Vaccine 33 (2015) 3354–3359

Contents lists available at ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

Exploring the presentation of HPV information online: A semantic network analysis of websites

Jeanette B. Ruiz*, George A. Barnett

Department of Communication, University of California, Davis, One Shields Ave., Davis, CA 95616, USA

ARTICLE INFO

Article history:
Received 19 February 2015
Received in revised form 22 April 2015
Accepted 8 May 2015
Available online 31 May 2015

Keywords:
HPV vaccine websites
Semantic network analysis
Concept co-occurrence

ABSTRACT

Context: Negative vaccination-related information online leads some to opt out of recommended vaccinations.
Objective: To determine how HPV vaccine information is presented online and what concepts co-occur.
Methods: A semantic network analysis of the words in first-page Google search results was conducted using three negative, three neutral, and three positive search terms for 10 base concepts such as *HPV vaccine*, and *HPV immunizations*. In total, 223 of the 300 websites retrieved met inclusion requirements. Website information was analyzed using network statistics to determine what words most frequently appear, which words co-occur, and the sentiment of the words.
Results: High levels of word interconnectivity were found suggesting a rich set of semantic links and a very integrated set of concepts. Limited number of words held centrality indicating limited concept prominence. This dense network signifies concepts that are well connected. Negative words were most prevalent and were associated with describing the HPV vaccine's side-effects as well as the negative effects of HPV and cervical cancer. A smaller cluster focuses on reporting negative vaccine side-effects. Clustering shows the words *women* and *girls* closely located to the words *sexually*, *virus*, and *infection*.
Discussion: Information about the HPV vaccine online centered on a limited number of concepts. HPV vaccine benefits as well as the risks of HPV, including severity and susceptibility, were centrally presented. Word cluster results imply that HPV vaccine information for women and girls is discussed in more sexual terms than for men and boys.

1. Introduction

HPV vaccine compliance rates are less than optimal [1]. Surveys conducted by the CDC found significantly lower rates of HPV vaccine uptake among young adult women (19– 26 years of age), with 35% reporting receiving at least one HPV vaccination dose in 2009 [2] compared with 53% of the adolescent (13–17 years old) population [3]. In 2011, HPV vaccination was recommended for boys aged 11 and 12. Among males aged 19–21, vaccination coverage for at least one dose of HPV vaccine was at 2.4% in 2012 [2].

Healthcare providers remain the primary source of medical advice, but when it comes to vaccination information, people are seeking additional information [4] outside of traditional sources and mainly from the Internet [5]. Much of the information available on the Internet is not verified [6] and the type of information accessed can depend on the search terms employed by the user [7]. Meanwhile, anti-vaccination sites often provides misleading vaccine information [8].

Studies have assessed news coverage of the HPV vaccine [9,10], online news stories [10,11], and online HPV vaccine information [12,13]. Madden et al. (2012) analyzed vaccine websites for their tone and content references to vaccine effectiveness, risk, and HPV susceptibility. In contrast, the present study uses semantic network analysis to analyze the vaccine-related information found online specifically, what concepts are presented and their co-occurrence.

While traditional content analysis relies on the researcher to set coding categories based on *priori* definitions on the basis of some theoretical framework [14], computer-based semantic network analysis (SNA) allows for the analysis of natural text rather than abstract a *priori* content categories [15]. It does so by identifying emergent clusters of potential meaning by analyzing relations among words instead of frequencies of isolated words. SNA can thus uncover meanings associated with a topic better than traditional methods [16,17]. Researchers can determine which concepts are interconnected as well as the sentiment of particular concepts

* Corresponding author. Tel.: +1 530 400 3102.
E-mail addresses: jbruiz@ucdavis.edu (J.B. Ruiz), gabarnett@ucdavis.edu (G.A. Barnett).

http://dx.doi.org/10.1016/j.vaccine.2015.05.017
0264-410X/

Table 1
Search terms used ($N = 223$ websites).

