $^{***}p < .001$. *TP* topic probability

Table 4 Effects of topic probabilities on mental health

	Life satisfaction									Depression								
	B	*SE*	[	95%*CL*		]	*β*	*t*	*p*	*B*	*SE*	[	95%*CL*		]	*β*	*t*	*p*
TP1: Mask																		
TP2: Physical symptoms										0.50	0.23	[	0.04	0.96	]	.08	2.13	.034
TP3: Children	3.12	0.38	[	2.39	3.86	]	.28	8.32	< .000	−0.60	0.20	[	−0.99	−0.22	]	−.11	−3.09	.002
TP4: Infection anxiety																		
TP5: Disinfection items																		
TP6: Economic influence	−2.12	0.40	[	−2.90	−1.33	]	−.18	−5.30	< .000	0.76	0.21	[	0.36	1.16	]	.13	3.69	< .000
TP7: Remote work	0.96	0.42	[	0.14	1.77	]	.08	2.31	.021	−0.71	0.22	[	−1.14	−0.29	]	−.12	−3.31	.001
TP8: Going out																		
TP9: Change of lifestyle																		
	$R^2 = .133, p < .000$									$R^2 = .059, p < .000$								
	Fear									Anger								
	B	*SE*	[	95%*CL*		]	*β*	*t*	*p*	*B*	*SE*	[	95%*CL*		]	*β*	*t*	*p*
TP1: Mask																		
TP2: Physical symptoms	0.54	0.21	[	0.12	0.96	]	.09	2.55	.011	0.42	0.20	[	0.02	0.82	]	.07	2.06	.040
TP3: Children																		
TP4: Infection anxiety																		
TP5: Dis-infection items	0.53	0.22	[	0.11	0.96	]	.09	2.47	.014									
TP6: Economic influence	0.59	0.19	[	0.22	0.96	]	.11	3.10	.002	0.49	0.18	[	0.14	0.85	]	.10	2.73	.006
TP7: Remote work	−0.72	0.20	[	−1.12	−0.32	]	−.13	−3.51	< .000									

Table 4 (continued)

	Life satisfaction									Depression								
	B	*SE*	[	95%*CL*		]	*β*	*t*	*p*	*B*	*SE*	[	95%*CL*		]	*β*	*t*	*p*
TP8: Going out	−0.38	0.22	[	−0.82	0.06	]	−.06	−1.69	.092									
TP9: Change of lifestyle																		
	R^2 = .058, p < .000									R^2 = .014, p = .004								
	Sadness									Disgust								
	B	*SE*	[	95%*CL*		]	*β*	*t*	*p*	*B*	*SE*	[	95%*CL*		]	*β*	*t*	*p*
TP1: Mask																		
TP2: Physical symptoms	0.63	0.21	[	0.22	1.03	]	.11	3.03	.003	0.59	0.21	[	0.18	1.00	]	.10	2.85	.005
TP3: Children																		
TP4: Infection anxiety																		
TP5: Dis-infection items	0.82	0.21	[	0.41	1.23	]	.14	3.93	< .000	0.55	0.21	[	0.14	0.97	]	.09	2.62	.009
TP6: Economic influence	1.22	0.18	[	0.86	1.58	]	.23	6.65	< .000	0.88	0.19	[	0.51	1.24	]	.17	4.70	< .000
TP7: Remote work	−0.54	0.19	[	−0.92	−0.16	]	−.10	−2.81	.005									
TP8: Going out																		
TP9: Change of lifestyle																		
	R^2 = .090, p < .000									R^2 = .039, p < .000								

TP = topic probability

Discussion

This study aimed to examine how individuals' descriptions of daily life during the pandemic were categorized and to clarify which topics were related to mental health. Topic modeling extracted 9 topics, which commonly appeared in previous studies (Abd-Alrazaq et al., 2020; Han et al., 2020; Kleinberg et al., 2020; Stokes et al., 2020). The most expressed topic in this study was "infection anxiety," which was consistent with previous reports that Japanese people experienced increased anxiety (Macromill, 2020; Shigemura & Kurosawa, 2020; Ueda et al., 2020).

