




ARTICLE



<https://doi.org/10.1057/s41599-023-02014-0>

OPEN

Public opinion in Japanese newspaper readers' posts under the prolonged COVID-19 infection spread 2019–2021: contents analysis using Latent Dirichlet Allocation

Hideaki Kasuga¹  , Shota Endo¹, Yusuke Masuishi¹, Tomoo Hidaka¹, Takeyasu Kakamu¹ & Tetsuhito Fukushima¹

During the COVID-19 pandemic, information on what people are interested in and what they are disseminating can be an important public health resource. Most studies on public opinion during the pandemic have focused on social networking services in the context of the early phases of the pandemic or on a short-term basis; research on public opinion outside of social networking services that consider long-term changes has not been conducted. To examine this gap in the research, we analyzed readers' posts on Japanese hardcopy newspaper articles in the public domain. A total of 1910 such posts published during four emergency declaration periods in Japan were included in the study. Latent Dirichlet Allocation was applied in our analysis to extract topics and Kendall rank correlation coefficients between the emergency declaration periods, and each topic was calculated to examine the effect size. We selected 10 topics and categorized them into three themes: (1) "Life" comprising "Family," "Daily Life in the COVID-19 Disaster," "Education in the COVID-19 Disaster," "The Importance of Humanity," and "Daily Life unrelated to COVID-19"; (2) "Awareness of the emergency" comprising, "Awareness of being a party to an emergency" and "Concerns about the medical environment,"; and (3) "Policy" comprising "Domestic and foreign policies," "Opposition to hosting the Tokyo Olympics," and "Criticisms of the Japanese Government." This research revealed that, as a result of the exposure to COVID-19-related material over a prolonged period, awareness of the emergency decreased ($r = -0.189$, $p < 0.000$), while other topics remained. This study showed that it is possible to extract more everyday public opinion topics from the reader's posts in newspapers and that these are issues that should be addressed from a long-term perspective since they do not change significantly in a stressful life.

¹Fukushima Medical University, Fukushima, Japan. ✉email: h-kasuga@fmu.ac.jp

Introduction

Since December 31, 2019, when the World Health Organization (WHO) issued a warning about an outbreak of pneumonia from unknown causes in Wuhan, China, people worldwide have been required to take measures to prevent the spread of the SARS-CoV-2 (COVID-19) infection. Under administrative directives, people must curtail economic activities, refrain from going out, wear masks, maintain social distancing, and change their lifestyles, all of which can result in stress during daily living (Nomura et al., 2020). In this context, what information people are interested in and what they are disseminating (public opinion) can be an important public health resource during both present emergency and future emergencies (Abd-Alrazaq et al., 2020; Chiptza et al., 2021; WHO, 2017).

Recent COVID-19-related studies on public opinion have drawn on questionnaires (Shen et al., 2021), interviews (Yunusa et al., 2021), and text data from social networking sites (SNS), such as Facebook, Group Chat, and microblogs like Twitter and Sin Weibo (Abd-Alrazaq et al., 2020; Maryam, 2022; Miao et al., 2020; Han et al., 2020; Wawrzuta et al., 2021; Zhong, 2021). These studies using SNS to analyze public opinion have examined attitudes (positive or negative) and emotions (fear, sadness, joy, and anger) related to specific topics, such as lockdown and other policies (Miao et al., 2020), vaccines (Wawrzuta et al., 2021), and COVID-19 generally (Maryam, 2022), as well as inductively explored topics that emerged from naturally occurring text (Abd-Alrazaq et al., 2020; Han et al., 2020; Zhong, 2021). Other predominant research areas that have emerged from analyzing SNS are the symptoms of COVID-19, the origin of confirmed cases of COVID-19, the impact of COVID-19 on people and countries, and preventive measures. In addition, public opinion in Japan has also compared local policies with global COVID-19 policies. SNS research has also examined fear and worry related to COVID-19, willingness to return to work, and public criticism of the government and media in relation to COVID-19. Apparent from the above summary of current research using SNS, public opinion can be considered useful for stakeholders, such as the government and other professionals, in designing interventions, measures, and policies related to COVID-19 (Abd-Alrazaq et al., 2020; WHO, 2017).

Although researchers were able to collect large amounts of data quickly using SNS information, two important limitations in that approach have been identified. First, the findings reported in such studies are based solely on data collected from those who have access to and use SNS. Since SNS use is dominated by younger people (Morgan-Lopez et al., 2017; Ministry of Internal Affairs and Communications (MIC), 2022), it is more difficult for SNS-based studies to capture the opinions of middle-aged and older people. Moreover, because anonymous users tend to be unsocial (Santana, 2014) when posting their opinions on the Internet, posts often emphasize one specific aspect of themselves, and SNS data alone cannot provide sufficient information for comprehensive understanding. To address this, this study focuses on readers' posts in hardcopy newspapers. The newspaper is a traditional media format used by many older age groups (MIC, 2022), and personal information, at least what is explicitly required by the newspaper, such as name, age, and occupation, is often specified when posts are submitted for publication; this data allows researchers to target different people and aspects than those targeted in existing studies. Newspapers retain a certain degree of credibility even today (Zhao et al., 2011), and, from the viewpoint that SNS, such as Twitter, and traditional media, such as newspapers, feature some topics in common, they also feature topics that do not overlap. Traditional media and new media are both interactive (Menzies and Menzies, 2020), and therefore, it is important to

consider the impact of SNS on traditional media and to look at newspapers as well.

Second, many studies seeking COVID-19-related public opinion collected data early in the pandemic on a short-term basis and therefore failed to capture the trends that occurred over longer periods of time and during later periods of the pandemic. COVID-19 is characterized by a comparatively long period of infection convergence. For example, the 2002 SARS outbreak lasted eight months, from November 16, 2002, to July 5, 2003 (WHO, 2003), and the 2009 influenza (A/H1N1) pandemic (April 2009 to August 10, 2010) was declared over by the WHO after a year and four months (WHO, 2010). By contrast, COVID-19 shows no sign of abatement as of February 2022, two years later. Public opinion during such a prolonged period of COVID-19 needs to be investigated in longitudinal research.

