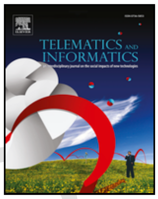
Contents lists available at ScienceDirect

Telematics and Informatics

journal homepage: www.elsevier.com



Exploring childhood vaccination themes and public opinions on Twitter: A semantic network analysis

Jieyu D. Featherstone*, Jeanette B. Ruiz, George A. Barnett, Benjamin J. Millam

University of California, Davis, One Shields Avenue, Kerr Hall 364, Davis, CA 95616, United States

ARTICLE INFO

Keywords:
Childhood vaccination
Twitter
Measles outbreak
Semantic network analysis
Social media

ABSTRACT

During the summer of 2018, a measles outbreak hit Europe, resulting in infections and deaths. Reports link the outbreak to increasing vaccine hesitancy fueled partly by anti-vaccine information propagated on social media. This study examines childhood vaccination themes on Twitter using semantic network analysis. Results showed the most prominent theme was HPV vaccination as a disease preventative. The MMR vaccine autism link was the second biggest theme, followed by measles outbreak rates. These themes reflect public opinions on popular vaccine topics and current issues. Results suggest social media may be an effective place to promote childhood vaccines.

1. Introduction

The recent resurgence of measles across both Europe and the United States (U.S.) is spurring a renewed discussion over childhood vaccination rates (Sun, 2019). In the first six months of 2018, measles outbreaks in Europe had taken 37 lives and infected over 41,000 people (WHO, 2018), hitting a record high over recent decades. The U.S. also saw an increase in measles with 372 infections in the latter half of 2018 with cases mostly in New York City, New York State, and New Jersey (CDC, 2019). Experts attributed these outbreaks to under-vaccinated children in both regions (Selim, 2019). Although children's immunization coverage remains fairly high and stable in the U.S. (Hill et al., 2019), there are pockets with lower than optimal vaccine rates leading to lowered herd immunity. Anti-vaccine sentiment and concerns over childhood vaccine safety have helped facilitate growing vaccine hesitancy online (Kata, 2010). The recent outbreaks, however, have exposed not only the problem of lowered measles, mumps, and rubella (MMR) vaccination rates but also the vulnerability of other childhood vaccinations as well, due to a false sense of security over vaccine-preventable diseases, vaccine hesitancy, and wide spread anti-vaccine sentiment online (Larson, 2018; Temoka, 2013).

Public perception of the severity and susceptibility of vaccine-preventable diseases has declined (Kata, 2010), while public concern about the adverse effects of childhood vaccination has increased (Temoka, 2013). Understandably, many people have trouble recognizing the benefits of vaccination since they have not had first-hand knowledge of these vaccine-preventable diseases (Glanz et al., 2011), drastically shifting public concern from disease prevention to vaccine safety (Freed et al., 2010). Past controversies about vaccine safety, specifically, an association between vaccines and rising autism rates, have further fueled the shift in public focus (Schwartz and Caplan, 2011). While the theory linking vaccines to autism has been discredited (Godlee et al., 2011; Kirkland, 2012), a recent Pew Research Center poll found that 43% of Americans believe that the MMR vaccine has medium to high risks to health (Funk et al., 2017).

* Corresponding author.
E-mail addresses: jding@ucdavis.edu (J.D. Featherstone); jbruiz@ucdavis.edu (J.B. Ruiz); gabarnett@ucdavis.edu (G.A. Barnett); bjmillam@ucdavis.edu (B.J. Millam)

https://doi.org/10.1016/j.tele.2020.101474
Received 21 January 2020; Received in revised form 16 July 2020; Accepted 21 July 2020
Available online xxx
0736-5853/© 2020.

Increased concern about the real or perceived risks of vaccination has created a rise in vaccine hesitancy in the developed world (Bloom et al., 2014). Vaccine hesitancy describes the extent of vaccine acceptance that includes the delay or refusal to vaccinate (Dubé et al., 2016). These concerns extend to vaccination decisions for children. Vaccine hesitant parents are often motivated to seek information from alternative sources online (Betsch, 2011). However, online vaccine information is often inaccurate and incomplete (Broniatowski et al., 2018; Guidry et al., 2015; Zimmerman et al., 2005), tied to vaccine conspiracy beliefs (Featherstone et al., 2019), and comprised of anti-vaccination rhetoric (Kata, 2010). For example, a review of Twitter vaccine posts finds the majority of posts to be vaccine-hesitant (Deiner et al., 2017). Negative vaccine information is also more widely circulated online, regardless of vaccine search terms used (Ruiz and Bell, 2014). An assessment of vaccine tweets found that tweets with negative sentiment attracted more attention than positive tweets and were more likely to be re-tweeted even though the number of negative vaccine tweets was much smaller (Blankenship et al., 2018; Massey et al., 2016; Schmidt et al., 2018).

An overview of vaccine discussions online found vaccine information to largely be negative in sentiment (Kata, 2010). Negative vaccine sentiment varied across social media platforms as 38% of Facebook posts were anti-vaccine and 52% for tweets (Deiner et al., 2019a). Anti-vaccine messages across social media platforms focused on the dangers of childhood vaccination with a concentration on conspiracy theories, the lack of free choice in vaccine decision-making for their children, and personal narratives about vaccine harms. For example, a widely circulated narrative is that vaccination recommendations serve as a means for big-pharma, often in conjunction with the government, to increase profits at the expense of children's health and safety (Guidry et al., 2015). On the other hand, online pro-vaccine messages emphasized the effectiveness of vaccines in protecting children from diseases (Featherstone et al., 2020; Love et al., 2013; Ruiz and Barnett, 2015) and focused on presenting data and statistics rather than personal narratives (Ruiz and Bell, 2014).

