Open Forum Infectious Diseases
MAJOR ARTICLE

Analyzing an Emerging Pandemic on Twitter: Monkeypox

Lauren N. Cooper,[1] Alexander P. Radunsky,[2] John J. Hanna,[1,2] Zachary M. Most,[3] Trish M. Perl,[2] Christoph U. Lehmann,[1] and Richard J. Medford[1,2]

[1]Clinical Informatics Center, University of Texas Southwestern Medical Center, Dallas, Texas, USA, [2]Department of Internal Medicine, Division of Infectious Diseases and Geographic Medicine, University of Texas Southwestern Medical Center, Dallas, Texas, USA, and [3]Department of Pediatrics, Division of Pediatric Infectious Disease, University of Texas Southwestern Medical Center, Dallas, Texas, USA

Background. Social media platforms like Twitter provide important insights into the public's perceptions of global outbreaks like monkeypox. By analyzing tweets, we aimed to identify public knowledge and opinions on the monkeypox virus and related public health issues.

Methods. We analyzed English-language tweets using the keyword "monkeypox" from 1 May to 23 July 2022. We reported gender, ethnicity, and race of Twitter users and analyzed tweets to identify predominant sentiment and emotions. We performed topic modeling and compared cohorts of users who self-identify as LGBTQ+ (an abreviation for lesbian, gay, bisexual, transgender, queer, and/or questioning) allies versus users who do not, and cohorts identified as "bots" versus humans.

Results. A total of 48 330 tweets were written by LGBTQ+ self-identified advocates or allies. The mean sentiment score for all tweets was −0.413 on a −4 to +4 scale. Negative tweets comprised 39% of tweets. The most common emotions expressed were fear and sadness. Topic modeling identified unique topics among the 4 cohorts analyzed.

Conclusions. The spread of mis- and disinformation about monkeypox was common in our tweet library. Various conspiracy theories about the origins of monkeypox, its relationship to global economic concerns, and homophobic and racial comments were common. Conversely, many other tweets helped to provide information about monkeypox vaccines, disease symptoms, and prevention methods. Discussion of rising monkeypox case numbers globally was also a large aspect of the conversation.

Conclusions. We demonstrated that Twitter is an effective means of tracking sentiment about public healthcare issues. We gained insight into a subset of people, self-identified LGBTQ+ allies, who were more affected by monkeypox.

Keywords. monkeypox; pandemic; social media; topic modeling; Twitter.

As cases of monkeypox (recently assigned the new preferred term "mpox" by the World Health Organization [WHO] [1]) started to spread in previously nonendemic countries in May 2022 [2], early public reactions compared the monkeypox virus to the ongoing coronavirus disease 2019 (COVID-19) pandemic [3]. As news proliferated on this new global infectious disease threat, people shared their experiences, opinions, and emotions on social media. Online social media communities have become increasingly important to people's understanding of the world around them [4]. Furthermore, social media is frequently used to discuss new health threats early, allowing people to share and partake in opinions and concerns about the latest "viral" (pun intended) pandemic [5–7]. We sought to obtain an understanding of public opinion on monkeypox using social media.

Our focus was on the discussion of monkeypox on the social networking service Twitter. Twitter has >240 million active users daily (as of the second quarter of 2022 [8]) and allows researchers easy access to its data and interfaces. Since the majority of Twitter users use the service to access news and pertinent information [9], we were able to study reactions to the release of information about monkeypox cases in nonendemic countries by analyzing the content, sentiment, and emotional responses of tweets throughout the months of May, June, and early July 2022. It was our hypothesis that Twitter would be valuable to identify and analyze public opinion and reactions to the spread of monkeypox and to identify potentially useful public health topics that could be addressed by government and public health officials.

Received 04 November 2022; editorial decision 13 March 2023; accepted 15 March 2023; published online 18 March 2023

Correspondence: Lauren N. Cooper, MS, Clinical Informatics Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75390 (Lauren.Cooper@UTSouthwestern.edu).

Open Forum Infectious Diseases®

https://doi.org/10.1093/ofid/ofad142

METHODS

Data Collection and Processing

Using the keyword "monkeypox," we identified and collected English-language tweets using the Twitter scraping tool, TWINT [10] from 1 May to 23 July 2022, encompassing the beginning of monkeypox cases in nonendemic countries until the date at which the WHO declared monkeypox a global health emergency. Utilizing the TWINT library allowed us to collect a large number of tweets from Twitter users quickly but limited metadata available to us. Using the Python (version 3.10.5) programming language [11] for processing and analysis of data, we extracted additional user metadata utilizing the Twitter API [12] and the Python library, Tweepy [13]. By utilizing Tweepy, we collected additional information such as user locations and profile information for each tweet, based on the unique tweet and user identifier, that we were unable to collect from the TWINT library. Using the Python library, Botometer by OSoMe [14], we determined the number of the unique accounts that possibly were "bots" instead of actual humans. The Botometer library utilizes random forest models to classify users as either human or a specific bot class, with a low bot score signifying a likely human account and a higher bot score signifying a likely bot account [14].

We processed the text of the tweets by converting tweets into plain text, removing any emojis and symbols such as "@" and "#", and expanding contractions. We then cleaned the text using the natural language processing library spaCy [15] and created bigrams and trigrams from the cleaned text using the phrase detection function from the Gensim library [16].

User Demographics

To infer the unique user demographics, we used the M3-Inference [17] and the Ethnicolr [18] libraries. The M3-Inference library is a deep learning system that infers the Tweeter's gender (binary, male or female), estimated age range, and whether the Tweeter is an organization (binary, yes or no) based on the user's image, full name, Twitter screen name, and Twitter biography [17]. Age ranges include <19 years, 19–29 years, 30–39 years, and ≥40 years. The Ethnicolr library utilizes machine learning based on the Florida state voter registrations to predict the ethnicity of a user based on their first and last name [19], which we gathered from the metadata included with each tweet. Ethnicity categories include non-Hispanic White, non-Hispanic Black, Asian, and Hispanic. Users were also categorized into 2 cohorts: (1) self-identified allies or advocates for the LGBTQ+ (an abreviation for lesbian, gay, bisexual, transgender, queer, and/or questioning) community and (2) users who did not self-identify as such. The cohorts were developed based on the presence of keywords in the user's profile that were indicative of support for the LGBTQ+ community [20, 21]. These keywords included but were not limited to such terms as "bisexual," "gay," "lesbian," "LGBTQ," and "nonbinary" and the use of pronouns such as "they/them," "he/him," and "she/her."