To address these limitations, it is appropriate to analyze the readers' posts in newspapers in Japan. The primary readership of newspapers in Japan is middle-aged and older adults, a different demographic from that targeted in previous studies. In addition, newspaper readers' posts have continued to record public opinion during the COVID-19 outbreak in a consistent format. These advantages will contribute to overcoming the limitations of previous studies.

The value of paper newsprint to the Japanese

Japan has one of the highest levels of newspapers used as an information source in the world (Haerpfer et al., 2022), and thus, the readers' columns of newspapers are possibly a useful source of data on public opinion (Hidaka et al., 2022). According to the data in 2016 (World Association of Newspapers and News Publishers (WAN-IFRA), 2016), the total circulation of Japan's representative daily newspapers, the Yomiuri, Asahi, and Mainichi Shimbun, was approximately 19 million, and these newspapers ranked in the top ten in terms of newspaper circulation globally. Japanese do not use newspapers at all for acute or entertainment topics, but they use newspapers relatively often for reliable information (Haerpfer et al., 2022; MIC, 2022). Both usage and trust are low among young people, which increase with age. Newspaper usage and trust tend to decline annually, however, as a whole, newspapers remain a valuable means of acquiring information.

COVID-19 in Japan

Japan had its first confirmed case of COVID-19 in January 2020, when people returning from Wuhan, China, and passengers on the Diamond Princess cruise ship that departed from Hong Kong tested positive. From those initial cases, COVID-19 spread throughout Japan. As Japanese law does not allow for the imposition of penalties or strict lockdowns (Okubo, 2020), the main form of infection control measures was government, administrator, and expert requests for public compliance in avoiding closed spaces, crowded places, and close-contact settings (Ministry of Health, Labor and Welfare, 2022). These measures included requests for restrictions in economic activities, such as shortened work hours, telecommuting, cancellation of events, and admission restrictions; educational activities, such as online classes, timetable changes, and cancellation of seasonal events; and personal activities, such as restricting dinners with large groups of people and provincial travel. The seriousness and range of requests varied, depending on the spread of infection. When the number of new infections increased to the extent that shortage of medical facilities became a concern, a state of emergency was declared. During the state of emergency, the strongest requests were made for a defined period throughout the country

or in areas with particularly large numbers of new infections. The state of emergency was declared by the government and a team of experts based on the infection rate and the transmissibility and severity of the COVID-19 variant. No objective or absolute numerical criteria were established by the government. A state of emergency was declared four times as of December 2021, including in Tokyo, the capital of Japan. The first emergency period was from April 7 to May 25, 2020, the second from January 8 to March 21, 2021, the third from April 25 to June 6, 2021, and the fourth from July 12 to September 30, 2021.

In Japan, the hosting of the Tokyo 2020 Olympic Games was also a subject of discussion (Dalton and Taylor, 2021). In March 2020, the Olympics was postponed to the summer of 2021, considering the global spread of COVID-19. However, the outbreak continued to fluctuate in Japan and abroad, and no consensus could be reached regarding the holding of the Olympics. In July 2021, the Japanese government and the International Olympic Committee hosted the Games under the slogan of “safety and security” amid the spread of infection in Japan. Civic and medical officials voiced opposition to the Olympics and criticized the government and the Committee, citing concerns about the spread of infection and the shortage of medical facilities (Dalton and Taylor, 2021). Although the Olympics was held without spectators and the outbreak of infection caused by the Olympics was limited as a result (Akashi et al, 2022), sufficient dialog with the public becomes necessary when a country decides to host a major event during an infectious disease pandemic.

Understanding trends in public opinion during a prolonged infectious disease outbreak may be useful for predicting how public opinion shows trends in the future, and for taking countermeasures when the next infectious disease epidemic occurs. Newspaper databases, which are likely to be more strictly controlled than SNS (wherein individual contributors can freely edit and delete data) were appropriate sources of long-term historical data. An added advantage is the ability to collect stable data in the context of an infectious disease outbreak, when research activities involving human subjects are restricted. The results, however analytical the Japanese newspapers may be, will help understand the world during a pandemic (that may occur again) and help implement more favorable policies.

Thus, the purpose of this study was to have a comprehensive understanding of COVID-19-related public opinion by exploring topics other than those encountered in relevant SNS studies, to clarify changes in people’s interests and issues over time, and provide clues to the types of assistance that should be provided from a medium- to long-term perspective. In addition, we extracted definite characteristics by targeting the periods during which the Japanese government recommended a particularly strong voluntary restraint on activities and preventive actions.

Methods

Data sampling. We conducted a document analysis of readers’ posts about newspaper articles in the public domain. Many Japanese newspapers have a section where readers can submit their opinions on any topic. Submissions can be made by mail, fax, or e-mail, and readers are asked to disclose their name, occupation, age, address, and contact information. All or part of these disclosed details are often published in the paper.

In this study, Yomiuri, Mainichi, and Asahi newspapers, the top three newspapers in Japan in terms of circulation (WAN-IFRA, 2016), were considered. Plugh (2018) and Taniguchi (2018) said that Japanese newspapers are not clearly divided into right-wing or left-wing, etc., and would never expressly state a political position, such as ‘this paper supports the Democratic/Liberal Democratic Party’. The age group of readers of each newspaper is

shown in Supplementary Table S1. The percentage of people in their 30s or younger accounts for less than 30% of the total readership of all of these newspapers, with readers in their 40s and older being the main readership, peaking in their 50s and 60s.