Anti-vaccine information on the internet and social media represents a threat to childhood vaccination rates (Larson, 2018) because research has shown that exposure to such information can negatively influence vaccine beliefs, attitudes, and vaccination intentions among parents (Betsch et al., 2010; Nan and Madden, 2012). Therefore, studying how childhood vaccination is portrayed on the internet and social media, in particular, can help understand the information to which people are exposed. Vaccine discussion online focuses on childhood vaccination (Getman et al., 2018). The top five vaccines discussed online are MMR, human papillomavirus (HPV), tetanus, diphtheria, and pertussis (Tdap), influenza, and polio (Kang et al., 2017; Love et al., 2013). We selected the top three childhood vaccination topics: MMR, HPV, and Tdap for assessment. We chose to focus on these vaccines for the following reasons: (1) the MMR vaccine and its link to Autism continues to be a highly controversial topic (Xu, 2019) and its lowered coverage in certain communities contributed to the measles outbreak in 2018 (Selim, 2019); (2) the HPV vaccine was newly introduced in 2006 (Madden et al., 2012) and has had relatively low vaccination coverage (CDC, 2018); (3) the Tdap vaccine continues to have lower vaccination coverage among adolescents because of concerns over its effectiveness (Acosta et al., 2015).

Twitter was selected to assess these three vaccines because it is a popular social media platform with 328 million users worldwide (Clement, 2019) and an even larger global reach due to its content coverage on traditional media (Graham, 2020). Of those users, 42% visit Twitter every day (Greenwood et al., 2016) and Twitter allows users to follow any user. Twitter also provides comprehensiveness data and allows for accurate data retrieval. Posts on Twitter are public and do not require private access (e.g. as do Facebook groups or subreddits), which allows researchers to obtain complete data. Also, Twitter's Premium API provides full access to the entire database of tweets which can be sorted by specified keywords and time ranges. Furthermore, vaccination is frequently mentioned on Twitter (Blankenship et al., 2018), with both pro- and anti-vaccine information discussed on the platform (Deiner et al., 2019a). Therefore, a better understanding of the dominant discourse about childhood vaccination on Twitter could provide critical information for public health professionals on how to better promote vaccination.

2. Themes of vaccine discourse on social media

Misinformation about childhood vaccines circulating on social media poses a threat to public health (Kata, 2010). Research has shown that exposure to vaccine information on social media resulted in lowered trust in vaccinations (Nan and Madden, 2012) and increased vaccine hesitancy in some populations (Dube et al., 2015). Furthermore, a large number of people from diverse demographics seek out vaccine information from social media and may use this information to help make vaccination decisions (Dredze et al., 2016). These studies raised the importance of assessing vaccine discourse on social media so that health professionals can better understand and address public opinions and concerns over childhood vaccination.

Broniatowski et al. (2018) applied thematic analysis to study tweets about vaccination. They discovered that trolls and bots amplified anti-vaccine information on Twitter through repeated anti-vaccine tweets. However, this study did not differentiate between vaccine topics and only analyzed around 10,000 tweets. Through this study, we sought to broaden the scope of assessment by increasing the number of tweets analyzed and by evaluating tweets during an outbreak using semantic network analysis (SNA).

Semantic networks have been used to deduce the themes used in texts. SNA describes the relationship between related concepts through the analysis of word co-occurrence. Deriving from the cognitive paradigm and the linguistic theory of frame semantics (Fillmore, 1982), SNA can highlight the most salient information in a body of text by assessing the networks that emerge. Previous studies have used this approach with other socio-scientific subjects, such as the representation of the HPV vaccine online (Ruiz and Barnett, 2015), genetically engineered foods (Jiang et al., 2018), and Twitter posts during an emerging outbreak (Tang et al., 2018). A better understanding of the eco-system for childhood vaccine themes on social media can help vaccine promotion by targeting what information social media discourse lacks. Our study seeks to understand online discourse for the top three most widely discussed childhood vaccine topics to evaluate similarities and differences in public perceptions.

Specifically, we assess how childhood vaccination is presented on Twitter from July to October 2018. This timeframe includes a period before, during, and after a measles outbreak in Europe. An assessment of tweets that include this time frame could provide important information about what themes emerge about vaccines around an outbreak. Although the outbreak was in Europe, media coverage was extensive in the U.S. Through SNA, childhood vaccine themes on Twitter can be evaluated to determine which concepts are interconnected. For example, does the word *autism* frequently co-occur with the word *vaccine*? The sentiment (negative, neutral, or positive) of concepts within the specific structure(s) can also be assessed.

We sought to assess the discourse about three childhood vaccines on Twitter before, during, and after a measles outbreak. The following research questions were posed:

RQ1: What central themes about childhood vaccines emerge on Twitter?
RQ2: What distinct concepts appear in the Twitter discourse on childhood vaccination?
RQ3: Is the sentiment of tweets negative, neutral, or positive?

3. Methods

This study employed computer-based SNA to evaluate Twitter content about childhood vaccinations from July 1, 2018 to October 15, 2018. This time period includes the measles outbreak in Europe and when the U.S. saw pockets with spikes in measles infections. It also includes the beginning of the U.S. school year when parents prepare to send kids to school. Public schools in the U.S. require proof of vaccination before enrollment and parents are likely faced with deciding whether or not to vaccinate during this period. Therefore, we believe this time frame could be critical for U.S. Twitter users to discuss childhood vaccination.

SNA allows for computer-based analysis of natural text (Rice and Danowski, 1993) and reveals the position and importance of words in relation to other words within a network based on word frequency and centrality measures (Freeman, 1979; Wasserman and Faust, 1994). Emergent clusters of potential meaning, or themes, are identified by analyzing relations among words based on word co-occurrence instead of frequencies of isolated words (Blondel et al., 2008). In this case, the co-occurrence of words is within a five-word window which is in line with established practice (Ruiz and Barnett, 2015). SNA provides an opportunity to uncover meanings associated with large datasets (Carley, 1993; Doerfel, 1998).

