

Gardasil on Twitter: A Content Mining Study Examining Message, Context, and Source Characteristics of Human Papilloma Virus (HPV) Vaccine-Related Tweets

Jessica R. Hawks¹, Atsushi Nara², Kristen J. Wells³, John L. Ferrand⁴, Eric R. Walsh-Buhi⁴ 

¹Division of Health Promotion and Behavioral Science, School of Public Health, San Diego State University, San Diego, CA, USA

²Department of Geography, San Diego State University, San Diego, CA, USA

³Department of Psychology, San Diego State University, San Diego, CA, USA

⁴Department of Applied Health Science, School of Public Health, Indiana University Bloomington, Bloomington, IN, USA

Email: erwals@iu.edu

How to cite this paper: Hawks, J.R., Nara, A., Wells, K.J., Ferrand, J.L. and Walsh-Buhi, E.R. (2020) Gardasil on Twitter: A Content Mining Study Examining Message, Context, and Source Characteristics of Human Papilloma Virus (HPV) Vaccine-Related Tweets. *Health*, 12, 1105-1120.

<https://doi.org/10.4236/health.2020.129081>

Received: July 15, 2020

Accepted: September 8, 2020

Published: September 11, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Genital human papilloma virus (HPV) is the most common sexually transmitted virus in the US. Effective HPV vaccines remain controversial and have poor uptake among many adolescent groups. Social media have potential for promoting the HPV vaccine, but little is known about the content or sentiment of social media discussions regarding the vaccine. The purpose of this study was to examine message, context, sentiment, and source characteristics of original HPV- and HPV vaccine-related Twitter posts (tweets) over a 3-month time period. This content mining study used publicly available data purchased from Twitter. One-thousand tweets were randomly selected from a sample of 45,260 English language tweets matching our search criteria and collected from December 2014-January 2015. Tweets were manually coded according to source, context, message, and sentiment characteristics using a coding document developed from previous research. Individuals posted just over half (55%) of tweets, whereas organizations posted 40%. More than half (57%) of all coded tweets were positive, and less than one-fifth were negative. Organizations' tweets were more likely to be positive than individuals' tweets. A temporal display of contextual events and tweet spikes demonstrates that positive tweets can increase during newsworthy events (e.g. FDA approval of Gardasil 9). The HPV vaccine was most often discussed under a cancer prevention message frame. In contrast with some existing research, discussions surrounding the HPV vaccine in original Twitter posts are largely positive. Our study reveals opportunities for adolescent health specialists to provide

accurate, positive HPV vaccination information via Twitter.

Keywords

Papillomavirus Vaccines, Twitter Messaging, Social Media, Information Dissemination, Twitter

1. Introduction

Genital human papillomavirus (HPV) is the most common sexually transmitted virus in the United States. It is estimated that more than 90% of all sexually active men and 80% of sexually active women will be infected with at least one type of HPV [1]. HPV is associated with cervical, vaginal, and vulvar cancers in women; penile cancer in men; and anal and oropharyngeal cancers in both men and women [2] [3] [4] [5] [6]. HPV can also cause genital warts and warts in the throat [7] [8] [9].

In 2014, the Food and Drug Administration (FDA) approved a vaccine that has the potential to prevent approximately 90% of anal, cervical, vaginal, and vulvar cancers [10]. The recombinant HPV 9-valent vaccine (Gardasil 9, Merck) covers 9 HPV types—5 more than Gardasil and 7 more than Cervarix (2 other FDA-approved HPV vaccines) [10]. Gardasil vaccines are frequently delivered in 3 separate doses over a short period of time. HPV vaccines have the potential to either reduce or exacerbate disparities in HPV-related diseases, depending on vaccine uptake which varies based on racial/ethnic group, income, and language [11].

While Gardasil 9, Gardasil, and Cervarix all have demonstrated positive effects in reducing multiple types of cancer, the vaccines remain controversial. Reasons include the fact that the vaccines prevent a virus that is transmitted only via sexual contact, differentiating it from most other recommended vaccines [12]. The cervical cancer prevalence and death rate in the US are also both relatively low. Furthermore, some parents believe giving their children this vaccine implies consent to engage in sexual activity, confers a false sense of protection from sexually transmitted infections (STIs), and will lead to earlier first sex [12]. There is also a vocal anti-vaccine movement in which individuals believe that children already have too many vaccinations on the immunization schedule; vaccines in general are “bad”; HPV vaccines are not safe; and/or long-term side effects are unknown [12].

Given these controversies, HPV vaccine uptake remains low, especially in key priority populations. Such health inequities remain a critical adolescent health concern. For example, in 2014, 56% of 13-year-old and 78% of 17-year-old girls reported completing the 3-dose vaccine series [13]. However, fewer males—47% of 13-year-olds and 62% of 17-year-olds—reported series completion [13]. Moreover, important racial disparities exist. Compared to their non-Hispanic White

peers, non-Hispanic Black females remain less likely to report 3-dose series completion [13]. Although HPV vaccination coverage among females has increased nationally, HPV coverage lags behind tetanus-diphtheria-acellular pertussis (Tdap) vaccine and meningococcal conjugate coverage [13].