All the newspapers clearly state that the submissions must be original and that corrections to the submissions are made only to the extent that the intent of the text is not compromised. Mainichi and Asahi disclose in their submission rules that rewards will be offered for published submissions. However, the publication criteria are not disclosed in either of the newspapers.

In this study, we focused on public opinion under the state of emergency declaration, which is considered to be more affected by the spread of COVID-19, and aimed to research the changes that occurred as a result of repeated states of emergency. We considered public opinion during states of emergency to be qualitatively different from that during a period of relatively less tension and restrictions on people’s lives and therefore excluded those from this study. Upon searching the database (Asahi Shimbun Cross-Search for Asahi, 2022; Maisaku for Mainichi, 2022; Yomidas Rekishikan for Yomiuri, 2022) for “coronavirus,” we found reader submissions during all four emergency declaration periods in Japan, during the dates mentioned previously.

We found 412 COVID-19-related readers’ posts during the first emergency declaration period: Yomiuri was 123, Mainichi was 123, and Asahi was 166; 521 during the second emergency declaration period: Yomiuri was 164, Mainichi was 165, and Asahi was 192; 458 during the third emergency declaration period: Yomiuri was 139, Mainichi was 151, and Asahi was 168; and 519 during the fourth emergency declaration period: Yomiuri was 133, Mainichi was 192, and Asahi was 194, with a total of 1910 posts, which is an average of 7.2 (range: 0–19) submissions per day. The mean number of characters per posting was 376.7 (SD = 107.3).

The subject terms comprised nouns and adjectival nouns (a Japanese lexical category to explain the trait or state of something) and can be analyzed in the same manner as nouns in English. There were 12,951 nouns with a total of 100,927 occurrences. In addition, some notational distortions were corrected, such as the unification of “world” and “all over the world” into “world.”

Analysis

Topic modeling using LDA. We used the Latent Dirichlet Allocation (LDA) algorithm for topic modeling to extract common topics from readers’ posts. LDA, the most widely used model for topic modeling (Lyu and Luli, 2021), was developed by Blei et al. (2003).

It provides an automatic or unsupervised means of summarizing large collections of documents. The exploratory algorithm is useful for discovering underlying topics within large bodies of text that is commonly referred to as *corpuses*. This inductive approach identifies topics that might not otherwise be anticipated (Chipidza et al., 2021). The goal is to compute the posterior probability given evidence, that is, the conditional distribution of topics, given documents within the corpus (Blei et al., 2003).

We applied the LDA algorithm to the data using the KH Coder (version 3. Beta. 04a), a software program for the Japanese language that has an R-language-based morphological system for statistical analysis (Higuchi, 2004). In the analysis, “COVID-19” and “Corona-Bruise” (Corona-Ka; コロナ禍), which is a term commonly used in Japan to encompass the negative impact of COVID-19 on people’s lives, appear too frequently, which expresses the difficult situation caused by the spread of COVID-19, and therefore, we did not include them in any of

the topics; these words were excluded because although they appeared in the database, they interfered with the feature-based topic extraction.

The LDA algorithm requires manually inputting the number of expected topics. The number of topics was determined by referring to the values in the evaluation indices of Arun et al. (2010), Cao et al. (2009), Deveaud et al. (2014), and Griffiths and Steyvers (2004) using the *ldatuning* package in R. For the Arun et al. (2010), Cao et al. (2009), and perplexity evaluation indices, the minimum value is the optimal one, while for the Deveaud et al. (2014) and Griffiths and Steyvers (2004) evaluation indices, the maximum value is the optimal one. KH Coder computes a topic score for each word. The ten words with the highest score for each topic were treated as words to be interpreted that represented that topic.

Statistical analysis. To examine how each topic mentioned differed during the emergency declarations, the Jonckheere-Terpstra test was applied using the topic ratio, which is a topic occurrence probability that sums to 1 for each post; afterward, Kendall rank correlation coefficients between the emergency declaration period and each topic were calculated to examine the effect size. Statistical analyses were performed using IBM SPSS Statistics 27 software.

In Japan, even with the prolonged spread of COVID-19, basic measures such as ensuring social distancing and hand sterilization remained the main countermeasures, and the only addition was vaccination. Additionally, in 2020 and 2021, Japanese infection and mortality rates were among the lowest in the developed world (Idogawa et al., 2020). This indicates that there was little change in social and infection contexts. Therefore, it is thought that Japan is an appropriate place to capture the trend of public opinion during the prolonged period of infection spread. In addition, there is the statistical challenge that *p*-values are more likely to be significant when sample sizes are large. Based on the above, in this study, Kendall rank correlation coefficients of 0.1 or higher were effective (Cohen, 1988; Field, 2005).

Results

The number of all posts related to COVID-19 published in each newspaper for each emergency declaration period is in Table 1. The age of the contributor was stated in 996 (52.2%) cases, with a median age of 62 (range: 8–98). We conducted a Chi-square test and the results revealed significant differences among newspapers and emergency declaration periods ($\chi^2(6) = 2.475, p = 0.116$).

The Griffiths and Steyvers (2004) and Deveaud et al. (2014) intersected at 10 to 11 topics, and Cao et al. (2009) had the lowest number of topics at 10; Arun et al. (2010) had the lower at 11 as compared to 10, but the climb range between 10 and 11 with the Cao et al. (2009) was greater. Therefore, the number of topics we selected was 10 (Fig. 1).

The ten topic names extracted were determined from the top ten words for each topic by discussion among the authors, considering the intention and context. To examine the trends in the occurrence of each topic during emergency declaration periods, the percentage of occurrence of each topic during each period was calculated. The topics were examined among the authors in

consideration of the top 10 words each, in the context of their treatment in the original text, and then summarized into three themes. The medians of each topic ratio by emergency periods, the results of the Jonckheere-Terpstra test, and Kendall rank correlation coefficients are indicated in Fig. 2.