Regression analyses demonstrated that those who wrote about economic influence, physical symptoms, and disinfection items experienced lower life satisfaction and higher depression and negative affect. In terms of economic influence, previous studies suggested that the economic situation is one of the most significant factors of poor mental health (Kikuchi et al., 2020; Luo et al., 2020; Sugawara et al., 2021; Ueda et al., 2020). For example, Li et al. (2020) revealed that anxieties about income and ability to pay loans were significantly associated with anxiety and depression. Given that this topic had a relatively strong negative relationship with mental health, it may be crucial to monitor and support those who have suffered economically due to the pandemic. Kawohl and Nordt (2020) also articulated that mental health services need to pay attention to those who are unemployed, because rising unemployment predicts an increased number of suicides (Nordt et al., 2015).

In addition, it has been reported that having physical symptoms can be a risk factor of not only mental illness (Gorrochategi et al., 2020; Özdin & Bayrak Özdin, 2020; Smith et al., 2020; Varshney et al., 2020) but also mortality, especially when organ dysfunctions occur (Zhou, et al., 2020a). Thus, the obtained results about topic 2 were consistent with these studies. Also, writing about disinfection items was mainly related to sadness. We assume that many people felt sad when the disinfection items were not in stock and had difficulties obtaining them, as Shigemura and Kurosawa (2020) reported. Furthermore, mentioning remote work was negatively associated with depression, fear, and sadness, which was also congruent with the previous findings (Choi et al., 2020) that individuals who were not able to work from home were more likely to have depression and anxiety.

Conversely, those who mentioned their children were likely to have high life satisfaction, which was contrary to previous findings that reported adults with children had higher levels of general stress than those without (American Psychological Association, 2020). This may be due to cultural differences, given that our sample comprises entirely of Japanese people. The words frequently mentioned in topic 3 included school closure, implying that participants' children spend more time at home. Perhaps parents feel relieved that their children's risk of being infected was largely reduced, even if school closure had a negative impact on their children's mental health due to the reduction of outdoor activities and social interactions (Xie et al., 2020; Zhou et al., 2020b). Future studies need to explain the underlying mechanisms of this difference observed between Japanese and American samples.

Strengths and Limitations

There are few surveys about COVID-19 in Japan; therefore, this study can contribute to accumulating worldwide data to promote comprehension about the unprecedented pandemic. This study showed that those at higher risk of depression were those who have concerns over their economic situation, their physical symptoms, and obtaining items for disinfection. Therefore, the government's support in these domains (e.g., providing disinfectant and monetary subsidy to the general population) may help reduce the risk of mental illness. Additionally, this is the first study to examine mental health during the COVID-19 pandemic using participant's text data from open-ended questions about their daily life. Although previous studies using text data have extracted some topics of the text, this study demonstrated the associations between the topics and participant-level variables by adopting topic modeling. This study highlighted that this exploratory bottom-up approach was useful to confirm previous findings and to offer some insights. For example, the results replicated and highlighted that monitoring the mental health of individuals with economic impacts and physical symptoms during the pandemic is necessary (Gorrochategi et al., 2020; Kawohl & Nordt, 2020; Li et al., 2020; Luo et al., 2020; Özdin & Bayrak Özdin, 2020; Smith et al., 2020; Varshney et al., 2020). Also, the finding that mentioning children or school closure was related to higher life satisfaction is worth reexamining in future studies.

The current study had several limitations that deserve mentioning. First, it is possible that the examples given to participants in the instruction may have been priming them to write about certain topics. It is vital to examine the influence of showing the example topics in the instruction, either by giving examples or no examples. Second, in some cases, the classification of topics was not clear, although the inclusion of different words is a common limitation of topic modeling. For example, topic 7 "change of lifestyle" included some different aspects of words such as "staying home," "daily necessities," or "working," which is difficult to summarize into only one topic. Future research needs to replicate these topics and demonstrate the topic validity. Lastly, the data sampling is conducted in online crowdsourcing platforms. It has been reported that samples from crowdsourcing services tend to have lower income than samples from face-to-face surveys (e.g., Berinsky et al., 2012), which might also be the case for Japanese crowdsourcing services. Also, full-time employees in this study were only 38.7%. Although we cannot rule out the possibilities that the pandemic caused this low employment rate, readers need to be mindful of the features of the sample in this study.

Conclusion

Results of this study revealed that, in Japan, those who wrote about economic influence, physical symptoms, and disinfection items during the pandemic experienced lower life satisfaction, higher depression, and more negative affect, whereas those who mentioned their children were likely to have higher life satisfaction. This study is the first to examine the association between daily life narratives via text mining and mental health during the pandemic. It highlighted the importance of monitoring the mental health of individuals with economic disadvantage and physical symptoms. While most efforts are put into treating COVID-19 infection symptoms, it is important to remember that the general population's mental health may be as much at risk as their physical health. Future research can thus aim at having interventions to improve people's psychological well-being.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11469-021-00587-y.