Social media represent a potential vaccine promotion modality, as Blacks, Hispanics, and adolescents and young adults who are within the indicated age-range for receiving the HPV vaccine are heavy users of social media [14]. Social media are interactive Web sites and applications that enable users to create, share, comment on, and, modify content [15]. One type of social media, the microblog, consists of the sharing of short pieces of information to the public. One of the best known microblogs is Twitter, started in 2006. Ten years later, Twitter averages 310 million (66 million US) monthly active users [16] [17], who post content in the form of “tweets”, or messages of ≤ 140 characters. Twitter is the social media platform of choice for 18% of teens, beating out Facebook (17%) [18]. Hispanic and Black individuals use Twitter more frequently than their White peers [19]. These young people represent important priority populations for HPV vaccine promotion efforts given both the generally inequitable promotion and uptake within these populations and the critical importance of effectively disseminating evidence-based information to those who are eligible to receive the vaccination.

In order to analyze social media posts, several studies have utilized tweets to examine myriad health issues such as influenza surveillance [20] [21] [22] and tobacco use [23], as well as to assess dissemination of information on antibiotics [24] and to examine attitudes toward HPV and the HPV vaccine [25] [26] [27] [28] [29]. Many of these studies utilized a popular method for analyzing the content of tweets known as sentiment analysis which entails using machine learning methods to analyze tweets in terms of the opinions they express, capturing public opinion about a concept or issue [30]. Investigators have theorized that knowledge gained about HPV and HPV vaccines through sentiment analysis could be harnessed in developing HPV vaccine promotion strategies and messaging through social media. In an effort examine this potential application, tweets about influenza vaccinations were mined to assess the sentiments, or opinions, toward vaccinations as well as how those sentiments could be mapped onto large networks for potential public health outreach [31]. That study found that while sentiments toward influenza vaccination vary, they are often shared within a virtual network of like-minded individuals that would make messages targeting vaccination behaviors difficult to effectively disseminate. Several studies found similar results when examining the content of tweets that discuss the HPV vaccine such that sentiments toward the vaccine are frequently shared within a social community that harbors those same beliefs [26] [32] [33]. These communities could be reached through several distinct sources such as doctors, organizations, celebrities, and fellow community members. This further highlights a research gap that exists between HPV vaccine information on Twitter

and adequately disseminating to the audiences that should consume this information.

The aim of this study was to examine message, context, and source characteristics related to original Twitter posts regarding the HPV vaccine over a 3-month time period. Specifically, our aims were to describe the source characteristics of twitter users (people and organizations) who post original messages on Twitter about HPV and the HPV vaccine, examine the types of HPV- and HPV vaccine-related information shared on Twitter, and assess the general sentiment toward the HPV vaccine among social media users who tweet about the vaccine. This study will add to the current literature by reporting sentiments expressed by organizations as well as by individuals.

2. Materials and Methods

2.1. Data Source

To address our study aims, we purchased social media data from Twitter. Those data included all tweets matching our search criteria posted in December 2014-February 2015. A total of 47,944 tweets met our criteria: 32,019 tweets matched the keywords HPV AND vaccine; 17,656 tweets matched the keyword Gardasil; 564 tweets matched the keyword Cervarix; and 5951 tweets matched the following criteria: (cervical OR vulvar OR vaginal OR anal OR genital warts) AND (vaccine). From this sample, we removed all non-English tweets, leaving 45,260 for this study.

2.2. Coding

Despite the growing use of sentiment analysis through machine learning methods, there is evidence that machine learning methods do not always align well with social science objectives [34]. For this study, investigators opted to use a qualitative coding methodology whereby we developed a coding document, adapted from sentiment analysis and other content analysis research (e.g. HPV vaccine information on YouTube [35] [36] [37], to manually extract information from this sample of tweets. To pilot test the coding document, we collected data via the Twitter streaming API between September 11 and November 20, 2014. Using 3 sets of keywords, tweets were searched within a 20-mile radius of each of the 100 largest US cities. We identified 5178 tweets: 3156 matching the keywords HPV and vaccine; 1190 matching the keyword Gardasil; and 842 matching the keywords “cervical cancer” and vaccine. Two researchers independently coded 4020 characteristics in a nonprobability sample of 96 tweets. Inter-rater reliability was high: 0.97 on 42 original tweets and 0.98 on 54 retweets.

After piloting the coding document, we then randomly selected 1000 original tweets, from the 45,260 tweets obtained from Twitter, for manual coding and analysis. Two researchers independently coded all tweets for source and message characteristics. Coded source characteristics included: whether the Twitter profile represented an individual, celebrity, verified user (indicated by a blue veri-

fied Twitter badge), and/or organization. Tweets posted by individuals were coded using information from the user's profile page indicating whether the individual self-identified as a parent, child, and/or spiritual/religious person; if they had a particular political affiliation; and their reported profession (e.g. journalist, physician). If the tweet was posted by an organization, organization type was coded using information from the user's profile (e.g. business, media outlet, nonprofit, government). Since we pulled data from the poster's profile, these data are not considered anonymous. For example, some "personal" information was recorded, such as poster name (e.g. Kinsey Institute) and affiliation.

All tweets were coded for sentiment using the following categories: "neutral", "negative", "ambiguous", "positive", and "other". "Neutral" tweets were defined as those using language indicating neither approval nor disapproval of the vaccine. "Negative" tweets used language disapproving of the HPV vaccine, whereas "positive" tweets used language approving of the vaccine. "Ambiguous" tweets contained both approving and disapproving language. Tweets coded as "other" did not fit into any of the other sentiment categories. Message tone was also coded as to whether tweets contained concerns about civil liberties, that the vaccine is a hoax, and that the vaccine is dangerous.