Search term	Valence
HPV vaccine risks	Negative
HPV vaccination risks	Negative
HPV immunization risks	Negative
Human papillomavirus vaccine risks	Negative
Human papillomavirus vaccination risks	Negative
Human papillomavirus immunization risks	Negative
Gardasil risks	Negative
Cervarix risks	Negative
Vaccine for HPV risks	Negative
Vaccine HPV risks	Negative
HPV vaccine	Neutral
HPV vaccination	Neutral
HPV immunization	Neutral
Human papillomavirus vaccine	Neutral
Human papillomavirus vaccination	Neutral
Human papillomavirus immunization	Neutral
Gardasil	Neutral
Cervarix	Neutral
Vaccine for HPV	Neutral
Vaccine HPV	Neutral
HPV vaccine benefits	Positive
HPV vaccination benefits	Positive
HPV immunization benefits	Positive
Human papillomavirus vaccine benefits	Positive
Human papillomavirus vaccination benefits	Positive
Human papillomavirus immunization benefits	Positive
Gardasil benefits	Positive
Cervarix benefits	Positive
Vaccine for HPV benefits	Positive
Vaccine HPV benefits	Positive

within the structure. For example, does the word *side-effect* co-occur with the word *mild* or *severe* most frequently? From this subjective information can be extracted to infer if the intent is neutral (*side-effect* + *mild*) or negative (*side-effect* + *severe*). This study addressed three broad questions: (a) Based on word co-occurrence, what concepts emerge from HPV vaccine websites? (b) What meanings are associated with HPV vaccines? (c) What sentiments about HPV vaccines prevail online?

2. Methods

2.1. Search terms

Web searches were conducted on November 30, 2014, via the Google search engine (www.google.com) that included negative, neutral, and positive search terms for 10 base concepts as shown in Table 1. Prior content analysis used seven of the 10 base terms to locate information about HPV vaccination online [13,18], and the additional three base terms were chosen because they were the most popular HPV vaccine search terms used, based on Google Trends data [19]. Previous research noted the importance of using both *vaccination* and *immunization* keywords [8], so these were included as base search terms. Also negative and positive searches were added to the base concepts in line with findings that note negative, neutral, and positive search terms impacts the valence of results [7]. Negatively valenced searches added *risks* to the base terms, neutral valenced search terms did not add a term to the base search word, and positively valenced search terms added *benefits* to the base terms. These 30 searches were conducted using both Google Chrome and Mozilla Firefox to confirm that the results are independent of the browser used.

2.2. Website retrieval

The first page of search results, comprising a list of 10 websites, was retrieved for each term because web users rarely go beyond first-page results [20–22]. The 30 searches used generated 300 potential websites for analysis (10 terms × 3 valences × 10 first-page websites). After eliminating 77 ineligible websites, 223 remained for analysis. A website was excluded from the study if it was a listserv website, a video result, a book preview and/or review, a directory devoted solely to listing other websites, a non-English website, a website focused exclusively on non-HPV vaccines, a website about veterinary vaccines, a journal access that required purchase for access, or a broken link that led to no active page. Some search terms duplicated previous search results, duplicate webpages were included only once in the analysis.

2.3. Meta-data network construction

ConText [23], a software tool designed to construct network data based on natural language text data, was used to extract the semantic network from the text. ConText allows for the organization of large bodies of text into meaningful groupings of concepts. Because the program identifies word co-occurrence, preconceived categories and intercoder reliability tests are unnecessary [24]. ConText reads the selected body of text, and then eliminates stop words (articles, prepositions, conjunctions, and transitive verbs that do not contribute to textual meaning). The remaining text (2483 independent words) was analyzed for the most frequently occurring words and word sentiment. Word sentiment analysis lists words and their polarity of sentiment under different parts of speech, whether the expressed opinion is positive, negative, or neutral. These were extracted using the ConText [23] sentiment dictionary. For example, the word *death* would be listed as negative, *woman* as neutral, and *healthy* as positive. Sentiment analysis relies on lists of words and phrases with specific connotations and takes into account the polarity of the words that precede and follow the words assessed to ensure an appropriate sentiment. For example, *not healthy* would be classified as negative, *stays healthy* as positive. A word-by-word matrix was created with each cell containing frequency co-occurrence of the words with the sentence as the unit of analysis and connectedness identified as words occurring within five adjacent words distance.

2.4. Semantic network analysis

SNA describes the structure of relationships among various entities, which could include people, organizations, or communication concepts [24–27]. In the present study, the entities studied were words within website postings [28]. SNA studies have incorporated social network analysis methods for assessing what topics are present and what concepts co-occur as well as frequency of topic presentation [24–26,28,29].