3.1. Data collection and analysis

Raw data were collected from Twitter's Premium API using Boolean search methods with keywords "vaccine", "vaccination", "vax", "shot", "immunization", "immunisation" in combination with childhood vaccine topics "MMR", "HPV", and "Tdap". The entire archive of English language tweets within a 15-week period was collected. Retweets and quotes within the established framework were also included in the dataset. Retweets are marked with "RT", meaning that the original tweet was directly shared by another user (not the original poster). Quotes are similar to retweets in that a user is sharing someone else's post but with added commentary from the person retweeting (Xiong et al., 2019). Tweets were restricted to users originating from U.S.-based locations. Both Python and R (Version 3.4.4) were used to organize the data. Using Python, tweet data was saved as text files in the Twitter API JSON format, then imported into R using the *jsonlite* (Ooms, 2014) package. Tweet text data was cleaned in R using both *tm* (Feinerer et al., 2008) and *qdap* (Rinker, 2019) packages, through which we removed URLs, converted to words to lowercase, expanded contractions, removed punctuation, and stripped whitespace. This tweet data was then saved into individual text files for further processing.

Preprocessing procedures were conducted through ConText (Diesner, 2014), which provides a method for organizing large bodies of text into meaningful groupings of concepts. First, syntactically functional words (articles, conjunctions, prepositions) were removed and different forms of the same word (e.g. signify and signifies) were stemmed. For this analysis, emojis were removed from the analysis because they have been found only to slightly improve the expressivity and overall sentiment scores for some post types (Ayvaz and Shiha, 2017). The remaining text was analyzed for word frequency and word sentiment. Words that occurred with frequencies above the mean were included in the analysis.

Next, semantic matrices were generated using the edited texts based on word co-occurrence. The basic network data set is an $n \times n$ matrix $\mathbf{S}$, where n equals the number of nodes (words) in the analysis and S_{ij} is the measured relationship between nodes i and j with the node serving as the unit of analysis. Here, the nodes are identified based on the weighted frequencies of the words. The measurement of word co-occurrence is the standard for creating links between words in a semantic network. Miller (1956) asserted that people can only process five to nine meaningful bits of information at a time; however, more recent studies suggest that this number may range from three to five (Cowan, 2016). Therefore, links were created for words that occurred within five words of one another within each tweet. The frequencies of word co-occurrence were then calculated and ranked.

The semantic networks were created using the network visualization software, Gephi (Bastian et al., 2009). The top 75 words by frequency were included in the network visualization. After importing the data, the network visualization was adjusted using the ForceAtlas2 layout (Jacomy et al., 2014) to examine the spatialization between words. The size of the word label indicated how frequently the word occurred. The thickness of each link represented the weight or number of co-occurrence between two words. The more closely related the words were, the shorter the link distance. Both retweets and quotes contribute to hypertextuality in which certain discourses are amplified in semantic network analysis because the repetition of the same tweets will increase the frequency of certain keywords and the relations between them (Wonneberger et al., 2020).

Modularity analysis and network measures were also conducted using Gephi (Bastian et al., 2009). *Modularity analysis* is a community detection method that reveals the different clusters within a network (Blondel et al., 2008). Each community is indicated by individual color. For example, if there are four different colors in the network visualization, that signifies the presence of four communities, or "clusters" in the corresponding network. Network measures, such as network density, degree, and eigenvector centrality, were also calculated. These indicators provide a depiction of how central and how connected words are within the network. *Network density* is the number of connections divided by the total number of potential connections in the network ($n*(n-1)/2$) and can range from 0 to 1.0. Network density refers to how intertwined the word concepts are, indicating how complex discussions are surrounding an issue. *Degree* is the number of links connecting each word. *Eigenvector centrality* indicates a word's relative influence or how central it is in the network. All measures of centrality were normalized, such that degree is the values divided by the maximum possible values expressed as percentages.

Lastly, sentiment analysis was conducted using IBM Watson *Natural Language Understanding (NLU)* (IBM, 2019). The IBM Watson *NLU* is a cloud service that derives semantic information from texts. The *NLU* sentiment analysis identifies attitude, opinions, or feelings in the text. The *NLU* analyzes sentiments not only based on the polarity of individual words but also the sequence of the text. The results are displayed through sentiment labels of positive (score 0–1), neutral (score 0), negative (score −1–0). The *NLU* model was trained based on Twitter data and other social media texts, so it is a good machine learning model to analyze our Twitter data (Vergara et al., 2017).

4. Results

In total, 139,433 tweets, retweets, and quotes about childhood vaccination were assessed. The average daily number of tweets during this period was 1315 and the average weekly tweets were 9296. The top five peak days of tweets were September 19 (4521 tweets), October 6 (4435 tweets), August 11 (3971 tweets), July 27 (3437 tweets), and September 21 (3321 tweets). The top peak week of tweets was from September 16 to 22 (17,189 tweets) (Fig. 1).

4.1. Semantic network

The average degree of the network was 194.27, and the average weighted degree was 25,034.86 with a network density of 0.80. The mean frequency of occurrence was 4,356. The modularity value was 0.48, indicating meaningful community detection (Blondel et al., 2008). Table 1 indicates the 50 most central terms based on both eigenvector centrality and degree. The most central words in the network were *HPV*, *not*, *vaccination, get,* and *cancer*. The most frequently occurring words in the network were *vaccination*, *HPV*, *not, cause*, and *measles*.

The semantic network for Twitter data is presented in Fig. 2. The network was constructed using the Fruchterman-Reingold algorithm in Gephi (Jacomy et al., 2014). The minimum tie strength between words was 5, indicating that the words co-occurred at least five times within five words of each other.

Four clusters were identified in the network (differentiated by color). The clusters are summarized in Table 2, which includes each cluster's theme, top word associations, and percentage share of the network. Themes were inferred based on the words that encompassed each cluster.

