Original Paper

Themes, Temporal Dynamics, and Public Engagement in COVID-19 Vaccination Disclosures on Twitter: Computational Content Analysis

Emily McKinley[1], PhD; Cuihua Shen[1], PhD; Kenji Sagae[2], PhD; Jingwen Zhang[1], PhD

[1]Department of Communication, University of California, Davis, Davis, CA, United States
[2]Department of Linguistics, University of California, Davis, Davis, CA, United States

Corresponding Author:

Jingwen Zhang, PhD
Department of Communication
University of California, Davis
469 Kerr Hall, One Shields Avenue
Davis, CA 95616-5270
United States
Phone: 1 530-752-0966
Email: jwzzhang@ucdavis.edu

Abstract

Background: Social media has become a primary infrastructure for vaccine communication. Personal vaccination disclosures, or first-person announcements of vaccination decisions shared publicly online, represent an underexamined channel through which peer influence may shape vaccine attitudes and behavior. Understanding what individuals disclose when they announce vaccination, and how those disclosures evolve over time, offers actionable insights for designing peer-based communication strategies in future health emergencies.

Objective: This study aimed to identify the themes present in COVID-19 vaccination disclosures on Twitter (rebranded as X in 2023), examine how those themes changed across 4 phases of the US vaccine rollout, and determine which themes were associated with greater public engagement.

Methods: Guided by the disclosure decision model (DDM), we analyzed 207,647 vaccination disclosure posts on Twitter from December 14, 2020, to January 1, 2022. Posts were identified using a robustly optimized BERT pretraining approach (RoBERTa) binary classifier (F_1-score=0.85) trained on 6348 human-coded posts. A mixed deductive-inductive coding approach, grounded in the DDM's motivational goals, yielded 11 disclosure themes (Krippendorff α=0.72–0.93). Eleven additional RoBERTa classifiers were trained to label themes across the full dataset. Network analysis examined co-occurrence of themes within each rollout phase. Negative binomial regression assessed the association between disclosure themes and an engagement metric (sum of likes, reposts, quotes, and replies), controlling for follower count, media presence, word count, sentiment, and other post-level factors.

Results: Eleven distinct disclosure themes were identified, including community identity (n=31,847, 15.34%), encouraging others (n=29,619, 14.26%), and gratitude (n=29,511, 14.21%). Themes shifted meaningfully across rollout phases: gratitude dominated phase 1, community identity rose in phases 2 and 4, and encouraging vaccination became prominent in phase 3. Gratitude was associated with a 25% increase in engagement (incidence rate ratio [IRR]=1.25; 95% CI 1.23-1.28; P<.001), alongside relief (IRR=1.26; 95% CI 1.19-1.33; P<.001) and community identity (IRR=1.19; 95% CI 1.17-1.21; P<.001). Side effects (IRR=0.85; 95% CI 0.83-0.87; P<.001), political (IRR=0.90; 95% CI 0.87-0.94; P<.001), and social benefits disclosures (IRR=0.92; 95% CI 0.90-0.95; P<.001) were associated with decreased engagement. Each additional theme in a post was associated with a 7% increase in engagement (IRR=1.07; 95% CI 1.06-1.08; P<.001).

Conclusions: Vaccination disclosures on Twitter were thematically patterned, theoretically interpretable via the DDM, and responsive to changing public health contexts. Identity-based and emotionally resonant disclosures, particularly gratitude and community identity, were among the themes most strongly associated with public engagement. These findings provide an empirically grounded framework for designing peer-based vaccine communication strategies that leverage voluntary self-disclosure in future immunization campaigns.

J Med Internet Res 2026;28:e99837; doi: 10.2196/99837

Keywords: COVID-19; social media; network analysis; semantic network analysis; vaccination; vaccines; self-disclosure; disclosure; vaccine disclosure

Introduction

Background

Vaccines remain one of the most powerful tools available for managing infectious diseases at a population level, and social media has become an increasingly central infrastructure through which vaccine-related attitudes, decisions, and experiences are communicated and shaped [1,2]. The COVID-19 pandemic, the most devastating pandemic of the social media era, posed new challenges for researchers and practitioners rethinking what constitutes effective health communication. The rapid development and distribution of COVID-19 vaccines generated varied public responses ranging from acceptance to skepticism and refusal, amid fear, hope, and widespread uncertainty, reflecting rising vaccine hesitancy and political polarization [3]. Social media emerged as a key platform for exchanging information and opinions about health behaviors, particularly COVID-19 vaccination [4]. As public health agencies now work to develop communication strategies for future health emergencies, understanding how individuals voluntarily share vaccine-related experiences online is critical. The COVID-19 vaccine rollout, which unfolded across distinct phases from December 2020 through January 2022, represents an unparalleled natural experiment in vaccine communication, providing a time-stamped record of how the public engaged with, responded to, and discussed vaccination in real time. Insights drawn from this period offer an empirical foundation for designing future immunization communication strategies at a time when vaccine hesitancy remains a persistent barrier to global health [2,5].

Although social media vaccine discourse has received growing scholarly attention, existing research has focused predominantly on sentiment analysis, misinformation dynamics, and the effects of narrative framing on vaccine intentions [3,4,6,7]. Far less attention has been paid to the specific content and temporal dynamics of organic, first-person vaccination disclosures, defined as messages in which individuals publicly announce their own decision to be vaccinated. The limited work that does exist, including studies of vaccination selfies [8,9] and general vaccine discourse content [4], does not provide a theoretically grounded, large-scale taxonomy of the themes that characterize these disclosures or examine how those themes shift across different phases of a rollout. This gap is consequential, as personal disclosures have been shown to shape peer attitudes and vaccination intentions [10,11], and understanding what people say when they disclose vaccination, and which types of disclosures generate the most public engagement, offers actionable guidance for health communicators seeking to leverage peer influence at scale.

This study addresses that gap by investigating the evolving digital discourse surrounding online vaccine disclosures during the COVID-19 vaccine rollout. Guided by the disclosure decision model (DDM) [12], which provides a framework for understanding how and why individuals share personal information, we use computational methods to analyze 207,647 vaccination disclosure posts on Twitter (rebranded as X in 2023) across 4 distinct rollout phases from December 14, 2020, to January 1, 2022. By identifying the themes present in these disclosures, tracking how those themes changed over time, and examining which themes were associated with greater public engagement, this study provides an empirically grounded understanding of how personal narratives, digital platforms, and public engagement interact to shape health information during a health crisis. The findings offer practical insights for designing vaccine communication strategies that build on peer disclosure in future public health emergencies.

Online Health Disclosures

Self-disclosure, or sharing personal details about oneself [13], plays a significant role in online health communication. Digital platforms have transformed health information dissemination, allowing people to share health experiences and decisions [14,15]. Such disclosures can enhance emotional well-being and social support [16,17]. Recent studies highlight how situational contexts and technology complicate online self-disclosure, affecting psychological well-being during crises [18]. The pandemic, with its focus on vaccines, offers a unique context to study online vaccine disclosures. Health care workers believed sharing personal vaccine stories positively influenced their social circles during this time [9]. This supports findings that personal anecdotes often resonate more than statistical data in health communication [19,20] and can be powerful in public health messaging about vaccines [21]. Sharing personal vaccination experiences online can shape public attitudes and intentions regarding vaccines [8,10,11,22]. However, a limited understanding of the content and dynamics of these types of messages hinders the ability to fully capitalize on their potential in online health communication and vaccine promotion.

DDM

The DDM [12] provides a framework for understanding how and why individuals disclose personal information across various contexts, including health disclosures. The DDM posits that disclosure decisions are influenced by both internal motivations and external contextual factors, with individuals engaging in a risk-reward assessment before disclosing personal information. According to the model, individuals disclose to achieve 1 or more of 5 primary social goals: identity clarification, relief of distress, social control, social approval, and relationship development [12,23,24]. For social

media contexts, the model was extended to include the goals of information storage/entertainment and information sharing to benefit others [23].

The pandemic created a unique disclosure environment characterized by uncertainty, fear, and a collective desire for normalcy, likely amplifying the salience of these social goals. The DDM indicates that individuals weigh potential risks against rewards when deciding whether to disclose personal information. During the COVID-19 vaccine rollout, these risk-reward calculations likely shifted as vaccination became more widespread, public discourse evolved, and the social meaning of vaccination changed. The 4 distinct phases of the vaccine rollout represent different contextual environments that might have influenced individuals' willingness to disclose their vaccination status. Hence, we pose the following research question (RQ): RQ1*:* How did the number of vaccination disclosures change as the vaccine rollout progressed?

The DDM identifies abstract motivational goals that drive self-disclosure, but the specific ways these goals manifest in COVID-19 vaccination disclosures require investigation. Each of the DDM's motivational goals might be expressed through specific, context-dependent themes in vaccination disclosures. For example, identity clarification might manifest when individuals express affiliation with specific vaccine communities or political groups (eg, "Team Moderna," "Vaccinated Republicans"), thereby reinforcing group identity [25]. Disclosures aimed at distress relief might involve sharing vaccine side effects or expressing stress-reducing emotions such as gratitude [26]. Social control, which involves influencing others' behavior for collective benefit, might appear as explicit encouragement for others to get vaccinated. By examining the themes that emerge in vaccination disclosures, we can understand how the abstract motivational goals in the DDM are concretely expressed in the context of the COVID-19 pandemic. Therefore, we ask: RQ2*:* What were the different themes of COVID-19 vaccination disclosures throughout the vaccine rollout phases from December 14, 2020, to January 1, 2022?