Guided by previous research [35], we coded tweets for specific information about the HPV vaccine. The following yes/no categories were coded for message frames, or mention of the vaccine being a: cancer prevention, STI prevention, or genital wart prevention tool. We also coded for other vaccine-related information (e.g. whether the tweet mentioned vaccine eligibility criteria, safety concerns, side effects) and the source that the tweet credited (e.g. Centers for Disease Control and Prevention [CDC], medical doctor). We coded for whether tweets gave a firsthand account of an experience with HPV or the vaccine, person writing about the personal experience was a child or adult, and personal experience came from a parent/relative of someone who had received or would receive the HPV vaccine. Finally, we coded messages for other HPV-related information (e.g. how HPV is transmitted, how to protect against infection, symptoms).

2.3. Analysis

We conducted all analyses using IBM® SPSS® Statistics (Release 22.0.0.0). Means and standard deviations were calculated for continuous variables. Frequencies and percentages were calculated for all categorical variables. We also conducted a descriptive temporal analysis of tweets, organized by date of original message posting and by tweet sentiment/tone. Lastly, a chi-square analysis was conducted to determine statistical differences between source types (*i.e.* individual versus organization) on tweet sentiment.

3. Results

3.1. Source Characteristics

Of the 1000 manually coded tweets, just over half ($n = 548$, 54.8%) were posted

by individuals, whereas organizations posted 404 tweets (40.4%). The remaining tweets came from profiles that were unclear in terms of representing an individual or organization. Among the 548 tweets posted by individuals, the largest identified types were parents (9.5%), journalists (8.2%), and physicians (7.3%; **Table 1**). Only one tweet came from a dental professional, and no tweets from children/teens were identified. Among the 404 tweets from organizations, the largest identified types were health information providers (43.6%), businesses (37.4%), non-profit/advocacy groups (16.6%), and health-care organizations (13.4%).

Only 3.6% of all tweets ($n=36$) were posted by a Twitter-verified user. The median number of followers at the time of tweet was 433 ($M = 2559.43$, $SD = 12,150.74$), ranging from a low of 0 to a high of 179,112 followers for The Kinsey Institute (<http://www.twitter.com/kinseyinstitute>). The median number of profiles the Twitter users followed at the time of tweet was 413 ($M = 3518.19$, $SD =$

Table 1. Source characteristics of tweets regarding the HPV vaccine.

<i>Of tweets posted by an individual (n = 548), the profile mentioned...</i>		
	n	%
Being a parent (e.g. mother or father)	52	9.5
Being a journalist or member of the press	45	8.2
Being a physician (e.g. medical doctor, M.D., D.O., resident)	40	7.3
Religion or spirituality (e.g. scripture, prayer, God)	24	4.4
Being a nurse or allied health worker	18	3.3
Being a health educator	12	2.2
Politics (e.g. political affiliation)	5	0.9
Being an epidemiologist	4	0.7
Being a dentist or dental professional	1	0.2
Being a child (e.g. son or daughter)	0	0.0
<i>Of tweets posted by an organization (n = 404), the profile explicitly mentioned being a...</i>		
	n	%
Health information provider	176	43.6
Business (e.g. company, franchise, store, product, or service)	151	37.4
Non-profit/advocacy group	67	16.6
Healthcare organization	54	13.4
City, state, or federal government	19	4.7
News or media organization	17	4.2
Non-health-related advocacy group	10	2.5
School or school district	8	2.0

a. Source characteristics do not add up to 100% due to information being unavailable on the individual poster's Twitter profile; b. Source characteristics exceed 100% because an organization may be included under multiple categories (e.g. a business and a healthcare organization).

17279.85), ranging from a low of 0 to a high of 369,846 profiles for The Toronto Star (<https://twitter.com/TorontoStar>) a news source that published a high-profile story on the HPV vaccine, which was later retracted and removed from its website.

3.2. Tweet Sentiment

More than half of tweets (n = 571, 57.1%) were coded as having a positive tone. Less than one-fifth of tweets were coded as having a negative tone (n = 188, 18.8%), whereas 195 (19.5%) were coded as neutral. Forty-six tweets were coded as either ambiguous (1.4%) or other tone (3.2%). **Table 2** displays examples of tweets that were coded across the various sentiment categories. Only a small percentage of tweets contained specific negative content about the vaccine, such as the vaccine being dangerous or a hoax, or civil liberties concerns (**Table 3**). **Figure 1** contains a temporal display of tweets, by sentiment/tone, across each day of the 3-month time period. The figure also contains some important contextual events that occurred on or near tweet spikes and demonstrates that there can be increases in positive tweets during newsworthy events, such as the FDA approval of Gardasil 9.

Table 2. Examples of tweets coded by sentiment.

<i>Positive</i>
<ul style="list-style-type: none"> Receiving The HPV Vaccine Won't Turn Eighth Graders Into Sluts http://t.co/bfd6Uxu2Xc#Flu#viruses#virus#disease Also Breaking: seat belts don't promote car accidents! "No increase in risky sexual activity w/HPV vaccine" @ThrupennyBit And now we know cervical cancer can be prevented by the HPV vaccine. Who wouldn't want THAT for their kids? Dear boys and gals, Get vaccinated for #HPV asap!! #Gardasil Well ladies, I just received my second and final Gardasil shot so I am free game. #ComeAndGetSome #CoEds #Parents: You don't open the door to #sex w/#HPV #vaccine. You close the door to #cancer! http://t.co/4qbs4nhXXQ
<i>Negative</i>
<ul style="list-style-type: none"> Gardasil: The decision we will always regret http://t.co/huYPkn5j06 12-year-old Girl Dies Hours After She is Injected with HPV Vaccine Warning: Gardasil injuries can have a permanent impact http://t.co/MBiPTnvM3b Gardasil: Be One Less to Get This Dangerous Vaccine!: http://t.co/wpocNPTd9#gardasil Please RT Wisconsin Sisters Say HPV Vaccine Caused Ovaries To Stop Producing Eggs
<i>Ambiguous</i>
<ul style="list-style-type: none"> Tonight on #ABC2InFocus why parents are still saying no to the HPV vaccine & why docs say getting kids vaccinated should be no-brainer Toronto Star claims HPV vaccine unsafe. Science says the Toronto Star is wrong. http://t.co/cmi9gxWBrW via @DrJenGunter Poor HPV vaccine getting #ecig-like coverage? Complaints of fear-mongering. Interesting, don't know who to believe http://t.co/hSJOn2Czvn
<i>Neutral</i>
<ul style="list-style-type: none"> I have no problem with an ad for the HPV vaccine. I'm just surprised they had #GoldenGlobes ad time money. There's a new HPV vaccine: Here's what you need to know @RobynUrback It's hard to tell what's worse—the Gardasil story or the way the criticism was handled