Theme: life. The first theme is “Life” comprising “Family,” “Daily life in the COVID-19 disaster,” “Education in the COVID-19 disaster,” “The Importance of Humanity,” and “Daily life unrelated to COVID-19.”

“Family” included the words *mother, family, daughter, home, husband, hand, father, grandchild, vigor, and child*. These words were used in the context of the reaffirmation of the value of family interaction while COVID-19 spread, family members’ concerns about infection and changes in their lives, and family members’ communication of their opinions. “Daily life during the COVID-19 disaster” included *TV, time, mind, place, cheer, high school, baseball, figure, last, and activity*. During COVID-19, information has often been acquired from television, and people may be cheered by thinking of times other than the present (now vs. then) and distant activities. “Education during the COVID-19 disaster” included *children, school, classes, impact, teachers, college, online, university students, anxiety, and high school students and younger*. There was much interest in the changes in student life, especially school closures and online schooling during the pandemic, and how children were affected by it. “The importance of humanity” involved the words *people, myself, life, words, voice, work, feelings, life, talk, and thoughts* in the context of the reaffirmation of the importance of people, including the self, and of our lives, activities, and communication during COVID-19. “Daily life unrelated to COVID-19” included *eyes, flowers, mind, body, likes, newspapers, friends, blood donation, books, and fun*. The respondents described thinking about things unrelated to the pandemic, such as natural scenery and their daily life before COVID-19.

p-values calculated using the Jonckheere-Terpstra test were 0.004, 0.044, 0.017, <0.000, and <0.000, respectively. Kendall rank correlation coefficients were −0.050, 0.034, −0.041, −0.057, and −0.060, respectively.

Theme: awareness of the emergency. “Awareness of the emergency” comprises “Awareness of being a party to an emergency” and “Concerns about the medical environment”.

The top ten words for “Awareness of being a party to an emergency” were *infection, spread, masks, self-restraint, emergency declaration, going out, infectious disease control, infected people, prevention, and nationwide*. This topic was named because it was the most common during the declaration of the first state of emergency, and it referred to information gathering and dissemination, and preventive actions under the emergency precipitated by the COVID-19 outbreak. “Concerns about the medical environment” included the words *vaccination, vaccine, medical, elderly, hospital, home, anxiety, patient, nursing, and appointment* in the context of the pressure on the medical care system, delays in vaccination, and the conditions faced by patients, especially elderly ones. We labeled this topic because it

Table 1 Number of COVID-19-related reader submissions/all reader submissions during the emergency declaration period.

	First period		Second period		Third period		Fourth period		p-value
Yomiuri Newspaper	123/229	(53.7%)	164/356	(46.1%)	139/281	(49.5%)	133/396	(33.6%)	0.116
Mainichi Newspaper	123/337	(36.5%)	165/519	(31.8%)	151/405	(37.3%)	192/546	(35.2%)	
Asahi Newspaper	166/490	(33.9%)	192/609	(31.5%)	168/439	(38.3%)	194/640	(30.3%)	

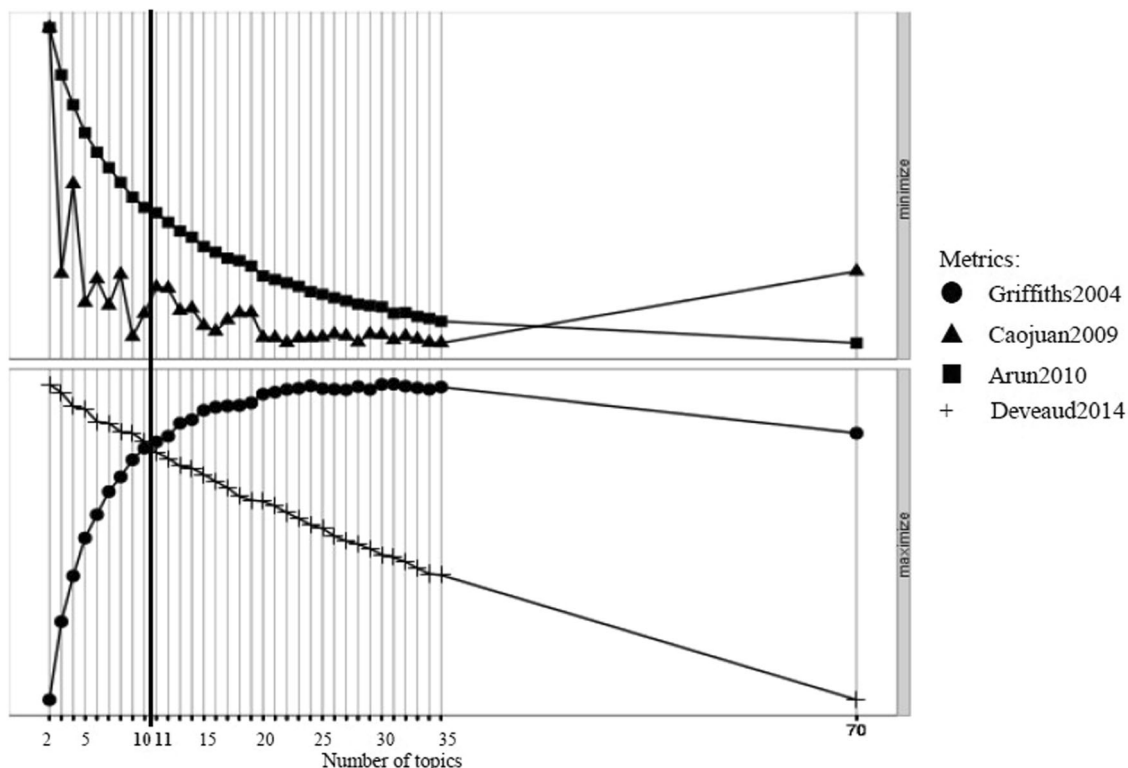


Fig. 1 Indicators of the number of topics by ldatuning. ●Griffiths and Steyvers (2004) ▲Cao et al. (2009) ■Arun et al. (2010) + Deveaud et al. (2014). Line 10 was thickened to increase readability.