The DDM further suggests that as contextual factors change, the prevalence of different social goals may shift as well. During the vaccine rollout, changes in vaccine availability, public sentiment, and political discourse likely altered the risk-reward calculations associated with different types of disclosures. Early disclosures may have emphasized different social goals than later ones as the social meaning of vaccination evolved. The vaccination rollout unfolded in four distinct US phases based on availability events: initial distribution to at-risk individuals and health care workers (December 14, 2020, to March 8, 2021), expanded eligibility to individuals aged 16 years and older (March 9, 2021, to May 9, 2021), further expansion to adolescents aged 12 to 15 years (May 10, 2021, to August 31, 2021), and the introduction of booster shots (September 1, 2021, to January 1, 2022). Each phase likely made certain social goals more salient, potentially impacting the themes present in vaccination disclosures. By examining how themes changed across phases, we can understand the dynamic interplay between context and motivation that the DDM predicts. This leads to our third RQ: RQ3: How did disclosure themes change as the vaccine rollout progressed?

Finally, the DDM indicates that anticipated social reactions influence disclosure decisions, with individuals more likely to disclose content they expect will receive positive responses [12]. On social media platforms, engagement metrics provide concrete evidence of these social responses, potentially reinforcing certain disclosure behaviors over time. Social media platforms have emerged as powerful conduits for shaping social norms and behaviors within health contexts. The dissemination of COVID-19 information on social media can impact an individual's vaccine acceptance [27]. Studies indicate public health disclosures can prompt informed discussions and behavioral changes [28-30]. On platforms like X, engagement metrics, such as likes and retweets, play a crucial role in amplifying message visibility and impact by indicating acceptance and relevance [31,32]. By analyzing which disclosure themes generate greater engagement, we can understand which types of disclosures are socially rewarded in this context. Therefore, our final RQ asks: RQ4: How are different vaccination disclosure themes associated with social media engagement metrics?

Methods

Vaccination Disclosure Data Collection and Cleaning

COVID-19 vaccination disclosure posts were collected from Twitter, initiating the data collection with 5 COVID-19 vaccine-related hashtags based on previously published literature [3,33]. Using a snowball sampling method starting with the initial hashtags (Multimedia Appendix 1), additional hashtags that appeared in collected posts were gathered iteratively until no new hashtags were collected. A comprehensive list of 66,059 unique hashtags was identified. Three researchers determined that hashtags for data collection should meet at least one of the following requirements: contain "covid," indicate a vaccine-related word (ie, "vaccine," "shot," "vax," "jab," "dose," "injection"), or mention a specific vaccine brand. A total of 1161 hashtags met these criteria. Using the Twitter Academic API and a custom Python script, we used the 1161 hashtags and Twitter's metadata language field, restricted to English, as search parameters. We excluded posts containing other vaccine keywords such as "flu," "hpv," "influenza" and "smallpox." A corpus of 644,540 posts was retrieved, posted between December 14, 2020, coinciding with the commencement of COVID-19 vaccine distribution in the United States, through January 1, 2022.

The collected posts underwent preprocessing to enhance dataset quality and relevance. Duplicate posts were eliminated based on textual content. Using named entity recognition from spaCy (version 2.2.4; en_core_web_sm model, version 2.2.5; Explosion) [34] applied to account display names, accounts whose display name contained a nonperson entity (eg, an organization) were filtered out, reducing the corpus

to 497,041 posts. In a post hoc validation, a random sample of 200 accounts from the final dataset was manually coded by the lead author; 190 (95%) were found to be individuals, 4 (2%) were organizations, and 6 (3%) were unclear. Posts from accounts with "bot" in the account description were removed, condensing the dataset to 494,045 posts.

A codebook for identifying vaccine disclosures in posts was compiled and used to train 4 research assistants. A post was considered a vaccine disclosure if it clearly indicated the user had received a COVID-19 vaccine dose. Each researcher individually hand-coded the same random sample of 400 posts from the dataset, achieving an intercoder reliability Krippendorff α of 0.826. Average pairwise agreement was 91.83%, with a 1-tailed 95% lower confidence bound of 89.65%, indicating that the sample size was adequate [35]. Each coder then independently coded their own random sample of 1587 distinct posts, resulting in a human-coded dataset of 6348 posts. After human coding was complete, the textual data of each post underwent multiple preprocessing procedures for classifier training using the *re* [36] and *pandas* [37] packages in Python, including removal of URLs, removal of email addresses, removal of date content, conversion to lowercase, elimination of the hashtag symbol, contraction mapping (eg, "I'll" was converted to "I will"), and conversion of the syringe emoji (💉), which users commonly included to signify vaccination, to the word "vaccine."

Using the human-coded dataset of 6348 posts, a binary robustly optimized BERT pretraining approach (RoBERTa) [38] text classifier was trained in Python to distinguish between vaccination disclosure posts and those that were not. The classifier achieved a precision of 0.83, a recall of 0.87, and an F_1-score of 0.85 for discerning vaccination disclosure content. Combining the machine-coded vaccination disclosure posts and the human-coded posts, and removing duplicate posts based on text, a final vaccination disclosure dataset of 207,647 posts spanning December 14, 2020, to January 1, 2022, was established.

Disclosure Theme Coding

The coding approach for identifying vaccination disclosure themes (RQ2) used a mixed deductive-inductive method to balance theoretical grounding with contextual discovery. Four research assistants underwent comprehensive training on the DDM framework. This training included close reading of the original DDM literature [12,23,24] and extended discussions about each of the model's disclosure goals: the 5 primary goals (identity clarification, relief of distress, social control, social approval, and relationship development) and the 2 social media-specific goals (information storage/entertainment and information sharing to benefit others). These goals served as sensitizing concepts in the initial coding framework. Rather than coding for abstract motivations, coders identified concrete content expressions reflecting these underlying theoretical goals. For example, posts expressing affiliation with specific vaccine communities were noted as expressing "identity clarification." To develop the theme taxonomy, coders independently applied these sensitizing concepts to the same random sample of 2000 posts, then convened to compare emergent categories, reconcile differences, and establish consensus definitions and decision rules for each theme. All posts in the dataset were considered baseline vaccination disclosures, meaning they inherently indicated an individual's decision to receive the COVID-19 vaccine. This process ultimately yielded 11 discrete vaccination disclosure themes: (1) community identity, (2) gratitude, (3) social benefits, (4) encouraging others to be vaccinated, (5) vaccination experience, (6) humor, (7) side effects, (8) health benefits, (9) civic duty, (10) political, and (11) relief (Table 1). To ensure theoretical grounding, each identified theme was systematically mapped back to one or more of the DDM disclosure goals, confirming that all 11 themes aligned with the theoretical framework (Multimedia Appendix 2). The 11 identified themes were conceptually distinct from one another, but multiple themes could appear within a single post, aligning with the DDM's recognition that individuals often disclose information to serve multiple social goals simultaneously. For example, a user might express both "gratitude" toward health care workers and "community identity" as a vaccine recipient within the same post, reflecting separate but co-occurring motivational goals. After reaching consensus on the themes and their definitions, each coder individually evaluated the same random sample of 100 posts, achieving intercoder reliability scores, measured using Krippendorff α, from 0.72 to 0.93 (Table 1). Average pairwise agreement ranged from 93.50% to 99.00% across themes, with 1-tailed 95% lower confidence bounds from 89.45% to 97.37% [35].

Table 1. Descriptions, example posts, intercoder reliability, and classifier performance for 11 vaccination disclosure themes in a computational content analysis of English-language COVID-19 vaccination disclosure posts on Twitter during the US vaccine rollout (December 14, 2020, to January 1, 2022; N=207,647)[a].