The basic network data set is an $n \times n$ matrix **S**, where n equals the number of nodes in the analysis and s_{ij} is the measured relationship between the specific nodes i and j. The node is the unit of analysis. In this study, nodes represent the most salient words, identified on the basis of cumulative (weighted) frequencies. Links are based on word co-occurrences within a five-word distance. There are a number of indicators of a node's position in the network. The Gini coefficient measures the inequality of words within a frequency distribution where a zero score expresses perfect equality and a score of one expresses maximum inequality among the words. Density refers to the interconnectedness of the concepts [30]. It is defined as the number of links among the words divided by the number of possible links ($n^*(n-1)/2$). It can range from 0.0 to 1.0. The share indicates the percentage of the text accounted for by each specific word.

Centrality denotes the importance, prominence, or power of a concept in the network [31,32]. In other words, it shows how central a concept is within the text. One measure of centrality is degree, which is a node's total number of links or the sum of the frequencies of co-mentions. Betweenness is a measure of centrality and illustrates the degree to which a word is directly connected to other words that are not directly connected to each other [33]. Another measure, eigenvector centrality measures the overall influence of a node in the network [34]. It assigns relative scores to all nodes in the network such that greater weight is placed on ties to more central words/concepts. All measures of centrality were normalized, such that degree and betweenness are their values divided by the maximum possible values expressed as percentages. Eigenvalue centrality is the scaled value divided by the maximum difference possible expressed as a percentage.

The procedures described above each provide a network description of online HPV vaccine information at one point in time. UCINET [35] was used to calculate network measures. The data networks were visualized in Gephi [36] where node colors indicate cluster affiliation (based on modularity)[1], node size is scaled by degree (number of direct ties), and tie width represents the frequency of co-occurrence of a pair of words. Visual representations are important to facilitate the understanding of networks and help illustrate the results of the analysis. The meaning of concepts in text is analyzed through a cluster analysis, or in terms of emerging clusters of nodes around a focal concept. Hierarchical cluster analysis groups members into subsets where the members are relatively similar or structurally equivalent [38]. Clusters are formed based on the distance between objects (in this case words), based on similarity between every pair of objects in the data set. This analysis employed hierarchical clustering using UCINET [35]. These multiple measures were employed to provide varied descriptions of the network's structure and to prevent biased interpretations based on the use of a single measure.

3. Results

The network structure reveals a very dense network where more than half of the words are connected to each other. The density of the network is .724. The Gini-coefficient for the distribution of words in the network is .60, indicating that the frequency of word use is concentrated among only a few words. As one would anticipate, the most frequently occurring words in the network are *vaccine* (2960 occurrences), followed by *HPV* (1842), *cancer* (1058), *cervical* (654), and the vaccine brand name *Gardasil* (628) (Table 2).

Table 2
Top 50 most frequently occurring words.

	Word	Frequency		Word	Frequency
1	Vaccine	2960	26	include	262
2	Hpv	1842	27	man	253
3	Cancer	1058	28	risk	248
4	Cervical	654	29	genital	241
5	Gardasil	628	30	recommend	237
6	Woman	581	31	immune	230
7	Infection	526	32	people	221
8	Year	473	33	disease	219
9	Type	431	34	cause	218
10	Report	425	35	human	215
11	Effect	415	36	study	211
12	Age	397	37	case	209
13	Prevent	385	38	event	203
14	Girl	383	39	state	202
15	Health	348	40	wart	199
16	Reaction	321	41	side	192
17	Protect	319	42	problem	191
18	Dose	318	43	virus	190
19	Adverse	286	44	death	187
20	Cervarix	286	45	papillomavirus	186
21	Sexually	273	46	child	182
22	Medical	271	47	active	181
23	Information	270	48	develop	180
24	Receive	270	49	young	167
25	Safety	268	50	research	161

Table 3 presents the network share and various centralities for the words in the online HPV vaccine information network. Ten words hold 40% of the network share. These words are also at the center of the network. In terms of betweenness centrality, the words *vaccine, HPV*, and *Gardasil* provided the most connections. The most central words (greater than one standard deviation than the mean) in terms of eigenvector centrality were *HPV, vaccine, cancer*, and *cervical*.

