Original Paper

Monitoring Mentions of COVID-19 Vaccine Side Effects on Japanese and Indonesian Twitter: Infodemiological Study

Kiki Ferawati, MStat; Kongmeng Liew, PhD; Eiji Aramaki, PhD; Shoko Wakamiya, PhD

Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, Japan

Corresponding Author:

Shoko Wakamiya, PhD Graduate School of Science and Technology Nara Institute of Science and Technology 8916-5, Takayama-cho Ikoma, 630-0192 Japan Phone: 81 743725250 Email: <u>wakamiya@is.naist.jp</u>

Abstract

Background: The year 2021 was marked by vaccinations against COVID-19, which spurred wider discussion among the general population, with some in favor and some against vaccination. Twitter, a popular social media platform, was instrumental in providing information about the COVID-19 vaccine and has been effective in observing public reactions. We focused on tweets from Japan and Indonesia, 2 countries with a large Twitter-using population, where concerns about side effects were consistently stated as a strong reason for vaccine hesitancy.

Objective: This study aimed to investigate how Twitter was used to report vaccine-related side effects and to compare the mentions of these side effects from 2 messenger RNA (mRNA) vaccine types developed by Pfizer and Moderna, in Japan and Indonesia.

Methods: We obtained tweet data from Twitter using Japanese and Indonesian keywords related to COVID-19 vaccines and their side effects from January 1, 2021, to December 31, 2021. We then removed users with a high frequency of tweets and merged the tweets from multiple users as a single sentence to focus on user-level analysis, resulting in a total of 214,165 users (Japan) and 12,289 users (Indonesia). Then, we filtered the data to select tweets mentioning Pfizer or Moderna only and removed tweets mentioning both. We compared the side effect counts to the public reports released by Pfizer and Moderna. Afterward, logistic regression models were used to compare the side effects for the Pfizer and Moderna vaccines for each country.

Results: We observed some differences in the ratio of side effects between the public reports and tweets. Specifically, fever was mentioned much more frequently in tweets than would be expected based on the public reports. We also observed differences in side effects reported between Pfizer and Moderna vaccines from Japan and Indonesia, with more side effects reported for the Pfizer vaccine in Japanese tweets and more side effects with the Moderna vaccine reported in Indonesian tweets.

Conclusions: We note the possible consequences of vaccine side effect surveillance on Twitter and information dissemination, in that fever appears to be over-represented. This could be due to fever possibly having a higher severity or measurability, and further implications are discussed.

(JMIR Infodemiology 2022;2(2):e39504) doi: 10.2196/39504

KEYWORDS

COVID-19; vaccine; COVID-19 vaccine; Pfizer; Moderna; vaccine side effects; side effects; Twitter; logistic regression

Introduction

Background

RenderX

Vaccinations have been proposed as one of the solutions to contain and end the COVID-19 pandemic [1-3]. Prior to their widespread deployment, an early Twitter poll suggested that

```
https://infodemiology.jmir.org/2022/2/e39504
```

public sentiment toward vaccinations was mostly positive, with many individuals indicating that they would seek vaccination, despite ongoing concerns about the safety of the vaccines [4]. However, these concerns persisted in the public eye, including issues like safety, the unusually quick development of vaccines, and possible side effects after administration [5]. These were also observed in vaccine-related search trends in early 2021,

when resulting side effects were identified as a significant area of concern [6].

The vaccine rollout evoked a diverse set of reactions from the general population, be it for or against vaccination. Bonnevie et al [7] investigated vaccine acceptance before the pandemic and in the middle of pandemic in 2020 and found stronger vaccine opposition on Twitter during the latter period. The trend appeared reversed in the study conducted by Lyu et al [8], who studied public perception and reactions toward COVID-19 vaccinations on Twitter through topic modeling and sentiment analysis and found that early discussions about vaccines stemmed from the development stage of the vaccines, and public sentiment leaned toward a positive outlook later on.

As one of the most popular social media platforms in use today, Twitter has been widely utilized as a source for research in COVID-19 infodemiology (see [9]), building on an extant body of literature on epidemic surveillance via that platform. For example, in the case of influenza, Twitter has been used as a detection tool to estimate individual diagnoses [10] and as social surveillance, functioning as an early warning tool for outbreak detection [11]. Despite the usefulness of Twitter for epidemiological surveillance studies, there are limitations, such as the spread of misinformation during the early periods of the COVID-19 pandemic [12] that may bias these data. In this paper, we proposed that Twitter can also be used in a similar fashion to monitor side effects from COVID-19 vaccination, by focusing on the Pfizer and Moderna vaccines in Japan and Indonesia.

Attitudes Toward COVID-19 Vaccination

Most studies on COVID-19 vaccination and Twitter have focused on general collective attitudes toward vaccination. Marcec and Likic [13] applied lexicon-based sentiment analysis to English tweets mentioning AstraZeneca, Pfizer, and Moderna and found that the sentiment for Pfizer and Moderna was generally more positive than that for AstraZeneca. Sattar and Arifuzzaman [14] analyzed tweets related to public sentiment about COVID-19 vaccination awareness and found strong positive sentiments despite the side effects of the vaccine. Kwok et al [15] used latent Dirichlet allocation topic modeling to identify topics in tweets related to COVID-19 vaccination in Australia, applied sentiment analysis to the tweets, and found that counts of tweets with positive sentiment were only slightly larger than counts of tweets with negative sentiment, thereby raising concerns over widespread vaccine acceptance. Yet, most of these studies were with Western and English samples, and there has been a considerable lack of similar research for tweets in non-English languages.