Sentiment and Emotion Analysis

After removing hashtags and symbols, URLs, and Twitter handles, the SentiStrength library classified processed tweets according to the text's sentiment. Sentiments were reported on a scale of −4 to +4, with −4 being an extremely negative sentiment and +4 being an extremely positive sentiment [22]. We chose the SentiStrength library because of its ability to classify sentiments with greater accuracy than other methods such as support vector machines, naive Bayes, and simple logistic regression on informal text samples similar to tweets. A comparison of sentiments between the ally and non-ally cohorts was made using Pearson χ^2 tests. Using the text2emotion library, we were able to categorize the processed tweets according to the probability of containing 1 of 5 emotions: happy, angry, sad, surprise, and fear [23].

Topic Modeling

After processing and cleaning the tweets, we used a latent Dirichlet allocation (LDA) model estimation algorithm from the Gensim library [16] for the topic modeling. We first created a corpus based on trigrams from our data set and then used that corpus to train the LDA model multiple times with the number of topics ranging from 1 to 20. For each trained model, we calculated a Umass coherence score to quantitatively compare models based on the measure of similarity between words in each topic [24]. Based on the optimal coherence score, a model with 6 topics for the cohort of self-identified allies and a model with 9 topics for the non-ally cohort were chosen. Running the models with 6 and 9 topics, respectively, we included the top 20 key words along with 10 randomly chosen tweets from each individual topic for analysis by a group with no prior assumptions about the topic model. We manually assigned labels/descriptions for each topic using a consensus approach. We repeated this process with models of 9 and 10 topics for the cohorts of identified bot-written tweets and human-written tweets.

RESULTS

During the study period, we collected 981 725 tweets from 436 157 unique users. Limiting our data to just English-language tweets; we retained 384 925 unique users, who generated 858 581 English-language tweets. Tweets that included media such as videos, photos, or URLs comprised 13%, 12%, and 33.4% of tweets, respectively (Table 1). Almost half (49.3%) of tweets had at least 1 like, 28.6% received at least 1 reply, and 21% were retweeted at least once. Figure 1 compares daily case counts of monkeypox cases with tweets and provides a high-level timeline of events relating to the monkeypox outbreak.

User Demographics

The majority of unique, English-speaking users (276 620 [73%]) were male. The largest group of users was <19 years old (37.9%). Using the Ethnicolr library, we determined that 76.3% of users were non-Hispanic White, 9.2% of users were

Table 1. Twitter User Demographics and Tweet Characteristics

Characteristic	No. (%)
User demographics	
Gender	
Male	276 620 (73)
Female	102 569 (27)
Age group, y	
<19	143 759 (37.9)
19–29	53 004 (14)
30–39	73 748 (19.4)
≥40	108 678 (28.7)
Race/ethnicity	
Non-Hispanic White	188 000 (76.3)
Non-Hispanic Black	22 588 (9.2)
Hispanic	12 888 (5.2)
Asian	22 588 (9.2)
Tweet characteristics	
Tweets with replies	245 390 (28.6)
Tweets with likes	423 583 (49.3)
Tweets with retweets	180 062 (21)
Tweets with videos	111 888 (13)
Tweets with photos	103 088 (12)
Tweets with URLs	286 815 (33.4)
Total No. of unique, English-language tweets	858 581 (100)

non-Hispanic Black, 5.2% were Hispanic, and 9.2% were Asian (Table 1).

Word Frequency

After excluding our tweet-identifying term "monkeypox," the most common word in our data set was "case" with 146 562 uses. The next 14 most used words were "vaccine" (81 543), "COVID" (80 826), "health" (75 327), "virus" (69 131), "spread" (66 920), "people" (62 434), "outbreak" (60 787), "like" (50 861), "say" (47 680), "new" (45 303), "know" (45 099), "go" (40 807), "disease" (38 453), and "get" (37 504).

Sentiment and Emotion Analysis

The overall sentiment of all tweets skewed negative with a mean sentiment score of −0.413 on a −4 to +4 scale. Of tweets, 39% were classified as negative with sentiment scores <0. Neutral tweets (scores of 0) comprised 46.5% of all tweets, while 14.5% were classified as positive with sentiment scores ranging from >0 to +4 (Figure 2). The mean sentiments for all topics from the self-identified allies and the non-allies were negative with the most negative topic of the allied cohort discussing "the spread of monkeypox with an emphasis that it is not a 'gay' disease" with a mean sentiment score of −0.48. The most negative topic of the non-allied cohort also had a score of −0.48 but discussed "the declaration of monkeypox as a global health emergency." While still negative, the most positive mean sentiment score (−0.20) for an ally topic discussed "information on and how to get the monkeypox vaccine." Similarly, the non-allied cohort had the most positive topic discussing the monkeypox "vaccine and the global monkeypox health emergency" with a mean sentiment score of −0.13. Most sentiment values resulted in a significant difference between the 2 cohorts with a P value less than the Bonferroni-corrected value of .0055 (n = 9). The sentiment values of −4, −1, and +4 did not show a significant difference with P values >.0055.

Congruent with the negative sentiment toward monkeypox exhibited in the tweets, the emotion analysis found mainly negative emotions. Of the 5 emotions (happy, surprise, anger, fear, and sadness), 67.8% of the tweets were labeled to contain negative emotions (anger, fear, sadness) with fear present most often in 38.6% of tweets (Figure 3). Examples of tweets with the predominant emotion of fear included "Has anyone published outcomes of Monkeypox infections yet? I have no sense of the seriousness of an infection, anyone else?" and "Warning signs ahead of Monkeypox outbreak went unheeded, experts say." Only 15.1% of the tweets were labeled as expressing happiness and 16.9% expressed surprise such as "Girl this monkeypox thing is actually happening huh …" and "Someone get me a smallpox vaccine right now, I need a re-up before monkeypox scales up." All emotions, except for anger, showed a significant difference between the ally and non-ally cohorts with P values from Pearson χ^2 tests with a Bonferroni correction (n = 5) of <.01. Comparing anger between the 2 groups resulted in a P value of .31.