Fig. 1 graphically displays the online HPV vaccine information network, showing the prominent words and the words they are connected to. The network figure was organized using the Fruchterman-Reingold algorithm [39]. The thicker the line connecting the words, the more often a pair of words co-occurred in the text. *Vaccine* and *HPV* are most central in the network as can be seen by their central placement in the network. Node and label size in the figure indicates the centrality and frequency of occurrence in the network. For example, *Gardasil* appears more frequently than *Cervarix*, as can be seen by the larger node and label size for *Gardasil*, but both words are closely connected to *vaccine* and *HPV* as can be seen by word placement in the figure. Thicker lines between words indicate greater interconnectedness of the words in the network.

The cluster analysis resulted in one major cluster that depicts the HPV vaccine as a preventative of HPV and cervical cancer. Two smaller clusters describe the HPV vaccine's negative side-effects and reporting of these adverse side-effects. The word *boy* is loosely linked in the network, which indicates that the concept of boys within the discussion of HPV vaccination is also further removed. *Girl* and *women* are close to *sexually* implying that the discussion about HPV vaccination is happening in more sexual terms than when the discussion centers on men and HPV vaccines.

Sentiment analysis revealed a higher percentage of negative to positive words; 54% of words are negative, 46% of words are positive. The most common negative words included: *cancer, adverse, risk, disease*, and *problem*. The highest occurring positive words were: *recommend, protect, effective, benefit*, and *important*.

4. Discussion

The findings from this study indicate a condensed network structure of concepts where the most central words appear most frequently and are highly interconnected. Such a dense network indicates that the scope of online information about the HPV vaccine is concentrated on a few concepts. These results are consistent with previous research that found U.S.-based websites about the HPV vaccine provided suboptimal information [40].

The primary frame present in the websites analyzed is that the HPV vaccine protects from HPV and cervical cancer. These findings are surprising since the assessment included websites retrieved on the basis of negative, neutral, and positive valenced search terms. However, results are in line with previous research that found less critical and negative coverage of the HPV vaccine online [12,13,41].

A larger number of negative words were found in the network. These negative words describe HPV vaccine side-effects and the negative effects of the diseases against which these vaccines immunize. Focusing on the costs of not vaccinating is particularly important as vaccine information that describes both the benefits of vaccination as well as the negative repercussions of not vaccinating, have been found to increase acceptance of the HPV vaccine [42].

[1] Modularity is the number of edges falling within groups minus the expected number in an equivalent network with edges placed at random [37] Newman MEJ. Modularity and community structure in networks. Proceedings of the National Academy of Sciences of the United States of America. 2006;103:8577–82.

Table 3
Degree, share and centralities for the top 100 words.

