



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



# Prospective associations of regional social media messages with attitudes and actual vaccination: A big data and survey study of the influenza vaccine in the United States



Man-pui Sally Chan<sup>a,\*</sup>, Kathleen Hall Jamieson<sup>b</sup>, Dolores Albarracín<sup>a,b</sup>

<sup>a</sup> Department of Psychology, University of Illinois at Urbana-Champaign, IL, United States

<sup>b</sup> The Annenberg Public Policy Center, University of Pennsylvania, PA, United States

## ARTICLE INFO

### Article history:

Received 25 April 2020

Received in revised form 24 July 2020

Accepted 25 July 2020

Available online 10 August 2020

### Keywords:

Influenza vaccine

Vaccine attitude

Vaccination

Social media

Discussions with family and friends

## ABSTRACT

**Objective:** Using longitudinal methods to assess regional associations between social media posts about vaccines and attitudes and actual vaccination against influenza in the US.

**Methods:** Geolocated tweets from U.S. counties ( $N = 115,330$ ) were analyzed using MALLET LDA (Latent Dirichlet allocation) topic modeling techniques to correlate with prospective individual survey data ( $N = 3005$ ) about vaccine attitudes, actual vaccination, and real-life discussions about vaccines with family and friends during the 2018–2019 influenza season.

**Results:** Ten topics were common across U.S. counties during the 2018–2019 influenza season. In the overall analyses, two of these topics (i.e., *Vaccine Science Matters* and *Big Pharma*) were associated with attitudes and behaviors. The topic concerning vaccine science in November–February was positively correlated with attitudes in February–March,  $r = 0.09$ ,  $BF_{10} = 3$ . Moreover, among respondents who did not discuss the influenza vaccine with family and friends, the topic about vaccine fraud and children in November–February was negatively correlated with attitudes in February–March and with vaccination in February–March, and April–May ( $r_s = -0.18$  to  $-0.25$ ,  $BF_{10} = 4-146$ ). However, this was absent when participants had discussions about the influenza vaccine with family and friends.

**Discussion:** Regional vaccine content correlated with prospective measures of vaccine attitudes and actual vaccination.

**Conclusions:** Social media have demonstrated strong associations with vaccination patterns. When the associations are negative, discussions with family and friends appear to eliminate them. Programs to promote vaccination should encourage real-life conversations about vaccines.

© 2020 Elsevier Ltd. All rights reserved.

## 1. Introduction

Although the influenza vaccine is one of the greatest achievements of public health and each year prevents millions of illnesses and thousands of deaths [1], vaccine messages expressing skepticism about vaccine safety and describing fake patient stories abound on social media [2,3]. Important research has advanced our understanding of social media contents, including how Russian linked-accounts have discussed vaccines in an attempt to influence the American public [4]. However, the ultimate associations between these contents and vaccination behavior remains an underexplored problem that we investigated in this study. To address this question, we analyzed regional variability in social

media messages in the US in combination with longitudinal survey measures from a representative sample of the U.S. adult population.

### 1.1. Regional associations between social media messages and attitudes and health behavior

As an example, even though more than a hundred thousand people had died of SAR-CoV-2 by the time this paper was written, and a vaccine may not be available until 2021, tweets opposing an eventual vaccine against the SARS-CoV-2 virus already populate the cyberworld. An example of these appears below.

“To each their own. I believe in medical freedom, not medical tyranny. If you want to get the flu shot, go ahead. But it shouldn't be mandatory. Same with any medicine or vaccine. I

\* Corresponding author at: Department of Psychology, University of Illinois at Urbana-Champaign, 603 East Daniel Street, Champaign, IL 61820, United States.

E-mail address: [sallycmp@illinois.edu](mailto:sallycmp@illinois.edu) (Man-pui Sally Chan).

personally would never take a Coronavirus vaccine. I know these sneaky snakes want it mandatory!" [5].

Other tweets, of course, support vaccination, such as the following example:

"The problem is when you make that choice [vaccine vs. chip] for yourself, you're also potentially making it for other people who are unable to get vaccinated. Polio, measles, diphtheria, etc are not things anyone should suffer through. Especially not when we have means of preventing them." [6]

A growing number of important studies have analyzed vaccine content on Facebook, Twitter, and YouTube. Although some studies have involved manual coding of posts [7–9], others have used computational techniques such as co-occurrence network analysis [10], and yet others have combined manual coding with computational methods, such as a support vector machine [11]. A recent study of anti- and pro-vaccine clips on YouTube identified words such as "chemical," "mercury," and "toxic" in anti-vaccine content and "hospital," "chronic," and "unvaccinated" in pro-vaccine material [9]. Similarly, an analysis of tweets showed factors containing words such as "cdcwhistleblower" and "coverup" posted by anti-vaccine users and topics containing "risk," "health," "chronic," and "science" posted by pro-vaccine users [11]. These studies have demonstrated that linguistic analyses can identify alternative types of vaccine content on social media. Yet, most of these studies covered only content obtained by means of pre-determined keywords. More exploratory methods of unsupervised learning (e.g., topic modeling) may provide further insights into vaccine social media contents [12].

The psychological theory has shed light on the nature of the relations between health messages and attitudes, as well as correspondence between attitudes and behaviors themselves [13–22]. Specifically, health contents can include information for people to act on, and include both persuasive posts by health officials as well as the personal opinions of lay people [18,23–25]. Health messages can then influence attitudes, which in turn guide people's decisions about what to do, including whether or not to vaccinate against the disease [14,25,26]. That is, people holding favorable attitudes toward the influenza vaccine are more likely to get vaccinated than those with unfavorable attitudes [27]. Recently, the relations between social media posts and vaccine attitudes and actual vaccination have also begun to be explored.

Prior studies have obtained cross-sectional evidence of relations between the use of or exposure to social media and vaccination status [28,29]. According to research on individuals, people who use either Twitter or Facebook, or both, as sources of health information are more likely to have received the influenza vaccine than are people who do not [28]. Moreover, across the US states, how common HPV vaccine topics are on Twitter accounts for 68% of the variance in the proportion of females receiving at least one dose of the HPV vaccine [29]. Although each of these studies offers useful evidence, both were cross-sectional, and the regional study [29] lacked individual-level data, which limits its conclusions and introduces the possibility of the ecological fallacy [30,31]. Thus, longitudinal data and a combination of individual- and regional-level data are necessary to understand the relations between social media messages posted within a region and individual attitudes and behaviors in that region.

### 1.2. The role of discussions with others as another influence source

Social media posts questioning the need for vaccines, for instance, may lead people to form negative attitudes toward vaccination and not vaccinate [32–34]. However, social media do not exist in a vacuum. Tweets posted in a region can be highly relevant

to the local context [35], and their impact may also depend on discussions of the influenza vaccine with family and friends in real life (for similar work in other areas, see [36]). In this regard, although the CDC encourages people to talk to friends and family about the influenza vaccine [37], the effectiveness of discussing vaccines with them is unclear. On the one hand, discussions with family and friends may further vaccine misinformation and hesitancy [2,38]. For example, people seek and circulate health information among family and friends [39], and this creates an information bubble (i.e., reading news or information that supports what people already believe in) that can pose challenges for public health announcements and physician recommendations to get through [40].

On the other hand, discussing vaccines with family and friends may correct inaccurate information found on social media. For example, discussing vaccination with family and friends predicts college students' positive dispositions toward vaccination [41], and norms can have positive influences on actual vaccination as well [42]. Moreover, discussions within one's network have been shown to reduce bias [43]. For example, Guilbeault and Centola [43] examined the idea of "networked collective intelligence" by assessing the impact of tobacco-prevention messages exchanged within the network. Participants who did and did not smoke were randomly assigned to either an experimental condition that allowed participants to have discussions with their networks or a control condition in which participants evaluated the health message by themselves. The results showed that smokers who discussed the message with others evaluated the health message more favorably than did smokers who evaluated the message alone. In short, introducing more diverse contents through discussions with family friends can be useful in correcting misinformation [44,45].

### 1.3. The present study

This study attempted to contribute to (a) provide a longitudinal assessment of the regional effects of social media messages on vaccine attitudes and vaccination, (b) incorporate individual-level data to minimize the threat of the ecological fallacy, and (c) examine the role of discussions with others as a factor that may strengthen or weaken the influences of social media. The guiding research questions included: What are the regional associations between social media messages and vaccine attitudes and vaccination? And how does discussing vaccines with family and friends influence the relation of social media content and vaccination? We used Twitter and the survey data corresponding to the 2018–2019 influenza season to examine these questions.