Topic Modeling

Based on the results of the LDA model estimation algorithm, we chose 6 topics to categorize our library of tweets from self-identified LGBTQ+ allies. The largest category discussed "informing prevention for monkeypox infection," which comprised 18.1% of all tweets, while the smallest topic, which could be described as "pandemic exasperation about vaccination," made up 15.1% of the data set. Other topics included the "reporting of monkeypox cases," "monkeypox vaccine information," "lessons from monkeypox and other epidemics," and "checking hetero-privilege about monkeypox transmission." For the collection of tweets from non-self-identified allies, our algorithm selected 9 topics. The largest category was "growing fear about monkeypox pandemic," which comprised 20% of all tweets. The smallest topic, with only 1.7% of this category's tweets, discussed "surprise at monkeypox spread outside Africa." Other topics for the non-ally tweets included "reporting of monkeypox cases," "sexual transmission of monkeypox," "anger about monkeypox," and the "declaration of monkeypox as a global health emergency." Additional topics, representative tweets, and top weighted keywords for each topic are listed in Table 2.

For the sample of tweets that were identified as being composed by bots, we chose 9 topics to categorize the tweets. The largest category, 19.5% of tweets, was described as "global reporting of monkeypox cases." Other topic categories included

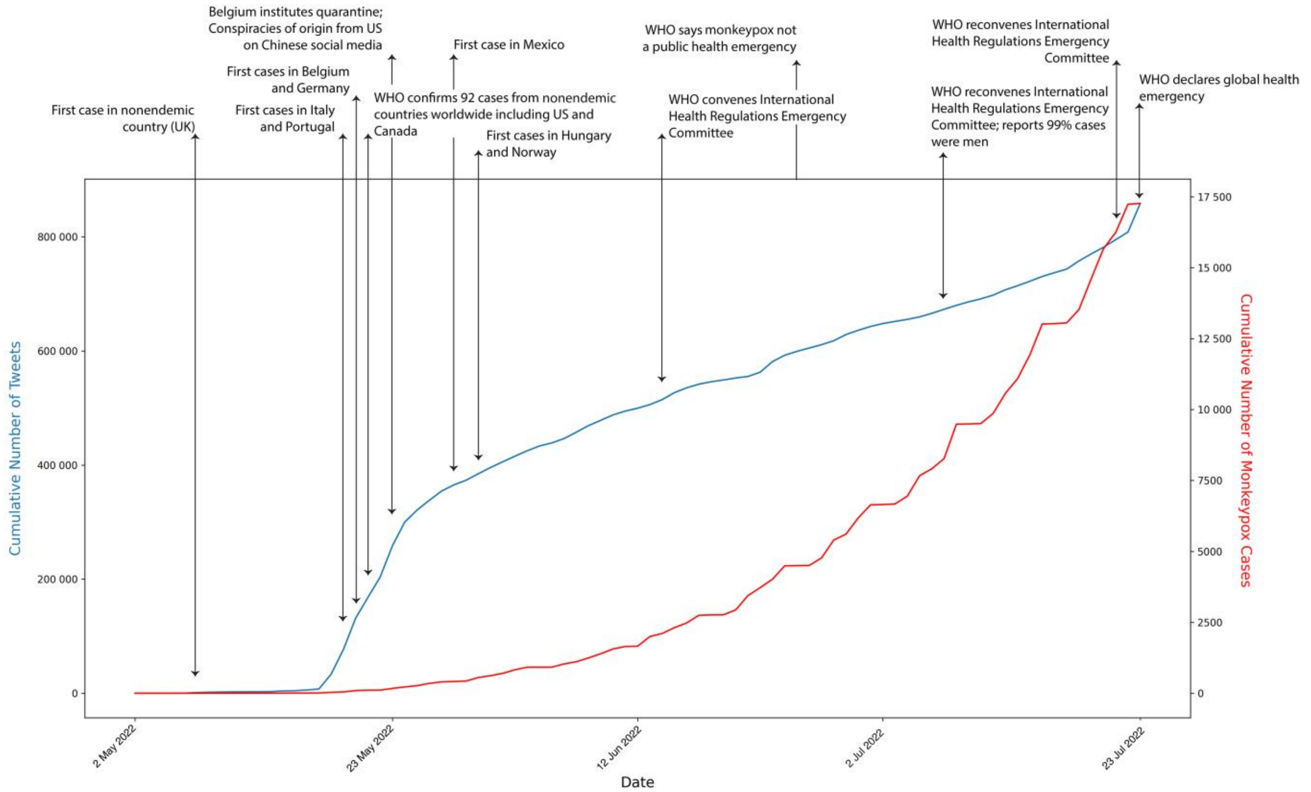

Figure 1. Cumulative count of daily monkeypox tweets and cumulative count of daily monkeypox cases [25]. Abbreviations: UK, United Kingdom; US, United States; WHO, World Health Organization.

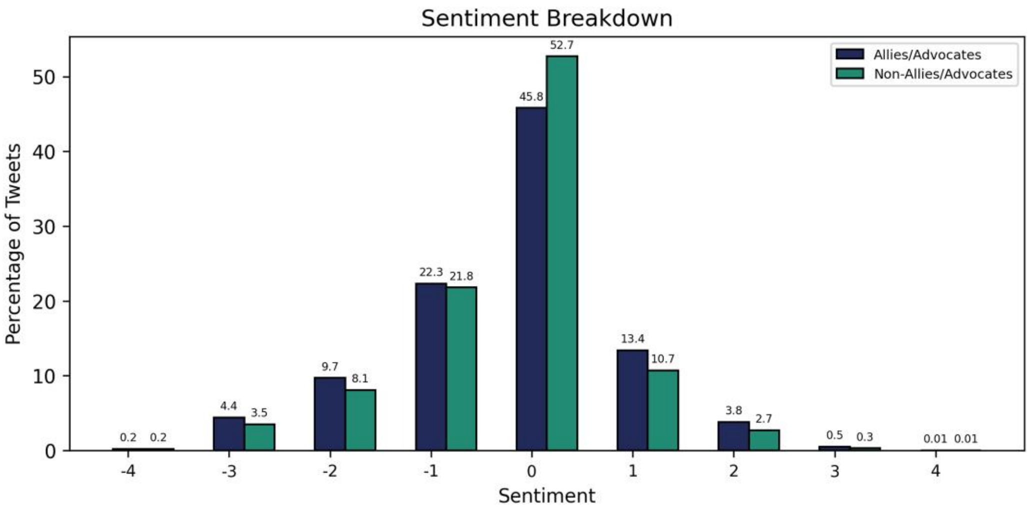

Figure 2. Sentiment analysis of tweets.

discussions about "comparing COVID to monkeypox," "vaccination against monkeypox," "public health response to monkeypox," "constructive sharing of monkeypox health information," and "news of monkeypox risk to gay men." We identified 10 categories for the tweets that were identified as composed by a human. For these tweets, the largest category discussed the comments made about "speculative discussion about COVID and monkeypox transmission." This topic