In this paper, we focused on tweets concerning widely available messenger RNA (mRNA) vaccines in Japan and Indonesia, 2 island countries located in the Pacific that have a large Twitter-using population (top 10 in terms of Twitter users globally [16]). As the Japanese and Indonesian languages are largely ubiquitously spoken in their respective countries [17,18], it provides for a relatively controlled environment to observe patterns unique to each country. This allows for a contextualization of Twitter usage to the wider society for added interpretations and behavioral analyses, especially in this pandemic.

https://infodemiology.jmir.org/2022/2/e39504

The year 2021 saw the adoption of COVID-19 vaccinations on a global scale. Vaccination for health care workers in Japan started in February 2021, and vaccinations for the elderly started in April 2021 [19]. The vaccination rate started picking up quickly in June 2021 and continued to rise until over 80% were fully vaccinated by the end of 2021 [20]. Meanwhile, vaccination in Indonesia started in January 2021, with health care workers as a priority, followed by the elderly and public officers, and finally for the general public [21]. Although the early vaccination campaign used the Coronavac and AstraZeneca vaccines, in August 2021, the Moderna and Pfizer vaccines started to be administered for vaccinations and boosters in the country. By the end of 2021, 46.7% of the population of Indonesia were fully vaccinated.

One key context behind vaccine hesitancy identified in both cultures is the role of side effects. In Japan, concerns about adverse side effects were arguably the main reason for vaccine hesitancy, alongside other factors like gender, living arrangements, economic status, and psychological issues [22]. Vaccine hesitancy was also found to be significantly more frequent in the younger generation than in the older generation [23]. Meanwhile, in Indonesia, concerns about vaccine safety, distrust toward the vaccine, and concerns about side effects were identified as a few common reasons for vaccine hesitancy [24]. As Twitter has been used to identify symptoms of COVID-19 [25], we proposed for this paper that it can also be utilized to examine side effects of COVID-19 vaccination. Moreover, by comparing side effect counts from Twitter with rates reported in phase 3 clinical trials of the Pfizer and Moderna vaccines, we can observe patterns in information dissemination on side effects that may be unique to Twitter (eg, are side effects overrepresented, appropriately, or underrepresented when mentioned on Twitter). These results may then determine the usefulness of Twitter for vaccine side effect monitoring or alternatively illustrate misinformation biases that are present on the platform.

Finally, we examined if there were differences in side effect reporting between the Pfizer and Moderna vaccines found in tweets by country. This is because publicly available research concerning vaccine (and maker-specific) side effects in both countries is still limited, and to our knowledge, only one study by Kitagawa et al [26] compared the side effects of the Moderna and Pfizer vaccines available in Japan through self-reported data. Analyses were conducted separately for Japan and Indonesia.

Methods

Data

Tweet Collection

To get a general sense of public opinions for the vaccination campaigns in Japan and Indonesia, tweets in Japanese (ja) and Indonesian (id; based on Twitter's language filter) were collected for the whole of 2021, from January 1, 2021, to December 31, 2021 (UTC). The search query comprised keywords for vaccines and side effects and excluded retweets (see Table 1). Of all the vaccines used in both countries, we

limited the query to Moderna and Pfizer because these 2 vaccines were used in both countries. Although AstraZeneca's vaccine was also used in both countries, it was much less common in Japan and was represented by a variety of names in the public sphere, so relevant tweets were even more difficult to obtain.

All data were obtained using the Python Twarc library (version 2.8.3) for Academic Research Access in Twitter [27]. As limitations on tweet quota restricted access to vast amounts of Japanese tweets, we were unable to obtain tweets that mentioned "vaccine" only or "side effects" only. Consequently, the vaccine keywords used for Japanese and Indonesian tweet scraping included the vaccine type ("Pfizer" and "Moderna"), and the side effect keywords followed the symptoms described by the Centers for Disease Control and Prevention in English: tiredness,

headache, muscle pain, chills, fever, and nausea [28]. Similarly, a list of symptoms was also available on the Indonesian government's official webpage for effect after COVID-19 vaccination, which is abbreviated as *kipi* (in Indonesian: *kejadian ikutan pasca imunisasi*) [29]. Corresponding sources from the Japanese government also mentioned the same symptoms, excluding nausea but including diarrhea [30]. We decided to exclude diarrhea since it was not listed on the Indonesian source and the detailed prevalence is not available in Moderna public reports. However, we decided to retain nausea in this study because of the availability of the corresponding statistics, and it was also referenced by some research about vaccine side effects in Japan [26,31]. This list of symptoms was then translated into Japanese and Indonesian, with additional keywords added from synonyms (see Table 1).

Table 1. List of keywords related to COVID-19 vaccines.

Terms ^a	Keywords (delimited by commas)						
	Japanese	Indonesian					
Vaccine		·					
Vaccine-related	ファイザー, モデルナ	pfizer, moderna					
Side effects							
Side effects	副反応	efek, kipi					
Tiredness	疲労, 疲れ, 倦怠感, だるい, だるさ	lelah, capai, capek, pegal, lemas, letih					
Headache	頭痛,頭が痛い	pusing, sakit kepala					
Muscle pain	筋肉痛	nyeri otot					
Chills	寒気, 悪寒, さむけ	meriang, menggigil					
Fever	熱, 高熱, 微熱, 発熱, 熱が高い, 熱があった, 熱がある, 熱 が出た, 熱風邪	demam, panas					
Nausea	悪心, 吐, 嘔吐, おう吐, 気分悪	mual					

^aTranslated into English.

Public Report Data

The comparison percentages listed in this paper were obtained from publicly available reports (press releases) published by Pfizer [32] and Moderna [33]. For Pfizer, these included data from participants aged 16 years to 55 years, and for Moderna, these included data from participants aged 18 years to 64 years. Both were collected for 7 days after the vaccination and classified as systemic adverse reactions.

Preprocessing of Tweet Data

For tweets from both languages, the initial preprocessing steps were removing usernames and web links. Afterward, for Japanese tweets, we removed emojis and special characters (such as Japanese punctuation). Tweets were then tokenized using mecab-ipadic-NEologd [34-36], which reduced terms into their simplest forms to facilitate further analyses. All keyword filtering was done using full-width characters. For Indonesian tweets, all characters were set into lowercase, and non-ASCII characters were removed.