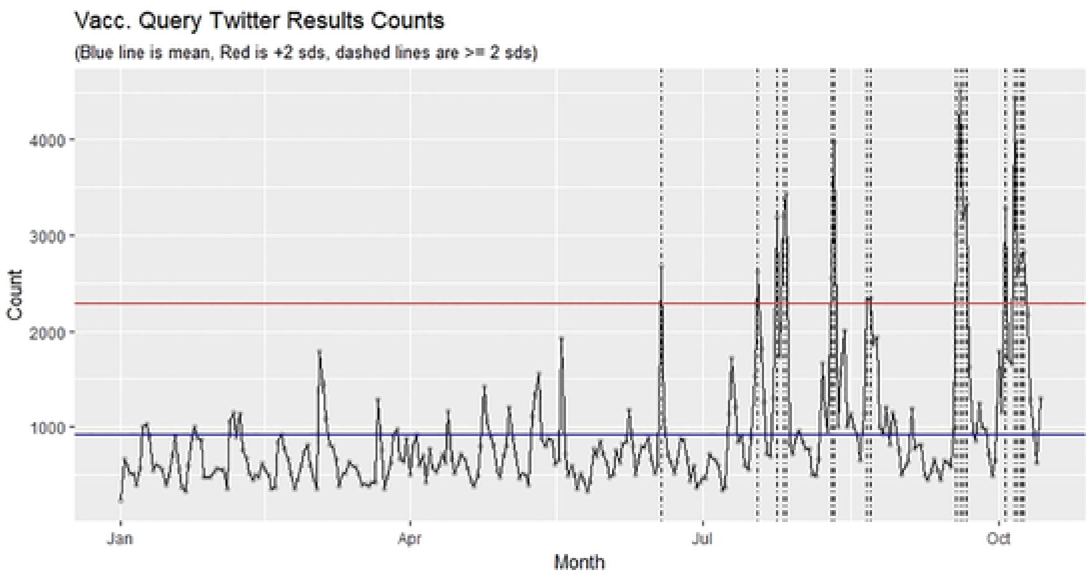


Fig. 1. Query Twitter Results Counts 2018.

Table 1
Summary output of the semantic network analysis (SNA).

Rank	Label	Degree	Eigencentrality	Rank	Label	Degree	Eigencentrality
1	vaccination	243	1	36	safe	226	0.956
2	hpv	243	1	37	age	226	0.954
3	not	243	1	38	gardasil	227	0.954
4	get	242	0.998	39	risk	225	0.953
5	will	241	0.997	40	show	225	0.953
6	can	240	0.996	41	research	225	0.953
7	cancer	241	0.995	42	increase	224	0.949
8	give	238	0.992	43	receive	223	0.948
9	prevent	237	0.989	44	news	224	0.946
10	child	238	0.989	45	life	222	0.946
11	mmr	239	0.987	46	report	223	0.944
12	health	236	0.986	47	good	224	0.943
13	cause	236	0.982	48	measles	224	0.942
14	woman	234	0.979	49	virus	218	0.937
15	need	233	0.976	50	autism	223	0.937
16	girl	232	0.974				
17	know	232	0.974				
18	case	232	0.973				
19	new	231	0.972				
20	go	232	0.972				
21	die	232	0.972				
22	many	232	0.972				
23	cervical	233	0.971				
24	boy	231	0.968				
25	boy	231	0.968				
26	young	230	0.968				
27	study	231	0.966				
28	disease	229	0.965				
29	protect	228	0.964				
30	use	228	0.963				
31	parent	228	0.962				
32	find	228	0.962				
33	vaccineswork	229	0.961				
34	rate	227	0.958				
35	work	226	0.967				

The theme of the largest cluster (71.43% of the network) focused on HPV vaccination for cervical cancer prevention. The most central words for this cluster included *vaccination*, *HPV*, *get*, *cancer*, *cervical*, *women*, *girls*, *boys*, and *prevent*. The largest word association within this cluster was *vaccination* and *HPV*. The second-largest cluster's theme revolved around the measles outbreak in Europe and the Centers for Disease Control and Prevention (CDC) scientists' data report, which covered 9.05% of the network. The most central words in this cluster were *CDC*, *report*, and *number*. The top association between words in this cluster was between *CDC* and *report*. The third-largest cluster (7.14% of the network) discussed MMR and autism. The most central words in this cluster were *doctor, cause*, *autism*, *measles*, *not,* and *science*. The top association between words in this cluster was *cause* and *autism*. The theme of the smallest cluster (4.29% of the network) was about the rising number of MMR cases in France and the consequence of low herd immunity. The most central words were *MMR*, *case*, and *rate*. The top association between words in this cluster was *MMR* and *rate*.

To better understand the themes of the largest cluster, we conducted a second-level semantic network analysis on this cluster alone. Fig. 3 shows the 50 most frequent words within the cluster. The average degree was 139.91 with a graph density of 0.86. The modularity value was just 0.11, indicating a single community, since a value of 0.4 or above should be obtained to generate meaningful clusters (Blondel et al., 2008). Two sub-clusters were detected with the orange cluster taking up 53.62% of the network and the green cluster 46.38%. In the orange cluster, the top five words are *cancer*, *get*, *cervical*, *woman*, *health* and the top five word co-occurrences are *HPV vaccination*, *HPV cancer*, *cancer vaccination*, *boy HPV*, and *boy vaccination*. In the green cluster, the top five words are *vaccination*, *HPV*, *boy*, *protect*, *girl* with the top five word co-occurrences: *get vaccination*, *cervical cancer*, *cervical vaccination*, *eliminate cancer*, *woman cancer*. Sub-clustering shows that the orange focuses on the HPV vaccination whereas the green emphasizes more on HPV diseases.