Acknowledgements The authors thank all participants.

Author Contribution Yuta Chishima: Conceptualization, Data collection and analysis, Writing — original draft. I-Ting Huai-Ching Liu: Conceptualization, Writing — review and editing.

Funding This work was supported by Grant-in-Aid for JSPS Fellows [Grant Number: 18J01883].

Declarations

Ethics Approval The study design was approved by the appropriate ethics review board in Kyoto University (Ref. 30-P-24).

Consent to Participate Informed consent was obtained from all individual participants included in the study.

Conflict of Interest The authors declare no conflict of interest.

References

Abd-Alrazaq, A., Alhuwail, D., Househ, M., Hamdi, M., & Shah, Z. (2020). Top concerns of tweeters during the COVID-19 pandemic: Infoveillance study. *Journal of Medical Internet Research, 22*(4), e19016. https://doi.org/10.2196/19016

American Psychological Association (2020, May). *Stress in the time of COVID-19: Volume one*. https://www.apa.org/news/press/releases/stress/2020/stress-in-america-covid.pdf

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis, 20*(3), 351–368. https://doi.org/10.1093/pan/mpr057

de Bruin. (2020) Age differences in COVID-19 risk perceptions and mental health: Evidence from a national US survey conducted in March 2020. *The Journals of Gerontology: Series b: Psychological Sciences and Social Sciences*. Advance Online Publication. https://doi.org/10.1093/geronb/gbaa074

Chao, M., Xue, D., Liu, T., Yang, H., & Halld, B. J. (2020). Media use and acute psychological outcomes during COVID-19 outbreak in China. *Journal of Anxiety Disorders*. Advance online publication. https://doi.org/10.1016/j.janxdis.2020.102248

Chishima, Y., Liu, I-T. H-C., & Wilson, A. E. (in press). Temporal distancing during the COVID-19 pandemic: Letter writing with future self can mitigate negative affect. *Applied Psychology: Health and Well-Being*.

Choi, E., Hui, B., & Wan, E. (2020). Depression and anxiety in Hong Kong during COVID-19. *International Journal of Environmental Research and Public Health, 17*(10), 3740. https://doi.org/10.3390/ijerph17103740

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Sage Publications.

Furukawa, T. A., Kawakami, N., Saitoh, M., Ono, Y., Nakane, Y., Nakamura, Y., Tachimori, H., Iwata, N., Uda, H., Nakane, H., Watanabe, M., Naganuma, Y., Hata, Y., Kobayashi, M., Miyake, Y., Takeshima, T., & Kikkawa, T. (2008). The performance of the Japanese version of the K6 and K10 in the World Mental Health Survey Japan. *International Journal of Methods in Psychiatric Research, 17*(3), 152–158. https://doi.org/10.1002/mpr.257

Gorrochategi, M. P., Munitis, A. E., Santamaria, M. D., & Etxebarria, N. O. (2020). Stress, anxiety, and depression in people aged over 60 in the COVID-19 outbreak in a sample collected in Northern Spain. *The American Journal of Geriatric Psychiatry.* Advance online publication. https://doi.org/10.1016/j.jagp.2020.05.022

Guo, J., Feng, X. L., Wang, X. H., van IJzendoorn, M. H. (2020). Coping with COVID-19: Exposure to COVID-19 and negative impact on livelihood predict elevated mental health problems in Chinese adults. *International Journal of Environmental Research and Public Health, 17*(11), 3875 https://doi.org/10.3390/ijerph17113857

Han, X., Wang, J., Zhang, M., & Wang, X. (2020). Using social media to mine and analyze public opinion related to COVID-19 in China. *International Journal of Environmental Research and Public Health, 17*(8), 2788. https://doi.org/10.3390/ijerph17082788

Holmes., E. A., O'Connor, R. C., Perry, V. H., Tracey, I., Wessely, S., Arseneault, L., Ballard, C., Christensen, H., Silver, R. C., Everall, I., Ford, T., John, A., Kabir, T., King, K., Madan, I., Michie, S., Przybylski, A. K., Shafran, R., Sweeney, A., … Bullmore, E. (2020) Multidisciplinary research priorities for the COVID-19 pandemic: A call for action for mental health science. *The Lancet Psychiatry, 7*(6), 547-560 https://doi.org/10.1016/S2215-0366(20)30168-1