Theme	Description	Example post	Krippendorff α	F_1-score (95% CI)
Community identity	User identifies with a specific vaccine community	*Just like that, I join the superior Team Moderna. #vaccinated*	0.86	0.98 (0.98-0.99)
Gratitude	User expresses gratitude related to vaccination	*Thank you to all the amazing health care and essential workers who made my vaccination possible! #vaccinated*	0.93	0.94 (0.93-0.95)
Social benefits	User refers to social benefit or benefits of the vaccine	*Got my 2nd shot of Pfizer today. Soon I can spend some long overdue time with my parents. #pfizerfam*	0.85	0.84 (0.82-0.86)
Encouraging others	User encourages others to get vaccinated	*Got my 1st shot of the Pfizer vaccine today. Be a good human and go get your shot! #imvaccinated*	0.82	0.87 (0.86-0.89)

Theme	Description	Example post	Krippendorff α	F_1-score (95% CI)
Vaccination experience	User provides further details regarding vaccination or its scheduling	*I had to play the back-and-forth game in the vaccine booking portal, but I finally snagged me an appointment yesterday! The nurse gave me an awesome vaccine sticker. #getvaccinated*	0.72	0.80 (0.78-0.82)
Humor	User incorporates humor in their vaccination disclosure	*I'm fully vaccinated! Still waiting on my powers. My 5G connection isn't faster yet!*	0.73	0.71 (0.68-0.73)
Side effects	User indicates the presence or absence of vaccine side effects	*Not gonna lie, I have felt like crap for the last 30 hours following my second #PfizerVaccine (fever, nausea, headache, etc.) Worth it though!*	0.92	0.91 (0.90-0.93)
Health benefits	User refers to health benefit or benefits of the vaccine	*I have been vaccinated, and still have not contracted COVID! Science works! #vaccinated.*	0.73	0.77 (0.74-0.80)
Civic duty	User communicates a sense of responsibility toward getting vaccinated	*I got vaccinated for my community. Proud to do my part to protect others and get back to normal. #vaccinatedandhappy*	0.92	0.87 (0.84-0.90)
Political	User includes some type of politically related content	*I voted for Biden, my parents voted for Trump, but today we all got #vaccinated.*	0.76	0.87 (0.84-0.90)
Relief	User indicates vaccination resulted in some sort of relief	*So relieved to finally be vaccinated. This pandemic has been a long journey, I finally feel a weight lift off my shoulders. #covidvaccination*	0.89	0.94 (0.91-0.97)

[a]The total sample of vaccination disclosures is 207,647. Krippendorff α is reported per theme. Eleven individual classifiers based on the robustly optimized BERT pretraining approach model were trained, one for each disclosure theme; classifier performance is reported as the F_1-score for the theme-present class from 5-fold cross-validation on the human-coded posts (n=7090), with 95% CIs from 2000 bootstrap resamples.

Each coder then independently coded a distinct random sample, resulting in a total of 7187 coded posts. Coders were instructed to flag any content that did not fit the 11 established themes under an "other" designation; however, no additional themes emerged with sufficient frequency or conceptual coherence to warrant inclusion as a 12th category. These human-coded posts served as training data for 11 RoBERTa text classifiers, one corresponding to each vaccination disclosure theme, following the same preprocessing steps described above. After removal of duplicate posts based on the cleaned text, 7090 posts remained. For each theme, these posts were divided into training (n=5,672, 80%), validation (n=709, 10%), and test (n=709, 10%) sets, stratified by the presence of the theme. Each classifier was fine-tuned from the RoBERTa-base model with a maximum sequence length of 128 tokens, a batch size of 16, a learning rate of 2e-5 with linear warm-up over the first 10% of training steps, a weight decay of 0.01, and up to 10 epochs, with early stopping after 3 epochs without improvement in validation F_1-score. Posts containing the theme were given additional weight in the loss function to account for class imbalance, and the decision threshold for each theme was selected to maximize F_1-score on the validation set. Classifier performance was estimated with stratified 5-fold cross-validation, in which every human-coded post was classified once by a model that had not been trained on it; 95% CIs were obtained from 2000 bootstrap resamples. The cross-validated performance metrics (F_1-scores with 95% CIs) for each classifier can be found in Table 1; theme frequencies by rollout phase are reported in Table 2. After processing the remaining dataset through these classifiers, a final dataset of 207,647 posts identified with themes was used for subsequent analyses.

Table 2. Frequency of 11 vaccination disclosure themes overall and by US vaccine rollout phase among English-language COVID-19 vaccination disclosure posts on Twitter (December 14, 2020, to January 1, 2022; N=207,647)[a].

Theme	Total, N (%)	Phase 1, n (%)	Phase 2, n (%)	Phase 3, n (%)	Phase 4, n (%)
Community identity	31,847 (15.34)	338 (1.91)	14,559 (21.42)	9487 (10.67)	7463 (22.54)
Gratitude	29,511 (14.21)	4343 (24.57)	10,066 (14.81)	12,106 (13.62)	2996 (9.05)
Social benefits	20,835 (10.03)	1380 (7.81)	7220 (10.62)	9423 (10.60)	2812 (8.49)
Encouraging others	29,619 (14.26)	1711 (9.68)	8132 (11.97)	15,030 (16.91)	4746 (14.33)
Vaccination experience	28,197 (13.58)	2938 (16.62)	9915 (14.59)	12,347 (13.89)	2997 (9.05)
Humor	18,262 (8.79)	989 (5.60)	6844 (10.07)	8433 (9.49)	1996 (6.03)
Side effects	23,038 (11.09)	2272 (12.86)	9632 (14.17)	8253 (9.28)	2881 (8.70)
Health benefits	15,560 (7.49)	1639 (9.27)	3466 (5.10)	6157 (6.93)	4298 (12.98)
Civic duty	7655 (3.69)	1217 (6.89)	1890 (2.78)	2925 (3.29)	1623 (4.90)
Political	8573 (4.13)	612 (3.46)	2265 (3.33)	4068 (4.58)	1628 (4.92)
Relief	2650 (1.28)	386 (2.18)	1133 (1.67)	845 (0.95)	286 (0.86)

[a]The total sample of vaccination disclosures is 207,647. Phase 1 (n=17,673): initial rollout, December 14, 2020-March 8, 2021. Phase 2 (n=67,960): expanded eligibility to individuals aged 16 years and older, March 9-May 9, 2021. Phase 3 (n=88,903): approval for adolescents aged 1215 years,

May 10-August 31, 2021. Phase 4 (n=33,111): booster availability, September 1, 2021-January 1, 2022. Percentages reflect the proportion of posts within each phase that contained the theme.

Network Analysis on Changing Dynamics of Disclosure Themes

Network analysis was used to explore relationships and dynamics between vaccination disclosure themes within individual posts over time (RQ3). This method visualizes and elucidates the interconnections and evolution of vaccination disclosure themes over rollout phases. Using the *igraph* package in R, a network of disclosure themes for each of the vaccination rollout phases was constructed, in which each node represented a vaccination disclosure theme, and an edge indicated the co-occurrence between 2 vaccination disclosure themes in a post. Node size represented the frequency of a theme, while edge thickness indicated the strength or frequency of co-occurrence between 2 vaccination disclosure themes.

Regression Analysis on Engagement

An engagement metric for each post was constructed by summing the likes, reposts, quotes, and replies a post received. This metric represents the overall responsiveness and attention garnered by each post (min=0, max=73,057, mean 19.40, SD 373.96).

A negative binomial regression was used to discern the relationship between vaccination disclosure themes and the engagement metric (RQ4). The model included engagement metric (count variable) as the dependent variable and 11 vaccination disclosure themes (binary variables) as independent variables. We controlled for media presence (ie, if the post had a picture or a video), log-transformed followers count, log-transformed following count, posting date, log-transformed word count, log-transformed hashtag count, log-transformed lifetime post count, age of the user account in weeks, and sentiment score of the post (Multimedia Appendix 3 provides additional descriptive statistics). Sentiment score was analyzed using the *syuzhet* package [39] in R, yielding scores from –7 to 9 (mean 0.45, SD 1.08). The continuous sentiment score was entered in the regression model. For descriptive purposes, 53.91% (n=111,943) of the posts had positive scores, 25.37% (n=52,680) had neutral, and 20.72% (n=43,024) had negative scores. It was anticipated that certain vaccination disclosure themes might have a stronger association with engagement, and the inclusion of control variables ensured a more robust analysis.

Ethical Considerations

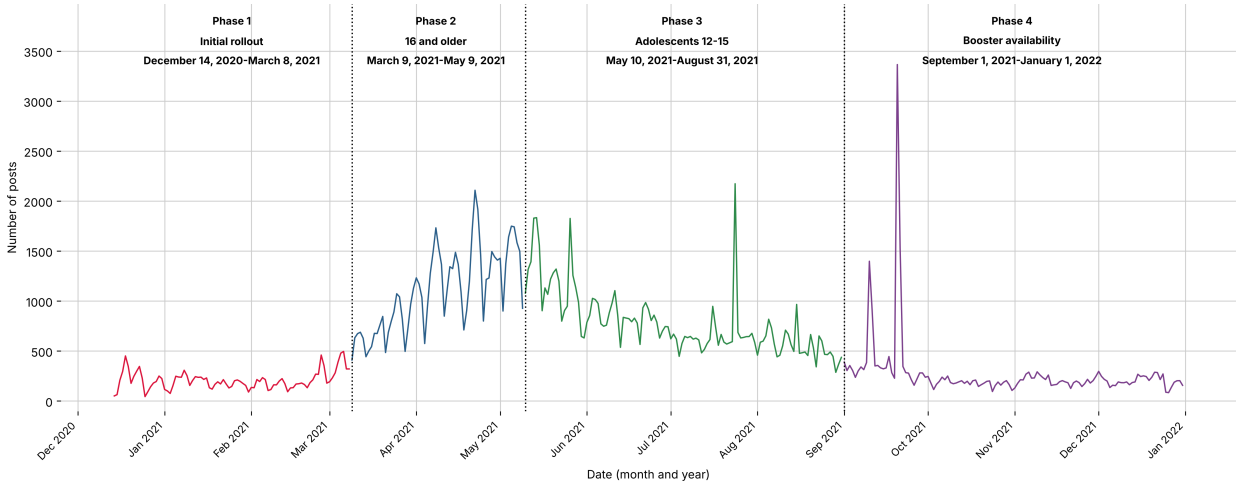
This study was approved by the University of California, Davis Institutional Review Board (IRBNet ID: 2074524-1). All data consisted of publicly available social media posts; no identifiable private information was collected. Informed consent was not required as the data were publicly accessible at the time of collection.