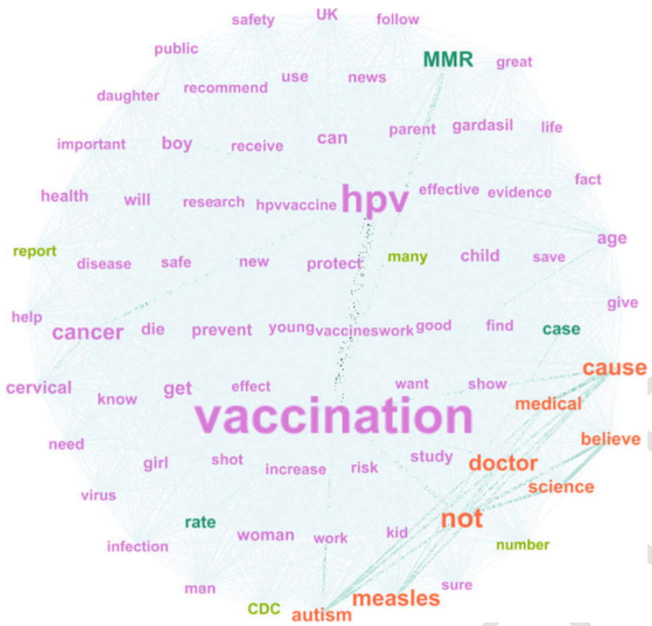


Fig. 2. Visualization of semantic network from Twitter.

4.2. Sentiment analysis

Sentiment analysis was conducted using IBM Watson's *NLU* which analyzed each tweet's sentiment. Sentiment frequencies were aggregated under negative, positive, and neutral sentiments. The average scores for each sentiment were calculated. Results showed that 70.9% of tweets were negative (M = −0.79), 12.6% were positive (M = 0.65), and 16.5% were neutral.

5. Discussion

These results denote a dense network, where the most central words are highly interconnected. High density indicates that the words in the text are closely related (Drieger, 2013). Density measures show that 80% of the words in the network occurred within five words of each other in the text. This network indicates a concentrated range of topics about vaccination on Twitter. During this period, the sentiment of tweets assessed were mostly negative. Surprisingly, most of the network was focused on the HPV vaccine even amidst a measles outbreak. Since the data was restricted to tweets generated in the U.S., it is likely that the outbreak in Europe did not receive significant primary attention in the U.S. social media environment. As expected, the MMR vaccine-autism link and the measles outbreak were the secondary focus in the network. Overall, Twitter discourse during this period focused on the HPV vaccine and reflected not only old debates about vaccines (the connections between the MMR vaccine and autism) but also recent disease trends (the measles outbreak).

The dominant theme present in the network links the HPV vaccine to cervical cancer prevention. In this cluster, keywords, such as "effective", "evidence", and "protect" suggested positive public discourse towards the HPV vaccine as a cancer preventive. These results were consistent with previous research using semantic network analysis for mental representations of the HPV vaccine (Smith and Parrott, 2012). It is noteworthy that the HPV vaccine discourse on Twitter also encouraged boys to get the vaccine, shown by keywords "boy", "important", and "receive". This indicates the success of HPV promotion as social media discourse looks favorably on the vaccine recommendation. Twitter discourse about the HPV vaccine is supportive, encouraging people to vaccinate boys, women, and girls. These results were quite contrary to past studies in which public discourse about the HPV vaccine were negative and controversial (Nan and Madden, 2012; Ruiz and Barnett, 2015). The change in discourse could be attributed to CDC efforts encouraging doctors to promote the HPV vaccine among teenagers (Jenco, 2019) and the active promotion of the HPV vaccine by health care professionals on social media (Massey et al., 2018).

Unsurprisingly, another theme emphasized the rising number of measles cases in France. Two clusters emerged here. One cluster (light green) centered on media reports of the outbreak along with information from the CDC. The other cluster (dark green) focused on the rising number of measles cases in France. In the latter theme, the current measles outbreak was connected to low herd immunity. Both clusters reflected traditional media reports on the European measles outbreak – providing important information from trusted sources (the CDC) while emphasizing causes (in this case, low herd immunity). The emphasis on scientific knowledge (CDC findings) and a proper connection between outbreaks and low herd immunity is a promising finding because it suggests that scientific knowledge is being shared on social media platforms. It is likely that news reports, with information from the CDC, about

Table 2
Summary output of the cluster analysis of the Twitter network.

Theme	Top associations		Association count	Cluster color	Share of network (%)
HPV vaccination and cervical cancer among women, girls, and boys	hpv	vaccination	70,610	Light purple	71.43
	hpv	cancer	17,235		
	cancer	vaccination	13,699		
	get	vaccination	11,320		
	cancer	cervical	11,155		
	boy	hpv	8999		
Beliefs about causes of measles and autism	not	science	15,818	Orange	5.75
	not	believe	15,774		
	cause	autism	11,198		
	cause	measles	10,944		
	not	doctor	10,549		
Data from CDC on cervical cancer, deaths, and disease	CDC	vaccination	1041	Light green	9.05
	CDC	hpv	969		
	CDC	measles	583		
	CDC	data	678		
	CDC	report	540		
Rising MMR cases in France and herd immunity	mmr	rate	5882	Dark green	4.29
	case	rise	5516		
	France	mmr	5390		
	Consequence	mmr	5376		

the outbreaks were shared widely on Twitter. These reports, or discussions surrounding the reports, were a focus of Twitter discourse. This is consistent with research that found people shared news updates, medical information, and scientific knowledge during a previous measles outbreak (Tang et al., 2018).

Although the cluster is small, especially in comparison to the HPV/cancer prevention cluster, the MMR autism debate is still a topic of discussion on Twitter. The frequency of the words “MMR” and “autism” is quite high. This particular Twitter conversation centered on the veracity of Wakefield’s claim that the MMR vaccine causes autism, showing that the controversy remains an active topic even after the medical and scientific communities have established no link between them. Conspiracy theorists believe, and promote the idea that Wakefield is being silenced by government agencies and the pharmaceutical industry, a common vaccine myth perpetuated on the internet (Jolley and Douglas, 2014) by influential users, organizations, and parents alike (Smith, 2017). This cluster highlights the importance of continuing to demystify the autism-vaccine link within anti-vaccine communities.