## 2. Methods

### 2.1. Source of data

The proposed study combined individual-level survey data with county-level tweets about vaccines, which were geolocated. The archival survey data involved a probability-based, nationally representative sample of American adults and assessed attitudes concerning the influenza vaccine, actual vaccination, and real-life discussions of the influenza vaccine in a five-wave panel (between September 21, 2018 and October 6, 2018; between October 7, 2018 and November 28, 2018; between November 29, 2018 and February 5, 2019; between February 6, 2019 and March 18, 2019; and between April 30, 2019 and May 5, 2019). In that flu season, most influenza vaccination occurred by December in 2018 although some were taken as late as early May in 2019. We collected both topically relevant tweets from the US and their metadata between

September 2017 and March 2019 using the Crimson Hexagon API. This social media analysis platform offers full access to all public tweets (i.e., Twitter Firehose). The survey data and the Twitter data are described next.

### 2.1.1. Survey data

**2.1.1.1. Design and methods.** We used five waves of a large panel survey assessing attitudes, knowledge, and health behaviors concerning infectious diseases and vaccines. The archival data included a probability-based, nationally representative sample of American adults who were randomly selected from AmeriSpeak panels of NORC at the University of Chicago in September–October and followed in October–November, November–February, February–March, and April–May (i.e., the 2018–2019 influenza season). This sampling frame covered 97% of U.S. households via a two-stage stratified sampling [46]. Respondents primarily completed the survey online, although a small percentage answered the survey over the phone (landlines and cell phones) to avoid a bias favoring people who use the Internet. This panel is the only probability panel that is available for hybrid online/phone administration in the US. After the data collection in September–October, weighted distributions of age, sex, race, and income showed only trivial differences compared to unweighted distributions. However, there was a discrepancy in education, with more of the survey sample having a college degree than the 2018 U.S. Census estimates. To mitigate this problem, respondents with a lower level of educational attainment were successfully targeted for retention in following waves with more incentives and reminders.

**2.1.1.2. Measures.** We analyzed responses to questions about vaccine attitudes, vaccination, and real-life discussions of the influenza vaccine. Respondents answered the questions about receiving the influenza vaccine in all five waves, whereas the questions about attitudes and real-life discussions were asked only in the first four waves of the study.

**2.1.1.2.1. Vaccination.** We used answers to the question, “Have you gotten the flu vaccine this season or not?” to indicate respondents’ influenza vaccination status. Answers were provided as “Yes” or “No.”

**2.1.1.2.2. Vaccine attitudes.** The survey included four attitude questions with 4-point scales: (a) “Based on what you know, how positive or negative do you feel about the flu vaccine?” (1: Very positive, 2: somewhat positive, 3: somewhat negative, 4: very negative); (b) “Just your best guess, how risky, if at all, do you think the flu vaccine is?” (1: Very risky, 2: somewhat risky, 3: not too risky, 4: not risky at all) (c) “Just your best guess, please indicate how effective, if at all, you think the flu vaccine will be at preventing the flu among those who get the vaccine this season?” (1: Very effective, 2: somewhat effective, 3: not too effective, 4: not effective at all), and (d) “How important, if at all, do you think it is that most people in your community get the flu vaccine?” (1: Very important, 2: somewhat important, 3: not too important, 4: not important at all). We reverse-scored all items except item (b), so a higher score indicates more favorable vaccine attitudes. The Cronbach’s alphas of the responses for all four waves were about 0.81–0.85, suggesting good internal consistencies. Therefore, we calculated a mean of these questions to represent respondents’ attitudes regarding the influenza vaccine in each wave.

**2.1.1.2.3. Discussing the influenza vaccine with family and friends.** Using a 4-point scale, participants also answered the question: “In the past month, how frequently have you discussed the flu vaccine with friends and family? (1: None, 2: rarely, 3: sometimes, 4: often). We rescored the responses of this question into a binary variable (i.e., 1: rarely, sometimes, or often as with discussions, and 0: none as without discussions of the influenza vaccine with family and friends) for the moderator analyses.

### 2.1.2. Tweets

Twitter continues to be one of the most popular social media platforms [47] and, unlike Facebook, provides open data access and has more precise and frequent location information. In this study, we collected Twitter data using Crimson Hexagon API. We adopted a top-down approach to collect relevant tweets by searching keywords related to vaccines posted between September 2017 and August 2018 to be the training data. We used additional keywords concerning influenza and the influenza vaccine with the keywords used for the training data to collect tweets between September 2018 and March 2019 to be the test data. In total, we collected 628,597 tweets and geotagged about 18% of the collected tweets for the analyses (see Table S1). This percentage is similar to the geotagging rates reported in previous studies [48,49].

We developed lists of keywords about influenza and flu by reviewing relevant literature [50–52] and consulted a team of health communication researchers to identify all keywords and hashtags relevant to the study period. Keywords used to collect the training data: (a) “vaccineswork” or “getvaccinated” or “gettheshotnottheflu” or “whyivax” or “vaccinessavelives” or “vaxyourkids” or “vaxyourself” or “fightflu” or “stoptheoutbreaks” or “vaxwithme”) and (b) (“vaccine” or “vax”) and (“cdcwhistleblower” or “hearthiswell” or “bigpharma” or “b1-less” or “nomandates” or “antivax” or “vaccineinjury” or “vaccineskill” or “vaccinesharm” or “thimerosal”). Additional keywords used to prepare the test data: (“flu” or “influenza”) and (“vaccine” or “vaccination” or “vaccinate” or “shot” or “nasal” or “spray” or “headache” or “cough” or “sick” or “fever” or “nose” or “throat” or “ill” or “outbreak” or “pandemic” or “jab” or “infected” or “infection” or “sneezing” or “temperature” or “thermometer” or “lung” or “pneumonia” or “recovered” or “cured” or “flufact” or “flufighter” or “fluseason” or “flushot” or “nivw” or “vaxxed” or “vaccineawareness” or “fluvax”).

## 2.2. Analytical procedure

We began with the geotagging procedure of tweets to US counties and applied topic modeling techniques to identify vaccine-related words that co-occurred (i.e., “topics”). We followed Maier et al.’s [53] recommended approach to select a final topic model and then computed each topic’s probabilities for U.S. counties during the 2018–2019 influenza season (see Fig. 1). Next, we combined topic probabilities with the survey data and performed correlations to examine the regional association of social media messages with vaccine attitudes and actual vaccination. Finally, we conducted the moderator analyses to assess if real-life discussions with friends and family affected the associations, and if so, what that influence was. Technical details about each step along with the topic modeling implementation follow.

### 2.2.1. Geotagging

Tweets, including retweets, are informative about popular topics and conversations within a community. Prior studies have shown that county-level tweets can provide signals for predicting HIV/STI infections [49,54] and are associated with attitudes, knowledge, and behaviors of infectious diseases such as Zika across U.S. counties [55]. Therefore, we decided to use county as the unit of analysis. We used the timestamp metadata to exclude tweets that do not originate from U.S. time zones. For each tweet, the user profile location and each tweet’s precise latitude and longitude coordinates were used to map tweets to U.S. counties using Geographic Information Systems (GIS) database operations (see [49] for details of the geotagging process). Tweets from the same county were organized into a single document, which constituted the unit of analysis. Thus, the tweets could not be used to re-identify individual users, and no personal information was



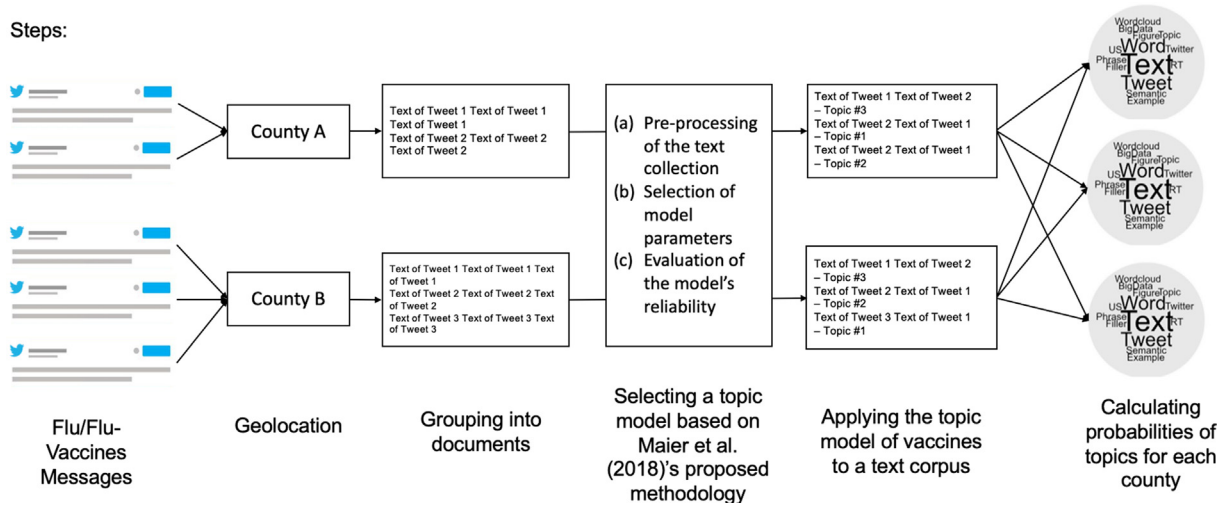


Fig. 1. Processing of Twitter Messages and Calculation of Topic Probabilities in Each Wave.

included in the analyses. Details about the training and test data appear in Table S1.