Table 3. Message characteristics and HPV-related content included in tweets regarding the HPV vaccine (n = 1000).

<i>Sentiment-related content: The tweet mentions...</i>		
	n	%
Something about the vaccine being dangerous (e.g. Gardasil kills, poisons, or injures)	51	5.1
Some concern regarding civil liberties (e.g. government wants mandate the vaccine, which threatens citizen freedom)	10	1.0
Something about the vaccine being a hoax	9	0.9
<i>Information about the HPV vaccine: The tweet mentions that the HPV vaccine can prevent...</i>		
Cancer	176	17.6
A sexually transmitted infection	66	6.6
Genital warts	4	0.4
<i>Information about the HPV vaccine: The tweet mentions...</i>		
Girls and women are approved for the HPV vaccine	48	4.8
Concerns about the long- and short-term safety and/or unknown risks associated with the vaccine	28	2.8
Boys and men are approved for the HPV vaccine	24	2.4
The appropriate age range approved for the vaccine	14	1.4
Side effects of the vaccine (e.g. soreness at the site of injection)	14	1.4
The cost of the 3-dose vaccine	8	0.8
Current or developing legislation to make the vaccine mandatory for schoolchildren	4	0.4
Issues concerning whether insurance companies will cover the newly approved vaccine	2	0.2
Concerns that the vaccine will cause sexual promiscuity or send a message to young people that sexual activity is acceptable at an early age	0	0.0
<i>Source attribution: The tweet cites information from...</i>		
A medical doctor	30	3.0
A member of the research community (e.g. researchers, scientists)	20	2.0
The CDC	13	1.3
Another source such as WebMD, Mayo Clinic	10	1.0
A state or local health department	8	0.8
Government officials (e.g. senators, governors, representatives) or from political organizations	7	0.7
The World Health Organization	4	0.4
An HPV vaccine manufacturer (e.g. Merck)	2	0.2
A celebrity	2	0.2
Cancer organizations (e.g. American Cancer Society)	1	0.1
<i>The tweet mentions...</i>		
The link between HPV and cervical cancer	76	7.6
Cervical Pap tests, anal Pap tests, or DNA tests can screen for HPV infection	11	1.1
The link between HPV and other types of cancers (e.g. head or neck)	11	1.1
The link between HPV and other types of genital cancers (e.g. vulvar, anal, penile)	8	0.8
The virus is transmitted through sexual contact	4	0.4
HPV often goes undetected or clears up on its own	3	0.3
The link between HPV and throat cancer	2	0.2
Safe sex practices (e.g. condom use or abstinence) to protect against acquiring HPV	2	0.2
Various symptoms associated with the virus including genital warts, dysplasia or abnormal cells, or no symptoms at all	1	0.1
Having multiple sexual partners increases the risk of HPV infection	0	0.0

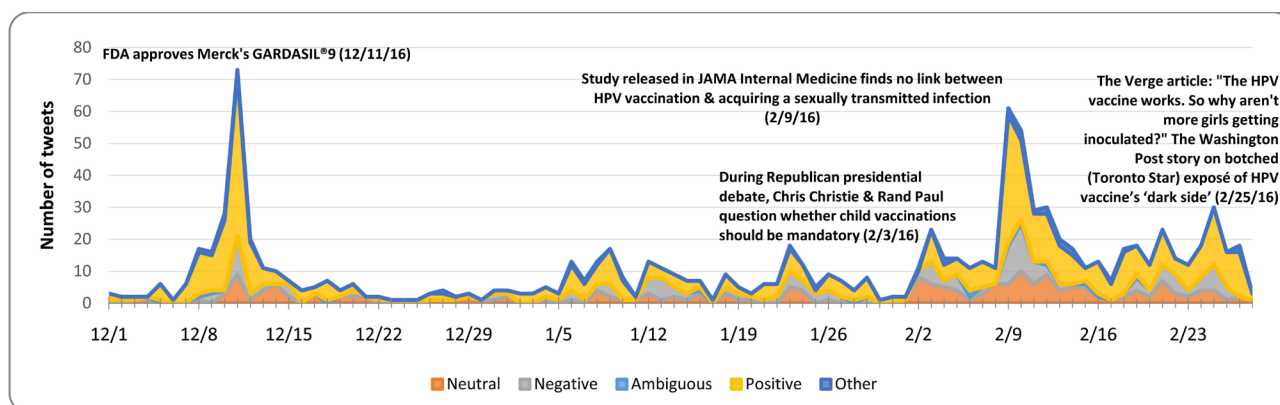


Figure 1. Number of tweets by coded sentiment/tone and date of posting.