Results

Vaccination Disclosures Throughout Vaccination Rollout Phases

Within the initial corpus of 644,540 vaccine-related posts, 207,647 (32.22%) were identified as vaccination disclosure posts. Figure 1 illustrates temporal variations in these disclosures, aligned with the 4 phases of the COVID-19 vaccine rollout. Notably, peaks in disclosure frequency appear to correspond with pivotal developments in the rollout phases. A notable uptick in disclosures occurred with the commencement of phase 2, as vaccine eligibility expanded. Phase 3 contained the most disclosures overall (88,903), but the daily rate fell from about 1096 posts per day in phase 2 to about 780 in phase 3, which may suggest that most willing participants had been vaccinated by this stage. Phase 4 showed a brief renewed surge in disclosures, aligning with the introduction of booster shots.

Figure 1. Daily number of English-language COVID-19 vaccination disclosure posts on Twitter across 4 US vaccine rollout phases (December 14, 2020, to January 1, 2022; N=207,647). A total of 17,673 posts were made in phase 1, 67,960 in phase 2, 88,903 in phase 3, and 33,111 in phase 4.



Vaccination Disclosure Themes

Eleven distinct vaccination disclosure themes were identified in the vaccination disclosure dataset. These themes are detailed in Table 1, which provides a brief description of each theme and an example post. In total, 32.52% (n=67,520) of the posts contained no themes, 40.38% (n=83,851) contained 1 theme, and 19.56% (n=40,609) contained 2 themes; the complete distribution across all posts is provided in Multimedia Appendix 4. The most prevalent themes overall were community identity, encouraging others to be vaccinated, and gratitude. Community identity posts articulated affiliation with specific vaccinated groups, often merging preexisting identities (eg, "vaccinated republicans" and "vaccinated gays") or indicating specific vaccine brand communities (eg, "Moderna gang" and "Pfizer family"). Posts encouraging others to be vaccinated directly urged readers to get their own shot (eg, "go get your shot"). Gratitude-themed posts frequently highlighted individuals' appreciation toward health care workers, the scientific community, and the overall accessibility and effectiveness of vaccines. The detailed conceptual mapping of all 11 themes to the motivations outlined in the DDM can be found in Multimedia Appendix 2.

Disclosure Theme Networks Over Vaccination Rollout Phases

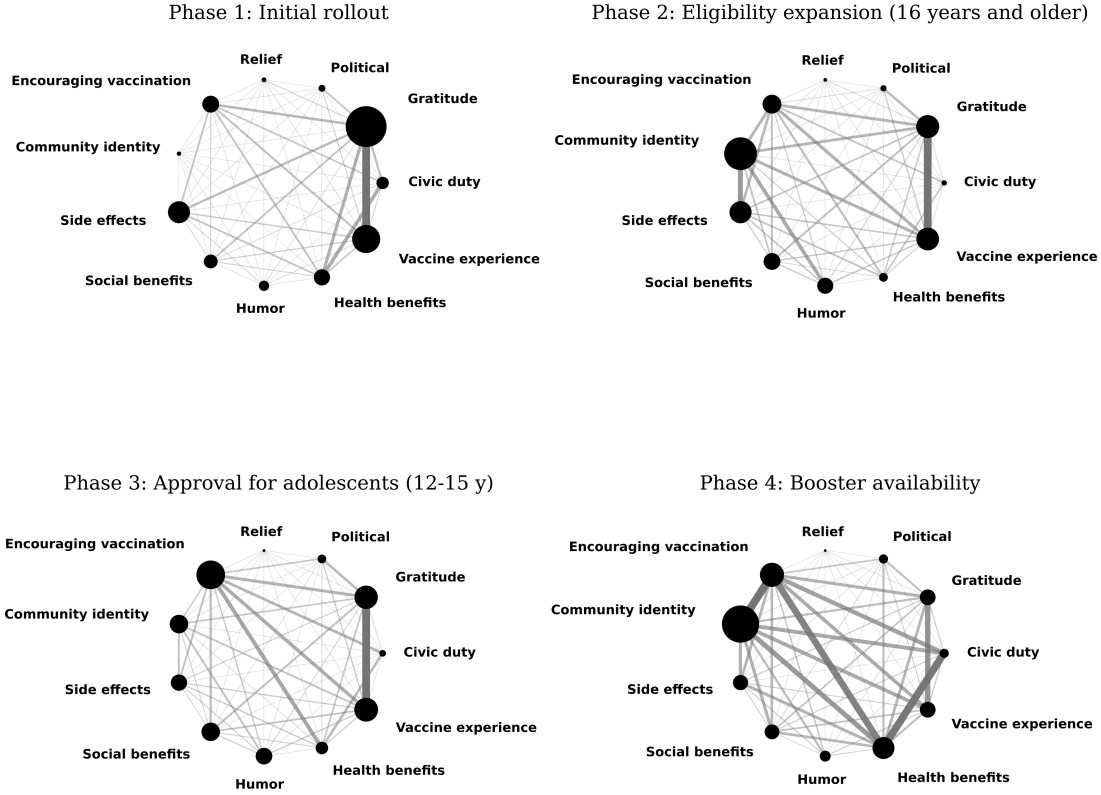
Figure 2 provides network visualizations of themes and their interrelationships during vaccine rollout phases. In all network phases, the size of the node represents the frequency of a theme's occurrence in the dataset during that period, while edge width signifies the number of co-occurrences between themes in a single post.

In phase 1, 17,673 (8.51% of the dataset) vaccination disclosures were recorded, with gratitude being the most frequent theme. During this time, 4810 posts contained 2 or more themes, producing 8478 pairwise co-occurrences. Gratitude paired with vaccination experience was the most common co-occurrence, showing up in 6.97% (n=1231) of disclosure posts in phase 1, followed by the pairing of health benefits with civic duty at 3.16% (n=558). Other co-occurrences ranged from 2.85% (n=503) to 0.01% (n=2).

During phase 2, the number of disclosures increased to 67,960, representing 32.73% of the total dataset. Community identity became a dominant theme, while gratitude remained prevalent. Gratitude and vaccination experience were the most prominent co-occurrence, appearing in 5.31% (n=3606) of the posts. The next highest co-occurrence was community identity and side effects, found in 3.38% (n=2,299) of the posts. All other co-occurrences ranged from 2.48% (n=1,688) to 0.05% (n=37). Total disclosures in phase 3 rose to 88,903, amounting to 42.81% of the entire dataset. Encouraging vaccination became the predominant theme, while vaccination experience and gratitude remained among the most frequent themes. Co-occurrence trends showed a pairing of gratitude and vaccination experience at 5.16% (n=4583), followed by encouraging vaccination and health benefits at 2.56% (n=2273). All other co-occurrences ranged from 2.26% (n=2007) to 0.02% (16).

In phase 4, 33,111 disclosures were recorded, contributing to 15.95% of the entire dataset. Community identity was the most dominant theme, followed by encouraging vaccination and health benefits. Community identity paired with encouraging vaccination appeared in 3.41% (n=1128) of the posts, and health benefits with civic duty in 3.22% (n=1067). All other co-occurrences ranged from 2.87% (n=949) to 0.02% (n=5). Complete pairwise co-occurrence counts for all themes across all 4 rollout phases are provided in Multimedia Appendix 4.

Figure 2. Co-occurrence networks of 11 vaccination disclosure themes in English-language COVID-19 vaccination disclosure posts on TwitterTwi, by US vaccine rollout phase (December 14, 2020, to January 1, 2022; N=207,647). Each node represents a vaccination disclosure theme, while an edge connecting 2 nodes means that the 2 themes appear together in a post. The thickness of the edges represents the co-occurrence frequency of the 2 themes, with thicker edges representing a greater number of co-occurrences between the 2 themes. The size of the node indicates the theme's frequency in the dataset during the indicated time period.