During the timeframe assessed, the Twitter-sphere disclosed general support of the HPV vaccine, an interest in the current measles outbreak, and an awareness of related vaccination issues. These themes showed the success and potential of vaccination promotion. The positive public discourse about the HPV vaccine indicates the success of its promotion. Also, the evident focus on news updates and reports about measles outbreaks suggest that social media provides a ready outlet for public health information dissemination. It seems, however, that the potential for vaccine promotion varies by vaccine topic. For example, the Tdap vaccine was not mentioned in the network, which may indicate low awareness of that particular vaccine, or that it may be less controversial and therefore, mentioned less. Either way, it is still important to discern who these vaccine-centered tweets have reached. Some anti-vaccine communities are generally more exposed to vaccine skeptical discourse, rather than to messaging that promotes childhood vaccines (Kata, 2012). Health professionals should seek these online communities and work to expose them to accurate vac-

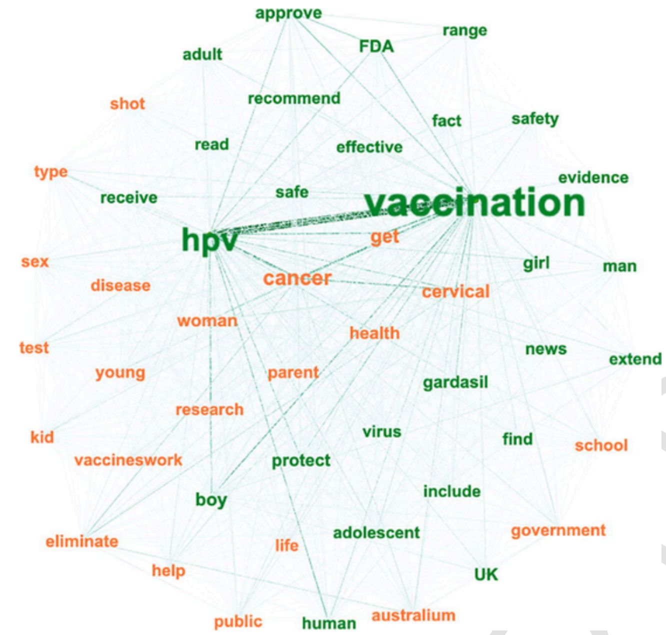


Fig. 3. Visualization of the Second-level Semantic Network Analysis on the Biggest Cluster.

cine information and provide targeted vaccine educational content. Specifically, as highlighted in this study, public health professionals must continue to work to dispel the autism-vaccine connection still employed by anti-vaccination proponents.

The sentiment analysis reveals that the majority of tweets (70.9%) were negative. It is possible that tweets about measles and vaccine hesitant tweets contributed to the negativity. Another study found that during a previous measles outbreak, Facebook posts and Twitter tweets were dominated by negative sentiments (Deiner et al., 2019a). Our study shows that social media has been a consistent platform for both expressing vaccine hesitancy and reporting on health-related news.

This study is not without limitations. First, a semantic network analysis can reveal only what is being discussed, but it cannot address questions about how these themes or discussions affect vaccine-related attitudes, beliefs, or behavior. Further research is needed to assess if and how the presentation of vaccine information on social media may impact vaccine decision making. Second, the analysis was limited to Twitter, one, although large and popular, social media platform among many. Therefore, the results did not reflect all social media discourse on childhood vaccine during this period. Third, we focused our assessment on a specific period of time. It could be that Twitter, or this period, allowed for these specific themes to emerge. Also, as mentioned earlier, retweets and quotes amplify certain discourses while diluting others. Therefore, in order to enhance our understanding of childhood themes and discourse on social media, we encourage studies to investigate further by assessing changes when adding or removing retweets and quotes.

People will continue to use social media as a means to gather information and to use these platforms as sounding boards for their thoughts and ideas about particular issues. These results indicate that social media can be an excellent place to promote childhood vaccines and discover current vaccine debates and topics of discourse. Public health officials could take advantage of these platforms to design more efficient vaccine campaigns tailored to specific vaccine topics and communities. Moreover, health professionals can use social media platforms, such as Twitter, to provide scientifically based vaccine information and recommendations that can be used and spread in a way that emphasizes the full potential of vaccines (Chen, 1999; Chen and Hibbs, 1998).

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.

Acknowledgements

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Innovative Genomics Institute Society Award [Grant number NA].

References

Acosta, A M, DeBolt, C, Tasslimi, A, Lewis, M, Stewart, L K, Misegades, L K, Messonnier, N E, Clark, T A, Martin, S W, Patel, M, 2015. Tdap vaccine effectiveness in adolescents during the 2012 Washington State pertussis epidemic. Pediatrics 135, 981–989.

Ayvaz, S, Shiha, M, 2017. The effects of emoji in sentiment analysis. Int. J. Comput. Electr. Eng. 9, 360–369.

Bastian, M., Heymann, S., Jacomy M. Gephi, 2009. An Open Source Software for Exploring and Manipulating Networks.

Betsch, C, 2011. Innovations in communication: the Internet and the psychology of vaccination decisions. Euro Surveill. 16, 4.

Betsch, C, Renkewitz, F, Betsch, T, Ulshofer, C, 2010. The influence of vaccine-critical websites on perceiving vaccination risks. J. Health Psychol. 15, 446–455.

Blankenship, E.B., Goff, M.E., Yin, J., Tse, Z.T.H., Fu, K.W., Liang, H., Saroha, N., Fung, I.C., 2018. Sentiment, contents, and retweets: a study of two vaccine-related twitter datasets Perm J. 22.

Blondel, V D, Guillaume, J L, Lambiotte, R, Lefebvre, E, 2008. Fast unfolding of communities in large networks. J. Stat. Mech.-Theory E 2008, P10008.