To assess the 2 vaccines separately, we filtered tweets to select tweets with the term Pfizer or Moderna only. Tweets mentioning

https://infodemiology.jmir.org/2022/2/e39504

both vaccines were removed and excluded from the analyses. Next, we defined user accounts with more than 10 tweets in our data as "high frequency users." We removed these high frequency users to avoid having data biased by excessive tweet counts from the same individual. We then grouped tweets by user account, focusing on user-level analyses. If a user had more than one tweet, they were merged into a single sentence for the analyses. This was to reduce bias arising from the same individual tweeting their side effects multiple times over different tweets.

Following that, tweets with the term "Pfizer" were coded as 1, while tweets with the term "Moderna" were coded as 0. As mentioned earlier and considering the respective timelines for vaccinations, we excluded tweets mentioning both types (Pfizer and Moderna) and filtered them out at this stage. The resulting variable from this step served as the outcome variable for the logistic regression analysis.

The sample of filtered tweets is shown in Textbox 1. The tweet samples were paraphrased due to Twitter's privacy policy. The Tweet ID for the data set processed in this study is available in Multimedia Appendix 1 (Japanese tweets) and Multimedia

```
XSL•FO
RenderX
```

Appendix 2 (Indonesian tweets). In each filtered tweet, we applied word matching of the specified keywords to the merged tweets. If the word for side effect was present, then the column was marked as "1," and if it was not, then it was marked as "0." There were 7 predictor variables in total: effect, tiredness, headache, muscle pain, chills, fever, and nausea. The presence

Textbox 1. Examples of filtered tweets.

Japanese tweet: こんばんは皆様。2回目のファイザーワクチンの接種後、夜から次の日にかけて、微熱と頭痛と接種部位の痛みを感じます。若い人の方が熱が出やすいと思われます。 # ファイザー # コロナワクチン (Good evening everyone. After the second dose of Pfizer vaccine, I feel a slight fever, headache and pain at the vaccination site during the night and the next day. It seems that younger people are more prone to fever. #pfizer #coronavaccine)

Indonesian tweet: *efek dosis kedua vaksin moderna membuat aku menangis karena sakit demam, menggigil, badan terasa nyeri dan sakit kepala* (side effect of second dose of Moderna made me cry in pain with fever, chills, sore and headache)

Logistic Regression Analysis

Logistic regression analysis is a statistical method to analyze associations with a binary outcome variable [37]. In this study, we conducted separate logistic regression models by country: Japan and Indonesia. The outcome variable was the vaccine type (Pfizer or Moderna), and the predictor variables were the identified side effects: side effect, tiredness, headache, muscle pain, chills, fever, and nausea (from the tweets). We then examined the likelihood of a specific side effect for each vaccine type by using the odds ratio obtained from the model parameter. A significance level of 5% was used to construct the confidence interval for the odds ratios. The model was evaluated using the Nagelkerke R^2 [38]. All statistical analyses were conducted using SPSS version 28.0.1 (IBM Corp, Armonk, NY). All variables included in the analyses were binary.

Ethical Considerations

This study did not require participants to be involved in any physical or mental intervention. As this research did not use personally identifiable information, it was exempt from institutional review board approval in accordance with the Ethical Guidelines for Medical and Health Research Involving Human Subjects stipulated by the Japanese national government.

Results

Comparisons With Public Report Data (Press Releases) From Clinical Trials

of each respective side effect was checked using exact word

matching. For example, based on the Japanese tweet in Textbox

1, the columns for fever and headache would be marked as 1,

while the rest of the side effects would be 0. Mentions of "pain" were not classified as "muscle pain," so it was marked as 0.

The final data set used in this research included 286,887 Japanese tweets from 214,165 users and 14,484 Indonesian tweets from 12,289 users. Table 2 shows the final tweet count after merging and the detailed breakdown for each side effect. For the final data set, the mean number of tweets per user for Japanese was less than the mean number of tweets per user for Indonesian, as shown in Table 2. However, since we aggregated tweets by user, we focused on user counts from merged tweets for subsequent analyses. The proportions of tweets about the Pfizer vaccine and Moderna vaccine in the Indonesian data set were 58.80% and 41.20%, respectively, with more tweets mentioning Moderna. For the Japanese data, 98.47% of the overall set of tweets mentioned Pfizer. The proportions of Pfizer and Moderna shots administered in Japan at the end of 2021 were 79.85% and 20.08%, respectively, and 0.08% for others (AstraZeneca) [39]. We were unable to access comparable statistics for Indonesia.

Table 2. Counts from the tweets.

	Japanese			Indonesian		
	Pfizer	Moderna	Total	Pfizer	Moderna	Total
Tweets, n	283,530	3357	286,887	5684	8800	14,484
Number of tweets per user, mean (SD)	1.344 (0.890)	1.028 (0.221)	1.339 (0.885)	1.122 (0.449)	1.218 (0.606)	1.178 (0.549)
Individual users (after merging tweets), n	210,899	3266	214,165	5063	7226	12,289
Individual users who mentioned side e	effects from any	vaccine, n				
Side effects	101,794	1224	103,018	1535	4628	6163
Tiredness	39,724	406	40,130	375	356	731
Headache	34,878	398	35,276	457	858	1315
Muscle pain	25,167	339	25,506	9	37	46
Chills	9361	143	9504	248	780	1028
Fever	131,172	1993	133,165	2089	2386	4475
Nausea	5334	49	5383	105	396	501

https://infodemiology.jmir.org/2022/2/e39504

JMIR Infodemiology 2022 | vol. 2 | iss. 2 | e39504 | p. 4 (page number not for citation purposes)

Figure 1 displays the side effects reported in press releases by Pfizer and Moderna, followed by the percentages of tweets observed in our study. The figure contains the side effects obtained from the word matching in tweets and the comparison with public report data to illustrate the ability of tweet data to capture the side effects. There appeared to be a difference between the percentage of side effects reported by press releases and that reported in tweet data and a slight difference between the percentages in Japanese and Indonesian tweets.