### 2.2.2. Topic modeling

We adapted Maier et al.'s [53] recommended approach to select a valid and reliable topic model, including (a) appropriate pre-processing of the text collection, (b) adequate selection of model parameters (i.e., the optimal number of topics of the topic model  $[K]$ ,  $\alpha$ , and  $\beta$ ), and (c) evaluation of the model's reliability. Before analysis, we removed unnecessary words and characters from the text corpus to retain only meaningful words and formed bi-grams (i.e., two-word phrases, see supplementary materials for details of the pre-processing step). Messages in each county were analyzed as a single document, and the words from all tweets from each county were combined into one single word-by-frequency matrix (see Fig. 1).

We used the Python package *gensim* to convert a collection of documents (i.e., each county's tweets) into a frequency matrix of token counts. The matrix of token counts was then analyzed using a well-established algorithm in computer science, namely Latent Dirichlet Allocation [56]. LDA is a Bayesian mixture model that groups words that often appear together to create topics (see Fig. 1). Then, we used a Python wrapper for LDA by implementing MALLET [57], and calculated coherence and log perplexity scores to determine the optimal number of topics (i.e.,  $K$ ) and hyperparameters (i.e.,  $\alpha$  and  $\beta$ , see supplementary materials for details). This process led to a topic model with ten topics (i.e.,  $K = 10$ ) and another with twenty topics (i.e.,  $K = 20$ ).

Next, we used the codes provided by Maier et al. (2018) to evaluate the reliability of two LDA topic models by conducting 33 experiments. Fig. 3 shows the reliability and coherence scores of the experiments, and a higher score indicates higher reliability and semantic coherence of the identified topics. As shown in Fig. 2, the topic model with a  $K$  of 10 had significantly higher reliability and semantic coherence than the topic model with a  $K$  of 20. Therefore, the selected topic model had ten topics, random initialization, and 1000 iterations.

Finally, we computed the probability of each topic for each county in each wave,  $p(\text{topic}, \text{county}) = \sum p(\text{topic} | \text{word}) \times p(\text{word} | \text{county})$ , where  $p(\text{word} | \text{county})$  represents the normalized proportion of words in a county document, and  $p(\text{topic} | \text{word})$  represents the probability of the topic given that word, using tweets that correspond to the survey period. We also calculated the standardized topic probabilities per county (i.e., the sum of all topic

probabilities per county is equal to 1) by replacing unrepresentable topic probabilities with "0.00001". The topic probabilities in each wave were then matched with the individual-level survey data by county and analyzed with Bayesian correlation analyses.

## 2.3. Statistical procedure

### 2.3.1. Intra-class correlation

The survey data involve individual participants who live in U.S. counties, and the Twitter topics belong to counties. Such a structure requires researchers to first determine the optimal way of analyzing the data, particularly whether participants should be nested within counties or whether the variance in survey responses can be assumed to vary as a function of participants. Hence, we calculated intra-class correlations (ICCs) to assess the overall variation in vaccination and vaccine attitudes explained by U.S. counties. Values of ICC can range from 0 to 1, and a value close to zero indicates that the observations within clusters are no more similar than observations from different clusters [58]. Values of  $\text{ICC} > 0.20$  is used as evidence of the need to use multi-level modeling, which in our case would imply modeling participants and counties as separate levels. Values above 0.20 would imply that participants must be modeled as nested within counties. We used the *performance* package [59] in R to calculate ICCs with vaccination and vaccine attitudes as the outcomes for all waves (see Table S3).

### 2.3.2. Bayesian correlation

We next used an R package *BayesMed* to calculate the Bayes factors for correlations between topics and attitudes or vaccination using a default uniform prior (i.e., 1) in which the possible correlation between  $-1$  and  $+1$  is equally plausible [60]. A Bayes factor (i.e.,  $\text{BF}_{10}$ ) between 3.2 and 10 provides substantial evidence that the data support the alternative hypothesis of a relation between two variables at least 3.2 to 10 times more than the null hypothesis of the absence of association [61].

## 3. Results

### 3.1. Identification of topics

Our selected topic model appears in Table 1. Two of the authors named these ten topics: *Vaccines Protect Children*, *Immunization Contributes to Global Health*, *Doctors Vaccinating Children*, *Fighting*





*Prevention Disease*



*Doctors Vaccinating Children*



*Big Pharma*



*Anti-Vaccine Story*

Fig. 2 (continued)



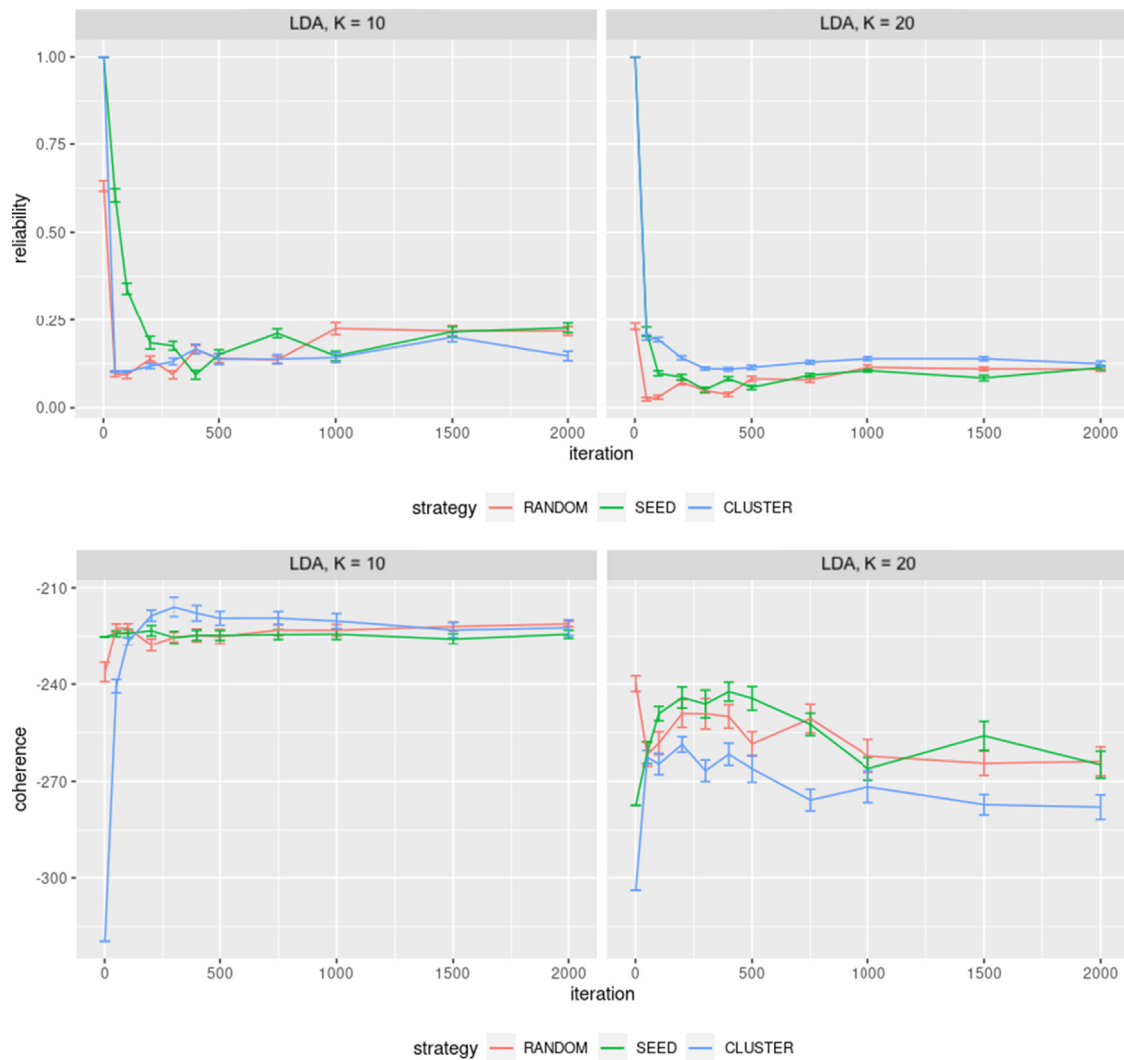
*Vaccine Science Matters*



*Fighting the Flu during the Flu Season*

Fig. 2 (continued)

### Reliability and Semantic Coherence Results of Two Topic Models ( $K_s = 10$ and 20)



**Fig. 3.** Reliability and Semantic Coherence Results of Two Topic Models ( $K_s = 10$  and 20). *Note.* Results were based on different initializations techniques (random = default random initialization; seed = fixed seed initialization; and cluster = semantic co-occurrence network initialization) and varying number of inference iterations (0–2000);  $K$  = number of topics (10 and 20).

between these topics and attitudes and vaccination, and the results showed moderate evidence of a prospective correlation with vaccine attitudes. Specifically, *Vaccine Science Matters* in November–February was positively correlated with attitudes in February–March,  $r = 0.09$ ,  $BF_{10} = 3.22$ . As for vaccination status, the results showed no evidence of the presence of associations with topics.