Bloom, B R, Marcuse, E, Mnookin, S, 2014. Addressing vaccine hesitancy. Science 344, 339.

Broniatowski, D A, Jamison, A M, Qi, S H, AlKulaib, L, Chen, T, Benton, A, Quinn, S C, Dredze, M, 2018. Weaponized health communication: twitter bots and russian trolls amplify the vaccine debate. Am. J. Public Health 108, 1378–1384.

Carley, K, 1993. Coding choices for textual analysis: a comparison of content analysis and map analysis. Sociol. Methodol. 23, 75–126.

CDC. Vaccination coverage for selected diseases among adolescents aged 13–17 years, by selected characteristics: United States, selected years 2008–2017. In: Statistics NCfH editor. Atlanta, GA, Centers for Disease Control and Prevention 2018.

CDC. Measles Cases in 2019. 2019.

Chen, R T, 1999. Vaccine risks: real, perceived and unknown. Vaccine 17 (Suppl 3), S41–S46.

Chen, R T, Hibbs, B, 1998. Vaccine safety: current and future challenges. Pediatr. Ann. 27, 445–455.

Clement, J., 2019. Twitter: Number of monthly active users worldwide from 1st quarter 2010 to 1st quarter 2019 (in millions). Statista.

Cowan, N, 2016. Working Memory Capacity. classic ed. Routledge.

ConText, D J, 2014. Software for the Integrated Analysis of Text Data and Network. International Communication Association (ICA), Seattle, WA, USA.

Deiner, M S, Fathy, C, Kim, J, Niemeyer, K, Ramirez, D, Ackley, S F, Liu, F, Lietman, T M, Porco, T C, 2019. Facebook and Twitter vaccine sentiment in response to measles outbreaks. Health Inf. J. 25 (3), 1116–1132. doi:10.1177/1460458217740723.

Doerfel, M L, 1998. What constitutes semantic network analysis? A comparison of research and methodologies. Connections 21, 16–26.

Dredze, M, Broniatowski, D A, Smith, M C, Hilyard, K M, 2016. Understanding vaccine refusal why we need social media now. Am. J. Prev. Med. 50, 550–552.

Drieger, P, 2013. Semantic network analysis as a method for visual text analytics. Proc. – Social Behav. Sci. 79, 4–17. doi:10.1016/j.sbspro.2013.05.053.

Dube, E., Gagnon, D., Ouakki, M., Bettinger, J.A., Guay, M., Halperin, S., Wilson, K., Graham, J., Witteman, H.O., MacDonald, S., Fisher, W., Monnais, L., Tran, D., Gagneur, A., Guichon, J., Saini, V., Heffernan, J.M., Meyer, S., Driedger, S.M., Greenberg, J., MacDougall, H., 2016. Understanding vaccine hesitancy in Canada: results of a consultation study by the Canadian immunization research network. Plos One 11.

Dube, E, Vivion, M, MacDonald, N E, 2015. Vaccine hesitancy, vaccine refusal and the anti-vaccine movement: influence, impact and implications. Exp. Rev. Vaccines 14, 99–117.

Featherstone, J D, Bell, R A, Ruiz, J B, 2019. Relationship of people's sources of health information and political ideology with acceptance of conspiratorial beliefs about vaccines. Vaccine 37, 2993–2997.

Featherstone, J.D., Ruiz, J.B., Barnett, G.A., 2020. Exploring childhood vaccination themes on twitter: a semantic network analysis. In: The 70th Annual Conference of the International Communication Association. Gold Coast, Australia.

Feinerer, I., Hornik, K., Meyer, D., 2008. Text mining infrastructure in R, 2008-03-18 ed., p. 54.

Fillmore, C J, 1982. Linguistics in the morning calm Linguistics Society of Korea. Hanshin, Frame Semantics. Seoul.

Freed, G L, Clark, S J, Butchart, A T, Singer, D C, Davis, M M, 2010. Parental vaccine safety concerns in 2009. Pediatrics 125, 654–659.

Freeman, L C, 1979. Centrality in social networks conceptual clarification. Soc. Networks 1, 215–239.

Funk, C, Kennedy, B, Hefferon, M, 2017. Vast majority of Americans say benefits of childhood vaccines outweigh risks. Pew Res. Center.

Getman, R, Helmi, M, Roberts, H, Yansane, A, Cutler, D, Seymour, B, 2018. Vaccine hesitancy and online information: the influence of digital networks. Health Educ. Behav. 45, 599–606.

Glanz, J M, McClure, D L, O'Leary, S T, Narwaney, K J, Magid, D J, Daley, M F, Hambidge, S J, 2011. Parental decline of pneumococcal vaccination and risk of pneumococcal related disease in children. Vaccine 29, 994–999.

Godlee, F, Smith, J, Marcovitch, H, 2011;342.. Wakefield's article linking MMR vaccine and autism was fraudulent. Br. Med. J..

J. Graham Graham, J., 2020. You would think Twitter is the most visited website, but it's not. By a long shot. SA Today.

Greenwood, S., Perrin, A., Duggan, M., 2016. Social media update 2016 [Internet]. Pew Research Center, Washington, DC.

Guidry, J P D, Carlyle, K, Messner, M, Jin, Y, 2015. On pins and needles: how vaccines are portrayed on Pinterest. Vaccine 33, 5051–5056.

Hill, H A, Singleton, J A, Yankey, D, Elam-Evans, L D, Pingali, S C, Kang, Y, 2019. Vaccination coverage by age 24 months among children born in 2015 and 2016 — national immunization survey-child, United States, 2016–2018. MMWR Morb. Mortal. Wkly. Rep. 68 (41), 913–918. doi:10.15585/mmwr.mm6841e2.

IBM, 2019. IBM Cloud API Docs: Natural Language Understanding. Online, IBM2019.

Jacomy, M., Venturini, T., Heymann, S., Bastian, M., 2014. ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software PLOS ONE 9, e98679.