		Degree	Share	Betweenness	Eigen vector
1	American	0.145	0.002	0.089	1.648
2	CDC	0.352	0.005	0.265	4.010
3	Cervarix	0.909	0.013	0.526	10.609
4	FDA	0.309	0.004	0.214	2.862
5	Gardasil	1.889	0.027	0.926	20.329
6	Merck	0.300	0.004	0.273	3.522
7	United	0.286	0.004	0.286	2.943
8	Active	0.281	0.004	0.121	3.276
9	Administer	0.189	0.003	0.114	2.869
10	Adverse	1.156	0.016	0.547	8.511
11	Age	1.173	0.017	0.659	11.891
12	Allergic	0.339	0.005	0.117	2.849
13	Approve	0.339	0.005	0.135	4.474
14	Benefit	0.247	0.003	0.128	3.252
15	Boy	0.334	0.005	0.150	3.701
16	Cancer	3.484	0.049	0.707	44.948
17	Case	0.618	0.009	0.431	7.982
18	Cause	0.801	0.011	0.431	11.906
19	Cervical	2.393	0.034	0.714	36.090
20	Child	0.384	0.005	0.302	4.625
21	Clinical	0.420	0.006	0.245	3.178
22	Common	0.494	0.007	0.399	6.573
23	Control	0.229	0.003	0.161	1.736
24	Datum	0.221	0.003	0.189	1.665
25	Death	0.530	0.008	0.259	6.175
26	Develop	0.505	0.007	0.314	7.747
27	Disease	0.640	0.009	0.462	7.759
28	Doctor	0.296	0.004	0.373	2.430
29	Dose	0.802	0.011	0.517	8.399
30	Drug	0.246	0.003	0.242	1.963
31	Effect	0.880	0.012	0.441	7.884
32	Event	0.691	0.010	0.272	5.982
33	Female	0.484	0.007	0.308	5.673
34	Fever	0.202	0.003	0.145	1.014
35	Find	0.456	0.006	0.448	5.561
36	Genital	1.011	0.014	0.312	13.113
37	Girl	1.286	0.018	0.687	15.046
38	Group	0.281	0.004	0.293	2.759
39	Headache	0.138	0.002	0.040	0.526
40	Health	0.750	0.011	0.665	6.598
41	High	0.148	0.002	0.160	2.024
42	Hpv	5.835	0.083	0.957	72.938
43	Human	0.733	0.010	0.279	7.793
44	Immune	0.186	0.003	0.101	1.815
45	Include	0.789	0.011	0.759	8.082
46	Infection	1.431	0.020	0.637	23.377
47	Information	0.473	0.007	0.315	5.215
48	Injection	0.344	0.005	0.211	2.485
49	Make	0.326	0.005	0.442	3.992
50	Man	0.625	0.009	0.303	7.712
51	Medical	0.308	0.004	0.311	2.499
52	Million	0.288	0.004	0.154	3.051
53	Month	0.270	0.004	0.133	2.184
54	National	0.283	0.004	0.092	3.352
55	Number	0.255	0.004	0.388	2.499
56	Occur	0.345	0.005	0.522	3.262
57	Pain	0.260	0.004	0.078	0.865
58	Pap	0.318	0.005	0.136	3.475
59	Papillomavirus	0.576	0.008	0.195	5.933
60	Patient	0.242	0.003	0.298	2.319
61	People	0.596	0.008	0.623	7.749
62	Percent	0.369	0.005	0.154	4.997
63	Person	0.247	0.003	0.304	2.185
64	Pregnant	0.181	0.003	0.063	3.061
65	Prevent	1.160	0.016	0.310	19.540
66	Problem	0.523	0.007	0.330	5.761
67	Program	0.442	0.006	0.225	5.618
68	Protect	0.844	0.012	0.240	14.936
69	Provide	0.283	0.004	0.201	3.698
70	Public	0.205	0.003	0.145	1.738
71	Rate	0.267	0.004	0.161	3.067
72	Reaction	1.026	0.015	0.399	8.113
73	Receive	1.015	0.014	0.745	12.437
74	Recommend	0.656	0.009	0.425	8.703
75	Report	1.194	0.017	0.607	9.353
76	Research	0.254	0.004	0.196	3.414
77	Risk	0.694	0.010	0.596	10.755
78	Safety	0.591	0.008	0.382	8.449
79	Screening	0.336	0.005	0.142	4.445
80	Severe	0.324	0.005	0.089	2.391
81	Sexually	0.564	0.008	0.270	6.143
82	Shot	0.293	0.004	0.386	2.966
83	Show	0.370	0.005	0.310	3.778
84	Side	0.739	0.010	0.279	6.494
85	Site	0.223	0.003	0.039	1.060
86	State	0.221	0.003	0.243	2.590
87	Study	0.593	0.008	0.673	6.387
88	System	0.347	0.005	0.231	3.287
89	Test	0.424	0.006	0.278	5.324
90	Time	0.258	0.004	0.335	3.609
91	Transmit	0.206	0.003	0.044	2.118
92	Treatment	0.163	0.002	0.114	1.668
93	Trial	0.474	0.007	0.215	4.558
94	Type	1.675	0.024	0.367	30.083
95	Vaccine	6.724	0.095	1.019	69.472
96	Virus	0.521	0.007	0.266	6.854
97	Wart	0.816	0.012	0.212	10.445
98	Woman	2.217	0.031	0.778	26.288
99	Year	1.470	0.021	0.838	14.796
100	Young	0.537	0.008	0.444	6.250

Although the negative sentiment concentrates on providing side-effect information, too much of a focus on this information without an equal or greater focus on the negative aspects of not vaccinating may discourage vaccination or cause inertia [43]. Furthermore, the focus on vaccine side-effects of makes the vaccine risks more salient for the reader as opposed to highlighting the dangers of the diseases vaccines prevent. A smaller cluster of negative words is focused on reporting negative side-effects. However, reporting adverse vaccine effects to central public health agencies can provide positive outcomes, (1) valuable information to advance the study of vaccine reactions, and (2) public reassurance of knowing that a vaccine safety infrastructure exists [44].