#### 3.3. Moderator analyses of discussing the influenza vaccine with friends and family

Next, we conducted moderator analyses to examine whether discussions with friends and family affected the correlation paths, and if so, what that influence was. We repeated the correlation analyses for respondents reporting real-life discussions of the influenza vaccine and those reporting no real-life discussions of the influenza vaccine, respectively (see Table 3). For respondents with no discussions with family and friends, we observed negative associations between *Vaccine Fraud and Children* in November–February and attitudes in February–March,  $r = -0.18$ ,  $BF_{10} = 4$ . *Vaccine Fraud and Children* in November–February was also inversely associated with vaccination in February–March and April–May,

$r_s = -0.24$  to  $-0.25$ ,  $BF_{10} = 49$ –146. Additionally, the results showed no evidence of any relations in respondents having real-life influenza vaccine discussions.

#### 4. Discussion

Combating an “infodemic” [62] on social media is important in preventing the spread of misinformation in the community. Prior research on social media and vaccines has focused mainly on either cross-sectional correlations between the use of or exposure to social media content and vaccination status, or cross-sectional correlations between the volume of tweets and state-level vaccination rates. Thus, the prior findings have not shed light on the paths between social media and vaccination, which require individual and longitudinal data. Filling this void, our study identified whether social media messages across U.S. counties are associated with vaccine attitudes and vaccination, and how vaccine discussions moderate the relations of social media messages with vaccine attitudes and with actual vaccination. We also found three prevailing vaccine topics: *Vaccine Science Matters*, *Big Pharma*, and *Vaccine Fraud and Children*. Furthermore, the probabilities of *Vaccine*



**Table 1**  
The final topic model for flu-related tweets in the U.S.

Topic number	Label	Top-10 word
<b>World Challenges</b>		
1	<i>Vaccines Protect Children</i>	measle, vaccination, disease, child, protect, immunization, health, year, world, kid
2	<i>Immunization Contributes to Global Health</i>	make, health, datum, global, great, immunization, learn, work, globalhealth, read
<b>Vaccine Administration</b>		
5	<i>Doctors Vaccinating Children</i>	child, vaccinate, measle, doctor, science, kid, make, hpv, parent, give
9	<i>Fighting the Flu during the Flu Season</i>	fightflu*, season, shoot, protect, learn, getvaccinate*, today, vaccination, people, late
<b>Vaccine Information</b>		
8	<i>Vaccine Science Matters</i>	vaxwithme*, ivax*, protect, anti, high, cancer, claim, evidence, sciencematter*, movement
<b>Vaccine Concerns</b>		
0	<i>Negative Vaccine Consequences</i>	vaxwithme*, vaxxed*, vaxxe, injury, ppl, shot, autism, safe, kill, explain
3	<i>Vaccine Fraud and Children</i>	vaccinate, child, kid, worldwide, call, late, commit, improving_access, validate, share
6	<i>Big Pharma</i>	health, child, bigpharma*, vaccinate, flushot, damage, autism, murder, share, alarm
7	<i>Anti-Vaccine Story</i>	autism, family, antivaccine, story, infection, battle, bad, fund, movement, st_place
<b>Infections and Diseases Prevention</b>		
4	<i>Prevention Disease</i>	whyivax*, share, adult, info, preventshingle*, patient, loved_one, change, follow, virus

Note. Asterisks indicate hashtags, “\_” indicate bi-grams.

**Table 2**  
Bayesian correlation results (r, Bayes Factors<sub>10</sub>) between Twitter topics and vaccine attitudes and vaccination in all waves.

Topic probability	September–October		October–November		November–February		February–March		April–May
	N = 1591–1662		N = 1591–1662		N = 1591–1662		N = 1591–1662		N = 881–1091
	Vaccine attitude	Vaccination	Vaccine attitude	Vaccination	Vaccine attitude	Vaccination	Vaccine attitude	Vaccination	Vaccination
<b>September–October</b>									
<i>Negative Vaccine Consequences</i>	0.03, 0	−0.03, 0	0.02, 0	0, 0	0.01, 0	−0.01, 0	0.01, 0	0, 0	−0.01, 0
<i>Vaccines Protect Children</i>	−0.01, 0	0.01, 0	−0.04, 0	−0.04, 0	−0.01, 0	0, 0	0.01, 0	0.02, 0	0.03, 0
<i>Immunization Contributes to Global Health</i>	0.01, 0	0.06, 1	0, 0	−0.01, 0	−0.02, 0	−0.02, 0	−0.03, 0	−0.02, 0	−0.04, 0
<i>Vaccine Fraud and Children</i>	0.01, 0	−0.02, 0	0.02, 0	0.02, 0	0, 0	0.03, 0	0.03, 0	0.04, 0	0.03, 0
<i>Prevention Disease</i>	0.05, 0	−0.04, 0	0.04, 0	0.04, 0	0.03, 0	0.02, 0	0.01, 0	0, 0	−0.01, 0
<i>Doctors Vaccinating Children</i>	−0.05, 0	0, 0	−0.02, 0	−0.05, 0	−0.05, 0	−0.07, 1	−0.03, 0	−0.06, 0	−0.05, 0
<i>Big Pharma</i>	−0.01, 0	0.05, 0	0, 0	0, 0	0, 0	0.03, 0	−0.03, 0	0, 0	0.01, 0
<i>Anti-Vaccine Story</i>	0, 0	−0.03, 0	−0.03, 0	0.03, 0	−0.01, 0	0.02, 0	−0.02, 0	0.01, 0	0.01, 0
<i>Vaccine Science Matters</i>	0, 0	0.02, 0	0.01, 0	0.01, 0	0.01, 0	0.01, 0	0, 0	0.01, 0	0.01, 0
<i>Fighting the Flu during the Flu Season</i>	0.02, 0	0.01, 0	0.04, 0	0.07, 1	0.07, 1	0.07, 1	0.08, 1	0.08, 1	0.08, 1
<b>October–November</b>									
<i>Negative Vaccine Consequences</i>			0.03, 0	0.02, 0	0.04, 0	0.06, 0	0.02, 0	0.05, 0	0.05, 0
<i>Vaccines Protect Children</i>			0, 0	−0.04, 0	0.01, 0	0, 0	0, 0	0, 0	−0.01, 0
<i>Immunization Contributes to Global Health</i>			−0.02, 0	−0.01, 0	−0.02, 0	−0.03, 0	−0.04, 0	−0.04, 0	−0.06, 0
<i>Vaccine Fraud and Children</i>			0.05, 0	0.05, 0	0.07, 0	0.06, 0	0.08, 0	0.08, 0	0.08, 0
<i>Prevention Disease</i>			−0.01, 0	0.01, 0	0, 0	−0.02, 0	0, 0	−0.01, 0	−0.02, 0
<i>Doctors Vaccinating Children</i>			0.03, 0	0.06, 0	0.05, 0	0.08, 0	0.08, 0	0.06, 0	0.07, 0
<i>Big Pharma</i>			0.02, 0	−0.02, 0	0.02, 0	−0.02, 0	0.03, 0	0.02, 0	0.02, 0
<i>Anti-Vaccine Story</i>			−0.02, 0	0.01, 0	−0.02, 0	−0.03, 0	−0.02, 0	−0.04, 0	−0.05, 0
<i>Vaccine Science Matters</i>			0, 0	−0.03, 0	−0.01, 0	−0.05, 0	−0.04, 0	−0.06, 0	−0.04, 0
<i>Fighting the Flu during the Flu Season</i>			−0.04, 0	−0.03, 0	−0.07, 0	−0.05, 0	−0.09, 0	−0.06, 0	−0.05, 0
<b>November–February</b>									
<i>Negative Vaccine Consequences</i>					−0.06, 0	−0.05, 0	−0.05, 0	−0.06, 0	−0.06, 0
<i>Vaccines Protect Children</i>					0.04, 0	0, 0	0.05, 0	0, 0	0, 0
<i>Immunization Contributes to Global Health</i>					0.03, 0	0.04, 0	0.01, 0	0.04, 0	0.05, 0
<i>Vaccine Fraud and Children</i>					−0.06, 1	−0.06, 1	−0.05, 0	−0.07, 1	−0.06, 0
<i>Prevention Disease</i>					0.01, 0	0.01, 0	0.02, 0	0.01, 0	0.01, 0
<i>Doctors Vaccinating Children</i>					−0.01, 0	0.01, 0	−0.04, 0	−0.01, 0	−0.02, 0
<i>Big Pharma</i>					0.03, 0	−0.01, 0	0.05, 0	0, 0	0.01, 0
<i>Anti-Vaccine Story</i>					−0.03, 0	−0.02, 0	−0.04, 0	−0.03, 0	−0.04, 0
<i>Vaccine Science Matters</i>					0.06, 1	0.01, 0	0.09, 3*	0.02, 0	0.01, 0
<i>Fighting the Flu during the Flu Season</i>					−0.01, 0	0.03, 0	−0.01, 0	0.05, 0	0.05, 0
<b>February–March</b>									
<i>Negative Vaccine Consequences</i>							−0.02, 0	−0.03, 0	−0.01, 0
<i>Vaccines Protect Children</i>							0.05, 0	0.05, 0	0.05, 0
<i>Immunization Contributes to Global Health</i>							0.05, 0	0.05, 0	0.05, 0