Tweets posted by organizations were more likely to be positive, compared to those posted by individuals ($\chi^2 [1, n = 955] = 60.89, p < 0.001$). Almost three-quarters (71.3%) of tweets from organizations were positive, compared with 57% from individuals. Alternatively, 23% of tweets posted by individuals were negative, compared with 13% of tweets from organizations.

3.3. Message Characteristics

The largest message frame related to the HPV vaccine as a cancer prevention tool, as 18% of tweets mentioned that the vaccine can prevent cancer (**Table 3**). Very few tweets in this sample included HPV vaccine-related information (e.g. girls are vaccine-eligible; **Table 3**). Attributing credit to sources of information in tweets was very uncommon in this sample, with the most commonly cited sources in tweets being medical doctors (3.0%; **Table 3**). Few tweets were delivered from a first person perspective. Of tweets posted by individuals, 23 (4.2%) included a personal account from individuals who discussed a firsthand experience with HPV or the vaccine. Only 0.6% originated from the perspective of a parent/relative of a potential or actual HPV vaccine recipient. Finally, few tweets included information about HPV, such as its link with cancers and how HPV can be transmitted (**Table 3**).

4. Discussion

The discussion surrounding the HPV vaccine in original Twitter posts appears to be positive, as more than half of tweets examined in our research contained language approving of the vaccine. These findings differ from some of the existing research on social media and the vaccine, including across 2 content analyses of YouTube videos [37] [38], a qualitative study [39] and content analysis [40] of online news stories, and a content analysis of top search engine results [36]. In more recent research, Bahk *et al.* [41] found that mainstream media were dominated by positive HPV vaccine sentiment (over a one-year period, the weekly average was 75.5% positive), but on Twitter the predominant sentiment toward the HPV vaccine was negative (over a one-year period, the weekly average was

74.9% negative). The Bahk *et al.* study included retweets in analyses, however, and it is possible that negative material is retweeted more frequently. In addition, our study indicates that the quantity, sentiment, and tone of tweets varies over time and is linked to “newsworthy” events, such as FDA approval for a new vaccination, research results being disseminated, and political debates.

Although HPV vaccine sentiment seems to vary across time, studies, and platforms, research indicates that negative sentiments seem to be more influential, at least regarding HPV vaccine sentiment. In a controlled experiment, for example, Nan and Madden [42] found that college students exposed to a negative blog post about the vaccine “perceived the vaccine as less safe, held more negative attitudes toward the vaccine, and had reduced intentions to receive the vaccine” compared to those who had no blog exposure. However, those exposed to the positive blog had no changes in perceptions, intentions, or attitudes regarding the vaccine [42]. Twitter-specific data on HPV vaccine sentiment suggest that we could stratify tweets based on social network connectedness to target large swaths of consumers that share specific sentiments regarding the vaccine [28] [43]. Targeted tweets or posts could then be disseminated to those community members to improve their sentiment regarding HPV vaccination. While this is one potential approach to effectively disseminating messages regarding HPV vaccination, there is still a critical need to optimize methods to counteract tweets providing negative and incorrect information, especially since HPV-related tweets often occur in quick spikes.

On Twitter, the HPV vaccine is framed largely as a cervical cancer prevention tool, with less focus on its potential to prevent genital warts or other cancers. In addition, there were more tweets describing how the HPV vaccine could help females, potentially making it seem less relevant for males. According to Bigman *et al.* [44], participants who learned about HPV vaccine effectiveness believed more in its effectiveness for preventing cervical cancer and felt more positively toward the vaccine for cervical cancer prevention. Adolescent medicine specialists and public health educators should expand communication regarding the HPV vaccine’s ability to prevent other types of cancers and warts, which can occur in males and females.

Very few tweets in this sample included HPV- and HPV vaccine-related information. Calloway *et al.* [45] and Habel *et al.* [40] found that US newspaper coverage and online news stories of the HPV vaccine lacked detailed information on HPV, which could contribute to people misunderstanding the complexity of cervical cancer and HPV. While the 140 character limit for tweets constricts the amount of information that can be shared, professionals can share detailed information via Twitter by including links with information about HPV and the vaccine.

Individuals (most commonly parents, journalists, and/or physicians) shared most (54.8%) tweets, but few tweets were written in first-person. Organizations (mostly health information providers, non-profit/advocacy groups, and health-care organizations) shared about two-fifths of tweets. Additionally, the majority

of negative sentiments were tweeted by individuals instead of by organizations. Thus, organizations may be a potential source of information that would reach more followers outside of social networks that share homogenous beliefs about the vaccine. Other types of healthcare personnel—nurses, allied health professionals, health educators, and dental professionals—might consider communicating correct and positive information about the HPV vaccine via Twitter. These practitioners could also be targeted as recipients of positive messages in the hopes that they would then disseminate the information to patients.

Study Limitations

This research should be considered within the context of its limitations. First, this research only examined data generated by Twitter users. Inclusion of data from other platforms is warranted, especially since the social media platforms of choice for 28% and 27% of teens are Snapchat and Instagram, respectively [18]. Second, our research did not examine the relations between Twitter users and their HPV vaccine sentiment. Salathé and Khandelwal [31] found that users with similar sentiments toward an issue share information more often with each other than users with dissimilar sentiments. Future research should explore how HPV vaccine information flows through and across social networks. Such research may guide adolescent health and medicine interventions and facilitate diffusion of information about the HPV vaccine across networks. Third, since only 0.4% of tweets contained exact geo-coordinates, we did not analyze these data from a geospatial perspective. According to Twitter, only 1% - 2% of all tweets are geo-tagged [46], so this limitation is inherent in most Twitter-based research studies that do not utilize a location inferencing technique [47] [48]. Fourth, our choice of keywords may have limited the population from which our sample emerged. For example, we searched for “vaccine” instead of “vax”, a term that may be more likely used among individuals opposing vaccination. Fifth, we only examined original tweets, excluding retweeted material. Because of this decision, we may have excluded potentially important information. However, according to Sysomos, 71% of tweets receive no reaction (compared to 23% of tweets that get a reply and 6% that get retweeted) [49]. Sixth, our research is limited to English language tweets and cannot be generalized to non-English tweets.

