Using Propensity Score Matching to Understand the Relationship Between Online Health Information Sources and Vaccination Sentiment

Nabeel Abdur Rehman, Jason Liu, and Rumi Chunara
Tandon School of Engineering and College of Global Public Health
New York University
{nabeel, jliu, rumi.chunara}.@nyu.edu

Abstract

There has been a strong rise of anti-vaccine sentiment in recent years. Human Papilloma Virus (HPV) vaccine has been particularly affected, and the effect of information campaigns from public health institutions, which is a modifiable intervention, is important. However, simply comparing sentiment and information sources via correlation is inadequate because there are underlying factors that could be connected to vaccination rates. Here we study the connection between HPV vaccination sentiment and information from "vetted" public health sources in a causal manner using propensity score matching. We control for factors known to be related to vaccination rates in the offline world using online metrics of language, location and network size. We find that following selected vetted source leads to a decrease in polarized sentiment on HPV vaccination. Further, we find that following a vetted source on Twitter results in a larger decrease in negative sentiment compared to positive sentiment.

The interplay of social, behavioral and health factors is particularly remarkable in infectious disease dynamics and vaccine control. Socio-political factors such as the peers or media campaigns can influence participation in a critical public health matter like vaccination even if it is scientifically validated. In particular, Human papillomavirus (HPV) vaccination coverage is lower than average vaccination rates and intent to vaccinate has actually decreased in recent years (Darden et al. 2013), despite HPV being one of the most commonly diagnosed sexually transmitted infections worldwide, and its association with multiple cancers in men and women. As well, HPV vaccination has been subject to rumors, which can propagate through the media, although to-date the effect of these influences on different groups has not been understood. Current data sources for understanding vaccine perception and behavior are limited in how well they can capture these nuanced effects; survey methods employed are labor-intensive, low-frequency, suffer from recall and information bias, and are limited in sample size.

Simultaneously, online social media has been used to assess health related events and behavior from influenza (Lamb et al. 2013) to obesity (Chunara et al. 2013) and depression (De Choudhury et al. 2013). Further, online sentiment has been shown to correspond to vaccination rates (Salathé et al. 2011), making sentiment drivers of interest. In particular in the online world sentiment can be assessed and sourcing of information can be measured; making it a constructive venue to further investigate the connection between public health information sources and vaccination sentiment.

Research using online social media to-date has incorporated multiple analytical techniques to demonstrate that these data have the potential to be used in public health in an anonymous and secure manner. To-date these studies have covered important areas including detecting trends early, monitoring the course of an epidemic and prediction of outcomes. As far as we can tell, there have been no studies using the data for understanding relation between treatments and health-related outcomes in a *causal* manner. Indeed there is opportunity to use the data systematically to quantitatively understand health behaviors and their drivers. This makes data from online social media a prime case study for understanding characteristics that relate to vaccination sentiment in a causal manner. Here we present a first implementation of an observational study using social media data to understand the relationship between online information sources and vaccine sentiment.

Specific contributions of this work are:

- First application of propensity score matching on social media data (to evaluate relationship between an online treatment and covariates)
- Assessment of the effect of vetted public health information on HPV vaccination sentiment

Inference Method

While there are many aspects of the relationship between information and sentiment that should be studied, as a first step in this regard we study the effect of information as a cross-sectional "treatment", over a relevant time-period for HPV vaccination in the United States. This motivated the use propensity scores, which we generate over the selected time period for each included user. Future studies that examine the time course of sentiment following such interventions may harness other causal inference methods for considering effects on sentiment over time.

Traditional observational data in public health, including reports of similar medical cases, observations of patients compared to healthy controls, cross-sectional epidemiological studies and cohorts followed over time prospectively or historically, have been used to derive mechanistic effects using propensity score methods. Matching participants using propensity scores is an approach that can create quasi-experimental data from otherwise non-experimental data. The method of propensity score is a statistical method; assuming a binary action (or treatment) X, and an arbitrary set S of measured covariates, the propensity score $R(s)$ is the probability that action $X = 1$ will be chosen by a participant with characteristics $S = s$, For example, articulating if certain demographics have higher propensity for behaviors or outcomes. As compared to other causal inference methods, it is simple to determine if propensity have been adequately specified and goodness of fit measures are well-defined (Lamb et al. 2013). Once propensity scores are calculated, there are various methodologies for inference, however several studies have demonstrated that matching eliminates a greater proportion of the systematic differences in baseline characteristics between treated and untreated groups than other propensity score methods and may be less sensitive to whether the propensity score has been accurately estimated.