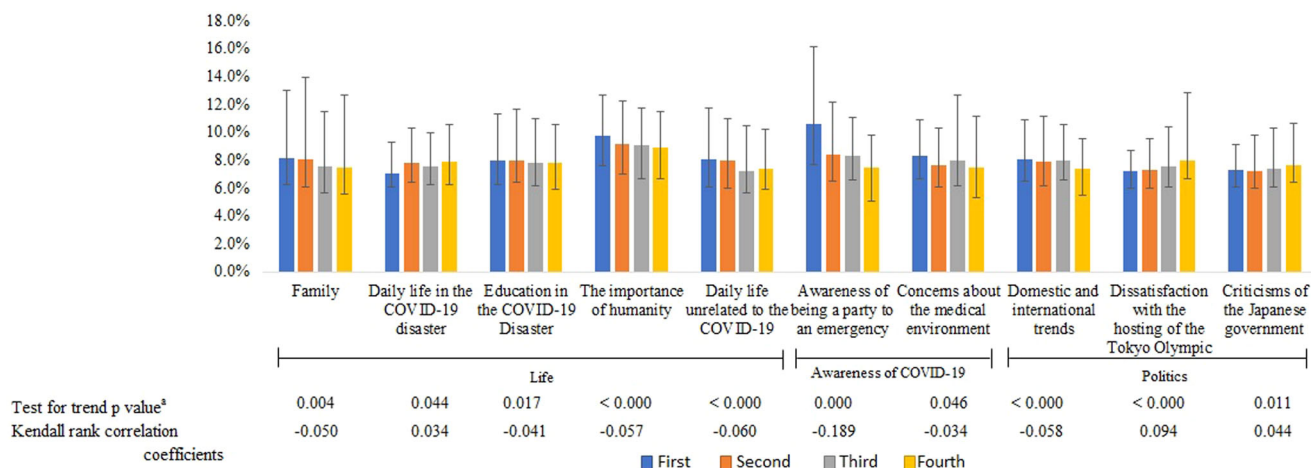


Fig. 2 Topic ratio of reader's posts per emergency declaration. p-values from the Jonckheere-Terpstra trend test for continuous variables. Error bar means 25-75%.

appears to refer to concerns about overall healthcare and was not limited to the pandemic.

p-values calculated using the Jonckheere-Terpstra test were 0.000 and 0.046, respectively. Kendall rank correlation coefficients were -0.189 and -0.034, respectively.

Theme: policy. “Policy” comprises “Domestic and foreign policies,” “Opposition to hosting the Tokyo Olympics,” and “Criticisms of the Japanese government.”

“Domestic and foreign policies” words included *Japan, U.S., society, world, economy, support, government, country, need, and company*, in the context of COVID-19 countermeasures and economic support in Japan and abroad, including in the U.S.

Accordingly, the topic was labeled. “Opposition to hosting the Tokyo Olympics” included the words *holding, Tokyo Olympics, athletes, canceling, Olympics, Tokyo, Japan, games, world, and spectators* in the context of the writers’ opposition to or concerns about the holding the Tokyo Olympics in 2021 or involving spectators. “Criticisms of the Japanese government” included the words *people, politics, prime minister, government, measures, country, diet, Suga (prime minister), information, and lawmakers*. Criticisms were primarily directed at the government and Prime Minister Suga personally.

p-values calculated using the Jonckheere-Terpstra test were <0.000, <0.000, and 0.011, respectively. Kendall rank correlation coefficients were -0.058, 0.094, and 0.044, respectively.

Public opinion during the prolonged spread of infection. The Jonckheere-Terpstra test showed a certain trend with the repetition of emergency declarations on all topics: “Family” was 0.004, “Daily life in the COVID-19 disaster” was 0.044, “Education in the COVID-19 disaster” was 0.017, “The importance of humanity” was <0.000 , “Daily life unrelated to the COVID-19” was <0.000 , “Awareness of being a party to an emergency” was 0.000, “Concerns about the medical environment” was 0.046, “Domestic and international trends” was >0.000 , “Dissatisfaction with the hosting of the Tokyo Olympic” was <0.000 , and “Criticisms of the Japanese government” was 0.011. Kendall rank correlation coefficients of “Family” was -0.050 , “Daily life in the COVID-19 disaster” was 0.034, “Education in the COVID-19 disaster” was -0.041 , “The importance of humanity” was -0.057 , “Daily life unrelated to the COVID-19” was -0.060 , “Awareness of being a party to an emergency” was -0.189 , “Concerns about the medical environment” was -0.034 , “Domestic and international trends” was -0.058 , “Dissatisfaction with the hosting of the Tokyo Olympic” was 0.094, and “Criticisms of the Japanese government” was 0.044. The results for each paper are shown in Supplementary Figs. S1 to S3.

Discussion

The purpose of this study was to provide new insights into people’s interests and communications during a particularly stressful period when restrictions were imposed on various activities to prevent the spread of COVID-19. We targeted readers’ posts in three major Japanese newspapers as opposed to SNS. Moreover, we sought to clarify how interests and communications changed over the prolonged pandemic. The results showed that more everyday topics could be extracted from the newspaper readers’ posts and that as the pandemic prolonged, people mentioned topics related to their awareness of the parties involved less frequently, however, their mentions of other topics remained unchanged.