Recommendations to vaccinate boys and young men for HPV came almost 5 years after the initial recommendation for girls and women so it is not surprising to find *boy* furthest removed in the network. While the word *men* is closer than *boy* in the network, it is still loosely connected in the network. It is interesting to find *girl* and *women* close to *sexually* as HPV is transmitted through sexual contact for both sexes. However, perhaps because the vaccine was initially recommended for women to protect against cervical cancer, the vaccine is still largely discussed as protection for women.

This study examined website content at one point in time, not longitudinally. However, when comparing results to earlier analyses there is some indication that online content about HPV vaccine safety and effectiveness may be framed more positively. For example, the sentiment analysis found only a slight variation in the number of positive versus negative words in the corpus. Also, the negative words tended to focus on information about mild vaccine side-effects rather than extreme vaccine results. Furthermore, the visualization displays largely positive meanings associated with the HPV vaccine. Another previous assessment found the majority of websites to be neutral or positive in tone [13]. This is a shift from earlier studies that found a greater amount of anti-vaccination information [18,45].

There are some limitations to this study. First, a semantic analysis reveals what people might learn about HPV vaccines online, but it cannot address concerns about the effects of the information on HPV-vaccine related attitudes, beliefs, and behaviors. Second, although a varied number of search terms were employed, all possible iterations of search terms used by individuals may not have been captured. It is likely, however, that other search terms would have

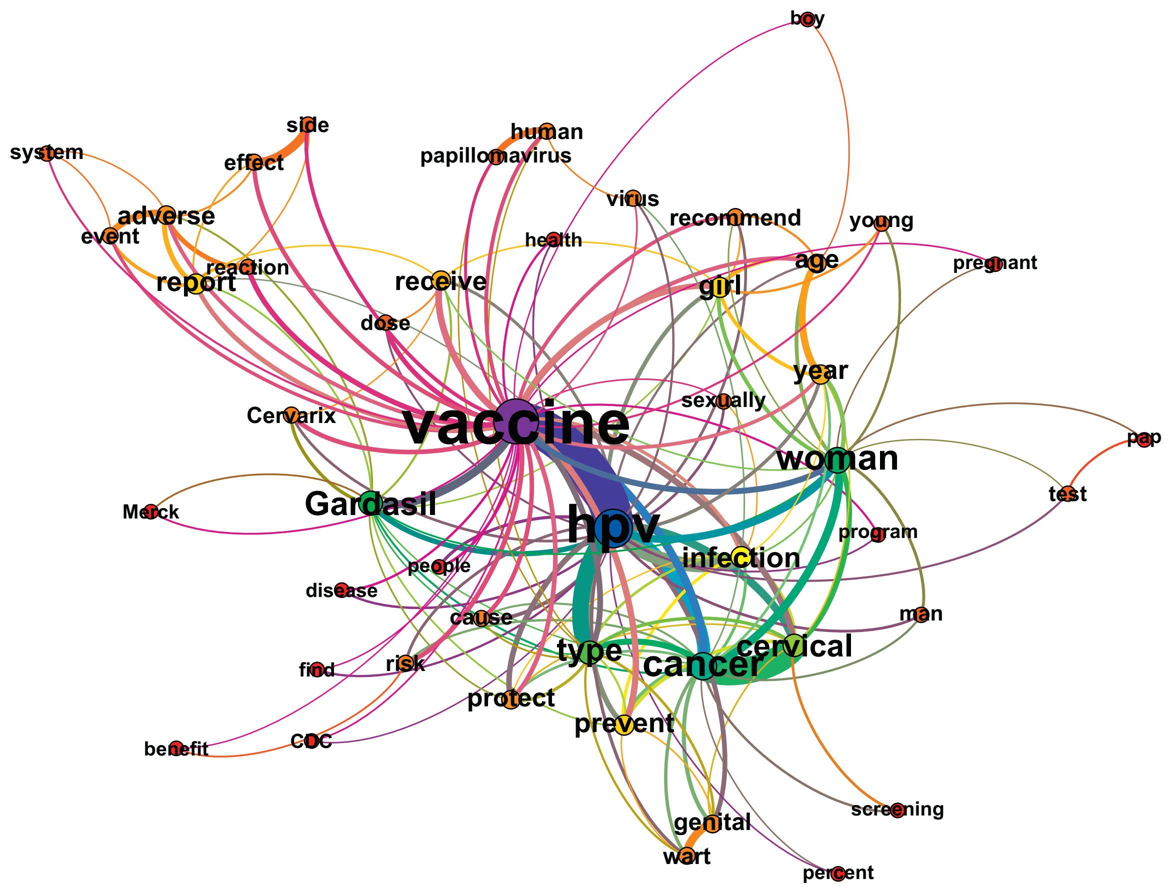

Fig. 1. HPV Vaccine Semantic Network. Top 100 words (one standard deviation from the mean). Node and label size in the figure indicate centrality and frequency of occurrence in the network. Line thickness between words demonstrates frequency of word co-occurrence.