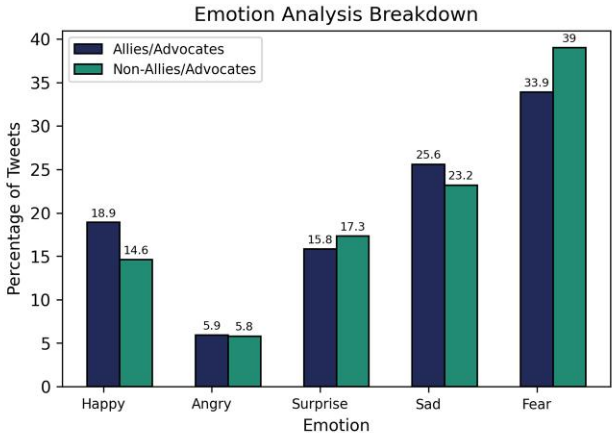

Figure 3. Emotion analysis of tweets.

comprised 22.2% of the tweet sample. The other 9 categories discussed topics such as "sharing informal information on monkeypox," "formal reporting of monkeypox cases," "alarmist discussion of monkeypox response," "fear about monkeypox vaccine shortage and vaccine effect," "testing for monkeypox," and finally "growing alarm about monkeypox risk." Additional topics, representative tweets, and top weighted keywords for each topic are listed in Table 3.

DISCUSSION

Engaging in social media is a popular pastime for billions of people worldwide [26]. During the COVID-19 pandemic, internet use and associated addictive behaviors increased in all age groups, resulting in increased levels of stress, anxiety, and depression [27]. Time spent using more recently developed types of social media (eg, Twitter, Instagram) was linked to personal insecurity and negative mental health outcomes [28–32]. Daily exposures to pandemic-focused social media have been associated with elevated levels of depression and posttraumatic stress disorder symptoms [33].

In response to the new infectious diseases threat of monkeypox, in our study the most common emotion expressed by Twitter users in their tweets was fear. Congruently, tweet sentiments were predominantly negative. Early in the COVID-19 pandemic, we found that 49.5% of COVID-19–related tweets expressed fear [5]. Fear seems to be a natural first response by the public to a new infectious diseases outbreak. However, not all the fear expressed about monkeypox reflected rational concerns expected in the light of an outbreak. Nearly a quarter (22%) of the tweets in our data set corresponded to various conspiracy theories related to monkeypox.

High prevalence of disinformation (defined as false information with the intent to harm) [6] forced social media companies to develop policies to combat the spread of mis- and disinformation [34]. In our study, we found dis- and misinformation about the origins of monkeypox and false claims that monkeypox was a side effect of COVID-19 vaccinations, as well as homophobic and racial comments in the context of the disease presented to a global audience via Twitter.

On 19 May 2022, Twitter introduced a crisis misinformation policy to tag tweets deemed as misinformation about crisis situations, including public health emergencies and natural disasters, and displayed warnings before users viewed the content provided by the most severe offenders of the policy [35]. Lanier et al showed that 20% of users of the hashtags #Scamdemic or #Plandemic in 2020 had been suspended or had left the Twitter platform in 2021 [36]. Future studies will determine whether this intervention will have any effect on the amount and level of dis- and misinformation spread on the platform.

There is evidence that throughout the COVID-19 pandemic, social media platforms such as Twitter were helpful to disseminate public health information in a positive and meaningful way that helped to alleviate anxieties and reduce the amount of harmful dis- and misinformation [5]. Many tweets categorized in topics that discuss public health issues such as monkeypox vaccine availability, symptoms, and preventions not only had the least negative sentiments but also were likely to add useful and relevant public health information to the debate. In the future, continued analysis of social media posts could allow public health officials to monitor public opinion on the next outbreak and spread relevant warnings and preventive measures to a wider audience sooner. Clinicians could also utilize future social media analysis to determine groups of patients requiring targeted disease education based on the amount of misinformation disseminated. Understanding how various demographic groups discuss a disease's symptoms or spread could allow targeted preventions such as vaccines to be mobilized to those populations quicker, preventing a greater spread of disease.

By dividing our library of collected tweets into various cohorts and analyzing using topic modeling as described above, we were able to identify several common themes that were present among all cohorts. The discussion of rising global monkeypox case numbers and the spread of monkeypox cases around the world was the second largest topic for all cohorts except for the group of humans. For humans, the topic of case numbers of disease spread was still heavily discussed but at a smaller percentage of tweets compared to the other cohorts.

The topics involving the emotions of anger, suspicion, and conspiracy theories were also present in all but the tweets that were from identified bots. The absence of these topics from this cohort suggests that the tool used for identifying bots worked well. Interestingly anger, suspicion, and conspiracy theories, while present in 3 of the 4 cohorts, were presented in different manners. Conspiracy theories such as the use of monkeypox as a distraction from the COVID-19 pandemic, global economic recessions, and

Table 2. Sentiment Analysis and Topic Modeling of Tweets per Topic for Tweets That Are From Self-identified LBGTQ+ Community Allies and Non-allies