In the public reports, side effects were more frequently observed with the second dose than with the first dose. However, we lacked comparable information from the tweets to similarly differentiate side effects between the first dose and the second dose in our data set. Except for fever, we noticed that all side effects were reported more frequently in the public reports (higher percentages) than in the tweets. A radar graph showing the comparison of each side effect for the first dose and second dose in the public reports versus those obtained from the tweets is available in Multimedia Appendix 3, with the value plotted corresponding to the side effect counts in Table 2 and percentages from the values shown in Figure 1. Looking into the tweet comparisons for Japanese and Indonesian tweets, for both Pfizer and Moderna, several side effects such as tiredness, muscle pain, and fever were reported at higher percentages in Japanese tweets than in Indonesian tweets. The percentages of headache, chills, and nausea were also different, with higher percentages in Indonesian tweets. Regardless of vaccine type, fever was by far the most reported side effect in tweets.

The percentages of all the side effects in Japanese tweets for the Pfizer vaccine were slightly higher than those for the Moderna vaccine, even when the total number of tweets for the 2 were notably different. On the other hand, in Indonesian tweets, the percentages of side effects with the Moderna vaccine were higher. We also noticed that the percentage reported for muscle pain was really small, which was probably caused by an inappropriate word choice used to represent this type of pain. For example, not many users may have been able to locate the exact part of the body from which the pain originated. A more general term "pain" may have been more suitable to represent this side effect than the specific term "muscle pain."

Figure 1. Percentage of side effects experienced after vaccination, as obtained from public reports by Pfizer and Moderna, compared with the percentages found in our Twitter sample. The overall percentage of general side effects was not mentioned in Pfizer's report, and the percentage of side effects in tweets was calculated from the number of tweets for specific side effects divided by the total number of filtered tweets. 1st: first vaccine dose; 2nd: second vaccine dose; ID: Indonesian tweets; JA: Japanese tweets.

		Pfizer			Moderna			
	Public	report	Tw	eets	Public	report	Two	eets
Side effects			48.27%	21.24%	57.00%	81.90%	37.48%	91.41%
Tiredness	1 st 47.00%	2 nd 59.00%	JA 18.84%	ID 5.19%	1 st 38.40%	2 nd 67.60%	JA 12.43%	ID 7.03%
Headache	1 st 42.00%	2 nd 52.00%	JA 16.54%	ID 6.32%	1 st 35.30%	2 nd 62.80%	JA 12.19%	ID 16.95%
Muscle pain	1 st 21.00%	2 nd 37.00%	JA 11.93%	ID 0.12%	1 st 23.70%	2 nd 61.60%	JA 10.38%	ID 0.73%
Chills	1 st 14.00%	2 nd 35.00%	JA 4.44%	ID 3.43%	1 st 9.20%	2 nd 48.60%	JA 4.38%	ID 15.41%
Fever	1 st	2 nd 16.00%	JA 62.20%	ID 28.91%	0.90%	2 nd 17.40%	JA 61.02%	ID 47.13%
Nausca	1 st	2 nd 2.00%	JA 2.53%	ID 1.45%	1 st 9.40%	2 nd 21.40%	JA 1.50%	ID 7.82%
	1 st	2 nd	JA	ID	1 st	2 nd	JA	ID

Logistic Regression Analysis of Vaccine Side Effects

We then compared Twitter mentions of side effects for the Pfizer and Moderna vaccines, with side effects as predictors and

https://infodemiology.jmir.org/2022/2/e39504

vaccine type (Pfizer or Moderna) as the outcome variable. Most of the predictor variables were significant (Table 3), suggesting that reported side effects differed significantly between Pfizer and Moderna vaccines. However, we note the high statistical

power resulting from a large sample size may have affected the calculation of statistical tests and the respective P value. We first report results for the Japanese data, followed by the Indonesian data separately.

The Nagelkerke \mathbb{R}^2 for Japanese tweets was 1.2%. In interpreting the model, we found that the odds of the umbrella term "side effect" appearing in a tweet about the Pfizer vaccine was about 1.907 times more. For specific terms (ie, muscle pain, fever, headache, and nausea), the odds ratios were close to each other: They were 1.338, 1.357, 1.362, and 1.458 times more likely, respectively, to be mentioned in Pfizer tweets, suggesting that those terms were more frequent in tweets about the Pfizer vaccine than in tweets about the Moderna vaccine. However, only chills had a small odds ratio. This is different from past results reported by Kitagawa et al [26], who compared the prevalence of the side effects of both vaccine types through a questionnaire study conducted in Japan and found that users receiving the Moderna vaccine reported more side effects than those receiving the Pfizer vaccine.

Indonesian tweets showed a different result, as other than tiredness and fever, all other side effects appeared less likely to be mentioned in tweets about the Pfizer vaccine and more likely to be mentioned in tweets about the Moderna vaccine. The Nagelkerke R^2 for the model for the Indonesian tweets was 17.4%. However, we noticed that tiredness appeared significantly more often in tweets about the Pfizer vaccine than in tweets about the Moderna vaccine. A closer look at the 95% CI for fever, which contained a value of 1, suggested that there may be little difference in those 2 side effects between the Pfizer and Moderna vaccines.

Table 3. Logistic regression analysis results for Japanese and Indonesian tweets.