Data, Treatment Variables and Covariates

Data for our study was initially selected from a 10% "gardenhose" access to the Twitter API between Sept. 1 2011 and Jan. 31 2014. We selected data starting Sept. 2011 due to key HPV-related events in the U.S. in that year. These events included: U.S. Presidential candidate Michele Bachman's claim that HPV vaccination might cause 'mental retardation' (Gostin 2011), CDC and American Academy of Pediatrics (AAP)'s recommendation of HPV vaccination for boys (Brady et al. 2012) and Advancement of HPV mandate repeal by the Virginia House (Washington Post 2012). Tweets containing English, Spanish or undisclosed in their language tag were separated for inclusion in our analysis as they made up 96% of the initial dataset. Based on analysis of Twitter posts and Internet vernacular, we carefully selected features (unigram and bigram) that would enable us to find Twitter posts that relate to HPV vaccination while minimizing the number of non-relevant Tweets; Gardasil, cervarix, hpv vaccin*, hpv shot, vph vac*. Based on these criteria, we were left with 108,325 Tweets.

Selection of covariates is important in propensity matching and there are a multitude of factors that can relate to the outcome and should be included, while data at hand can limit what variables are included. In order to assess relevant factors, is important to consider what is relevant to the relation being tested. To reiterate, our study examines the effect of online public health information sources on *online* vaccine sentiment. We selected covariates from the online Twitter profiles of people Tweeting about HPV that relate to parameters that are known in the offline world to correspond to HPV vaccine uptake thus are precisely relevant (Wei et al. 2013, Chow et al. 2003). Covariates that were available from our data and relevant to the above-described relationship to be tested include: location (via time zone of the account), language (English, Spanish, undisclosed), network size (number of followers and number of friends). The focus on online information and sentiment means that even though we do not know about the offline behavior of the people Tweeting, it is not relevant to our study.

Treatment was defined as "following" a vetted public health source. We chose a set of large public health institutions in the U.S, each known to generate peer-reviewed research and issue policy recommendations. As well, each chosen source has a substantial Twitter presence. The sources (Twitter accounts) chosen were: the American Public Health Association (@PublicHealth), the CDC (@CDCgov), the American Academy of Pediatrics (@AmerAcadPeds), CDC Global Health (@CDCGlobal), CDC Emergency (@CDCemergency), CDC eHealth (@CDC_eHealth) and the Harvard School of Public Health (@HarvardChanSPH).

Sentiment Classification

For classification of Tweet sentiment, we implemented a supervised approach using a multiclass SVM classifier, and a bag of words feature set. The Support Vector Machine (SVM) was chosen as the main classifier for evaluation purposes due to its known ability in text classification (Joachims 1998) and best performance in benchmark text categorization. To serve as our training set, we labeled a total of 9,104 Tweets manually and via Amazon Mechanical Turk, assigning each Tweet a sentiment of positive, negative or neutral. Definitions for sentiment were developed in line with existing work regarding sentiment about vaccination on Twitter (Salathé et al.

2011). A Tweet that is positive about HPV vaccination was defined as one in which the author is likely to get or recommend the HPV vaccine. A negative sentiment means the author is unlikely to get the HPV vaccine. Neutral was assigned if no clear sentiment can be detected. For example, a positive Tweet reads *"Got my HPV shot so I don't get cervical cancer #YOLO"*. A negative Tweet reads *"Gardasil shot 40X more deadly than cervical cancer. http://fb.me/6vAvn005aÂ "*. A neutral Tweet reads *"The HPV vaccine - what do you really know about it? http://t.co/locFw5EO via @HealthRanger"*. And an irrelevant Tweet reads "*my mom just interrupted me doing my homework to tell me that her friends son has the hpv virus THANKS MOM"*. Based on time constraints we only assigned one label to each Tweet; future work should include multiple labels and find consensus on labels. A group of SVM classifiers based on radial basis function and polynomial kernel were trained and the classifier with the highest accuracy was chosen. English stop-words were removed, and trigram features were considered as part of the bag of words strategy. The best performing SVM classifier provided an accuracy of 0.82 with 3-fold validation on the training set and was used to predict the remainder of the data set. Once all of the Tweets were classified by sentiment, a list of distinct twitter users was generated. Users with multiple Tweets with different sentiments were excluded resulting in a total of 46,927 distinct users.

Propensity Analysis

The objective of our study was to identify, under controlled conditions, and for those who have Tweeted about HPV, if subscription to public health sources affects sentiment towards HPV vaccine. We considered covariates describing the user's online: network size (friends; number of accounts following and followers; number of accounts followed), language, and location. To generate covariates representing a user's online network size two mutually exclusive features, namely "popular" and non-popular based on the 'followers' count of users were introduced using a cut-off value of 7,800: based on the weighted mean number of followers. Similarly, variables for 'friends' were also introduced with a cut-off value of 2,000 (Figure 1). For location, 6 mutually exclusive features to represent the location of users were introduced. In order to bin user locations we used the time zone field. 4 of these 6 features represented set of users present in Eastern Daylight Time, Central Daylight Time, Mountain Daylight Time, Pacific Daylight Time, while the remaining two were assigned to users who were outside these time zones and who had not disclosed their time zone. Finally, language was categorized as English, Spanish and undisclosed. Language categories were selected based on language of keywords used to filter data and known underlying factors that have been studied relate to HPV vaccination rates (Brewer et al. 2007). Following any of the selected vetted public health sources was the binary treatment variable. For matching, we used a nearest neighbor method due to the number of covariates wherein a logit method was used to calculate distance between elements in treatment and control group. The R matchit package was used for analysis, and all analyses were performed in R version 3.2.2.