Topic	Tweets in Topic, No. (%)	Mean Sentiment of Topic	Topic Keywords	Representative Tweets
Self-identifying allies/advocates				
Informing prevention for monkeypox infection	8792 (18.1)	−0.28	go, think, like, want, know, people, start, day, good, right, time, test, COVID, try, mask, maybe, way, spread, thread, tell	"@[…] No mention of masking and not a mask in sight. On the plus side, washing hands might help slow Monkey https://t.co/WMM2uYYlkf"
Geographic reporting of monkeypox cases	8438 (17.5)	−0.29	case, virus, outbreak, cdc, new, spread, report, uk, state, know, gay_man, health, far, symptom, europe, update, child, country, nyc, confirm_case	"25–30 Cases Of Rare Monkeypox Reported in UK, Portugal; Reported Predominantly Among Gay, MSM; Somewhat Treatable https://t.co/IEverFkSqN"
Sharing of monkeypox vaccine information	8366 (17.3)	−0.20	vaccine, say, people, look, need, gay, get, great, s***, thing, know, find, worried, see, f***king, info, new, vaccination, news, sure	"A reminder to people in NYC who qualify for #monkeypox vaccination: the latest appointment slots open up for boo https://t.co/RForQEzFNR"
Lessons from monkeypox and other epidemics	8008 (16.6)	−0.31	COVID, thing, pandemic, mean, catch, get, smallpox, wait, let, vax, real, long, well, polio, deal, actually, year, know, virus, sex	"u will not catch me in any clubs any time soon, im not trying to catch the double deluxe (COVID-19 â€¢ Monkeypox)Â² pandemic special."
Checking hetero-privilege about monkeypox transmission	7447 (15.4)	−0.48	people, spread, hear, bad, std, sti, COVID, way, response, yes, hiv, information, gay_disease, gay, man, infect, homophobic, think, believe, transmit	"Monkeypox spreads by contact with surfaces infected people with sores have touched. Treating it like an STD is insane https://t.co/Uzm5sP82RJ"
Pandemic exasperation about vaccination	7279 (15.1)	−0.40	get, like, f***, know, time, need, thank, come, stop, smallpox_vaccine, spread, literally, take, COVID, world, year, wtf, public_health, protect, tweet	"The sad truth is Monkey Pox was avoidable. Another example of our public health systems failing. It was ignored https://t.co/roubOeKBjU"
Non-self-identifying allies/advocates				
Growing fear about monkeypox pandemic	153 663 (20)	−0.43	COVID, get, go, people, time, say, think, right, pandemic, polio, start, world, happen, bad, new, come, worry, know, like, hope	"COVID, monkeypox, and now someone in NY has Polio. My oh my"
Geographic reporting of monkeypox cases	137 823 (18)	−0.27	case, confirm, report, uk, say, country, spread, new, outbreak, virus, cdc, state, rise, news, detect, identify, europe, india, infection, total	"Monkeypox is spreading among gay men worldwide—aidsmap https://t.co/N1rvoMun2c via @[…]"
Sexual transmission of monkeypox	121 745 (15.9)	−0.34	spread, know, people, say, think, virus, sex, stop, cdc, gay, want, gay_man, like, airborne, tell, man, disease, mask, go, way	"Lol monkey pox is spread through sex. They going to ban gay people as monkey pox is linked by gay men sex. So basically it;s a STD"
Learning from smallpox vaccine for monkeypox	115 827 (15.1)	−0.20	vaccine, people, COVID, need, go, know, virus, vax, new, smallpox, smallpox_vaccine, work, shingle, think, get, protect, jab, vaccinate, hear, take	"@[…] Jabs versus monkeypox are smallpox jabs. ACAM2000 is live virus and fully able to replicate!! JYNNEOS is liveâ€¦ https://t.co/N7nqxVEuFS"
Anger about monkeypox	110 808 (14.4)	−0.25	like, come, thing, s***, go, pandemic, f***, COVID, think, get, know, look, want, new, look_like, good, oh, catch, lol, need	"F*** off with the monkey pox bulls***. What the f*** is going on on this world? #newmediafear #justinsplantoreset #sideeffectsofvax"
Concerns about monkeypox as a global health emergency	64 825 (8.5)	−0.48	outbreak, virus, declare_global_health_emergency, spread, symptom, global, need_know, africa, concern, disease, contain, public_health_emergency, human, global_emergency, find, know, declare_outbreak, world_health_organization, world_health_organization_declare, youtube	"The World Health Organization has called for "urgent" action to prevent the spread of monkeypox in Europe. https://t.co/QqlaHUV5qR"
Global health response and monkeypox vaccine	26 190 (3.4)	−0.13	vaccine, declare_public_health_emergency, test, international_concern, nyc, news, demand, wait, appointment, today, man, supply, child_diagnose, new_york_city, start, run, nychealthy, read, offer, update	"Toronto monkeypox vaccine clinic targets high-risk communities â€"CBC News: TheÂ National https://t.co/qph8KtFvsj"
Declaration of monkeypox as a global health emergency	22 174 (2.9)	−0.29	declare_global_emergency, emergency, health_official, india_report, kerala, expert, consider_declare_global_emergency, vaccine_available, un_health_agency_chief, news, declares_spread_global_health, child_diagnose_official, un_health_agency, cdc_raise_alert_level, infection, prevention, say, testing, consider_declare_global_health, health	"They have been talking about it for a month now … WHO considers declaring monkeypox a global health emergency https://t.co/81hnl6ZFzZ"
Surprise at monkeypox spread outside Africa	13 129 (1.7)	−0.27	new, worried, concerned, europe, say, biden_say, virus, std, cdc, spread, virus_entrench, spread_sex_rave, join, likely, expert, african_scientist_baffle, vaccine_appointment, case, pm, revealed_fauci_recent_grant	"Monkeypox spreading in Europe, US, has African scientists baffled https://t.co/W2D8tpoowB"

Tweets are shown in an unedited form.

Table 3. Sentiment Analysis and Topic Modeling of Tweets by Topic for Tweets Labeled as Either Bots or Humans