Variable	Coefficient	Standard error	P value	Odds ratio	95% CI
Japanese Tweets			· · · ·		
(Intercept)	3.551	0.043	<.001	34.831	a
Side effects	0.646	0.043	<.001	1.907	1.755-2.073
Tiredness	0.489	0.054	<.001	1.631	1.468-1.814
Headache	0.309	0.055	<.001	1.362	1.223-1.517
Muscle pain	0.291	0.059	<.001	1.338	1.191-1.503
Chills	-0.101	0.087	.25	0.904	0.762-1.072
Fever	0.305	0.042	<.001	1.357	1.250-1.473
Nausea	0.377	0.146	.01	1.458	1.095-1.941
ndonesian Tweets					
(Intercept)	0.455	0.036	<.001	1.576	
Side effects	-1.461	0.042	<.001	0.232	0.214-0.252
Tiredness	0.238	0.084	.004	1.269	1.077-1.495
Headache	-0.294	0.068	<.001	0.745	0.653-0.851
Muscle pain	-0.98	0.398	.01	0.375	0.172-0.818
Chills	-1.029	0.080	<.001	0.358	0.306-0.418
Fever	0.003	0.043	.94	1.003	0.922-1.092
Nausea	-0.821	0.120	<.001	0.440	0.348-0.557

^aNot applicable.

Discussion

Principal Findings

Our results highlight a large gap between expressions of side effects on Twitter and percentages reported in public press releases: For most of the side effects, we found that the percentages of Twitter users who reported them were far lower than the percentages reported in the public reports. Our first result is focused on the descriptive comparison of counts reported on Twitter and counts of observation in public press releases. Considering that several studies make use of symptom reporting on Twitter for epidemiological surveillance, our study shows that, at least for vaccine side effect surveillance, we may be at risk of overrepresenting "fever" relative to other (milder) side effects. Although our study did not explicitly examine the reasons behind this phenomenon, we speculate on a few possible explanations.

First, this could reflect a difference in how lay people and health professionals perceive and talk about vaccinations [40]. Our study may be relying too much on relatively specialized terminology (eg, muscle pain) that may not be the most salient term accessible to the broader lay population, at least in Japan and Indonesia. Hence, the difference in focus and word choice between lay people and professionals may result in different

RenderX

expressions used to describe side effects experienced after vaccination. Given that our study used more "professional" terms, our results for side effects like muscle pain or chills could thus have been an underrepresentation of available tweets.

Second, the higher reporting rates for fever could be due to its ease of measurement by lay members of the public. Thermometers are widely available and widely used, and there are general conventions (thresholds) for determining if a person has a fever. On the other hand, other side effects such as chills, headaches, and tiredness sometimes might not have clear, objective thresholds and measurement methods that are of common knowledge to the lay person. With the ambiguity and subjectivity around these side effects, Twitter users may hesitate to update their statuses, especially compared with fever, which comes with a clear and objective threshold. Accordingly, Twitter users who may experience more than one side effect may then decide to only report the clearer, more observable one.

Finally, another possible reason could be the age difference between people observed in the studies (public reports) and Twitter users who share their experiences in their tweets. In a survey conducted by Statista, close to 80% of Japanese respondents aged 20 years to 29 years reported using the microblogging and social networking service Twitter. Although this suggests that the penetration rate among Japanese youths was also on a high level, it was much less widely used by older age groups [41]. A past study also suggested that systemic incidence of side effects from the Pfizer vaccine was significantly higher in young participants than in older adults [31], a finding that was also previously extended to the Japanese context [42]. The rate of Twitter users in the country also shows that there are people who did not use Twitter, which means they will not share their side effects through tweets. As the incidence of the vaccine side effects is higher in older age groups while the Twitter penetration rate is lower, this might also influence the number of side effects that can possibly be found in Twitter.

Our findings from Japanese and Indonesian tweets were also different from reported (vaccine) side effects in English tweets in the United States, where soreness, fatigue, and headache were listed as the top 3 side effects for the Pfizer and Moderna vaccines [43]. One probable reason could be cultural differences in how people express themselves on Twitter, which might stem from their respective cultural background and habits. People in collectivist cultures (like Japan and Indonesia) may be less open and active on social media, as compared with individuals in individualistic cultures (like the United States) [44]. Accordingly, users in the United States could be reporting their symptoms with more detail and frequency on social media, whereas Japanese and Indonesian users may be "saving" their posts for worse side effects (ie, fever). In any case, we suggest that future studies on infodemiological surveillance of vaccine side effects may consider focusing primarily on fever-related keywords in these countries.

Regardless of interpretation, our study appears to suggest that "fever," as a subjectively stronger side effect of vaccination, is discussed disproportionately more on Twitter in Japan and Indonesia. One possible consequence could be in the echo chamber effect on Twitter [45], which could contribute to

vaccine hesitancy or other aversive behaviors. As an illustration, due to this disproportionate reporting, consider Marie, a Twitter user, who is currently considering vaccination. She may observe that many users on her Twitter feed discuss their experiences with fever as a side effect of vaccination, which could lead to a perceived overrepresentation of fever risks that may dissuade her from receiving the vaccine. In contrast, if tweets had discussed side effects in a more representative manner, Marie would have had an accurate representation of the risks and may not have been discouraged from vaccination for this reason. We note again that side effects were a strong reason for vaccine hesitancy in Japan [22] and postulate that this overrepresentation of strong side effects on Twitter may have had a role to play in contributing toward hesitancy, although follow-up research is needed to test this hypothesis.

Limitations

Although we limited our investigation to only Moderna and Pfizer vaccines, these received relatively late approval in Indonesia, and we did not examine tweets on side effects from other vaccine makers (eg, Sinovac, AstraZeneca). Consequently, some of the discourse surrounding vaccines and their side effects was not captured in the earlier tweets. Second, some tweets also elaborated on side effects without mentioning the specific vaccine type received. Third, the search query for tweets was limited to the specified keywords (Table 1) and did not include other possible words not listed in the table, and we did not consider the positive or negative sentiment expressed by the tweets. Last, we focused only on tweets that mentioned one type of vaccine only and removed tweets mentioning both. Nevertheless, there was little chance of people getting both Pfizer and Moderna vaccinations administered in the observed period, and tweets mentioning both mainly referred to news articles and related discussions, not the actual side effects experienced by the public.