Topic overview. In this study, topics such as *outbreak, pandemic*, and *source of novel coronavirus*, which were found in previous studies (Abd-Alrazaq et al., 2020; Maryam, 2022; Miao et al., 2020; Han et al., 2020; Wawrzuta et al., 2021; Zhong, 2021), were not found. Those topics are acute and were of particular interest in the early stages of COVID-19. Newspapers often publish information from more specialized and reliable sources such as specialists and doctors, compared to SNS, and due to the functional limitation of the reader’s contribution section of a newspaper, it does not allow for citation of scientific information in a simple way, such as posting a link on an SNS. As such, it may be more difficult to have a scientific discussion in the reader’s post column of a hardcopy newspaper. This suggests that SNS are more appropriate than newspapers for gathering public opinion data on rapidly occurring or specialized topics.

In contrast, topics such as “Family,” “Daily life during the COVID-19 disaster,” “The importance of humanity,” and “Daily life unrelated to COVID-19” were not found in studies targeting posts on SNS. Possible reasons for this include differences between newspaper and SNS users as mentioned in the Introduction, functional differences such as the number of characters in a post and the citation function, and whether posts are made anonymously or not. The average of 376.7 characters in the posts analyzed in this study is more than 1.3 times greater than the 280 characters per post allowed on Twitter; that is, more information can be placed in a single post by a newspaper reader. In other words, SNS may not be able to include the keywords “COVID-19” and a description of their daily life in a single post, thus omitting them from the search. It is possible that anonymity

may influence the content of posts, as anonymous Internet posts are more likely to be unsocial (Santana, 2014). The general public often withholds their personal information on SNS, however, in the case of reader’s posts to the newspapers in this study, most of their personal information, such as names, occupations, and ages was clearly stated. In a newspaper reader’s column, everyday topics may be more likely to be chosen because contributors consider content that will not cause problems if their names are known to others. In view of the above, to capture more of people’s daily lives, it may be appropriate to analyze newspaper readers’ contributions.

Long-term trends in public opinion. This research showed that public opinion changes over prolonged infection duration, as all topics showed significant increases or decreases. There were decreasing trends in “Daily life in the COVID-19 disaster,” “Opposition to hosting the Tokyo Olympics,” and “Criticisms of the Japanese government,” while there were increasing trends in the other seven topics. These changes may be due to the fact that people have lived in a COVID-19 environment for a prolonged period of time. It is also possible that the discussion became more active as the Tokyo Olympics approached, and that the prolonged pandemic itself led to a growing distrust of the government, which was leading the countermeasures. However, this study used 1, 910 submissions, which could easily have been statistically significant. Therefore, changes were examined using effect sizes and it was determined whether they were significant or not. The effect sizes of topics were examined by the Kendall rank sum correlation coefficient, which was less than 0.1 except for “Awareness of being party to an emergency,” making it difficult to say whether there was a meaningful trend and that public opinion did not change significantly with prolonged infection. In other words, the opinions of people’s daily concerns and interests, which were observed from the early stages of the outbreak, did not change much as the pandemic progressed, with the exception of “Awareness of being a party to an emergency.”

“Awareness of being party to an emergency” declined, reflecting a decrease in mentions of preventive action and infection status, possibly reflecting pandemic fatigue. Pandemic fatigue is the demotivation to take recommended preventive actions, which emerged gradually (WHO, 2020), and a decrease in awareness of the parties involved is also related to a decrease in preventive action (Bandura, 1977; Forsyth and McMillan, 1981). This study suggests that only “Awareness of being a party to an emergency” decreased, while overall interest in COVID-19 itself did not decrease. Pandemic fatigue is affected by a variety of emotions, experiences, perceptions, gender, age, and psychological loneliness (Sulemana et al., 2023; Rodriguez-Blazquez et al., 2022). While maintaining awareness is important for infection prevention, we must be aware that it does not always stem from indifference.

In contrast, increases and decreases in other everyday topics were limited. In Japan, various restrictions were imposed after the infection spread, including long-distance travel restrictions, restrictions on restaurants and large-scale commercial facilities, and the simultaneous closure of educational institutions. However, over time and with an increased understanding of COVID-19, the country experienced a relaxation of travel restrictions, limited reopening of facilities and events, and an increase in online classes and teleworking (Karako et al., 2021), and the situation was considered stable. However, nearly 60% of the respondents still felt some form of increased anxiety than before the pandemic (Supplementary Fig. S2: Cabinet Office, Government of Japan, 2021¹), suggesting that stability in a stressful environment did not solve or eliminate the challenges and

concerns that people had. This study, which examined such long-term public opinions, highlighted that even after repeated emergency declarations, people confirmed the importance of family and humanity and their interest in education. Concerns about the medical environment did not decrease, and their dissatisfaction with the government did not dissipate. In other words, the everyday public opinion since the early stages of COVID-19 may not be resolved as long as the pandemic continues. While SNS may be more effective in identifying acute opinions currently, analysis of newspaper readers' columns can extract public opinions about everyday aspects of people's lives that have existed since the early stages of the outbreak and over a long period of time. Therefore, in addition to immediate responses to acute opinions, it is also important to extract everyday opinions and take measures to deal with them from a long-term perspective during a prolonged pandemic.

Limitations. This study has limitations concerning the generalizability of the findings. The data used were readers' posts in three major Japanese newspapers which were selected by each newspaper's editors. Not all submissions were guaranteed to be published, and therefore it is possible that posts may have been biased toward a given newspaper's editorial policies or political perspective in an attempt to ensure publication. However, this process has not been clarified. For these reasons, the findings cannot be generalized to Japanese society, newspaper readers, or even to those who sent their comments. There is also concern regarding the analysis as LDA is dependent on various factors such as data quality and quantity, appropriate parameter selection, and accurate interpretation of results which can lead to variations in the topics extracted. Further accumulation of research is required for resolving these limitations.

Conclusion

The purpose of this study was to obtain clues for a comprehensive understanding of COVID-19-related public opinion from a long-term analysis of newspaper readers' post columns. The results of this study indicate that, in contrast to research that used SNS data which can extract acute topics, the analysis of newspaper readers' posts can extract everyday topics, and that these topics exist continuously during the pandemic. To comprehensively understand public opinion during the pandemic and appropriately respond to it, we think that we can collect information through various methods using both SNS and other media, analyze it from acute/short-term and daily/long-term perspectives, and respond more effectively.