Topic	Tweets in Topic, No. (%)	Mean Sentiment of Topic	Topic Keywords	Representative Tweets
Suspected to be bots				
Global reporting of monkeypox cases	47 759 (19.5)	−0.31	case, confirm, report, news, health, global, say, outbreak, suspect, emergency, world_health_organization, canada, country, rise, world, india, health_emergency, total, detect, record	"The discovery of the sixth case of monkeypox infection in "Israel", according to a statement by the #Israeli Ministry of Health."
Comparing COVID to monkeypox	45 214 (18.4)	−0.30	COVID, come, go, think, get, need, people, like, vaccine, new, tell, spread, good, want, let, live, watch, mask, look, thing	"LMAO now we got monkeypox in a nation of fools who don't wash their hands and believe masks and vaccines are bad."
Public health reporting on new monkeypox cases	38 543 (15.7)	−0.30	case, report, uk, spread, new, confirm, state, country, health_official, europe, outbreak, official, identify, say, detect, number, cases, probable, health, massachusetts	"#Monkeypox spreading in UK through community transmission, with new cases identified "daily", says senior doctor |ΓÇª https://t.co/r4GfYYlDkL"
Encouragement to vaccinate against monkeypox	24 418 (9.9)	−0.37	vaccine, COVID, people, pandemic, say, time, spread, smallpox_vaccine, start, way, government, give, cdc, aids, fight, effect, vaccines, week, hear, racist	"Yes. The smallpox vax is also made to fight monkey pox. We have vaccines stored. The vaccine is effective and fairlâ€¦ https://t.co/TL5tmgPsQg"
Public health response to monkeypox	19 859 (8.1)	−0.54	outbreak, expert, health, world_health_organization, fear, protect, contain, say, declare_global_emergency, spread, control, agency, public_health_emergency, virus, community, response, see, like, find, COVID	"UN health agency declares monkeypox a global emergency | WTOP News Moneypox is declared an emergency. Great ≡ƒÖäΓÇª coviΓÇª https://t.co/Mp1okMuBgl"
Constructive sharing of monkeypox health information	19 281 (7.9)	−0.46	symptom, outbreak, vaccine, call, test, world, spread, country, know, new, sex, patient, time, transmission, read, case, africa, stop, nearly, prevent	"U.S. expands monkeypox testing in bid to better track scope of outbreak https://t.co/7x04gJjAhb"
Broad discussion of monkeypox virus	19 155 (7.8)	−0.43	virus, man, outbreak, disease, need_know, infection, uk, youtube, health, spread, say, symptom, cdc, new, like, explain, thing, biden, case, issue	"TW: Monkeypox Virus Explained: How contagious is Monkeypox virus and how to prevent it?: ETHealthWorld Updated: Jâ€¦ https://t.co/T4Z2xLm7Xg"
CDC and public health responses to monkeypox	16 097 (6.6)	−0.35	cdc, report, case, say, death, today, far, worried, airborne, recent, update, story, urge, suggest, week, public_health_official, disease, infect, new, COVID	"America is done worrying about COVID-19, but CDC says monkeypox is new reason to fret-Washington TimesΓÇª https://t.co/QHBYox2D1b"
News of monkeypox risk to gay men	15 186 (6.2)	−0.29	know, spread, say, year, doctor, catch, europe, case, get, risk, emergency, outbreak, begin, ready, contract, talk, vaccine_dose, nigeria, man_sex_man, vaccine	"@[…] "And while anyone can catch monkeypox, the outbreak is overwhelmingly concentrated in gay and bisexualâ€¦ https://t.co/G4gSqklw5k"
Not suspected to be a bot (humans)				
Speculative discussion about COVID and monkeypox transmission	43 128 (22.2)	−0.32	COVID, time, like, go, new, say, people, work, catch, way, airborne, want, mask, think, start, cdc, happen, vax, come, bad	"Online bad info and misinformed hot takes are bad enough for COVID, but it looks like they are going to be off the charts for monkeypox."
Sharing informal information on monkeypox	23 442 (12)	−0.35	people, virus, know, need, case, go, like, today, human, think, c, infect, pandemic, say, outbreak, point, n, probably, youtube, lot	"Did they not learn anything from HIV/AIDS? Framing monkeypox as something only geigh people need to be cautious aboâ€¦ https://t.co/LSYAy5NOWh"
Expletive filled sharing of informal monkeypox information	22 764 (11.6)	−0.29	get, spread, know, right, wait, f***, read, people, COVID, like, shingle, let, jab, actually, look_like, thing, question, think, leave, aids	"Monkey pox is a cover up for the massive amount of people that are getting shingles because of the gene therapy shots."
Comparing other pandemic vaccines to monkeypox vaccine	21 286 (10.9)	−0.32	vaccine, outbreak, pandemic, look, year, smallpox, COVID, smallpox_vaccine, yes, find, think, maybe, news, like, good, give, biden, avoid, problem, small	"Yes please. U.K. to offer smallpox vaccine shots to health workers as monkeypox spreads in Europe | CBC News https://t.co/ZXdlt5HIro"
Formal reporting of monkeypox cases	20 243 (10.3)	−0.32	case, report, disease, virus, man, cdc, patient, symptom, health, infection, cdc_gov, vaccination, state, transmission, canada, nyc, suspect_case, confirm, death, article	"Oklahoma Health Dept. identifies second probable case of monkeypox in state - KOCO Oklahoma City https://t.co/wrIhDe6yPM via @[…]"

Table 3. Continued

Topic	Tweets in Topic, No. (%)	Mean Sentiment of Topic	Topic Keywords	Representative Tweets
Early cases and reporting of monkeypox outbreak	18 678 (9.5)	−0.28	case, spread, uk, new, outbreak, mean, report, hear, issue, good, ukhsa, siga, man, europe, emergency, high, live, go, run, real	"The new #monkeypox outbreak continues to spread in the UK and worldwide, mainly affecting men who have sex with men https://t.co/XyJe2oAkkj"
Alarmist discussion of monkeypox response	14 982 (7.7)	−0.27	come, stop, say, confirm_case, government, lockdown, country, need, wtf, thing, uk, plan, quarantine, hit, case_confirm, racist, think, woman, skynews, monkeypoxalypse_outbreak_quarantine_globalpandemic	"WTF is monkey pox and how is it suddenly a thing? This is sus AF!"
Fear about monkeypox vaccine shortage and vaccine effect	14 181 (7.2)	−0.25	vaccine, tell, get, think, want, go, day, worried, kill, great, interesting, read, business, long, thing, line, sex, got, kid, expose	"got my monkey pox vaccine today! thank you @[…] for supplying the staff and vaccines! ðŸ'ðŸ¦ ðŸ'‰"
Encouraging monkeypox testing	9715 (5)	−0.26	test, s***, vaccine, lol, week, go, call, release, effect, cdc, treatment, die, protect, tweet, suppose, send, want, update, COVID, video	"Inbox: White House announces new CDC partnership with Sonic Healthcare USA to boost capacity for monkeypox testing. https://t.co/pqqJDBpVOh"
Growing alarm about monkeypox risk	7395 (3.8)	−0.21	concern, world, vaccine, oh, speak, declare_global_health_emergency, stock, see, outbreak, approve, global_emergency, coincidence, ape, go, plandemic, world_health_organization, bill, public_health_emergency_international, pathogen, sorry	"In the past week my monkeypox concern meter has moved from 'not at all concerned' to 'not concerned' to 'not concerned, yet' to 'hmmm …' "

Tweets are shown in an unedited form.
Abbreviations: CDC, Centers for Disease Control and Prevention; COVID, coronavirus disease 2019.

government control were frequently mentioned in the cohort of humans, while similar suspiciousness is present in tweets that discussed monkeypox as a result of COVID-19 vaccination. Anger was also present in the majority of the cohorts in various ways. One such noteworthy example was the use of anger in the self-identified allies cohort when describing how "monkeypox outbreaks could have been prevented through the use of vaccinations for smallpox" in younger individuals, who were not required to receive the smallpox vaccine as part of their childhood vaccination regimen. In contrast, anger was also present in the non-allied cohort through the use of explicative language and emotionally charged statements about the fact that there was another significant global viral outbreak so soon after the COVID-19 pandemic.

The role that the LGBTQ+ community played in the spread of the monkeypox virus around the globe was present in the topic modeling for all 4 cohorts. Self-identified allies and non-allies both identified the gay community as at increased risk of monkeypox, but allies more explicitly highlighted the generalized risk, with many especially concerned about stigmatization of their community in relation to this virus. Comparing the tweets of bots and humans, humans were more likely to share conspiratorial, stigmatizing, and homophobic content in their tweets.