Parents, adolescents, and young adults all play a role in deciding whether or not to receive the HPV vaccine. Dempsey and Zimet [50] found that social media may influence adolescents' intentions and decisions regarding vaccinations. It may be that different approaches to promoting the vaccine are necessary for different population segments. In addition, sentiment toward the vaccine appears to be swayed by media and political events, but it is not clear how people move back and forth from having a positive or negative attitude toward the HPV vaccine. It is also unclear how Twitter can be used to change that or motivate parents, adolescents, or young adults to choose to get the vaccine. Future research

should investigate tweets and Twitter-users to determine how reflective tweets are of underlying beliefs and knowledge about the HPV vaccine and whether these behavioral constructs can be changed via Twitter.

While it is clear that some healthcare providers—mainly physicians—disseminate HPV vaccine information via Twitter, our study reveals an opportunity for a greater variety of healthcare personnel to provide accurate and positive HPV vaccination information on this platform. In addition, healthcare providers should be aware of messages that parents, adolescents, and young adults are receiving about the vaccine, especially in times of substantial coverage of an HPV vaccine news topic. Future directions could also target these various healthcare practitioners to then disseminate positive messages regarding HPV vaccinations to appropriate recipients. Finally, it is critical that public health professionals learn more about how to harness the power of social media like Twitter to better inform the decision making of parents, adolescents, and young adults related to the HPV vaccine.

Declarations

Funding: This work was supported in part by the Center for Human Dynamics in the Mobile Age (for use of its SMART Dashboard), and seed money provided by San Diego State University.

Ethical approval: None required.

Guarantor: EWB.

Contributorship: EWB, JRH, and JLF researched literature and/or conceived the study. EWB, AN, JLF, and KJW was involved in protocol development and data analysis. EWB and JRH wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Acknowledgements

We thank Ming-Hsiang (Ming) Tsou, Ph.D. and the Center for Human Dynamics in the Mobile Age for use of the SMART Dashboard, and San Diego State University for the seed money that funded the acquisition of Twitter data and supported personnel efforts. We also thank Marcus Lewis, Roxana Rezai, and Pegah Sabzi for their work as tweet coders.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Chesson, H.W., Dunne, E.F., Hariri, S. and Markowitz, L.E. (2014) The Estimated Lifetime Probability of Acquiring Human Papillomavirus in the United States. *Sexually Transmitted Diseases*, **41**, 660-664.
<https://doi.org/10.1097/OLQ.000000000000193>

- [2] Forman, D., *et al.* (2012) Global Burden of Human Papillomavirus and Related Diseases. *Vaccine*, **30**, 12-23. <http://dx.doi.org/10.1016/j.vaccine.2012.07.055>
- [3] Viens, L., *et al.* (2016) Human Papillomavirus-Associated Cancers—United States, 2008-2012. *Morbidity and Mortality Weekly Report*, **65**, 661-666. <https://doi.org/10.15585/mmwr.mm6526a1>
- [4] Gillison, M., *et al.* (2012) Prevalence of Oral HPV Infection in the United States, 2009-2010. *JAMA*, **307**, 693-703. <https://doi.org/10.1001/jama.2012.101>
- [5] International Agency for Research on Cancer (2007) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Volume 90: Human Papillomaviruses. International Agency for Research on Cancer, Lyon.
- [6] Shiels, M.S., Kreimer, A.R., Coghill, A.E., Darragh, T.M. and Devesa, S.S. (2015) Anal Cancer Incidence in the United States, 1977-2011: Distinct Patterns by Histology and Behavior. *Cancer Epidemiology, Biomarkers & Prevention*, **24**, 1548-1556. <https://doi.org/10.1158/1055-9965.EPI-15-0044>
- [7] Greer, C.E., *et al.* (1995) Human Papillomavirus (HPV) Type Distribution and Serological Response to HPV Type 6 Virus-Like Particles in Patients with Genital Warts. *Journal of Clinical Microbiology*, **33**, 2058-2063. <https://doi.org/10.1128/JCM.33.8.2058-2063.1995>
- [8] Gissmann, L., Wolnik, L., Ikenberg, H., Koldovsky, U., Schnurch, H.G. and Hausen, H.Z. (1983) Human Papillomavirus Types 6 and 11 DNA Sequences in Genital and Laryngeal Papillomas and in Some Cervical Cancers. *Proceedings of the National Academy of Sciences of the United States of America*, **80**, 560-563. <https://doi.org/10.1073/pnas.80.2.560>
- [9] Dinh, T., Sternberg, M., Dunne, E. and Markowitz, L. (2008) Genital Warts among 18- to 59-Year-Olds in the United States, National Health and Nutrition Examination Survey, 1999-2004. *Sexually Transmitted Diseases*, **35**, 357-360. <https://doi.org/10.1097/OLQ.0b013e3181632d61>
- [10] Merck (2014) Highlights of Prescribing Information: Gardasil 9 (Human Papillomavirus 9-Valent Vaccine, Recombinant).
- [11] Jeudin, P., Liveright, E., Del Carmen, M.G. and Perkins, R.B. (2014) Race, Ethnicity, and Income Factors Impacting Human Papillomavirus Vaccination Rates. *Clinical Therapeutics*, **36**, 24-37. <https://doi.org/10.1016/j.clinthera.2013.11.001>
- [12] Haber, G., Malow, R.M. and Zimet, G.D. (2007) The HPV Vaccine Mandate Controversy. *Journal of Pediatric and Adolescent Gynecology*, **20**, 325-331. <https://doi.org/10.1016/j.jpag.2007.03.101>
- [13] Reagan-Steiner, S., *et al.* (2015) National, Regional, State, and selected Local Area Vaccination Coverage among Adolescents Aged 13-17 Years—United States, 2014. *MMWR Morb Mortal Wkly Rep*, **64**, 784-792. <https://doi.org/10.15585/mmwr.mm6429a3>
- [14] Duggan, M., Ellison, N.B., Lampe, C., Lenhart, A. and Madden, M. (2015) Demographics of Key Social Networking Platforms. *Pew Research Center*, 1-13.
- [15] Ahlqvist, T., Bäck, A., Heinonen, S. and Halonen, M. (2010) Road-Mapping the Societal Transformation Potential of Social Media. *Foresight*, **12**, 3-26. <https://doi.org/10.1108/14636681011075687>
- [16] Statista (2016) Number of Monthly Active Twitter Users in the United States from 1st Quarter 2010 to 2nd Quarter 2016 (In Millions).
- [17] Statista (2016) Number of Monthly Active Twitter Users Worldwide from 1st Quarter 2010 to 1st Quarter 2016 (In Millions).