Kawohl, W., & Nordt, C. (2020). COVID-19, unemployment, and suicide. *The Lancet Psychiatry, 7*(5), 389–390. https://doi.org/10.1016/S2215-0366(20)30141-3.
Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S.-L.T., Manderscheid, R. W., Walters, E. E., & Zaslavsky, A. M. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry, 60*(2), 184–189. https://doi.org/10.1001/archpsyc.60.2.184
Kikuchi, H., Machida, M., Nakamura, I., Saito, R., Odagiri, Y., Kojima, T., Watanabe, H., Fukui, K., & Inoue, S. (2020). Changes in psychological distress during the COVID-19 pandemic in Japan: A longitudinal study. *Journal of Epidemiology, 30*(11), 522–528. https://doi.org/10.2188/jea.JE20200271
Kleinberg, B., van der Vegt, I., & Mozes, M. (2020). Measuring emotions in the COVID-19 real world worry dataset. *Proceedings of ACL 2020 Workshop on Natural Language Processing for COVID-19*. Advance online publication. https://openreview.net/pdf?id=JQCYcdHfXyJ
Kudo, T. (2006). MeCab: Yet another part-of-speech and morphological analyzer. https://taku910.github.io/mecab/. Accessed 31 August 2020.
Li, J., Yang, Z., Qiu, H., Wang, Y., Jian, L., Ji, J., & Li, K. (2020). Anxiety and depression among general population in China at the peak of the COVID-19 epidemic. *World Psychiatry, 19*(2), 249–250. https://doi.org/10.1002/wps.20758
Luo, M., Guo, L., Yu, M., & Wang, H. (2020). The psychological and mental impact of coronavirus disease 2019 (COVID-19) on medical staff and general public: A systematic review and meta-analysis. *Psychiatry Research*. Advance online publication. https://doi.org/10.1016/j.psychres.2020.113190
Macromill. (2020) Macromill Weekly Index. https://www.macromill.com/data_and_insights/weeklyindex/. Accessed 31 August 2020.
Mimno, D., Wallach, H., Talley, E., Leenders, M., & McCallum, A. (2011). Optimizing semantic coherence in topic models. *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, 262–272.
Nordt, C., Warnke, I., Seifritz, E., & Kawohl, W. (2015). Modelling suicide and unemployment: A longitudinal analysis covering 63 countries, 2000–11. *The Lancet Psychiatry, 2*(3), 239–245. https://doi.org/10.1016/S2215-0366(14)00118-7
Oda, Y., Takano, R., Abe, T., & Kikuchi, K. (2015). Development of the Emotion and Arousal Checklist (EACL). *The Japanese Psychological Research, 85*(6), 579–589. https://doi.org/10.4992/jjpsy.85.13231
Özdin, S., & Bayrak Özdin, Ş. (2020). Levels and predictors of anxiety, depression and health anxiety during COVID-19 pandemic in Turkish society: The importance of gender. *International Journal of Social Psychiatry, 66*(5), 504–511. https://doi.org/10.1177/0020764020927051.
Prenda, K. M., & Lachman, M. E. (2001). Planning for the future: A life management strategy for increasing control and life satisfaction in adulthood. *Psychology and Aging, 16*(2), 206–216. https://doi.org/10.1037/0882-7974.16.2.206
Roberts, M. E., Stewart, B. M., Airoldi, E. M. (2016) A model of text for experimentation in the social sciences. *Journal of the American Statistical Association ,111*(515), 988-1003 https://doi.org/10.1080/01621459.2016.1141684
Roberts, M. E., Stewart, B. M., Tingley, D. (2019). stm: R package for structural topic models. *Journal of Statistical Software, 91*(2), 1–4. https://doi.org/10.18637/jss.v091.i02
Satici, B., Gocet-Tekin, E., Deniz, M.E., Satici, S. A. (2020) Adaptation of the fear of COVID-19 scale: Its association with psychological distress and life satisfaction in Turkey. *International Journal of Mental Health and Addiction*. Advance online publication. https://doi.org/10.1007/s11469-020-00294-0
Shigemura, J., Kurosawa, M. (2020). Mental health impact of the COVID-19 pandemic in Japan. *Psychological Trauma: Theory, Research, Practice, and Policy*. Advance online publication https://doi.org/10.1037/tra0000803
Smith, L., Jacob, L., Yakkundi, A., McDermott, D., Armstrong, N. C., Barnett, Y., López-Sánchez, G. F., Martin, S., Butler, L., & Tully, M. A. (2020). Correlates of symptoms of anxiety and depression and mental wellbeing associated with COVID-19: A cross-sectional study of UK-based respondents. *Psychiatry Research*. Advance online publication. https://doi.org/10.1016/j.psychres.2020.113138
Stokes, D. C., Andy, A., Guntuku, S. C., Ungar, L. H., & Merchant, R. M. (2020). Public priorities and concerns regarding COVID-19 in an online discussion forum: Longitudinal topic modeling. *Journal of General Internal Medicine*. Advance online publication. https://doi.org/10.1007/s11606-020-05889-w
Sugawara, D., Masuyama, A., & Kubo, T. (2021). Socioeconomic impacts of the COVID-19 lockdown on the mental health and life satisfaction of the Japanese population. *International Journal of Mental Health and Addiction*. (in press).