Contrasting This Study Adds With the Literature

A survey in January 2023 of the literature involving Twitter and monkeypox demonstrated limited availability of publications. Of the 18 articles identified, 4 were related to our study. Ortiz-Martínez and colleagues' [37] analysis of the top 100 tweets from late May 2022 discussed the spread of misinformation and the rate at which misinformation is shared through replies and retweets. Farahat et al [38] discussed tweets relating to the monkeypox virus and analyzed a collection of >8000 tweets through topic modeling and sentiment analysis. A study completed by Ng et al also focused on topic modeling and sentiment analysis but used a larger dataset of >350 000 tweets [39]. A dataset of monkeypox tweets was presented and made available for analysis by Thakur [7]. Thakur presented the methods by which the tweets were collected, along with brief quantitative descriptions of the data and sentiment analysis revealing a majority neutral sentiment toward monkeypox.

One major similarity between the articles identified in the literature search and our study was the use of an LDA model algorithm to complete topic modeling for the shorter texts typical for Twitter. The LDA algorithm is one of the more widely used topic model methods with a recent study demonstrating that the LDA algorithm "produced higher-quality topics" and was "more flexible" than other methods [40]. While the 4 articles identified in the literature search have similar methods, our study is unique in its depth of demographical analysis and investigation of the contrasting opinions between LGBTQ+ allies and non-allies. We analyzed the age, gender, and ethnicity of the Twitter users in our dataset and compared the sentiments, emotions, and topics that were discussed by users who self-identified as LGBTQ+ allies to those who did not.

Limitations

This study has several limitations to generalizability that require discussion. Using social media data limited the study to the expressed views of people who have the means and access to technology and internet capabilities. We also limited our study to English-language tweets. According to a study completed by the Pew research center of Twitter users in the United States (US), 25% of users post 97% of all US-located tweets [9]. This indicates that even though our study retrieved a large number of tweets over the period, we likely collected the viewpoints of a narrow subset of the population.

Another limitation to the study was the lack of location metadata available from Twitter. Based on a study presented in 2019, in an analysis of >40 billion tweets, only 2.31% of those tweets were geotagged [41]. As of 18 June 2019, the option of automatically logging precise geotagging of tweets became unavailable to most Twitter users [42]. This left Twitter users limited options such as general place tagging (tagging a general location as opposed to using longitude and latitude coordinate), geotagging pictures in the tweets, or using a third-party app to include tweet locations [43]. Only a small percentage of tweets collected were precisely geotagged, and we therefore did not include location in our analysis.

Our classification of LGBTQ+ allies and non-allies was based on self-identification in the user profile. We reference several LGBTQ+ resources to devise a list of relevant and current terms. However, language that references the LGBTQ+ communities can be fluid and changes over time. Also, not all users who are allies to the LGBTQ+ community identify as such publicly in their user profiles. Thus, there is likely misclassification of allies that would underestimate any differences between allies and non-allies in this study.

CONCLUSIONS

After analyzing >850 000 English-language tweets that utilized the keyword "monkeypox," we obtained a collection of tweets that exhibited mainly negative emotions such as fear and sadness over yet another global outbreak. While some tweets exhibited misinformation and helped fuel conspiracy theories about the monkeypox virus, its origins, and vaccinations, many tweets were able to disseminate vital public health information to a global audience such as locations where vaccinations were available to tracking global case counts. By identifying LGBTQ+ allies out of a large group of Twitter users, we were able to examine the unique viewpoints of a community that had largely been affected by monkeypox; in the event of future public health events, this type of analysis could be repeated to provide better healthcare opportunities to other affected populations.

Notes

Patient consent. Only publicly available data were used for this manuscript; therefore, this study did not include factors necessitating patient consent.

Disclaimer. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Financial support. This work was supported by grant funding from the Centers for Disease Control and Prevention and research funding from Verily Life Sciences, the Sergey Brin Family Foundation, and the Texas Health Resources Clinical Scholar program to R. J. M. Research reported in this publication was also supported by the National Center for Advancing Translational Sciences of the National Institutes of Health (award number UL1 TR003163).

Potential conflicts of interest. The authors: No reported conflicts of interest.

References

1. World Health Organization. WHO recommends new name for monkeypox disease. **2023**. Available at: https://www.who.int/news/item/28-11-2022-who-recommends-new-name-for-monkeypox-disease. Accessed 30 January 2023.
2. World Health Organization. Multi-country monkeypox outbreak in non-endemic countries. **2022**. Available at: https://www.who.int/emergencies/disease-outbreak-news/item/2022-DON385. Accessed 22 July 2022.
3. University Hospitals. Monkeypox and COVID-19: here are key differences. **2022**. Available at: https://www.uhhospitals.org/Healthy-at-UH/articles/2022/05/monkeypox-and-covid-19-here-are-key-differences. Accessed 26 July 2022.
4. Kozinets R. Netnography: redefined. 2nd ed. Thousand Oaks, CA: SAGE Publications; **2015**.
5. Medford RJ, Saleh SN, Sumarsono A, Perl TM, Lehmann CU. An "infodemic": leveraging high-volume Twitter data to understand early public sentiment for the coronavirus disease 2019 outbreak. Open Forum Infect Dis **2020**; 7:ofaa258.
6. Saleh SN, Lehmann CU, McDonald SA, Basit MA, Medford RJ. Understanding public perception of coronavirus disease 2019 (COVID-19) social distancing on Twitter. Infect Control Hosp Epidemiol **2021**; 42:131–8.
7. Thakur N. Monkeypox2022tweets: a large-scale twitter dataset on the 2022 monkeypox outbreak, findings from analysis of tweets, and open research questions. Infect Dis Rep **2022**; 14:855–83.
8. Twitter announces second quarter 2022 results. **2022**. Available at: https://s22.q4cdn.com/826641620/files/doc_financials/2022/q2/Final-Q2'22-Earnings-Release.pdf. Accessed 22 July 2022.
9. McClain C, Widjaya R, Rivero G, Smith A. The behaviors and attitudes of U.S. adults on Twitter. **2021**. Available at: https://www.pewresearch.org/internet/2021/11/15/the-behaviors-and-attitudes-of-u-s-adults-on-twitter/. Accessed 22 July 2022.
10. Github. Twint. **2022**. Available at: https://github.com/twintproject/twint. Accessed 22 July 2022.
11. Python Software Foundation. Python. Version 3.10.5. **2022**. Available at: https://www.python.org/. Accessed 22 July 2022.
12. Twitter, Inc. Getting started with the Twitter API. **2022**. Available at: https://developer.twitter.com/en/docs/twitter-api/getting-started/about-twitter-api. Accessed 8 September 2022.
13. Roesslein J. Tweepy documentation. **2022**. Available at: https://docs.tweepy.org/en/latest/. Accessed 8 September 2022.
14. Sayyadiharikandeh M, Varol O, Yang K-C, Flammini A, Menczer F. Detection of novel social bots by ensembles of specialized classifiers. In: Proceedings of the 29th Association for Computing Machinery International Conference on Information and Knowledge Management, Virtual/Ireland, **2020**. Available at: https://doi.org/10.1145/3340531.3412698. Accessed 8 September 2022.
15. ExplosionAI GmbH. spaCy: Industrial-Strength Natural Language Processing. **2022**. Available at: https://spacy.io/. Accessed 22 July 2022.
16. Sojka P, Řehůřek R. Software framework for topic modelling with large corpora. LREC 2010 Workshop on New Challenges for NLP Frameworks; **2010**:45–50.
17. Wang Z, Hale S, Adelani DI, et al. Demographic inference and representative population estimates from multilingual social media data. In: World Wide Web Conference, San Francisco, CA, **2019**. Available at: https://doi.org/10.1145/3308558.3313684. Accessed 22 July 2022.
18. Laohaprapanon S, Sood G, Naji B. ethnicolr: predict race and ethnicity from name. **2022**. Available at: https://ethnicolr.readthedocs.io/ethnicolr.html. Accessed 22 July 2022.
19. Sood G, Laohaprapanon S. Predicting race and ethnicity from the sequence of characters in a name. arXiv [Preprint]. Posted online 5 May 2018. doi:10.48550/ARXIV.1805.02109