(continued on next page)

Table 2 (continued)

Topic probability	September–October		October–November		November–February		February–March		April–May
	N = 1591–1662		N = 1591–1662		N = 1591–1662		N = 1591–1662		N = 881–1091
	Vaccine attitude	Vaccination	Vaccine attitude	Vaccination	Vaccine attitude	Vaccination	Vaccine attitude	Vaccination	Vaccination
<i>Vaccine Fraud and Children</i>							0.01, 0	–0.01, 0	0.01, 0
<i>Prevention Disease</i>							–0.07, 0	–0.04, 0	–0.01, 0
<i>Doctors Vaccinating Children</i>							0.01, 0	–0.02, 0	–0.02, 0
<i>Big Pharma</i>							–0.1, 6*	–0.06, 0	–0.05, 0
<i>Anti-Vaccine Story</i>							–0.02, 0	–0.06, 0	–0.06, 0
<i>Vaccine Science Matters</i>							–0.01, 0	0.02, 0	0.01, 0
<i>Fighting the Flu during the Flu Season</i>							0.01, 0	0.04, 0	0.02, 0

\* BF<sub>10</sub> > 3.2, \*\* BF<sub>10</sub> > 10, \*\*\* BF<sub>10</sub> > 32, \*\*\*\* BF<sub>10</sub> > 100.

Table 3

Bayesian correlation results (r, Bayes Factors<sub>10</sub>) between Twitter topics and vaccine attitudes and vaccination for two subgroups.

Topic probability	September–October		October–November		November–February		February–March		April–May
	Vaccine attitude	Vaccination	Vaccine attitude	Vaccination	Vaccine attitude	Vaccination	Vaccine attitude	Vaccination	Vaccination
Subgroup analyses: Having no discussions of the influenza vaccine with family and friends									
September–October									
<i>Negative Vaccine Consequences</i>	0.02, 0	–0.06, 0	0, 0	–0.03, 0	0, 0	–0.04, 0	–0.04, 0	–0.03, 0	–0.03, 0
<i>Vaccines Protect Children</i>	–0.01, 0	–0.02, 0	–0.04, 0	–0.04, 0	0, 0	0.02, 0	–0.03, 0	0.01, 0	0.01, 0
<i>Immunization Contributes to Global Health</i>	0.01, 0	0.1, 1	–0.03, 0	0, 0	0.01, 0	–0.01, 0	–0.02, 0	–0.01, 0	–0.03, 0
<i>Vaccine Fraud and Children</i>	–0.02, 0	–0.01, 0	0.01, 0	–0.01, 0	–0.01, 0	0.01, 0	0.13, 0	0.06, 0	0.05, 0
<i>Prevention Disease</i>	0.02, 0	–0.05, 0	0, 0	–0.03, 0	–0.04, 0	–0.05, 0	–0.07, 0	–0.08, 0	–0.08, 0
<i>Doctors Vaccinating Children</i>	0.01, 0	0.08, 0	0.06, 0	0.04, 0	0.02, 0	0.01, 0	0.05, 0	0.02, 0	0.02, 0
<i>Big Pharma</i>	–0.06, 0	–0.01, 0	–0.01, 0	0.01, 0	–0.02, 0	–0.01, 0	0, 0	0, 0	0, 0
<i>Anti-Vaccine Story</i>	0.06, 0	0, 0	–0.04, 0	0.08, 0	0, 0	0.08, 0	0.03, 0	0.07, 0	0.08, 0
<i>Vaccine Science Matters</i>	0, 0	–0.02, 0	0.06, 0	0.02, 0	0.06, 0	0.04, 0	0.08, 0	0.07, 0	0.05, 0
<i>Fighting the Flu during the Flu Season</i>	–0.03, 0	–0.01, 0	–0.03, 0	0, 0	0.03, 0	0.03, 0	0.05, 0	0.03, 0	0.05, 0
October–November									
<i>Negative Vaccine Consequences</i>			0.05, 0	0, 0	0.02, 0	–0.03, 0	–0.05, 0	–0.02, 0	–0.02, 0
<i>Vaccines Protect Children</i>			–0.1, 0	–0.05, 0	–0.01, 0	0.03, 0	–0.03, 0	0.03, 0	0, 0
<i>Immunization Contributes to Global Health</i>			–0.03, 0	0.01, 0	–0.02, 0	–0.03, 0	–0.06, 0	–0.06, 0	–0.08, 0
<i>Vaccine Fraud and Children</i>			0.08, 0	0.06, 0	0.01, 0	0, 0	0.11, 0	0.1, 0	0.1, 0
<i>Prevention Disease</i>			–0.12, 0	–0.11, 0	–0.13, 0	–0.15, 0	–0.09, 0	–0.14, 0	–0.15, 0
<i>Doctors Vaccinating Children</i>			–0.01, 0	–0.03, 0	0, 0	0.06, 0	0.01, 0	0.07, 0	0.09, 0
<i>Big Pharma</i>			0.02, 0	–0.09, 0	0, 0	–0.1, 0	0.02, 0	–0.09, 0	–0.1, 0
<i>Anti-Vaccine Story</i>			0.11, 0	0.07, 0	0.09, 0	–0.06, 0	0, 0	–0.07, 0	–0.1, 0
<i>Vaccine Science Matters</i>			0, 0	–0.04, 0	–0.03, 0	–0.05, 0	0.02, 0	–0.02, 0	0, 0
<i>Fighting the Flu during the Flu Season</i>			0.06, 0	0.13, 1	0.02, 0	0.06, 0	0.02, 0	0.02, 0	0.05, 0
November–February									
<i>Negative Vaccine Consequences</i>					–0.07, 0	–0.09, 0	0.01, 0	–0.07, 0	–0.08, 0
<i>Vaccines Protect Children</i>					–0.05, 0	0.02, 0	–0.04, 0	0.03, 0	0.03, 0
<i>Immunization Contributes to Global Health</i>					0.05, 0	0.08, 0	0.03, 0	0.07, 0	0.07, 0
<i>Vaccine Fraud and Children</i>					–0.2, 36***	–0.2, 32***	–0.18, 4*	–0.25, 164****	–0.24, 49***
<i>Prevention Disease</i>					0.03, 0	0.05, 0	0, 0	0.06, 0	0.06, 0
<i>Doctors Vaccinating Children</i>					–0.06, 0	–0.05, 0	–0.04, 0	–0.07, 0	–0.09, 0
<i>Big Pharma</i>					0.03, 0	0.03, 0	0.11, 0	0.05, 0	0.05, 0
<i>Anti-Vaccine Story</i>					0.02, 0	–0.04, 0	–0.07, 0	–0.05, 0	–0.07, 0
<i>Vaccine Science Matters</i>					0.05, 0	–0.03, 0	0.07, 0	0, 0	–0.03, 0
<i>Fighting the Flu during the Flu Season</i>					0.13, 1	0.11, 0	0.08, 0	0.11, 0	0.13, 1
February–March									
<i>Negative Vaccine Consequences</i>							–0.02, 0	–0.02, 0	0, 0
<i>Vaccines Protect Children</i>							0.1, 0	0.05, 0	0.05, 0
<i>Immunization Contributes to Global Health</i>							0.09, 0	0.06, 0	0.06, 0
<i>Vaccine Fraud and Children</i>							–0.02, 0	–0.09, 0	–0.07, 0