Finally, we did not control for negation in tweets. However, we sampled 100 tweets for each vaccine and language. Of all sampled tweets, we focused our observation on fever, which was the most frequent side effect found in our Twitter data set and found that only a minority of tweets contained negation. Based on a manual inspection of the sample, we found that, for Japanese tweets, negation (for fever) was observed in 15 of 63 tweets mentioning fever, or 23.80% of relevant tweets about the Pfizer vaccine. By doing the same process, we obtained 32.76% negation in Moderna tweets. Meanwhile, in Indonesian tweets, the negation for fever was 21.43% for the Pfizer vaccine and 16.67% for the Moderna vaccine. Negation is a difficult challenge in Twitter analyses, as there are many ways to express negation and may not necessarily be easily filtered out through designated negation words [46]. Although negation handling may improve the final results, it did not appear to hinder the utility of tweets in displaying consistent patterns as real-world (unfiltered) data in past research [47]. Considering that the observed percentage of negation for both vaccines in each language was similar in our random sample of tweets, we decided to retain all tweets for these analyses.

We also lacked the means to verify whether the tweets were from a personal or nonpersonal (eg, corporate) account and

```
XSL•FO
RenderX
```

whether the said individual behind the account actually received vaccination and follow-up confirmations of reported side effects. Finally, we also lacked sufficient information about whether the side effects were from the first, second, or third dose of the vaccination, as we were limited to the side effects shared in the tweets by the users that matched the language filter of our Twitter API query without any deeper demographic or contextual information.

Conclusions

We found that fever was the most prevalent side effect reported in Japanese and Indonesian tweets, and this may be a reflection of bias on social media toward reporting severe or measurable side effects (like fever). Furthermore, in examining side effects from different vaccine makers, we found that Twitter yielded inconsistent information from Japan and Indonesia, in that side effects were reported relatively more in tweets about the Pfizer vaccine in Japan but more in tweets about the Moderna vaccine in Indonesia. As such, given the inconsistencies and gaps in findings from Twitter and the vaccine press releases, we present cautious optimism that Twitter can prove useful for infodemiological surveillance for vaccine side effects that is best suited for detecting prevalences of fever symptoms in Japanese and Indonesian populations.

Acknowledgments

This work was supported by Japan Science and Technology Agency (JST) Strategic International Research Cooperative Program (SICORP) Grant Number JPMJSC2107, Japan and Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number JP22K12041.

Authors' Contributions

KF collected the data and performed the analysis, with the assistance of KL. KF and KL wrote the manuscript, and SW and EA provided critical comments. All authors contributed to the conceptualization of the study and to the initial study design. EA and SW supervised the project.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Tweet ID for Japanese tweets. [TXT File , 5603 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Tweet ID for Indonesian tweets. [TXT File , 283 KB-Multimedia Appendix 2]

Multimedia Appendix 3

Radar graph for Pfizer (A) and Moderna (B), showing comparisons of the percentages of each side effect. [PNG File , 213 KB-Multimedia Appendix 3]

References

- Dhama K, Sharun K, Tiwari R, Dadar M, Malik Y, Singh K, et al. COVID-19, an emerging coronavirus infection: advances and prospects in designing and developing vaccines, immunotherapeutics, and therapeutics. Hum Vaccin Immunother 2020 Jun 02;16(6):1232-1238 [FREE Full text] [doi: 10.1080/21645515.2020.1735227] [Medline: 32186952]
- Fadda M, Albanese E, Suggs LS. When a COVID-19 vaccine is ready, will we all be ready for it? Int J Public Health 2020 Jul;65(6):711-712 [FREE Full text] [doi: 10.1007/s00038-020-01404-4] [Medline: 32529534]
- 3. Yamey G, Schäferhoff M, Hatchett R, Pate M, Zhao F, McDade KK. Ensuring global access to COVID-19 vaccines. Lancet 2020 May 02;395(10234):1405-1406 [FREE Full text] [doi: 10.1016/S0140-6736(20)30763-7] [Medline: 32243778]
- Eibensteiner F, Ritschl V, Nawaz FA, Fazel SS, Tsagkaris C, Kulnik ST, et al. People's willingness to vaccinate against COVID-19 despite their safety concerns: Twitter poll analysis. J Med Internet Res 2021 Apr 29;23(4):e28973 [FREE Full text] [doi: 10.2196/28973] [Medline: 33872185]
- Latkin CA, Dayton L, Yi G, Konstantopoulos A, Boodram B. Trust in a COVID-19 vaccine in the U.S.: A social-ecological perspective. Soc Sci Med 2021 Feb;270:113684 [FREE Full text] [doi: 10.1016/j.socscimed.2021.113684] [Medline: 33485008]
- 6. An L, Russell DM, Mihalcea R, Bacon E, Huffman S, Resnicow K. Online search behavior related to COVID-19 vaccines: Infodemiology study. JMIR Infodemiology 2021;1(1):e32127 [FREE Full text] [doi: 10.2196/32127] [Medline: 34841200]