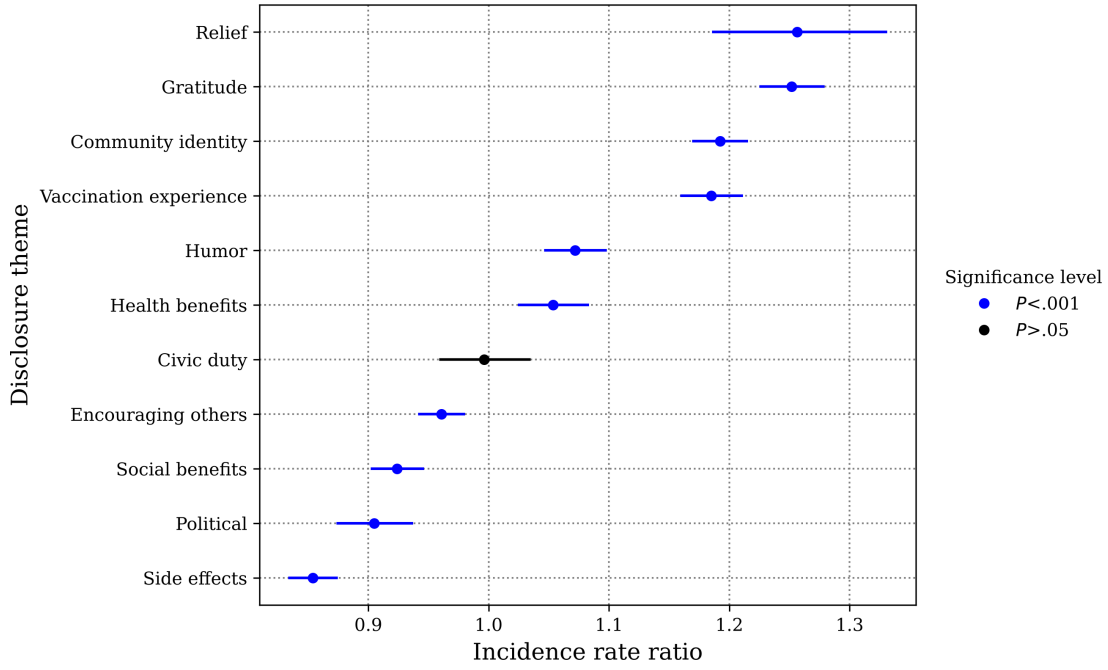
Engagement Metrics Associated With Disclosure Themes

To address the last RQ, which sought to understand the impact of the vaccination disclosure themes on engagement metrics, a negative binomial regression model was used. The analysis revealed significant associations between specific themes and engagement metrics, which can be found in Figure 3. Notably, compared to posts without gratitude, posts incorporating gratitude exhibited a 25% increase in engagement (incidence rate ratio [IRR]=1.25, 95% CI 1.23-1.28; *P*<.001). Similarly, relief (IRR=1.26, 95% CI 1.19-1.33; *P*<.001) and community identity (IRR=1.19, 95% CI 1.17-1.21; *P*<.001) themes were linked to statistically significant increases in engagement, of 26% and 19% respectively. Humor (IRR=1.07, 95% CI 1.05-1.10; *P*<.001) also displayed a positive, albeit smaller, association with engagement, increasing engagement by 7%. Posts describing the vaccination experience were also associated with higher engagement (IRR=1.19, 95% CI 1.16-1.21; *P*<.001).

Conversely, political disclosure posts were associated with a 10% decrease in engagement (IRR=0.90, 95% CI 0.87-0.94; *P*<.001). Interestingly, posts containing social benefits experienced an 8% decline in engagement (IRR=0.92, 95% CI 0.90-0.95; *P*<.001). Posts encouraging vaccination were also associated with slightly lower engagement, a 4% decrease (IRR=0.96, 95% CI 0.94-0.98; *P*<.001). Posts mentioning side effects were associated with a 15% decrease in engagement (IRR=0.85, 95% CI 0.83-0.87; *P*<.001). These findings shed light on the nuanced relationship between diverse disclosure themes and audience engagement within the context of vaccination communication.

In subsequent exploratory analysis, negative binomial regression was used to examine how the number of vaccination disclosure themes in a post relates to engagement. Controlling for the same variables as the previous model, each additional vaccination disclosure theme in a tweet was associated with a 7% increase in engagement (IRR=1.07, 95% CI 1.06-1.08; *P*<.001). This finding shows that posts with more vaccination disclosure themes had significantly higher engagement. The output for both regression models can be found in Multimedia Appendix 3.

Figure 3. Incidence rate ratios (IRRs) from a negative binomial regression of post engagement on 11 vaccination disclosure themes among English-language COVID-19 vaccination disclosure posts on Twitter (December 14, 2020, to January 1, 2022; N=207,647). Each point represents the IRR for a given vaccination disclosure theme from a negative binomial regression model predicting post engagement (N=207,647). Engagement was operationalized as the sum of likes, reposts, quotes, and replies. IRRs greater than 1.0 indicate a positive association with engagement relative to posts without that theme; IRRs less than 1.0 indicate a negative association. Horizontal lines represent 95% CIs. Blue points indicate themes with statistically significant associations ($P<.001$); black points indicate nonsignificant associations ($P>.05$). The model controlled for media presence, log-transformed follower count, log-transformed following count, posting date, log-transformed word count, log-transformed hashtag count, log-transformed lifetime post count, user account age, and sentiment score.



Discussion

Principal Findings

This study provides an empirically grounded account of how individuals publicly shared their vaccination decisions during one of the most consequential public health events of the social media era. Analyzing 207,647 vaccination disclosure posts across 4 phases of the US COVID-19 vaccine rollout, we identified 11 theoretically grounded disclosure themes, documented meaningful shifts in those themes as the rollout progressed, and found that identity-based and emotionally resonant disclosures, particularly gratitude and community identity, were associated with substantially higher public engagement. These findings extend the DDM to a large-scale, real-world health communication context, showing that themes consistent with the DDM's motivational goals appear in patterned, measurable ways in organic social media discourse. More practically, they suggest that the content of peer vaccination disclosures is not random, but patterned, theoretically interpretable, and responsive to changing public health contexts in ways that communicators can anticipate and leverage.

Themes in Vaccination Disclosures: Identity Clarification and Emotional Relief

Two prevalent themes in this study, community identity and gratitude, are consistent with 2 DDM goals: identity clarification and distress relief. The emergence of community identity as a prevalent theme is consistent with identity clarification in digital spaces and suggests that individuals may use vaccination status as a means of social identification during uncertainty. Users aligned their vaccination status with specific vaccine communities such as "Pfizer Gang" or "Vaccinated Lesbians," asserting social identities beyond mere health decisions.

The inclination toward identity discussions in vaccination disclosures can be attributed to the fundamental human need for belonging and community, a tendency spanning both offline and online settings [40,41]. Uncertainty-identity theory [42] suggests that group membership reduces feelings of self-uncertainty by prescribing social identities and behaviors, providing validation for perceptions, attitudes, and actions. This theory elucidates why, amid the global pandemic and a time of unprecedented uncertainty, individuals are more inclined to engage in identity clarification through vaccination disclosures. The alignment with vaccine-specific communities or ideologies not only serves as a means of reducing personal uncertainty but also reinforces one's sense of belonging and identity within these newly formed in-groups [42].

The findings also resonate with the social identity model of deindividuation effects, which explains how online interactions, with minimized personal identity cues, amplify social identity aspects [43]. During the COVID-19 crisis, declaring vaccination status online was an expression of group affiliation, not just a personal health choice, emphasizing social identity in digital spaces [44,45].

Gratitude emerged as a significant theme across all phases of the vaccine rollout, consistent with the DDM's recognition that self-disclosure can serve as a cathartic outlet, providing emotional relief during periods of heightened stress. Expressing gratitude has been associated with improved well-being and reduced stress [26], aiding in coping with COVID-19-related stress [46], diminishing adversity's negative impacts [47], and alleviating stress-induced psychological disturbances [48]. The widespread expression of gratitude in vaccination disclosures contributes to individual relief and serves as a collective coping mechanism, fostering communal optimism [49].

Temporal Evolution of Vaccination Disclosure Themes

The temporal analysis of disclosure themes across the vaccine rollout phases offers insights into the public's ongoing assessment of risks and rewards associated with disclosing vaccination decisions, a key process described in the DDM [12].

Phase 1 was largely characterized by gratitude, indicating strong public appreciation toward health care workers and the vaccine. During this phase, strict eligibility criteria focused on high-risk groups and health care workers, and the period was marked by an overwhelming expression of gratitude. The public's eagerness to disclose this milestone highlights the perceived rewards outweighing risks, consistent with the DDM's risk-reward calculation.

As vaccine accessibility expanded in phase 2, a shift to community identity occurred, suggesting an evolution in disclosure rewards toward expressing belonging and solidarity. This phase underscored the role of social cohesion and collective identity in the face of a global health crisis, as observed through the lens of the social identity model of deindividuation effects [50].

In phase 3, encouraging others to vaccinate became the most frequent theme. With vaccines widely available and eligibility extended to adolescents, the emphasis of disclosures appeared to shift from announcing one's own vaccination toward urging others to follow. Interpreted through the DDM, this trend suggests that individuals disclosed to influence others' behavior toward a collective good [12].

Phase 4 saw a resurgence of community identity, likely influenced by evolving pandemic dynamics and the introduction of booster shots. This resurgence, along with continued encouragement of others and greater attention to the health benefits of vaccination, reflects a complex interplay of individual and collective motivations. The attention to health benefits is consistent with survey evidence that people who received boosters were motivated largely by vaccine effectiveness and by protecting those around them [51]. Renewed confidence in vaccine safety after previous concerns [52,53] and the choice between booster brands may have also contributed to this renewed focus on community identity. Vaccination status can also function as a group identity in its own right, and among vaccinated individuals, stronger identification has been associated with a greater sense of collective responsibility [54]. This may help explain why community identity so often appeared alongside encouragement of others in this phase.