20. University of California, San Francisco LGBT Resource Center. General definitions. Available at: https://lgbt.ucsf.edu/glossary-terms. Accessed 8 September 2022.
21. University of Florida. LGBTQ terms and definitions. Available at: https://lgbtq.multicultural.ufl.edu/programs/speakersbureau/lgbtq-terms-definitions/. Accessed 8 September 2022.
22. Thelwall M, Buckley K, Paltoglou G, Cai D, Kappas A. Sentiment strength detection in short informal text. J Am Soc Inform Sci Tech **2010**; 61:2544–58.
23. Aman Gupta AB, Sharma S, Bilakhiya K. Text2emotion. **2022**. Available at: https://shivamsharma26.github.io/text2emotion/. Accessed 22 July 2022.
24. Stevens K, Kegelmeyer P, Andrzejewski D, Buttler D. Exploring topic coherence over many models and many topics. Proceedings of the 2012 joint conference on empirical methods in natural language processing and computational natural language learning. Assoc Comput Linguist **2012**:952–61.
25. Monkeypoxtally.info. Monkeypox tally. **2022**. Available at: https://www.monkeypoxtally.info/. Accessed 2 August 2022.
26. Wikipedia. Social media. **2021**. Available at: https://en.wikipedia.org/w/index.php?title=Social_media&oldid=1028133356. Accessed 13 June 2021.
27. Petrović Z K, Peraica T, Kozarić-Kovačić D, Palavra IR. Internet use and internet-based addictive behaviours during coronavirus pandemic. Curr Opin Psychiatry **2022**; 35:324–31.
28. Coyne SM, Rogers AA, Zurcher JD, Stockdale L, Booth M. Does time spent using social media impact mental health? An eight year longitudinal study. Comput Human Behav **2020**; 104:106160.
29. Huang C. A meta-analysis of the problematic social media use and mental health. Int J Soc Psychiatry **2022**; 68:12–33.
30. Haddad JM, Macenski C, Mosier-Mills A, et al. The impact of social media on college mental health during the COVID-19 pandemic: a multinational review of the existing literature. Curr Psychiatry Rep **2021**; 23:70.
31. Schønning V, Hjetland GJ, Aarø LE, Skogen JC. Social media use and mental health and well-being among adolescents—a scoping review. Front Psychol **2020**; 11:1949.
32. Cunningham S, Hudson CC, Harkness K. Social media and depression symptoms: a meta-analysis. Res Child Adolesc Psychopathol **2021**; 49:241–53.
33. Price M, Legrand AC, Brier ZMF, et al. Doomscrolling during COVID-19: the negative association between daily social and traditional media consumption and mental health symptoms during the COVID-19 pandemic. Psychol Trauma **2022**; 14:1338–46.
34. Nandita Krishnan JG, Rebekah Tromble, Lorien C. Abroms. Research note: Examining how various social media platforms have responded to COVID-19 misinformation. Harvard Kennedy School Misinformation Rev **2021**; 2(6). https://doi.org/10.37016/mr-2020-85. Accessed 27 March 2023.
35. Roth Y. Introducing our crisis misinformation policy. **2022**. Available at: https://blog.twitter.com/en_us/topics/company/2022/introducing-our-crisis-misinformation-policy. Accessed 22 July 2022.
36. Lanier HD, Diaz MI, Saleh SN, Lehmann CU, Medford RJ. Analyzing COVID-19 disinformation on Twitter using the hashtags #scamdemic and #plandemic: retrospective study. PLoS One **2022**; 17:e0268409.
37. Ortiz-Martínez Y, Sarmiento J, Bonilla-Aldana DK, Rodríguez-Morales AJ. Monkeypox goes viral: measuring the misinformation outbreak on Twitter. J Infect Dev Ctries **2022**; 16:1218–20.
38. Farahat RA, Head MG, Tharwat S, et al. Infodemic and the fear of monkeypox: call for action. Trop Med Health **2022**; 50:63.
39. Ng QX, Yau CE, Lim YL, Wong LKT, Liew TM. Public sentiment on the global outbreak of monkeypox: an unsupervised machine learning analysis of 352,182 Twitter posts. Public Health **2022**; 213:1–4.
40. Albalawi R, Yeap TH, Benyoucef M. Using topic modeling methods for short-text data: a comparative analysis. Front Artif Intell **2020**; 3:42.
41. Huang B, Carley KM. A large-scale empirical study of geotagging behavior on Twitter. In: Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Vancouver, British Columbia, Canada, 2019.
42. Twitter, Inc. Twitter support. Available at: https://twitter.com/TwitterSupport/status/1141039841993355264. Accessed 22 July 2022.
43. Hu Y, Wang R-Q. Understanding the removal of precise geotagging in tweets. Nat Hum Behav **2020**; 4:1219–21.