RenderX

- Bonnevie E, Gallegos-Jeffrey A, Goldbarg J, Byrd B, Smyser J. Quantifying the rise of vaccine opposition on Twitter during the COVID-19 pandemic. Journal of Communication in Healthcare 2020 Dec 15;14(1):12-19. [doi: 10.1080/17538068.2020.1858222]
- Lyu JC, Han EL, Luli GK. COVID-19 vaccine-related discussion on Twitter: Topic modeling and sentiment analysis. J Med Internet Res 2021 Jun 29;23(6):e24435 [FREE Full text] [doi: <u>10.2196/24435</u>] [Medline: <u>34115608</u>]
- 9. Banda JM, Tekumalla R, Wang G, Yu J, Liu T, Ding Y, et al. A large-scale COVID-19 Twitter chatter dataset for open scientific research -- an international collaboration. ArXiv 2020 Apr 07:1. [Medline: <u>32550247</u>]
- Wakamiya S, Kawai Y, Aramaki E. Twitter-based influenza detection after flu peak via tweets with indirect information: Text mining study. JMIR Public Health Surveill 2018 Sep 25;4(3):e65 [FREE Full text] [doi: 10.2196/publichealth.8627] [Medline: 30274968]
- Yousefinaghani S, Dara R, Poljak Z, Bernardo TM, Sharif S. The assessment of Twitter's potential for outbreak detection: Avian influenza case study. Sci Rep 2019 Dec 03;9(1):18147 [FREE Full text] [doi: 10.1038/s41598-019-54388-4] [Medline: 31796768]
- Kouzy R, Abi Jaoude J, Kraitem A, El Alam MB, Karam B, Adib E, et al. Coronavirus goes viral: Quantifying the COVID-19 misinformation epidemic on Twitter. Cureus 2020 Mar 13;12(3):e7255 [FREE Full text] [doi: 10.7759/cureus.7255] [Medline: 32292669]
- Marcec R, Likic R. Using Twitter for sentiment analysis towards AstraZeneca/Oxford, Pfizer/BioNTech and Moderna COVID-19 vaccines. Postgrad Med J 2022 Jul;98(1161):544-550 [FREE Full text] [doi: 10.1136/postgradmedj-2021-140685] [Medline: 34373343]
- 14. Sattar NS, Arifuzzaman S. COVID-19 vaccination awareness and aftermath: Public sentiment analysis on Twitter data and vaccinated population prediction in the USA. Applied Sciences 2021 Jun 30;11(13):6128. [doi: 10.3390/app11136128]
- Kwok SWH, Vadde SK, Wang G. Tweet topics and sentiments relating to COVID-19 vaccination among Australian Twitter users: Machine learning analysis. J Med Internet Res 2021 May 19;23(5):e26953 [FREE Full text] [doi: 10.2196/26953] [Medline: <u>33886492</u>]
- 16. Dixon S. Countries with the most Twitter users 2022. Statista. 2022 Jul 27. URL: <u>https://www.statista.com/statistics/242606/</u> number-of-active-twitter-users-in-selected-countries/ [accessed 2022-09-23]
- 17. Bernier A. An Introduction To The Japanese Language. Babbel Magazine. 2021 Jul 28. URL: <u>https://www.babbel.com/en/</u> <u>magazine/guide-to-japanese-language</u> [accessed 2022-09-23]
- 18. Devlin TM. How Many People Speak Indonesian, And Where Is It Spoken? Babbel Magazine. 2021 May 11. URL: <u>https://www.babbel.com/en/magazine/how-many-people-speak-indonesian-where-is-it-spoken</u> [accessed 2022-09-23]
- 19. Wang S. Japan begins vaccine rollout on Wednesday. CNN. 2021 Feb 17. URL: <u>https://edition.cnn.com/world/live-news/</u> <u>coronavirus-pandemic-vaccine-updates-02-16-21/h_45ac6dc81faf4905a94fb4c8ab93092a</u> [accessed 2022-09-23]
- 20. Hasell J, Mathieu E, Beltekian D, Macdonald B, Giattino C, Ortiz-Ospina E, et al. A cross-country database of COVID-19 testing. Sci Data 2020 Oct 08;7(1):345 [FREE Full text] [doi: 10.1038/s41597-020-00688-8] [Medline: 33033256]
- 21. Jiao C, Aditya A. Indonesia Begins COVID-19 Vaccine Rollout and Jokowi Is First to Get Jab. Time. 2021 Jan 13. URL: https://time.com/5929160/indonesia-vaccine-rollout-jokowi/ [accessed 2022-09-23]
- 22. Okubo R, Yoshioka T, Ohfuji S, Matsuo T, Tabuchi T. COVID-19 vaccine hesitancy and its associated factors in Japan. Vaccines (Basel) 2021 Jun 17;9(6):662. [doi: 10.3390/vaccines9060662] [Medline: 34204465]
- 23. Khan MSR, Watanapongvanich S, Kadoya Y. COVID-19 vaccine hesitancy among the younger generation in Japan. Int J Environ Res Public Health 2021 Nov 07;18(21):11702 [FREE Full text] [doi: 10.3390/ijerph182111702] [Medline: 34770216]
- 24. Indonesian Ministry of Health, National Immunization Technical Advisory Group (NITAG), UNICEF, World Health Organization (WHO). COVID-19 Vaccine Acceptance Survey in Indonesia. 2020 Nov. URL: <u>https://covid19.go.id/storage/app/media/Hasil%20Kajian/2020/November/vaccine-acceptance-survey-en-12-11-2020final.pdf</u> [accessed 2022-09-23]
- Boon-Itt S, Skunkan Y. Public perception of the COVID-19 pandemic on Twitter: Sentiment analysis and topic modeling study. JMIR Public Health Surveill 2020 Nov 11;6(4):e21978 [FREE Full text] [doi: 10.2196/21978] [Medline: 33108310]
- 26. Kitagawa H, Kaiki Y, Sugiyama A, Nagashima S, Kurisu A, Nomura T, et al. Adverse reactions to the BNT162b2 and mRNA-1273 mRNA COVID-19 vaccines in Japan. J Infect Chemother 2022 Apr;28(4):576-581 [FREE Full text] [doi: 10.1016/j.jiac.2021.12.034] [Medline: 35058126]
- 27. twarc. URL: <u>https://twarc-project.readthedocs.io/en/latest/</u> [accessed 2022-09-23]
- 28. What to Expect after Getting a COVID-19 Vaccine. Centers for Disease Control and Prevention. URL: <u>https://www.cdc.gov/</u> <u>coronavirus/2019-ncov/vaccines/expect/after.html</u> [accessed 2022-09-23]
- 29. Information about side effects or reaction after COVID-19 vaccination. kipi. URL: <u>https://kipi.covid19.go.id/</u> [accessed 2022-09-23]
- 30. COVID-19 Vaccines. Ministry of Health, Labour and Welfare Japan. URL: <u>https://www.mhlw.go.jp/stf/covid-19/vaccine.</u> <u>html</u> [accessed 2022-09-23]
- Izumo T, Kuse N, Awano N, Tone M, Sakamoto K, Takada K, et al. Side effects and antibody titer transition of the BNT162b2 messenger ribonucleic acid coronavirus disease 2019 vaccine in Japan. Respir Investig 2021 Sep;59(5):635-642 [FREE Full text] [doi: 10.1016/j.resinv.2021.06.003] [Medline: 34210623]