The pairing of gratitude with vaccination experience was the most common co-occurrence in the first 3 phases, suggesting that individuals often expressed thanks while recounting the details of getting vaccinated. In phase 4, the most common pairing shifted to community identity with encouraging vaccination, suggesting that later disclosures more often joined group affiliation with appeals to others. Interpreted through the DDM, these patterns suggest that a single disclosure often served more than 1 social goal and that the combination of goals shifted as the rollout progressed [12].

Disclosure Themes and Engagement

The analysis shows a strong positive association between gratitude disclosures and engagement, suggesting that expressions of thankfulness resonate deeply with audiences. Disclosures expressing relief showed a positive association of similar size. Gratitude, which is linked to positive personal, social, and physiological effects [55], likely served as a positive anchor, facilitating connection and shared sentiment. This aligns with findings on the amplifying effect of positive emotions such as gratitude on online engagement [56] and its role in fostering meaningful conversations and deeper social connections on social media [57].

Interestingly, posts themed around social benefits were associated with lower engagement. This finding could stem from varying emotional resonance. Messages that evoke strong emotions, such as those expressing gratitude for overworked health care workers or highlighting the health benefits for those at risk, might be more engaging than those focused on the return to travel or in-person dining [58]. Therefore, social benefits-themed discussions might not resonate as profoundly as more personal health or gratitude narratives, resulting in lower engagement. A complementary possibility is that some of these posts carry mixed emotional valence, pairing anticipation of social contact with references to separation or loss during the pandemic, which may elicit a less uniform response from readers. Posts encouraging others to be vaccinated were also associated with slightly lower engagement. One possible explanation is psychological reactance, in which appeals that direct others toward a health behavior can be perceived as a threat to freedom and met with resistance [59,60]. By contrast, disclosures that recount a personal experience may come across as less directive, and messages told as stories tend to provoke less reactance than messages that instruct [60]. This association was small, however, and engagement counts do not show how readers evaluated these posts.

Additionally, posts featuring multiple vaccination disclosure themes showed increased engagement, indicating that audiences value multifaceted presentations. The integration of diverse content in a single post provides a richer, more comprehensive narrative, potentially appealing to various interests and fostering broader discussions and interactions.

The variable impact of different themes on engagement underscores the complexities inherent in online health communication, particularly during a global health crisis. These insights are crucial for public health communicators, who must adapt to the shifting dynamics of public sentiment to develop resonant and effective campaigns.

Implications for Health Campaign Design

The findings from this study offer concrete guidance for public health communicators designing vaccine campaigns, particularly in contexts where peer influence and social media play central roles in shaping uptake.

The strong positive association between gratitude-themed disclosures and engagement suggests that campaign strategies encouraging vaccinated individuals to publicly express appreciation for health care workers, scientific progress, or community protection may be especially effective at generating organic reach. Gratitude content was among the themes most strongly associated with engagement, increasing it by 25%, and appeared in every rollout phase, suggesting that its resonance is not limited to a particular moment in a campaign cycle. Public health communicators may consider prompting this type of disclosure through campaign messaging that explicitly invites expressions of thanks.

Community identity disclosures were similarly associated with higher engagement and grew in prominence after the initial rollout, when eligibility expanded to the broader public. The tendency of individuals to affiliate their vaccination with specific communities, whether defined by political identity, demographics, or vaccine brand, suggests that campaigns emphasizing belonging and shared identity may resonate more broadly than those centered solely on individual health benefits. Messaging that frames vaccination as an act of group membership, rather than only personal protection, may therefore encourage both disclosure and uptake.

Conversely, politically framed disclosures were associated with a 10% decrease in engagement, and disclosures focused on social benefits saw an 8% decline. This suggests that campaign messages emphasizing political dimensions of vaccination or the social benefits of being vaccinated may generate less peer reinforcement. Communicators should be cautious about framing that politicizes vaccination or leads with distal benefits such as travel and social activities, as these appear to resonate less with general audiences in terms of social reward.

Finally, the finding that posts containing multiple disclosure themes were associated with higher engagement points to the value of multidimensional messaging. Encouraging individuals to share disclosures that combine, for example, community identity with gratitude or personal experience, may amplify the social reach of peer-based campaigns. Together, these findings support a communication approach that is emotionally grounded, identity-affirming, and attentive to the motivational dynamics that drive voluntary public disclosure.

Limitations

This study presents unique insights despite several limitations that should be considered when interpreting the findings. The motivations underlying disclosures were inferred from post content rather than derived from self-report data, meaning the mapping of themes to DDM goals represents an interpretive step rather than a direct test of motivation. Content-based motivation inference is an inherent limitation of computational approaches to large-scale social media data, where surveying individual posters is not feasible. That said, the mixed deductive-inductive coding approach allowed themes to emerge from the data iteratively rather than being imposed directly from the DDM, and the post hoc mapping of themes to DDM goals served as a theoretical grounding check rather than a primary analytic step. Future research combining computational content analysis with qualitative interviews or survey-based measures of disclosure motivation would strengthen causal claims about the psychological drivers of vaccination disclosure behavior.

Relying on Twitter data limits generalizability to the broader population, though Twitter was a leading platform for analyzing public attitudes about the pandemic [61] and effective for real-time detection of COVID-19-related discussions [62]. Generalizability is further constrained by the restriction to English-language posts, which excludes a substantial portion of global vaccine discourse. Although the keyword-based search may have missed some disclosures, the approach was designed to capture a broad array of relevant conversations and effectively represented the predominant discourse. Additionally, several themes appeared in few posts within the 100-post intercoder sample, so reliability estimates for these themes rest on a small number of positive cases.

Bot filtering relied on a keyword heuristic that removed posts from accounts whose descriptions contained the term “bot.” This approach identifies only self-labeled automated accounts; automated accounts involved in spreading content on Twitter are often deceptive [63] and can be difficult to distinguish from human users in vaccine-related discourse [64], so some may remain in the dataset. Because the DDM describes an individual’s decision to reveal personal information, posts from any remaining automated accounts are unlikely to reflect such decisions, yet they may have contributed to theme frequencies. Engagement counts also cannot be separated by account type, so associations between themes and engagement may partly reflect automated amplification, which has been documented for other content on Twitter [63], rather than human response alone.

The engagement findings should also be interpreted with caution. The engagement metric aggregates likes, reposts, quotes, and replies without distinguishing between positive and negative reactions, meaning high engagement on a post may reflect controversy rather than endorsement. Stance toward vaccination was not coded; because a post qualified as a disclosure whenever it indicated receipt of a vaccine dose, the dataset may include reluctant or critical disclosures, and the association between side-effects content and engagement may partly reflect such posts. Additionally, X’s

recommendation algorithm influences which posts receive engagement in ways that cannot be fully controlled for, meaning observed associations between themes and engagement reflect a combination of audience preference and algorithmic amplification. Finally, the data collection period ended on January 1, 2022, prior to the peak of the Omicron wave and the normalization of routine boosters, and the patterns documented here may not fully generalize to later pandemic phases or future immunization campaigns conducted in different social and political contexts. Addressing these limitations in future work will strengthen the generalizability and causal interpretability of findings in this area.

Conclusions

This investigation into the changing landscape of COVID-19 vaccination disclosures on Twitter reveals the significant role of social media in fostering public engagement and discourse during a health crisis. By mapping disclosure themes throughout the vaccine rollout, this study shows how disclosure themes shifted across the rollout and how they related to public engagement. The emergence of community identity and gratitude as dominant and high-engagement themes underscores the importance of social identity and emotional resonance in vaccine communication. The COVID-19 rollout was a historically unique natural experiment, but the communication dynamics it surfaced are not unique to COVID-19. As public health systems prepare for future immunization campaigns and health emergencies, the patterns documented here provide a replicable analytical framework and a practical evidence base for designing peer-based communication strategies. The specific content of what people disclose and what resonates with audiences can inform campaign design in ways that generic sentiment analysis cannot. Personal disclosures on social media may be a promising and underused resource for vaccine communication at scale.

Generative AI Use

During the preparation of this work, the authors used Claude (Anthropic) to assist with grammar checking, identification of typographical errors, and language editing. During the revision, Claude was also used to review the analysis code, identify errors, improve the code's efficiency and organization, and assist with formatting the reference list according to journal guidelines. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Funding

The authors declared no financial support was received for this work.

Data Availability

The data and computational code supporting the findings of this study are available from the corresponding author upon reasonable request. Data and analysis code will be shared upon request to enable rehydration and replication. To request access, contact ekmckinley@ucdavis.edu or jwzzhang@ucdavis.edu.

Authors' Contributions

Conceptualization: EM (lead), CS (equal), JZ (equal)
Data curation: EM
Formal analysis: EM (lead), CS (supporting), JZ (supporting)
Funding acquisition: not applicable
Investigation: EM
Methodology: EM (lead), CS (supporting), KS (supporting), JZ (supporting)
Project administration: EM
Supervision: CS (equal), KS (equal), JZ (equal)
Visualization: EM
Writing – original draft: EM
Writing – review & editing: EM (lead), CS (supporting), KS (supporting), JZ (supporting)
EM is the co-corresponding author and can be reached at ekmckinley@ucdavis.edu.