Table 3 (continued)

Topic probability	September–October		October–November		November–February		February–March		April–May
	Vaccine attitude	Vaccination	Vaccine attitude	Vaccination	Vaccine attitude	Vaccination	Vaccine attitude	Vaccination	Vaccination
<i>Prevention Disease</i>							−0.15, 0	−0.12, 0	−0.1, 0
<i>Doctors Vaccinating Children</i>							0.02, 0	0.02, 0	0.03, 0
<i>Big Pharma</i>							−0.1, 0	−0.13, 1	−0.15, 1
<i>Anti-Vaccine Story</i>							−0.13, 1	−0.1, 0	−0.11, 0
<i>Vaccine Science Matters</i>							−0.07, 0	−0.03, 0	−0.04, 0
<i>Fighting the Flu during the Flu Season</i>							0.04, 0	0.08, 0	0.09, 0
Subgroup analyses: Having discussions of the influenza vaccine with family and friends									
September–October									
<i>Negative Vaccine Consequences</i>	0.06, 0	0, 0	0.05, 0	0.04, 0	0.04, 0	0.03, 0	0.05, 0	0.03, 0	0.02, 0
<i>Vaccines Protect Children</i>	−0.01, 0	0.02, 0	−0.03, 0	−0.04, 0	−0.02, 0	−0.01, 0	0.03, 0	0.02, 0	0.04, 0
<i>Immunization Contributes to Global Health</i>	0.01, 0	0.04, 0	0.01, 0	−0.03, 0	−0.04, 0	−0.04, 0	−0.04, 0	−0.03, 0	−0.04, 0
<i>Vaccine Fraud and Children</i>	−0.01, 0	−0.02, 0	−0.04, 0	0, 0	−0.05, 0	0.01, 0	−0.06, 0	0, 0	−0.01, 0
<i>Prevention Disease</i>	0.08, 1	−0.02, 0	0.07, 0	0.08, 1	0.08, 1	0.07, 0	0.06, 0	0.06, 0	0.03, 0
<i>Doctors Vaccinating Children</i>	−0.08, 2	−0.03, 0	−0.06, 0	−0.09, 2	−0.08, 1	−0.10, 3	−0.07, 0	−0.09, 1	−0.09, 1
<i>Big Pharma</i>	0, 0	0.06, 0	−0.01, 0	−0.02, 0	−0.01, 0	0.03, 0	−0.06, 0	−0.02, 0	−0.01, 0
<i>Anti-Vaccine Story</i>	−0.03, 0	−0.04, 0	−0.02, 0	0.02, 0	−0.01, 0	0, 0	−0.03, 0	−0.01, 0	−0.02, 0
<i>Vaccine Science Matters</i>	0, 0	0.03, 0	−0.02, 0	−0.01, 0	−0.03, 0	−0.02, 0	−0.05, 0	−0.03, 0	−0.02, 0
<i>Fighting the Flu during the Flu Season</i>	0.04, 0	0.02, 0	0.06, 0	0.09, 2	0.09, 1	0.09, 1	0.09, 1	0.1, 1	0.09, 1
October–November									
<i>Negative Vaccine Consequences</i>			0.03, 0	0.04, 0	0.04, 0	0.08, 0	0.04, 0	0.07, 0	0.07, 0
<i>Vaccines Protect Children</i>			0.04, 0	−0.02, 0	0.03, 0	0.01, 0	0.01, 0	0, 0	0, 0
<i>Immunization Contributes to Global Health</i>			−0.01, 0	0, 0	−0.02, 0	−0.02, 0	−0.02, 0	−0.03, 0	−0.04, 0
<i>Vaccine Fraud and Children</i>			0.04, 0	0.04, 0	0.08, 0	0.06, 0	0.06, 0	0.06, 0	0.07, 0
<i>Prevention Disease</i>			0.02, 0	0.03, 0	0.03, 0	0.01, 0	0.02, 0	0.01, 0	0, 0
<i>Doctors Vaccinating Children</i>			0.04, 0	0.07, 0	0.05, 0	0.07, 0	0.09, 0	0.06, 0	0.06, 0
<i>Big Pharma</i>			0.02, 0	0, 0	0.02, 0	0, 0	0.03, 0	0.05, 0	0.05, 0
<i>Anti-Vaccine Story</i>			−0.06, 0	0, 0	−0.05, 0	−0.02, 0	−0.02, 0	−0.03, 0	−0.03, 0
<i>Vaccine Science Matters</i>			0, 0	−0.02, 0	0, 0	−0.04, 0	−0.05, 0	−0.06, 0	−0.04, 0
<i>Fighting the Flu during the Flu Season</i>			−0.08, 0	−0.07, 0	−0.1, 0	−0.08, 0	−0.12, 0	−0.08, 0	−0.08, 0
November–February									
<i>Negative Vaccine Consequences</i>					−0.05, 0	−0.03, 0	−0.06, 0	−0.05, 0	−0.05, 0
<i>Vaccines Protect Children</i>					0.07, 1	0, 0	0.07, 1	−0.01, 0	−0.01, 0
<i>Immunization Contributes to Global Health</i>					0.03, 0	0.03, 0	0.01, 0	0.03, 0	0.04, 0
<i>Vaccine Fraud and Children</i>					−0.02, 0	−0.02, 0	−0.01, 0	−0.02, 0	−0.03, 0
<i>Prevention Disease</i>					0.01, 0	0.01, 0	0.03, 0	0.01, 0	0.01, 0
<i>Doctors Vaccinating Children</i>					0, 0	0.02, 0	−0.05, 0	0, 0	0, 0
<i>Big Pharma</i>					0.03, 0	−0.02, 0	0.03, 0	−0.01, 0	0, 0
<i>Anti-Vaccine Story</i>					−0.04, 0	−0.01, 0	−0.04, 0	−0.02, 0	−0.04, 0
<i>Vaccine Science Matters</i>					0.06, 0	0.02, 0	0.09, 2	0.02, 0	0.01, 0
<i>Fighting the Flu during the Flu Season</i>					−0.06, 0	0.01, 0	−0.03, 0	0.03, 0	0.02, 0
February–March									
<i>Negative Vaccine Consequences</i>							−0.03, 0	−0.05, 0	−0.04, 0
<i>Vaccines Protect Children</i>							0.03, 0	0.04, 0	0.05, 0
<i>Immunization Contributes to Global Health</i>							0.04, 0	0.05, 0	0.04, 0
<i>Vaccine Fraud and Children</i>							0.01, 0	0, 0	0.02, 0
<i>Prevention Disease</i>							−0.04, 0	−0.02, 0	0.01, 0
<i>Doctors Vaccinating Children</i>							0.01, 0	−0.03, 0	−0.04, 0
<i>Big Pharma</i>							−0.09, 1	−0.02, 0	0, 0
<i>Anti-Vaccine Story</i>							0.01, 0	−0.06, 0	−0.06, 0
<i>Vaccine Science Matters</i>							0, 0	0.02, 0	0.01, 0
<i>Fighting the Flu during the Flu Season</i>							0.01, 0	0.03, 0	0.01, 0

\* BF<sub>10</sub> > 3.2, \*\* BF<sub>10</sub> > 10, \*\*\* BF<sub>10</sub> > 32, \*\*\*\* BF<sub>10</sub> > 100.

*Science Matters* were prospectively associated with vaccine attitudes, and the probabilities of *Vaccine Fraud and Children* were prospectively associated with vaccine attitudes and actual vaccination. Importantly, in line with the idea of networked collective intelligence [43], discussing the influenza vaccine with others in real-life appeared to mitigate the negative regional relations between *Vaccine Fraud and Children* and vaccine attitudes and vaccination.

Even though our results showed strong to very strong evidence of associations between topics (i.e., *Vaccine Science Matters*, *Big Pharma*, and *Vaccine Fraud and Children*) and vaccine attitudes and behavior, these associations do not necessarily imply causation and await experimental results. However, the findings can provide important insights into social media effects on health attitudes and behaviors. For example, given the prospective associations, it is possible that posts about scientific evidence of

vaccines promote the spread of factual information of vaccines, which contribute to more favorable attitudes towards the flu vaccine. Future studies may assess whether the presence of an unobserved or third variable, such as the change of regional patients' minimum age for convenient immunization at pharmacies [63], increases discussions on Twitter while also increasing vaccine attitudes and actual vaccination. Another direction of research may be to study the spread of conspiracy theories (e.g., *Big Pharma*) on social media, which may raise concerns about vaccine safety. People who are high in needle sensitivity may be more prone to the influences of these posts, and a downward spiral may develop, which may result in more negative attitudes toward vaccines [64].