led to one of the websites included in the present sample. Third, this research did not assess website sources or time of last update. Finally, the study was limited to search results returned from U.S.-based sites generated through traditional online searches, not including other online information sources, such as social media.

This study suggests that online information about HPV vaccines presents information about vaccine side-effects and the negative aspects of HPV and cervical cancer. Public health communications should continue to concentrate on improving online vaccine messages by focusing on the benefits of vaccination, highlighting the risks of not vaccinating [4,44], and addressing the importance and benefits of vaccinating boys and young men [46]. Future studies should focus on widening the online search strategy to include monitoring of blogs, forums, commentaries following news articles, and various social media platforms. Semantic network analytic methods provide a quick method for extracting and analyzing information from large amounts of text, which is especially useful for assessing fast changing online content. Health professionals who aim to increase HPV vaccine compliance should consider periodic reviews of the information available online in order to understand what information a general search will provide online information seekers. A current understanding of what information individuals may encounter online can help health professionals devise messaging that counteracts any false and or negative information about the HPV vaccine.

Conflicts of interest

The authors declare no conflict of interest in the research for this study.

Acknowledgments

We are grateful to Robert A. Bell for his comments on this paper. The authors thank the two anonymous reviewers for insightful comments that improved this paper.

References

[1] CDC. NIS-Teen Vaccination Coverage Table Data. Centers for Disease Control and Prevention. CDC; 2014.
[2] Williams WW, Lu P-J, O'Halloran A, Bridges CB, Pilishvili T, Hales CM, et al. Noninfluenza vaccination coverage among adults – United States, 2012. Morb Mortal Wkly 2014;63:95–102.
[3] Curtis CR, Dorell C, Yankey D, Jeyarajah J, Chesson H, Saraiya M, et al. National human papillomavirus vaccination coverage among adolescents aged 13-17 years – National Immunization Survey – Teen, United States, 2011. Morb Mortal Wkly 2014;63:61–70.
[4] Chen RT, Hibbs B. Vaccine safety: current and future challenges. Pediatr Ann 1998;27:445–55.
[5] Blume S. Anti-vaccination movements and their interpretations. Social Sci Med 2006;62:628–42.
[6] Davies P, Chapman S, Leask J. Antivaccination activists on the world wide web. Arch Dis Child 2002;87:22–5.
[7] Ruiz JB, Bell RA. Understanding vaccination resistance: vaccine search term selection bias and the valence of retrieved information. Vaccine 2014;32:5776–80.
[8] Wolfe RM, Sharp LK. Vaccination or immunization? The impact of search terms on the Internet. J Health Commun 2005;10:537–51.
[9] Calloway C, Jorgensen CM, Saraiya M, Tsui J. A content analysis of news coverage of the HPV vaccine by US newspapers, January 2002–June 2005. J Womens Health 2006;15:803–9.
[10] Abdelmutti N, Hoffman-Goetz L. Risk messages about HPV, cervical cancer, and the HPV vaccine Gardasil: a content analysis of Canadian and US national newspaper articles. Women Health 2009;49:422–40.
[11] Habel MA, Liddon N, Stryker JE. The HPV vaccine: a content analysis of online news stories. J Womens Health 2009;18:401–7.