Jenco, M., 2019. CDC: Teens' HPV vaccination rates improve slightly. AAP News. Online, American Academy of Pediatrics.

Jiang, K, Anderson, B N, Ronald, P C, Barnett, G A, 2018. Semantic network analysis reveals opposing online representations of the search term "GMO". Global Challenges 2, 1–8.

Jolley, D., Douglas, K.M., 2014. The effects of anti-vaccine conspiracy theories on vaccination intentions. Plos One 9.

Kang, G J, Ewing-Nelson, S R, Mackey, L, Schlitt, J T, Marathe, A, Abbas, K M, Swarup, S, 2017. Semantic network analysis of vaccine sentiment in online social media. Vaccine 35, 3621–3638.

Kata, A, 2010. A postmodern Pandora's box: anti-vaccination misinformation on the Internet. Vaccine 28, 1709–1716.

Kata, A, 2012. Anti-vaccine activists, Web 2.0, and the postmodern paradigm – an overview of tactics and tropes used online by the anti-vaccination movement. Vaccine 30, 3778–3789.

Kirkland, A, 2012. The legitimacy of vaccine critics: what is left after the autism hypothesis? J. Health Polit. Policy Law 37, 69–97.

Larson, H J, 2018. The biggest pandemic risk? Viral Misinformation Nature 562, 309.

Love, B, Himelboim, I, Holton, A, Stewart, K, 2013. Twitter as a source of vaccination information: content drivers and what they are saying. Am. J. Infect. Control 41, 568–570.

Madden, K, Nan, X L, Briones, R, Waks, L, 2012. Sorting through search results: a content analysis of HPV vaccine information online. Vaccine 30, 3741–3746.

Massey, P M, Budenz, A, Leader, A, Fisher, K, Klassen, A C, Yom-Tov, E, 2018. What drives health professionals to tweet about #HPVvaccine? Identifying strategies for effective communication. Prev. Chronic Dis. 15, 1–10.

Massey, P.M., Leader, A., Yom-Tov, E., Budenz, A., Fisher, K., Klassen, A.C., 2016. Applying multiple data collection tools to quantify human papillomavirus vaccine communication on Twitter. J. Med. Internet Res. 18.

Miller, G A, 1956. The magical number seven plus or minus two: some limits on our capacity for processing information. Psychol. Rev. 63, 81–97.

Nan, X L, Madden, K, 2012. HPV vaccine information in the blogosphere: how positive and negative blogs influence vaccine-related risk perceptions. Attitudes Behav. Intentions Health Commun. 27, 829–836.

Ooms, J., 2014. The jsonlite package: A practical and consistent mapping between json data and r objects. arXiv preprint arXiv:1403.2805.

Rice, R E, Danowski, J A, 1993. Is it really just like a fancy answering machine? Comparing semantic networks of different types of voice mail users. Int. J. Bus. Commun. 30, 369–397.

Rinker, T W, 2019. Quantitative Discourse Analysis Package. Buffalo, New York.

Ruiz, J B, Barnett, G A, 2015. Exploring the presentation of HPV information online: a semantic network analysis of websites. Vaccine 33, 3354–3359.

Ruiz, J B, Bell, R A, 2014. Understanding vaccination resistance: vaccine search term selection bias and the valence of retrieved information. Vaccine 32, 5776–5780.

Schmidt, A L, Zollo, F, Scala, A, Betsch, C, Quattrociocchi, W, 2018. Polarization of the vaccination debate on Facebook. Vaccine 36, 3606–3612.

Schwartz, J L, Caplan, A L, 2011. Vaccination refusal: ethics, individual rights, and the common good. Prim. Care 38 (717–728), ix.

Selim, L, 2019. Measles explained: what's behind the recent outbreaks? UNICEF.

Smith, T C, 2017. Vaccine rejection and hesitancy: A review and call to action. Open Forum Infectious Diseases 4 (3), ofx146. doi:https://doi.org/10.1093/ofid/ofx146.

Smith, R A, Parrott, R L, 2012. Mental representations of HPV in Appalachia: gender, semantic network analysis, and knowledge gaps. J. Health Psychol. 17, 917–928.

Sun, L.H., 2019. Anti-vaxxers face backlash as measles cases surge. The Washington Post.

Tang, L, Bie, B J, Zhi, D G, 2018. Tweeting about measles during stages of an outbreak: a semantic network approach to the framing of an emerging infectious disease. Am. J. Infection Control 46, 1375–1380.
Temoka, E., 2013. Becoming a vaccine champion: evidence-based interventions to address the challenges of vaccination. S. D. Med. Spec no, 68–72.
Vergara, S., El-Khouly, M., El Tantawi, M., Marla, S., Sri, L., 2017. Building cognitive applications with IBM watson services: Volume 7 Natural Language Understanding. ibm.com/redbooks, IBM Readbooks, p. 110.
Wasserman, S, Faust, K, 1994. Social Network Analysis: Methods and Applications. Cambridge University Press.
WHO, 2018. Measles cases hit record high in the European Region. In: Europe WHOROf editor. Copenhagen, Denmark.
Wonneberger, A., Hellsten, I.R., Jacobs, S.H.J., 2020. Hashtag activism and the configuration of counterpublics: Dutch animal welfare debates on Twitter. Inf. Commun. Soc.
Xiong, Y, Cho, M, Boatwright, B, 2019. Hashtag activism and message frames among social movement organizations: Semantic network analysis and thematic analysis of Twitter during the #MeToo movement. Public Relat. Rev. 45, 10–23.
Xu, Z, 2019. Personal stories matter: topic evolution and popularity among pro- and anti-vaccine online articles. J. Comput. Social Sci. 2, 207–220.
Zimmerman, R K, Wolfe, R M, Fox, J R, Nowalk, M P, Troy, J A, Sharp, L K, 2005. Vaccine criticism on the World Wide Web. J. Med. Internet Res. 7, 8.