This study has a number of noteworthy limitations. Twitter users may not be representative of social media messaging in general or of communication in the non social media world. First, although Twitter continues to be a popular social media platform [47], other social media should be analyzed in future research. Second, those on Twitter use that platform to connect to friends and acquaintances who live in their area, cross-regional influences also may be important and deserve future consideration [65]. Third, the present study identified a topic model based on a limited albeit large set of tweets (i.e., tweets with location metadata that allow geotagging to U.S. counties). Fourth, in addition to social media, other factors such as health insurance coverage are critical determinants of vaccination, and their influence in interaction with social media should be investigated in the future. Despite these limitations, our findings echo the notion that Twitter can have an influence on the local context [35,66].

#### CRedit authorship contribution statement

**Man-pui Sally Chan:** Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing - original draft, Visualization, Project administration, Funding acquisition. **Kathleen Hall Jamieson:** Resources, Writing - review & editing. **Dolores Albarracín:** Conceptualization, Writing - review & editing, Supervision, Funding acquisition.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgement

Research reported in this publication was supported by the National Institute of Mental Health under Award Number R01MH114847, the National Institute on Drug Abuse under Award Number DP1 DA048570, and the National Institute Of Allergy And Infectious Diseases under Award Number R01AI147487. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We would like to thank the APPC ASK Group for reviewing keywords and hashtags about influenza and flu that are relevant to the study period.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.vaccine.2020.07.054>.

#### References

- Rolfes MA, Flannery B, Chung JR, et al. Effects of influenza vaccination in the United States during the 2017–2018 influenza season. *Clin Infect Dis* 2019;69(11):1845–53. <https://doi.org/10.1093/cid/ciz075>.
- Igoe KJ. Establishing the truth: Vaccines, social media, and the spread of misinformation. *Risk Commun Anal*; 2020. <https://www.hsph.harvard.edu/ecpe/vaccines-social-media-spread-misinformation/> [accessed March 5, 2020].
- Gunaratne K, Coomes EA, Haghbayan H. Temporal trends in anti-vaccine discourse on Twitter. *Vaccine* 2019;37(35):4867–71. <https://doi.org/10.1016/j.vaccine.2019.06.086>.
- Walter D, Ophir Y, Jamieson KH. Russian Twitter accounts and the partisan polarization of vaccine discourse, 2015–2017. *Am J Public Health* 2020;110(5):718–24. <https://doi.org/10.2105/ajph.2019.305564>.
- @LegendaryEnergy. To each their own. I believe in medical freedom, not medical tyranny. If you want to get the flu shot, go ahead. But it shouldn't be mandatory. Same with any medicine or vaccine. I personally would never take a Coronavirus vaccine. I know these sneaky sna; 2020. <https://twitter.com/LegendaryEnergy/status/1247162417894785024>.
- @Chickentetrazin. The problem is when you make that choice for yourself, you're also potentially making it for other people who are unable to get vaccinated. Polio, measles, diphtheria, etc are not things anyone should suffer through. Especially not when we have means of p; 2020. <https://twitter.com/Chickentetrazin/status/1242940175765901313>.
- Basch CH, MacLean SA. A content analysis of HPV related posts on Instagram. *Hum Vaccines Immunother* 2019;15(7–8):1476–8. <https://doi.org/10.1080/21645515.2018.1560774>.
- Hoffman BL, Felter EM, Chu KH, et al. It's not all about autism: The emerging landscape of anti-vaccination sentiment on Facebook. *Vaccine* 2019;37(16):2216–23. <https://doi.org/10.1016/j.vaccine.2019.03.003>.
- Yiannakoulis N, Slavik CE, Chase M. Expressions of pro- and anti-vaccine sentiment on YouTube. *Vaccine* 2019;37(15):2057–64. <https://doi.org/10.1016/j.vaccine.2019.03.001>.
- Porreca A, Scozzari F, Di Nicola M. Using text mining and sentiment analysis to analyse YouTube Italian videos concerning vaccination. *BMC Public Health* 2020;20(1). <https://doi.org/10.1186/s12889-020-8342-4>.
- Mitra T, Counts S, Pennebaker JW. Understanding anti-vaccination attitudes in social media. Proceedings of the 10th International Conference on Web and Social Media, ICWSM 2016, 2016.
- Chan MS, Morales A, Farhadloo M, Palmer RP, Albarracín D. Social media harvesting. In: Blanton H, LaCroix JM, Webster GD, editors. *Measurement in Social Psychology*. New York, NY: Taylor & Francis; 2019. p. 228–64.
- Ajzen I, Albarracín D, Hornik R. Prediction and change of health behavior: Applying the reasoned action approach. *Predict Chang Heal Behav Appl Reason action approach* 2007;xi:309. <https://doi.org/10.4324/9780203937082>.
- Albarracín D, Johnson BT, Fishbein M, Muellerleile PA. Theories of reasoned action and planned behavior as models of condom use: A meta-analysis. *Psychol Bull* 2001;127(1):142–61. <https://doi.org/10.1037/0033-2909.127.1.142>.
- Rimer BK, Kreuter MW. Advancing tailored health communication: A persuasion and message effects perspective. *J Commun* 2006;56(suppl.1):S184–201. <https://doi.org/10.1111/j.1460-2466.2006.00289.x>.
- Frank LB, Murphy ST, Chatterjee JS, Moran MB, Baezconde-Garbanati L. Telling stories, saving lives: Creating narrative health messages. *Health Commun* 2015;30(2):154–63. <https://doi.org/10.1080/10410236.2014.974126>.
- Bushar JA, Kendrick JS, Ding H, Black CL, Greby SM. Text4baby influenza messaging and influenza vaccination among pregnant women. *Am J Prev Med* 2017;53(6):845–53. <https://doi.org/10.1016/j.amepre.2017.06.021>.
- Albarracín D, Gillette JC, Earl AN, Glasman LR, Durantini MR. A test of major assumptions about behavior change: A comprehensive look at the effects of passive and active HIV-prevention interventions since the beginning of the epidemic. *Psychol Bull* 2005;131(6):856–97. <https://doi.org/10.1037/0033-2909.131.6.856.A>.
- Albarracín D, Wyer RS. The cognitive impact of past behavior: influences on beliefs, attitudes, and future behavioral decisions. *J Pers Soc Psychol* 2000;79(1):5–22. <https://doi.org/10.1037/0022-3514.79.1.5>.
- Albarracín D, McNatt PS, Klein CTF, Ho RM, Mitchell AL, Kumkale GT. Persuasive communications to change actions: An analysis of behavioral and cognitive impact in HIV prevention. *Heal Psychol* 2003;22(2):166–77. <https://doi.org/10.1037/0278-6133.22.2.166>.
- Albarracín D, Sunderrajan A, Dai W. Action, inaction, and actionability: Definitions and implications for communications and interventions to change behaviors. In: Bevins RA, Hope D, editors. *The 65th Annual Nebraska Symposium on Motivation - Change and Maintaining Change*. Springer; 2018.
- Albarracín D, Sunderrajan A, Dai W, White BX. The social creation of action and inaction: From concepts to goals to behaviors. *Adv Exp Soc Psychol* 2019. <https://doi.org/10.1016/bs.aesp.2019.04.001>.
- Albarracín D. Cognition in persuasion: An analysis of information processing in response to persuasive communications. In: Zanna MP, editors. *Advances in Experimental Social Psychology*, vol 34; 2002, p. 61–130. <http://www.sciencedirect.com/science/article/pii/S0065260102800041>.
- Albarracín D, Wyer Jr RS. Elaborative and nonelaborative processing of a behavior-related communication. *Personal Soc Psychol Bull* 2001;27(6):691–705. <https://doi.org/10.1177/0146167201276005>.