[12] Keelan J, Pavri-Garcia V, Tomlinson G, Wilson K. YouTube as a source of information on immunization: a content analysis. J Am Med Assoc 2007;298: 2481–4.
[13] Madden K, Nan X, Briones R, Waks L. Sorting through search results: a content analysis of HPV vaccine information online. Vaccine 2012;30:3741–6.
[14] Krippendorff K. Content anlaysis an introduction to its methodology. 3rd ed. Thousand Oaks, CA: Sage; 2012.
[15] Rice RE, Danowski JA. Is it really just like a fancy answering machine? Comparing semantic networks of different types of voice mail users. Int J Bus Commun 1993;30:369–97.
[16] Doerfel ML. What constitutes semantic network analysis? A comparison of research and methodologies. Connections 1998;21:16–26.
[17] Carley K. Coding choices for textual analysis: a comparison of content analysis and map analysis. Sociol Methodol 1993;23:75–126.
[18] Kata A. A postmodern Pandora's box: anti-vaccination misinformation on the Internet. Vaccine 2010;28:1709–16.
[19] Google. Google Trends; 2014.
[20] Eysenbach G, Köhler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. Br Med J 2002;324:573.
[21] Eysenbach G, Powell J, Kuss O, Sa E-R. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. J Am Med Assoc 2002;287:2691–700.
[22] Malaga RA. Worst practices in search engine optimization. Commun ACM 2008;51:147–50.
[23] Diesner J. ConText. Software for the integrated analysis of text data and network data. Seattle, WA: International Communication Association (ICA); 2014.
[24] Doerfel ML, Barnett GA. A comparison of the semantic and affiliation networks of the International Communication Association. Hum Commun Res 1999;25:589–603.
[25] Danowski JA. Network analysis of message content. In: Richards WD, Barnett GA, editors. Progress in communication science, XII. Norwood, NJ: Ablex; 1993. p. 197–221.
[26] Danowski JA. A network-based content analysis methodology for computer mediated communication: an illustration with a computer board. Commun Yearbk 1982;6:904–25.
[27] Monge PR, Eisenberg ER. Emergent communication networks. In: Jablin F.M.P. LL, Roberts KH, Porter LW, editors. Cycles and dynamic patterns in communiation processes. New York: Ablex; 1987.
[28] Jang HY, Barnett GA. Cultural differences in organizational communication: a semantic network analysis. Bull Methodol Sociol 1994;44:31–59.
[29] Doerfel ML, Barnett GA. The use of CATPAC for textual analysis. Cult Anthropol Methods 1996;8:4–7.
[30] Wasserman S, Faust K. Social network analysis: methods and applications. New York: Cambridge University Press; 1994.
[31] Knoke D, Yang S. Social network analysis. Thousand Oaks, CA: Sage; 2008.
[32] Monge PR, Contractor NS. Theories of communication networks. New York: Oxford University Press; 2003.
[33] Freeman LC. Set of measures of centrality based on betweenness. Sociometry 1977;40:35–41.
[34] Bonacich P. Power and centrality. A family of measures. Am J Sociol 1987;92:1170–82.
[35] Borgatti SP, Everett MG, Freeman LC. UCINET for Windows: software for social network analysis. Harvard, MA: Analytic Technologies; 2002.
[36] Bastian M, Heyman S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. In: International AAAI Conference on Weblogs and Social Media. 2009.
[37] Newman MEJ. Modularity and community structure in networks. Proc Natl Acad Sci USA 2006;103:8577–82.
[38] Aldenderfer MS, Blashfield RK. Cluster analysis. Beverly Hills, CA: Sage University Press; 1984.
[39] Fruchterman TJ, Reingold EM. Graph drawing by force-directed placement. Softw Pract Exp 1991;21:1129–64.
[40] Betsch C, Brewer NT, Brocard P, Davies P, Gaissmaier W, Haase N, et al. Opportunities and challenges of Web 2.0 for vaccination decisions. Vaccine 2012;30:3727–33.
[41] Ache KA, Wallace LS. Human papillomavirus vaccination coverage on YouTube. Am J Prev Med 2008;35:389–92.
[42] Gerend MA, Shepherd JE. Using message framing to promote acceptance of the human papillomavirus vaccine. Health Psychol 2007;26:745.
[43] Reyna VF. Risk perception and communication in vaccination decisions: a fuzzy-trace theory approach. Vaccine 2012;30:3790–7.
[44] Chen RT. Vaccine risks: real, perceived and unknown. Vaccine 1999;17(Suppl. 3):S41–6.
[45] Wolfe RM, Sharp LK, Lipsky MS. Content and design attributes of antivaccination web sites. J Am Med Assoc 2002;287:3245–8.
[46] McRee A, Reiter PL, Brewer NT. Parents' Internet use for information about HPV vaccine. Vaccine 2012;30:3757–62.