- [18] Chew, J. (2016) Snapchat Now Beats Instagram among Teen Users. *Fortune*.
- [19] Krogstad, J. (2015) Social Media Preferences Vary by Race and Ethnicity.
- [20] Culotta, A. (2010) Towards Detecting Influenza Epidemics by Analyzing Twitter Messages. *Proceedings of the 1st Workshop on Social Media Analytics*, Washington DC, July 2010, 115-122. <https://doi.org/10.1145/1964858.1964874>
- [21] Signorini, A., Segre A.M. and Polgreen, P.M. (2011) The Use of Twitter to Track Levels of Disease Activity and Public Concern in the US during the Influenza a H1N1 Pandemic. *PLoS ONE*, **6**, e19467. <https://doi.org/10.1371/journal.pone.0019467>
- [22] Paul, M.J., Dredze, M. and Broniatowski, D. (2014) Twitter Improves Influenza Forecasting. *PLoS Currents Outbreaks*, **6**. <https://doi.org/10.1371/currents.outbreaks.90b9ed0f59bae4ccaa683a39865d9117>
- [23] Prier, K.W., Smith, M.S., Giraud-Carrier, C. and Hanson, C.L. (2011) Identifying Health-Related Topics on Twitter. In: Salerno, J., Yang, S.J., Nau, D. and Chai, SK., Eds., *Social Computing, Behavioral-Cultural Modeling and Prediction*, Springer, Berlin, 18-25. https://doi.org/10.1007/978-3-642-19656-0_4
- [24] Scansfeld, D., Scansfeld, V. and Larson, E.L. (2010) Dissemination of Health Information through Social Networks: Twitter and Antibiotics. *American Journal of Infection Control*, **38**, 182-188. https://doi.org/10.1007/978-3-642-19656-0_4
- [25] Shapiro, G.K., Surian, D., Dunn, A.G., Perry, R. and Kelaheer, M. (2017) Comparing Human Papillomavirus Vaccine Concerns on Twitter: A Cross-Sectional Study of Users in Australia, Canada and the UK. *BMJ Open*, **7**, e016869. <https://doi.org/10.1136/bmjopen-2017-016869>
- [26] Keim-Malpass, J., Mitchell, E.M., Sun, E. and Kennedy, C. (2017) Using Twitter to Understand Public Perceptions Regarding the #HPV Vaccine: Opportunities for Public Health Nurses to Engage in Social Marketing. *Public Health Nursing*, **34**, 316-323. <https://doi.org/10.1111/phn.12318>
- [27] Massey, P.M., Leader, A., Yom-Tov, E., Budenz, A., Fisher, K. and Klassen, A.C. (2016) Applying Multiple Data Collection Tools to Quantify Human Papillomavirus Vaccine Communication on Twitter. *Journal of Medical Internet Research*, **18**, e318. <https://doi.org/10.2196/jmir.6670>
- [28] Surian, D., Nguyen, D.Q., Kennedy, G., Johnson, M., Coiera, E. and Dunn, A.G. (2016) Characterizing Twitter Discussions about HPV Vaccines Using Topic Modeling and Community Detection. *Journal of Medical Internet Research*, **18**, e232. <https://doi.org/10.2196/jmir.6045>
- [29] Du, J., Xu, J., Song, H., Liu, X. and Tao, C. (2017) Optimization on Machine Learning Based Approaches for Sentiment Analysis on HPV Vaccines Related Tweets. *Journal of Biomedical Semantics*, **8**, Article No. 9. <https://doi.org/10.1186/s13326-017-0120-6>
- [30] Giachanou, A. and Crestani, F. (2016) Like It or Not: A Survey of Twitter Sentiment Analysis Methods. *ACM Computing Surveys*, **49**, Article No. 28. <https://doi.org/10.1145/2938640>
- [31] Salathé, M. and Khandelwal, S. (2011) Assessing Vaccination Sentiments with Online Social Media: Implications for Infectious Disease Dynamics and Control. *PLoS Computational Biology*, **7**, e1002199. <https://doi.org/10.1371/journal.pcbi.1002199>
- [32] Zhou, X.J., *et al.* (2015) Using Social Connection Information to Improve Opinion Mining: Identifying Negative Sentiment about HPV vaccines on Twitter. *Studies in Health Technology and Informatics*, **216**, 761-765.
- [33] Dunn, A.G., Leask, J., Zhou, X., Mandl, K.D. and Coiera, E. (2015) Associations