- [25] Albarracín D. *Action and Inaction in a Social World: Predicting and Changing Attitudes and Behaviors*. Cambridge: Cambridge University Press; 2020.
- [26] Ajzen I, Fishbein M, Lohmann S, Albarracín D. The Influence of Attitudes on Behavior. In: Albarracín D, Johnson BT, editors. *The Handbook of Attitudes*, Volume 1: Basic Principles. Routledge; 2018. <https://doi.org/10.1007/97802294218>.
- [27] Maurer J, Uscher-Pines L, Harris KM. Perceived seriousness of seasonal and A (H1N1) influenzas, attitudes toward vaccination, and vaccine uptake among U. S. adults: Does the source of information matter? *Prev Med (Baltim)* 2010;51(2):185–7. <https://doi.org/10.1016/j.ypmed.2010.05.008>.
- [28] Ahmed N, Quinn SC, Hancock GR, Freimuth VS, Jamison A. Social media use and influenza vaccine uptake among White and African American adults. *Vaccine* 2018;36(49):7556–61. <https://doi.org/10.1016/j.vaccine.2018.10.049>.
- [29] Dunn AG, Surian D, Leask J, Dey A, Mandl KD, Coiera E. Mapping information exposure on social media to explain differences in HPV vaccine coverage in the United States. *Vaccine* 2017;35(23):3033–40. <https://doi.org/10.1016/j.vaccine.2017.04.060>.
- [30] King G, Rosen O, Tanner M. Information in ecological inference: An introduction. In: King G, Rosen O, Tanner M, editors. *Ecological Inference: New Methodological Strategies*. Cambridge: Cambridge University Press; 2004. p. 1–12.
- [31] Brewer P, Venaik S. The ecological fallacy in national culture research. *Organ Stud* 2014;35(7):1063–86. <https://doi.org/10.1177/0170840613517602>.
- [32] Glasman LR, Albarracín D. Forming attitudes that predict future behavior: A meta-analysis of the attitude-behavior relation. *Psychol Bull* 2006;132(5):778–822. <https://doi.org/10.1037/0033-2909.132.5.778>.
- [33] Lehmann BA, Ruiter RA, Wicker S, Chapman G, Kok G. Medical students' attitude towards influenza vaccination; 2015. <https://doi.org/10.1186/s12879-015-0929-5>.
- [34] Chen GM. Tweet this: A uses and gratifications perspective on how active Twitter use gratifies a need to connect with others. *Comput Human Behav* 2011;27(2):755–62. <https://doi.org/10.1016/j.chb.2010.10.023>.
- [35] Singh JP, Dwivedi YK, Rana NP, Kumar A, Kapoor KK. Event classification and location prediction from tweets during disasters. *Ann Oper Res* 2019;283(1–2):737–57. <https://doi.org/10.1007/s10479-017-2522-3>.
- [36] Middlestadt SESE, Fishbein M, Albarracín D, et al. Evaluating the Impact of a National AIDS Prevention Radio Campaign in St. Vincent and the Grenadines 1. *J Appl Soc Psychol* 1995;25(1):21–34. <https://doi.org/10.1111/j.1559-1816.1995.tb01581.x>.
- [37] Centers for Disease Control and Prevention. Everyone can be a flu vaccine advocate! Centers for Disease Control and Prevention. <https://blogs.cdc.gov/publichealthmatters/2017/11/fluvox/> [published 2017].
- [38] Broniatowski DA, Jamison AM, Qi S, et al. Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. *Am J Public Health* 2018;108(10):e1–7. <https://doi.org/10.2105/AJPH.2018.304567>.
- [39] Wahowiak L. Public health working to fight misinformation through trust, relationships: Facts not enough. *Nations Health* 2018;48(5).
- [40] Yaqub O, Castle-clark S, Sevdalis N, Chataway J. Attitudes to vaccination: A critical review. *Soc Sci Med* 2014;112:1–11. <https://doi.org/10.1016/j.socscimed.2014.04.018>.
- [41] Nyhan B, Reifler J, Richey S. The role of social networks in influenza vaccine attitudes and intentions among college students in the southeastern united states. *J Adolesc Heal* 2012;51(3):302–4. <https://doi.org/10.1016/j.jadohealth.2012.02.014>.
- [42] Quinn SC, Hilyard KM, Jamison AM, et al. The influence of social norms on flu vaccination among African American and White adults. *Health Educ Res* 2017;32(6):473–86. <https://doi.org/10.1093/her/cyx070>.
- [43] Guilbeault Douglas, Centola Damon, Zollo Fabiana. Networked collective intelligence improves dissemination of scientific information regarding smoking risks. *PLoS ONE* 2020;15(2):e0227813. <https://doi.org/10.1371/journal.pone.0227813>.
- [44] Scheufele DA, Krause NM. Science audiences, misinformation, and fake news. *Proc Natl Acad Sci USA* 2019;116(16):7662–9. <https://doi.org/10.1073/pnas.1805871115>.
- [45] Laughlin PR. *Group Problem Solving*. Princeton University Press; 2011.
- [46] Dennis JM. Documentation for NORC's AmeriSpeak for Institutional Review Boards; 2016.
- [47] Aaron S, Anderson M. Social Media Use in 2018.; 2018. doi:10.1007/bf02067487
- [48] Ireland ME, Schwartz HA, Chen Q, Ungar LH, Albarracín D. Future-oriented tweets predict lower county-level HIV prevalence in the United States. *Heal Psychol* 2015;34(Suppl):1252–60. <https://doi.org/10.1037/hea0000279>.
- [49] Chan MS, Lohmann S, Morale A, et al. An Online Risk Index for the cross-sectional prediction of new HIV, chlamydia, and gonorrhea diagnoses across U. S. counties and across years. *AIDS Behav* 2018;22(7):2322–33. <https://doi.org/10.1007/s10461-018-2046-0>.
- [50] Signorini A, Segre AM, Polgreen PM. The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. *PLoS ONE* 2011;6(5):e19467. <https://doi.org/10.1371/journal.pone.0019467>.
- [51] Mowery J. Twitter Influenza Surveillance: Quantifying Seasonal Misdiagnosis Patterns and their Impact on Surveillance Estimates. *Online J Public Health Inform* 2016;8(3):e198. <https://doi.org/10.5210/ajph.v8i3.7011>.
- [52] Nagar R, Yuan Q, Freifeld CC, et al. A case study of the New York City 2012–2013 influenza season with daily geocoded Twitter data from temporal and spatiotemporal perspectives. *J Med Internet Res* 2014;16(10):e236. <https://doi.org/10.2196/jmir.3416>.
- [53] Maier D, Waldherr A, Miltner P, et al. Applying LDA topic modeling in communication research: Toward a valid and reliable methodology. *Commun Methods Meas* 2018;12(2–3):93–118. <https://doi.org/10.1080/19312458.2018.1430754>.
- [54] Morales A, Gandhi N, Chan MS, et al. Multi-attribute topic feature construction for social media-based prediction. In 2018 IEEE International Conference on Big Data (Big Data). Seattle (WA, USA): Association for Computational Linguistics; 2018, p. 1073–8. <https://doi.org/10.1109/BigData.2018.8622347>.
- [55] Farhadloo M, Winnege K, Chan MS, Jamieson KH, Albarracín D. Associations of topics of discussion on Twitter with survey measures of attitudes, knowledge, and behaviors related to Zika: Probabilistic study in the United States. *JMIR Public Health Surveill* 2018;4(1):. <https://doi.org/10.2196/publichealth.8186>e16.
- [56] Blei DM, Ng AY, Jordan MI. Latent Dirichlet Allocation. *J Mach Learn Res* 2003;3(4–5):993–1022. <https://doi.org/10.1162/jmlr.2003.3.4.5.993>.
- [57] McCallum AK. MALLET: A machine learning for language toolkit; 2002. <http://mallet.cs.umass.edu>.
- [58] Hox JJ. *Multilevel Analysis: Techniques and Applications*. 2nd Ed. New York, NY, US: Routledge/Taylor & Francis Group; 2010.
- [59] Lüdecke D, Makowski D, Philip Waggoner. performance: Assessment of regression models performance; 2020. <https://easystats.github.io/performance/>.
- [60] Marsman M, Wagenmakers E-J. Bayesian benefits with JASP. *Eur J Dev Psychol* 2017;14(5):545–55. <https://doi.org/10.1080/17405629.2016.1259614>.
- [61] Spiegelhalter DJ, Abrams KR, Myles JP. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. John Wiley & Sons; 2004. <https://doi.org/10.1002/0470092602>.
- [62] World Health Organization. Novel Coronavirus(2019-NCov) Situation Report - 13; 2020.
- [63] Cohen E, Bonifield J. Most states make it difficult for children to get a flu shot. CNN. <https://edition.cnn.com/2019/11/26/health/flu-shots-children-pharmacy/index.html> [published 2019. accessed July 5, 2020].
- [64] Callaghan T, Motta M. Conspiracy theories and fear of needles contribute to vaccine hesitancy for many parents. *The Conversation*. <https://theconversation.com/conspiracy-theories-and-fear-of-needles-contribute-to-vaccine-hesitancy-for-many-parents-1208> [published 2018. accessed July 8, 2020].
- [65] Phillips NE, Levy BL, Sampson RJ, Small ML, Wang RQ. The social integration of American cities: Network measures of connectedness based on everyday mobility across neighborhoods. *Sociol Methods Res.* July 2019;004912411985238. <https://doi.org/10.1177/0049124119852386>.
- [66] Boulianne S. Social media use and participation: A meta-analysis of current research. *Inf Commun Soc* 2015;18(5):524–38. <https://doi.org/10.1080/1369118X.2015.1008542>.