Taddy, M. (2013). Multinomial inverse regression for text analysis. *Journal of the American Statistical Association, 108*(503), 755–770. https://doi.org/10.1080/01621459.2012.734168

Teti, M., Schatz, E., & Liebenberg, L. (2020). Methods in the time of COVID-19: The vital role of qualitative inquiries. *International Journal of Qualitative Methods*. Advance online publication. https://doi.org/10.1177/1609406920920962

Tull, M. T., Edmonds, K. A., Scamaldo, K., Richmond, J. R., Rose, J. P., & Gratz, K. L. (2020). Psychological outcomes associated with stay-at-home orders and the perceived impact of COVID-19 on daily life. *Psychiatry Research*. Advance online publication. https://doi.org/10.1016/j.psychres.2020.113098

Ueda, M., Stickley, A., Sueki, H., & Matsubayashi, T. (2020). Mental health status of the general population in Japan during the COVID-19 pandemic. *Psychiatry and Clinical Neurosciences, 74*(9), 505–506. https://doi.org/10.1111/pcn.13105

Varshney, M., Parel, J. T., Raizada, N., & Sarin, S. K. (2020). Initial psychological impact of COVID-19 and its correlates in Indian community: An online (FEEL-COVID) survey. *PLoS ONE, 15*(5), e0233874. https://doi.org/10.1371/journal.pone.0233874

Verma, S., & Mishra, A. (2020). Depression, anxiety, and stress and socio-demographic correlates among general Indian public during COVID-19. *International Journal of Social Psychiatry*. Advance online publication. https://doi.org/10.1177/0020764020934508

Wallach, H. M., Murray, I., Salakhutdinov, R., & Mimno, D. (2009). Evaluation methods for topic models. *Proceedings of the 26th Annual International Conference on Machine Learning*, *26*, 1105–1112.

Xie, X., Xue, Q., Zhou, Y., Zhu, K., Liu, Q., Zhang, J., Song, R. (2020) Mental health status among children in home confinement during the coronavirus disease 2019 outbreak in Hubei Province. *JAMA Pediatrics*. Advance online publication https://doi.org/10.1001/jamapediatrics.2020.1619

Yamamoto, T., Uchiumi, C., Suzuki, N., Yoshimoto, J., & Murillo-Rodriguez, E. (2020). The psychological impact of 'mild lockdown' in Japan during the COVID-19 pandemic: A nationwide survey under a declared state of emergency. *International Journal of Environmental Research and Public Health, 17*(24), 9382. https://doi.org/10.3390/ijerph17249382

Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., Xiang, J., Wang, Y., Song, B., Gu, X., Guan, L., Wei, Y., Li, H., Wu, X., Xu, J., Tu, S., Zhang, Y., Chen, H., & Cao, B. (2020a). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: A retrospective cohort study. *The Lancet, 395*(10229), 1054–1062. https://doi.org/10.1016/S0140-6736(20)30566-3

Zhou, S. J., Zhang, L. G., Wang, L. L., Guo, Z. C., Wang, J. Q., Chen, J. C., Liu, M., Chen, X., & Chen, J. X. (2020b). Prevalence and socio-demographic correlates of psychological health problems in Chinese adolescents during the outbreak of COVID-19. *European Child & Adolescent Psychiatry, 29*, 749–758. https://doi.org/10.1007/s00787-020-01541-4