Conflicts of Interest

None declared.

Multimedia Appendix 1

COVID-19 vaccine-related hashtags used for data collection.
[DOCX File (Microsoft Word File), 23 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Disclosure decision model goal mappings for vaccination disclosure themes.
[DOCX File (Microsoft Word File), 18 KB-Multimedia Appendix 2]

Multimedia Appendix 3

Descriptive statistics and regression model outputs.
[DOCX File (Microsoft Word File), 20 KB-Multimedia Appendix 3]

Multimedia Appendix 4

Vaccination disclosure theme frequencies and co-occurrences.
[DOCX File (Microsoft Word File), 29 KB-Multimedia Appendix 4]

References

1. Odame EA, Dada O, Nelson J, Ogunyiola A, Haley J. Preparing for future pandemics: a qualitative exploration of social media in light of the COVID-19 pandemic and vaccine hesitancy. PLoS Glob Public Health. 2025;5(7):e0004317. [doi: 10.1371/journal.pgph.0004317] [Medline: 40622951]
2. Viswanath K, Lee EJ, Dryer E. Communication inequalities and incomplete data hinder understanding of how social media affect vaccine uptake. BMJ. Jun 20, 2024;385:e076478. [doi: 10.1136/bmj-2023-076478] [Medline: 38901868]
3. Lyu JC, Han EL, Luli GK. COVID-19 vaccine-related discussion on Twitter: topic modeling and sentiment analysis. J Med Internet Res. Jun 29, 2021;23(6):e24435. [doi: 10.2196/24435] [Medline: 34115608]
4. Scannell D, Desens L, Guadagno M, et al. COVID-19 vaccine discourse on Twitter: a content analysis of persuasion techniques, sentiment and mis/disinformation. J Health Commun. Jul 3, 2021;26(7):443-459. [doi: 10.1080/10810730.2021.1955050] [Medline: 34346288]
5. Nelson V, Bashyal B, Tan PN, Argyris YA. Vaccine rhetoric on social media and COVID-19 vaccine uptake rates: a triangulation using self-reported vaccine acceptance. Soc Sci Med. May 2024;348:116775. [doi: 10.1016/j.socscimed.2024.116775] [Medline: 38579627]
6. Borah P, Xiao X, Vishnevskaya A, Su Y. Narrative versus statistical messages: the interplay of perceived susceptibility and misperceptions on vaccine intention. Curr Psychol. Jun 2, 2023;43(20):1-16. [doi: 10.1007/s12144-023-04770-3] [Medline: 37359635]
7. Ko Y, Kim H, Seo Y, et al. The persuasive effects of social media narrative PSAs on COVID-19 vaccination intention among unvaccinated young adults: the mediating role of empathy and psychological reactance. J Soc Mark. Nov 10, 2023;13(4):490-509. [doi: 10.1108/JSOCM-09-2022-0185]
8. Ittefaq M, Ahmad Kamboh S, Abwao M. COVID-19 vaccine selfie: a modest endeavor to increase vaccine acceptance. Psychol Health. Feb 2023;38(2):209-213. [doi: 10.1080/08870446.2021.1957888] [Medline: 34339323]
9. Markovitz NH, Strome AL, Patel PK. Commentary: “the vaccine selfie” and its influence on COVID-19 vaccine acceptance. Vaccine. May 20, 2022;40(23):3085-3086. [doi: 10.1016/j.vaccine.2022.04.063] [Medline: 35487813]
10. Attwell K, Freeman M. I Immunise: an evaluation of a values-based campaign to change attitudes and beliefs. Vaccine. Nov 17, 2015;33(46):6235-6240. [doi: 10.1016/j.vaccine.2015.09.092] [Medline: 26458802]
11. Ji Y. Exposure to “vaccination selfies” relate to readers’ vaccination intention: the moderating role of partisan media use and the mediating role of affective responses. J Health Commun. Dec 2, 2021;26(12):867-876. [doi: 10.1080/10810730.2021.2025174] [Medline: 35100947]
12. Omarzu J. A disclosure decision model: determining how and when individuals will self-disclose. Pers Soc Psychol Rev. May 2000;4(2):174-185. [doi: 10.1207/S15327957PSPR0402_05]
13. Derlega VJ, Berg JH, editors. Self-Disclosure: Theory, Research, and Therapy. Springer; 1987. [doi: 10.1007/978-1-4899-3523-6]
14. Cline RJ, Haynes KM. Consumer health information seeking on the Internet: the state of the art. Health Educ Res. Dec 2001;16(6):671-692. [doi: 10.1093/her/16.6.671] [Medline: 11780707]
15. Oh A, Gaysynsky A, Winer RL, Lee HY, Brewer NT, White A. Considerations and opportunities for multilevel HPV vaccine communication interventions. Transl Behav Med. Feb 16, 2022;12(2):343-349. [doi: 10.1093/tbm/ibab129] [Medline: 35171276]
16. Shim M, Cappella JN, Han JY. How does insightful and emotional disclosure bring potential health benefits?: study based on online support groups for women with breast cancer. J Commun. Jun 2011;61(3):432-464. [doi: 10.1111/j.1460-2466.2011.01555.x] [Medline: 25568496]
17. Yan L, Tan Y. Feeling blue? Go online: an empirical study of social support among patients. Inf Syst Res. Dec 2014;25(4):690-709. [doi: 10.1287/isre.2014.0538]
18. Matthes J, Koban K, Neureiter A, Stevic A. Longitudinal relationships among fear of COVID-19, smartphone online self-disclosure, happiness, and psychological well-being: survey study. J Med Internet Res. Sep 27, 2021;23(9):e28700. [doi: 10.2196/28700] [Medline: 34519657]
19. Betsch C, Ulshöfer C, Renkewitz F, Betsch T. The influence of narrative v. statistical information on perceiving vaccination risks. Med Decis Making. 2011;31(5):742-753. [doi: 10.1177/0272989X11400419] [Medline: 21447730]

20. Betsch C, Renkewitz F, Haase N. Effect of narrative reports about vaccine adverse events and bias-awareness disclaimers on vaccine decisions: a simulation of an online patient social network. Med Decis Making. Jan 2013;33(1):14-25. [doi: 10.1177/0272989X12452342] [Medline: 22875721]
21. Shelby A, Ernst K. Story and science: how providers and parents can utilize storytelling to combat anti-vaccine misinformation. Hum Vaccin Immunother. Aug 2013;9(8):1795-1801. [doi: 10.4161/hv.24828] [Medline: 23811786]
22. Ittefaq M, Abwao M, Ahmad Kamboh S. COVID-19 vaccine selfie: why does it matter? Am J Health Educ. 2021;52(6):360-363. [doi: 10.1080/19325037.2021.1973615]
23. Bazarova NN, Choi YH. Self-disclosure in social media: extending the functional approach to disclosure motivations and characteristics on social network sites. J Commun. Aug 2014;64(4):635-657. [doi: 10.1111/jcom.12106]
24. Derlega VJ, Grzelak J. Appropriateness of self-disclosure. In: Chelune GJ, editor. Self-Disclosure: Origins, Patterns, and Implications of Openness in Interpersonal Relationships. Jossey-Bass; 1979:151-176. ISBN: 9780875894331
25. Lane DS, Lee SS, Liang F, et al. Social media expression and the political self. J Commun. Feb 1, 2019;69(1):49-72. [doi: 10.1093/joc/jqy064]
26. Wood AM, Maltby J, Gillett R, Linley PA, Joseph S. The role of gratitude in the development of social support, stress, and depression: two longitudinal studies. J Res Pers. Aug 2008;42(4):854-871. [doi: 10.1016/j.jrp.2007.11.003]
27. Ouyang H, Ma X, Wu X. The prevalence and determinants of COVID-19 vaccine hesitancy in the age of infodemic. Hum Vaccin Immunother. Dec 31, 2022;18(1):2013694. [doi: 10.1080/21645515.2021.2013694] [Medline: 35172676]
28. Balatsoukas P, Kennedy CM, Buchan I, Powell J, Ainsworth J. The role of social network technologies in online health promotion: a narrative review of theoretical and empirical factors influencing intervention effectiveness. J Med Internet Res. Jun 11, 2015;17(6):e141. [doi: 10.2196/jmir.3662] [Medline: 26068087]
29. Entwistle VA, France EF, Wyke S, et al. How information about other people's personal experiences can help with healthcare decision-making: a qualitative study. Patient Educ Couns. Dec 2011;85(3):e291-8. [doi: 10.1016/j.pec.2011.05.014] [Medline: 21652162]
30. Yan L, Yan X, Tan Y, Sun SX. Shared minds: how patients use collaborative information sharing via social media platforms. Prod Oper Manag. Jan 2019;28(1):9-26. [doi: 10.1111/poms.12895]
31. Sundar SS. The MAIN model: a heuristic approach to understanding technology effects on credibility. In: Metzger MJ, Flanagin AJ, editors. Digital Media, Youth, and Credibility. MIT Press; 2008:73-100. ISBN: 9780262562324
32. Twitter's recommendation algorithm. Twitter Engineering Blog. Mar 31, 2023. URL: https://web.archive.org/web/20231209221528/https://blog.twitter.com/engineering/en_us/topics/open-source/2023/twitter-recommendation-algorithm [Accessed 2026-10-07]
33. Yousefinaghani S, Dara R, Mubareka S, Papadopoulos A, Sharif S. An analysis of COVID-19 vaccine sentiments and opinions on Twitter. Int J Infect Dis. Jul 2021;108:256-262. [doi: 10.1016/j.ijid.2021.05.059] [Medline: 34052407]
34. Honnibal M, Montani I. SpaCy 2: natural language understanding with bloom embeddings, convolutional neural networks and incremental parsing. Open Research Knowledge Graph (ORKG). 2017. URL: https://orkg.org/resources/R11007?noRedirect= [Accessed 2026-10-07]
35. Lacy S, Riffe D. Sampling error and selecting intercoder reliability samples for nominal content categories. Journal Mass Commun Q. Dec 1996;73(4):963-973. [doi: 10.1177/107769909607300414]
36. The Python standard library. Python. URL: https://docs.python.org/3.8/library/ [Accessed 2026-10-07]
37. McKinney W. Data structures for statistical computing in Python. Proc 9th Python Sci Conf. 2010:56-61. [doi: 10.25080/Majora-92bf1922-00a]
38. Liu Y, Ott M, Goyal N, et al. RoBERTa: a robustly optimized BERT pretraining approach. arXiv. Preprint posted online on Jul 26, 2019. [doi: 10.48550/arXiv.1907.11692]
39. Jockers ML. syuzhet. GitHub. 2015. URL: https://github.com/mjockers/syuzhet [Accessed 2026-10-07]
40. Baumeister RF, Leary MR. The need to belong: desire for interpersonal attachments as a fundamental human motivation. Psychol Bull. May 1995;117(3):497-529. [doi: 10.1037/0033-2909.117.3.497] [Medline: 7777651]
41. McKenna KYA, Green AS. Virtual group dynamics. Group Dyn Theory Res Pract. Mar 2002;6(1):116-127. [doi: 10.1037/1089-2699.6.1.116]
42. Hogg MA. Uncertainty–identity theory. In: Zanna MP, editor. Advances in Experimental Social Psychology. Academic Press; 2007:69-126. [doi: 10.1016/S0065-2601(06)39002-8]
43. Brewer MB. The social self: on being the same and different at the same time. Pers Soc Psychol Bull. Oct 1991;17(5):475-482. [doi: 10.1177/0146167291175001]
44. Lee EJ. Effects of visual representation on social influence in computer-mediated communication: experimental tests of the social identity model of deindividuation effects. Hum Commun Res. Apr 1, 2004;30(2):234-259. [doi: 10.1111/j.1468-2958.2004.tb00732.x]