RenderX

- Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, C4591001 Clinical Trial Group. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. N Engl J Med 2020 Dec 31;383(27):2603-2615 [FREE Full text] [doi: 10.1056/NEJMoa2034577] [Medline: 33301246]
- Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, COVE Study Group. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. N Engl J Med 2021 Feb 04;384(5):403-416 [FREE Full text] [doi: 10.1056/NEJMoa2035389] [Medline: 33378609]
- 34. Sato T, Hashimoto T, Okumura M. Implementation of word-spacing dictionary mecab-ipadic-NEologd and its effective use in information retrieval. 2017 Presented at: The 23rd Annual Conference of the Association for Natural Language Processing; March 13-17, 2017; Tsukuba, Ibaraki, Japan URL: <u>https://www.anlp.jp/proceedings/annual_meeting/2017/pdf_dir/B6-1.pdf</u>
- 35. Sato T. Neologism dictionary based on the language resources on the Web for Mecab. neologd / mecab-ipadic-neologd. 2015. URL: <u>https://github.com/neologd/mecab-ipadic-neologd</u> [accessed 2022-09-23]
- 36. Sato T, Hashimoto T, Manabu O. Operation of a word segmentation dictionary generation system called NEologd (in Japanese). 2016 Presented at: Information Processing Society of Japan, Special Interest Group on Natural Language Processing (IPSJ-SIGNL) Information Processing Society of Japan; 20-22 December 2016; Tokyo p. NL-229-NL-215.
- 37. Hosmer Jr DW, Lemeshow S, Sturdivant RX. Applied Logistic Regression. Hoboken, NJ: John Wiley & Sons; 2013.
- 38. Nagelkerke NJD. A note on a general definition of the coefficient of determination. Biometrika 1991;78(3):691-692. [doi: 10.1093/biomet/78.3.691]
- 39. COVID-19 Vaccines. Prime Minister's Office of Japan. URL: <u>https://www.kantei.go.jp/jp/headline/kansensho/vaccine.html</u> [accessed 2022-09-23]
- 40. Luo C, Ji K, Tang Y, Du Z. Exploring the expression differences between professionals and laypeople toward the COVID-19 vaccine: text mining approach. J Med Internet Res 2021 Aug 27;23(8):e30715 [FREE Full text] [doi: 10.2196/30715] [Medline: 34346885]
- 41. Twitter's penetration rate in Japan FY 2021, by age group. Statista. 2022 Sep 07. URL: <u>https://www.statista.com/statistics/</u>1077538/japan-twitter-penetration-rate-by-age-group/ [accessed 2022-09-23]
- 42. Saita M, Yan Y, Ito K, Sasano H, Seyama K, Naito T. Reactogenicity following two doses of the BNT162b2 mRNA COVID-19 vaccine: Real-world evidence from healthcare workers in Japan. J Infect Chemother 2022 Jan;28(1):116-119 [FREE Full text] [doi: 10.1016/j.jiac.2021.09.009] [Medline: 34580011]
- Lian AT, Du J, Tang L. Using a machine learning approach to monitor COVID-19 vaccine adverse events (VAE) from Twitter data. Vaccines (Basel) 2022 Jan 11;10(1):103 [FREE Full text] [doi: <u>10.3390/vaccines10010103</u>] [Medline: <u>35062764</u>]
- 44. Jackson LA, Wang J. Cultural differences in social networking site use: A comparative study of China and the United States. Computers in Human Behavior 2013 May;29(3):910-921. [doi: 10.1016/j.chb.2012.11.024]
- 45. Colleoni E, Rozza A, Arvidsson A. Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. J Commun 2014 Mar 19;64(2):317-332. [doi: 10.1111/jcom.12084]
- 46. Kharde VA, Sonawane S. Sentiment analysis of Twitter data: A survey of techniques. IJCA 2016 Apr 15;139(11):5-15. [doi: <u>10.5120/ijca2016908625</u>]
- 47. Aramaki E, Maskawa S, Morita M. Twitter Catches The Flu: Detecting Influenza Epidemics using Twitter. 2011 Presented at: Conference on Empirical Methods in Natural Language Processing; July 27-31, 2011; Edinburgh, United Kingdom.

Abbreviations

JSPS: Japan Society for the Promotion of Science JST: Japan Science and Technology Agency mRNA: messenger RNA SICORP: Strategic International Research Cooperative Program

Edited by T Mackey; submitted 12.05.22; peer-reviewed by S Kwok, J Banda, S Matsuda; comments to author 24.06.22; revised version received 29.07.22; accepted 19.09.22; published 04.10.22

<u>Please cite as:</u> Ferawati K, Liew K, Aramaki E, Wakamiya S Monitoring Mentions of COVID-19 Vaccine Side Effects on Japanese and Indonesian Twitter: Infodemiological Study JMIR Infodemiology 2022;2(2):e39504 URL: <u>https://infodemiology.jmir.org/2022/2/e39504</u> doi: <u>10.2196/39504</u> PMID: <u>36277140</u>



©Kiki Ferawati, Kongmeng Liew, Eiji Aramaki, Shoko Wakamiya. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 04.10.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO RenderX