- between Exposure to and Expression of Negative Opinions about Human Papillomavirus Vaccines on Social Media: An Observational Study. *Journal of Medical Internet Research*, **17**, e144. <https://doi.org/10.2196/jmir.4343>
- [34] Chen, N.-C., Drouhard, M., Kocielnik, R., Suh, J. and Aragon, C.R. (2018) Using Machine Learning to Support Qualitative Coding in Social Science: Shifting the Focus to Ambiguity. *ACM Transactions on Interactive Intelligent Systems*, **8**, Article No. 9. <https://doi.org/10.1145/3185515>
- [35] Johnson, J.Q., Sionean, C. and Scott, A.M. (2011) Exploring the Presentation of News Information about the HPV Vaccine: A Content Analysis of a Representative Sample of U.S. Newspaper Articles. *Health Communication*, **26**, 491-501. <https://doi.org/10.1080/10410236.2011.556080>
- [36] Madden, K., Nan, X., Briones, R. and Waks, L. (2012) Sorting through Search Results: A Content Analysis of HPV Vaccine Information Online. *Vaccine*, **30**, 3741-3716. <https://doi.org/10.1016/j.vaccine.2011.10.025>
- [37] Briones, R., Nan, X., Madden, K. and Waks, L. (2012) When Vaccines Go Viral: An Analysis of HPV Vaccine Coverage on YouTube. *Health Communication*, **27**, 478-485. <https://doi.org/10.1080/10410236.2011.610258>
- [38] Ache, K. and Wallace, L. (2008) Human Papillomavirus Vaccination Coverage on YouTube. *American Journal of Preventive Medicine*, **35**, 389-392. <https://doi.org/10.1016/j.amepre.2008.06.029>
- [39] Casciotti, D., Smith, K., Tsui, A. and Klassen, A. (2014) Discussions of Adolescent Sexuality in News Media Coverage of the HPV Vaccine. *Journal of Adolescent Health*, **37**, 133-143. <https://doi.org/10.1016/j.adolescence.2013.11.004>
- [40] Habel, M.A., Liddon, N., and Stryker, J.E. (2009) The HPV Vaccine : A Content Analysis of Online News Stories. *Journal of Women's Health*, **18**, 401-407. <https://doi.org/10.1089/jwh.2008.0920>
- [41] Bahk, C.Y., Cumming, M., Paushter, L., Madoff, L., Thomson, A. and Brownstein, J.S. (2016) Publicly Available Online Tool Facilitates Real-Time Monitoring of Vaccine Conversations and Sentiments. *Health Affairs*, **35**, 341-347. <https://doi.org/10.1377/hlthaff.2015.1092>
- [42] Nan, X. and Madden, K. (2012) HPV Vaccine Information in the Blogosphere: How Positive and Negative Blogs Influence Vaccine-Related Risk Perceptions, Attitudes, and Behavioral Intentions. *Health Communication*, **27**, 829-836. <https://doi.org/10.1080/10410236.2012.661348>
- [43] Backstrom, L., Huttenlocher, D., Kleinberg, J. and Lan, X.Y. (2006) Group Formation in Large Social Networks: Membership, Growth, and Evolution. *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Philadelphia, August 2006, 44-54. <https://doi.org/10.1145/1150402.1150412>
- [44] Bigman, C.A., Cappella, J.N. and Hornik, R.C. (2010) Effective or Ineffective: Attribute Framing and the Human Papillomavirus (HPV) Vaccine. *Patient Education and Counseling*, **81**, S70-S76. <https://doi.org/10.1016/j.pec.2010.08.014>
- [45] Calloway, C. Jorgensen, C.M., Saraiya, M. and Tsui, J. (2006) A Content Analysis of News Coverage of the HPV Vaccine by U.S. Newspapers, January 2002-June 2005. *Journal of Women's Health*, **15**, 803-809. <https://doi.org/10.1089/jwh.2006.15.803>
- [46] Gnip (2014) Twitter Geographical Metadata. <http://support.gnip.com/articles/geo-intro.html>
- [47] Ajao, O., Hong, J. and Liu, W.R. (2015) A Survey of Location Inference Techniques

on Twitter. *Journal of Information Science*, **41**, 855-864.

<https://doi.org/10.1177/0165551515602847>

- [48] Laylavi, F., Rajabifard, A. and Kalantari, M. (2016) A Multi-Element Approach to Location Inference of Twitter: A Case for Emergency Response. *ISPRS International Journal of Geo-Information*, **5**, 56. <https://doi.org/10.3390/ijgi5050056>
- [49] Sysomos (2010) Replies and Retweets on Twitter. <https://sysomos.com/inside-twitter/twitter-retweet-stats>
- [50] Dempsey, A.F. and Zimet, G.D. (2015) Interventions to Improve Adolescent Vaccination: What May Work and What Still Needs to Be Tested. *American Journal of Preventive Medicine*, **49**, S445-S454. <https://doi.org/10.1016/j.amepre.2015.04.013>