Data availability

The datasets generated during and/or analyzed during this study are available in the Dryad repository, <https://doi.org/10.5061/dryad.8w9ghx3q2>.

Received: 29 December 2022; Accepted: 4 August 2023;

Published online: 14 August 2023

Note

1 The 4th Survey on Changes in Attitudes and Behaviors under the Influence of New Coronavirus Infections (Cabinet Office, Government of Japan, 2021) is published only in Japanese, so we have translated the survey into English as Supplementary Fig. S2.

References

Abd-Alrazaq A, Alhuwail D, Househ M et al. (2020) Top concerns of tweeters during the COVID-19 pandemic: infoveillance study. *J Med Intern Res* 22:e19016. <https://doi.org/10.2196/preprints.19016>

- Akashi H, Shimada S, Tamura T et al. (2022) SARS-CoV-2 infections in close contacts of positive cases in the olympic and paralympic village at the 2021 Tokyo olympic and paralympic games. *J Am Med Assoc* 327:978–980. <https://doi.org/10.1001/jama.2022.0818>
- Arun R, Suresh V, Veni Madhavan CE et al. (2010) On finding the natural number of topics with Latent Dirichlet Allocation: Some observations. In: Zaki MJ, Yu JX, Ravindran B, et al. (eds) *Advances in knowledge discovery and data mining*. Springer, Berlin, pp. 391–402
- Asahi Newspaper database (2022) <http://database.asahi.com/index.shtml>. Accessed 1 Apr 2022
- Bandura A (1977) *Self-efficacy: the exercise of control*. W. H. Freeman and Company, New York
- Blei DM, Ng A, Jordan MI (2003) Latent Dirichlet Allocation. *J Mach Learn Res* 3:993–1022. <https://doi.org/10.1162/jmlr.2003.3.4-5.993>
- Cabinet Office, Government of Japan (2021) The 4th Survey on Changes in Attitudes and Behaviors under the Influence of New Coronavirus Infections (In Japanese). p. 26. https://www5.cao.go.jp/keizai2/wellbeing/covid/pdf/result4_covid.pdf. Accessed 5 Jun 2023
- Cao J, Xia T, Li J (2009) A density-based method for adaptive LDA model selection. *Neurocomputing* 72:1775–1781. <https://doi.org/10.1016/j.neucom.2008.06.011>
- Chipidza W, Akbaripourdebazar E, Gwanzura T et al. (2021) Topic analysis of traditional and social media news coverage of the early COVID-19 pandemic and implications for public health communication. *Disast Med Public Health Preparedness* 1–8. <https://doi.org/10.1017/dmp.2021.65>
- Cohen J (1988) *Statistical power analysis for the behavioral sciences*, 2nd edn. Lawrence Erlbaum, NJ
- Dalton CB, Taylor J (2021) Are COVID-19-safe Tokyo Olympics and Paralympics really possible? *Med J Aust* 215:54–55. <https://doi.org/10.5694/mja2.51159>
- Deveaud R, Sanjuan E, Bellot P (2014) Accurate and effective latent concept modeling for ad hoc information retrieval. *Document numérique* 17:61–84. <https://doi.org/10.3166/dn.17.1.61-84>
- Field A (2005) *Discovering statistics using SPSS*, 2nd edn. Sage Publications, London
- Forsyth DR, McMillan JH (1981) Attributions, affect, and expectations: a test of Weiner's three-dimensional model. *J Educ Psychol* 73:393–403. <https://doi.org/10.1037/0022-0663.73.3.393>
- Griffiths TL, Steyvers M (2004) Finding scientific topics. *Proc Natl Acad Sci USA* 101:5228–5235. <https://doi.org/10.1073/pnas.0307752101>
- Haerpfer C, Inglehart R, Moreno A et al. (2022) In: Haerpfer C et al. (eds.). *World Values Survey: Round Seven - Country-Pooled Datafile Version 5.0*. Madrid, Spain & Vienna, Austria: JD Systems Institute & WVSA Secretariat. p. 564. <https://doi.org/10.14281/18241.20>
- Han X, Wang J, Zhang M et al. (2020) Using social media to mine and analyze public opinion related to COVID-19 in China. *Int J Environ Res Public Health* 17:2788. <https://doi.org/10.3390/ijerph17082788>
- Hidaka T, Endo S, Kasuga H (2022) Visualizing the decline of public interest in the Great East Japan Earthquake and Fukushima Daiichi nuclear power plant accident by analyzing letters to the editor in Japanese newspapers (in press). *Fukushima J Med Sci*. <https://doi.org/10.5387/fms.2021-18>
- Higuchi K (2004) Quantitative analysis of textual data: differentiation and coordination of two approaches. *Sociol Theor Method* 19:101–115
- Idogawa M, Tange S, Nakase H et al. (2020) Interactive web-based graphs of coronavirus disease 2019 cases and deaths per population by country. *Clin Infect Dis* 71:15. <https://doi.org/10.1093/cid/ciaa500>
- Karako K, Song P, Chen Y et al. (2021) Overview of the characteristics of and responses to the three waves of COVID-19 in Japan during 2020–2021. *BioScience Trends* 15(1):1–8. <https://doi.org/10.5582/bst.2021.01019>
- Lyu JC, Luli GK (2021) Understanding the public discussion About the Centers for Disease Control and Prevention During the COVID-19 pandemic using twitter data: text mining analysis study. *J Med Intern Res* 23:e25108. <https://doi.org/10.2196/25108>
- Mainichi Newspaper database (2022). <https://dbs.g-search.or.jp/WMNP/NPerr5.html>. Accessed 1 Apr 2022
- Maryam M (2022) Predicting the popularity of tweets by analyzing public opinion and emotions in different stages of Covid-19 pandemic. *International Journal of Information Management Data Insights* 2. <https://doi.org/10.1016/j.jjime.2021.100053>
- Menzies RE, Menzies RG (2020) Death anxiety in the time of COVID-19: Theoretical explanations and clinical implications. *Cogn Behav Therap* 13:e19. <https://doi.org/10.1017/S1754470X20000215>
- Miao L, Last M, Litvak M. (2020) Twitter Data Augmentation for Monitoring Public Opinion on COVID-19 Intervention Measures. In *Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020*, Online. Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.nlp-covid19-2.19>
- Ministry of Health, Labor and Welfare (2022) Novel Coronavirus (COVID-19). https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000164708_00079.html. Accessed 1 Apr 2022