45. Lee EJ. Deindividuation effects on group polarization in computer-mediated communication: the role of group identification, public-self-awareness, and perceived argument quality. J Commun. Jun 1, 2007;57(2):385-403. [doi: 10.1111/j.1460-2466.2007.00348.x]
46. Jans-Beken L. A perspective on mature gratitude as a way of coping with COVID-19. Front Psychol. 2021;12:632911. [doi: 10.3389/fpsyg.2021.632911] [Medline: 33828504]
47. Fredrickson BL, Tugade MM, Waugh CE, Larkin GR. What good are positive emotions in crises? A prospective study of resilience and emotions following the terrorist attacks on the United States on September 11th, 2001. J Pers Soc Psychol. Feb 2003;84(2):365-376. [doi: 10.1037//0022-3514.84.2.365] [Medline: 12585810]
48. Yukhymenko-Lescroart MA. Student academic engagement and burnout amidst COVID-19: the role of purpose orientations and disposition towards gratitude in life. J Coll Stud Retent Res Theory Pract. Aug 2024;26(2):544-566. [doi: 10.1177/15210251221100415]
49. Puente-Díaz R, Cavazos-Arroyo J. Feeling grateful to be optimistic: the influence of recalling special moments on feelings of gratitude and optimism during the COVID-19 pandemic. Int J Psychol. Jun 2022;57(3):336-340. [doi: 10.1002/ijop.12830] [Medline: 34993962]
50. Spears R, Postmes T. Group identity, social influence, and collective action online: extensions and applications of the SIDE model. In: Sundar SS, editor. The Handbook of the Psychology of Communication Technology. Wiley; 2015:23-46. [doi: 10.1002/9781118426456]
51. Léger C, Deslauriers F, Gosselin Boucher V, Phillips M, Bacon SL, Lavoie KL. Prevalence and motivators of getting a COVID-19 booster vaccine in Canada: results from the iCARE study. Vaccines (Basel). Jan 28, 2023;11(2):291. [doi: 10.3390/vaccines11020291] [Medline: 36851169]
52. Mahase E. COVID-19: Johnson and Johnson vaccine trial is paused because of unexplained illness in participant. BMJ. Oct 13, 2020;371:m3967. [doi: 10.1136/bmj.m3967] [Medline: 33051241]
53. Østergaard L. SARS CoV-2 related microvascular damage and symptoms during and after COVID-19: consequences of capillary transit-time changes, tissue hypoxia and inflammation. Physiol Rep. Feb 2021;9(3):e14726. [doi: 10.14814/phy2.14726] [Medline: 33523608]
54. Henkel L, Sprengholz P, Korn L, Betsch C, Böhm R. The association between vaccination status identification and societal polarization. Nat Hum Behav. Feb 2023;7(2):231-239. [doi: 10.1038/s41562-022-01469-6] [Medline: 36329314]
55. Henning M, Fox GR, Kaplan J, Damasio H, Damasio A. A potential role for mu-opioids in mediating the positive effects of gratitude. Front Psychol. 2017;8:868. [doi: 10.3389/fpsyg.2017.00868] [Medline: 28680408]
56. Berger J, Milkman KL. What makes online content viral? J Mark Res. Apr 2012;49(2):192-205. [doi: 10.1509/jmr.10.0353]
57. Maheux AJ, Nesi J, Galla BM, Roberts SR, Choukas-Bradley S. #Grateful: longitudinal associations between adolescents' social media use and gratitude during the COVID-19 pandemic. J Res Adolesc. Sep 2021;31(3):734-747. [doi: 10.1111/jora.12650] [Medline: 34448294]
58. Schreiner M, Fischer T, Riedl R. Impact of content characteristics and emotion on behavioral engagement in social media: literature review and research agenda. Electron Commer Res. Jun 2021;21(2):329-345. [doi: 10.1007/s10660-019-09353-8]
59. Dillard JP, Shen L. On the nature of reactance and its role in persuasive health communication. Commun Monogr. Jun 2005;72(2):144-168. [doi: 10.1080/03637750500111815]
60. Reynolds-Tylus T. Psychological reactance and persuasive health communication: a review of the literature. Front Commun. Oct 31, 2019;4:56. [doi: 10.3389/fcomm.2019.00056]
61. Tsao SF, Chen H, Tisseverasinghe T, Yang Y, Li L, Butt ZA. What social media told us in the time of COVID-19: a scoping review. Lancet Digit Health. Mar 2021;3(3):e175-e194. [doi: 10.1016/S2589-7500(20)30315-0] [Medline: 33518503]
62. Amen B, Faiz S, Do TT. Big data directed acyclic graph model for real-time COVID-19 twitter stream detection. Pattern Recognit. Mar 2022;123:108404. [doi: 10.1016/j.patcog.2021.108404] [Medline: 34744186]
63. Shao C, Ciampaglia GL, Varol O, Yang KC, Flammini A, Menczer F. The spread of low-credibility content by social bots. Nat Commun. Nov 20, 2018;9(1):4787. [doi: 10.1038/s41467-018-06930-7] [Medline: 30459415]
64. Broniatowski DA, Jamison AM, Qi S, et al. Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. Am J Public Health. Oct 2018;108(10):1378-1384. [doi: 10.2105/AJPH.2018.304567] [Medline: 30138075]

Abbreviations

DDM: disclosure decision model
IRR: incidence rate ratio
NER: Named Entity Recognition

RoBERTa: robustly optimized BERT pretraining approach
RQ: research question

Edited by Amaryllis Mavragani; peer-reviewed by Chandler Carter, Yelena Mejova; submitted 03.May.2026; final revised version received 08.Oct.2026; accepted 11.Oct.2026; published 17.Nov.2026

Please cite as:
McKinley E, Shen C, Sagae K, Zhang J
Themes, Temporal Dynamics, and Public Engagement in COVID-19 Vaccination Disclosures on Twitter: Computational Content Analysis
J Med Internet Res 2026;28:e99837
URL: https://www.jmir.org/2026/1/e99837
doi: 10.2196/99837

© Emily McKinley, Cuihua Shen, Kenji Sagae, Jingwen Zhang. Originally published in the Journal of Medical Internet Research (https://www.jmir.org), 17.Nov.2026. 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 the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on https://www.jmir.org/, as well as this copyright and license information must be included.