- Ministry of Internal Affairs and Communications (2022) Digital usage trends in the daily life of the public, Information and Communications in Japan (White Paper 2022), pp. 79–81. <https://www.soumu.go.jp/johotsusintokei/whitepaper/eng/WP2022/2022-index.html>
- Morgan-Lopez AA, Kim AE, Chew RF et al. (2017) Predicting age groups of Twitter users based on language and metadata features. *PLoS ONE* 12:e0183537. <https://doi.org/10.1371/journal.pone.0183537>
- Nomura S, Yoneoka D, Shi S et al. (2020) An assessment of self-reported COVID-19 related symptoms of 227,898 users of a social networking service in Japan: has the regional risk changed after the declaration of the state of emergency? *Lancet Region Health West Pac* 1:100011. <https://doi.org/10.1016/j.lanwpc.2020.100011>
- Okubo T (2020) Spread of COVID-19 and telework: evidence from Japan Covid Economics: Vetted and Real-Time Papers 32:1–25
- Plugh H (2018) Born again Yokozuna: sports and national identity. In: Darling-Wolf F (ed) *Routledge handbook of Japanese media*. Routledge, London, p 101–120
- Rodriguez-Blazquez C, Romay-Barja M, Falcon M, Ayala A, Forjaz MJ (2022) Psychometric properties of the COVID-19 pandemic fatigue scale: cross-sectional online survey study. *JMIR Public Health Surveill* 8(9):e34675. <https://doi.org/10.2196/34675>
- Santana AD (2014) Virtuous or vitriolic: the effect of anonymity on civility in online newspaper reader comment boards. *J Pract* 8:18–33. <https://doi.org/10.1080/17512786.2013.813194>
- Shen X, Li J, Dong T et al. (2021) Public opinion and expectations: development of public health education in China after COVID-19 pandemic. *Front Public Health* 9:702146. <https://doi.org/10.3389/fpubh.2021.702146>
- Sulemana A-S, Lal S, Nguyen TXT et al. (2023) Pandemic fatigue in Japan: factors affecting the declining COVID-19 preventive measures. *Sustainability* 15(7):6220. <https://doi.org/10.3390/su15076220>
- Taniguchi M (2018) Ghanging political communication in Japan. In: Darling-Wolf F (ed.) *Routledge handbook of Japanese media*. Routledge, London, pp. 121–135
- Wawrzuta D, Jaworski M, Gotlib J (2021) What arguments against COVID-19 vaccines run on facebook in Poland: content analysis of comments. *Vaccines* 9:481. <https://doi.org/10.3390/vaccines9050481>
- World Association of Newspapers and News Publishers (2016) World press trends 2016: facts and figures. Available from: <http://www.wptdatabase.org/world-press-trends-2016-facts-and-figures>. Accessed 1 Apr 2022
- World Health Organization (2010) H1N1 in post-pandemic period. <https://www.who.int/news/item/10-08-2010-h1n1-in-post-pandemic-period>. Accessed 1 Apr 2022
- World Health Organization (2017) Communicating risk in public health emergencies: a WHO guideline for emergency risk communication (ERC). <https://www.who.int/publications/i/item/9789241550208>. Accessed 1 Apr 2022
- World Health Organization (2020) Pandemic fatigue -Reinvigorating the public to prevent COVID-19. Available from: <https://apps.who.int/iris/bitstream/handle/10665/335820/WHO-EURO-2020-1160-40906-55390-eng.pdf>. Accessed 10 Aug 2023
- World Health Organization: severe acute respiratory syndrome (SARS): status of the outbreak and lessons for the immediate future (2003). http://www.who.int/csr/media/sars_waha.pdf. Accessed 1 Apr 2022
- Yomiuri Newspaper database (2022) <https://database.yomiuri.co.jp/rekishikan/>. Accessed 1 Apr 2022
- Yunusa I, Iloanusi S, Mgbere O et al. (2021) Public opinion regarding government response to COVID-19: case study of a large commercial city in Nigeria. *Pan Afr Med J* 38:282. <https://doi.org/10.11604/pamj.2021.38.282.26361>
- Zhao WX, Jiang J, Werng J et al. (2011) Comparing twitter and traditional media using topic models. In: Clough P, Foley C, Gurrin C, et al. (eds) *Advances in information retrieval*. ECIR 2011. Lecture notes in computer science, 6611. Springer, Berlin
- Zhong Z (2021) Internet public opinion evolution in the COVID-19 event and coping strategies. *Disast Med Public Health Prep* 15:e27–e33. <https://doi.org/10.1017/dmp.2020.299>

Acknowledgements

We would like to thank Editage [<http://www.editage.com>] for editing and reviewing this manuscript for the English language.

Competing interests

The authors declare no competing interests.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

Informed consent

This article does not contain any studies with human participants performed by any of the authors. We have neither attempted to reveal the personally identifiable information of contributors to the newspaper nor combined multiple sources to disclose some private information that is not publicly available on the newspapers' databases.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1057/s41599-023-02014-0>.

Correspondence and requests for materials should be addressed to Hideaki Kasuga.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



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

© The Author(s) 2023