$$Subscription_{Public\ Health} = L.english + L.Spanish + L.Undisclosed + Fl.Popular + Fl.nPolpular + Fr.Popular + Fr.nPolpular + T.none + T.EDT + T.CDT + T.MDT + T.PDT + T.other \qquad (1)$$

The propensity score was calculated using Equation 1, where all variables are binary; and *L, Fl, Fr, T* represent features related to language, followers, friends and time zone respectively. After propensity scores for each user were calculated, a t-test was performed to identify difference in mean of treatment and control groups before and after matching.

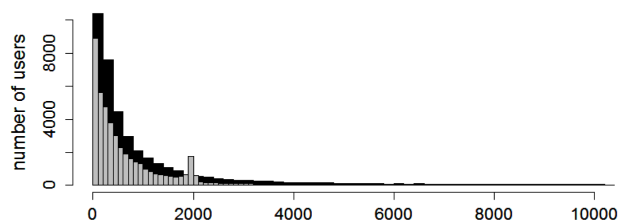

Figure 1: Frequency distribution of followers count (black) and friends count (grey) of users

Results

All covariates were well represented across positive, negative and neutral sentiment Tweets before and after matching (Table 1). Subscription to Public health sources causes a decrease in both negative and positive sentiment while reinforcing the neutral sentiment. Moreover, the decrease is negative sentiment (shown as an inverse in unmatched data), is greater than the decrease in positive sentiment after matching. Finally, the difference in means of the treatment and control groups after matching increases for the neutral sentiment group while it decreases for the positive and negative sentiment groups (Table 2).

Discussion

108,325 Tweets related to HPV vaccination were classified as conveying a positive, negative or neutral sentiment about the HPV vaccine. A propensity score was then generated for the resulting 46,927 distinct twitter users and their sentiment towards the vaccine, incorporating

		Number of Tweets (before matching)			Number of Tweets (after matching)		
		pos	neg	neu	pos	neg	neu
Language	EN	11474	7211	18472	3202	1898	6187
	U	2725	1492	3442	512	300	1012
	ES	938	112	1036	118	20	160
Time Zone	EDT	2866	1888	4480	993	617	1986
	CDT	2997	1734	4050	835	497	1585
	MDT	457	281	761	129	76	243
	PDT	1702	995	2493	469	284	860
	none	3614	2063	5345	657	369	1186
	others	3528	1854	5821	749	375	1517
Followers	high	652	495	1736	249	203	760
	low	14485	8320	21214	3583	2015	6599
Friends	high	908	1101	2762	374	426	1141
	low	14229	7714	20188	3458	1729	6218

Table 1: Breakdown of (pos)itive, (neg)ative) and (neu)tral sentiment users across all covariates before and after matching

	Unmatched	Matched
Positive Sentiment	-0.0825*	-0.0676*
Negative Sentiment	-0.0756*	-0.0846*
Neutral	0.1583*	0.1525*

*Table 2: Difference in mean between treatment and control group before and after matching. *p-value<0.001*

covariates of language, location and network size. We then tested if subscription to select public health vetted sources had an effect on the sentiment of users their sentiment towards the vaccine, incorporating covariates of language, location and network size. Results show that subscription to Public Health sources reduces the polar sentiment about the vaccine in users while reinforcing a neutral stance towards vaccine. The magnitude and direction change of the difference in means for matched samples is informative in comparison to a sole correlation analysis.

Our study adds to the existing literature first through demonstrating a novel methodology of causal inference on social media data. Given the plethora of new, observational data sources of which social media are just one type, we anticipate that propensity matching amongst other techniques will be useful to articulate relationships in a causal manner. Second, our study also for the first time examines a causal relationship between online information from public health institutions and sentiment, and the centering of sentiment based on public health information input. Results in this regard are in line with other studies of online sentiment and exposure (Salathé et al. 2013).

As with any study there are many limitations that must be considered when examining the results of the work. First, when considering exposure to public health sources, there could be other methods for ensuring that not only is a vetted public health source followed, but something from that source regarding HPV vaccination was read (i.e. the user Re-Tweeted content in this regard). However, just having the source in a network could indicate broad awareness, so was a relevant treatment to examine in this study. Another limitation of any propensity matching study is in regards to unmeasured covariates. The bias-reducing potential of propensity scores depends critically on the specific choice of S or, more accurately, on the cause–effect relationships among variables inside and outside S. Future work should include more covariates from the online environment and potentially use inference to include those of the offline world (e.g. age, gender).

Overall, human generated content provides opportunity to understand many aspects of our life and in particular health that are not measured through traditional survey data methods. The study here, albeit a proof of concept method, findings are confirmed by other studies of related topics, and has potential for application in many other studies.

Acknowledgements

We thank C. Corley for providing data used in this study.

References

Brady, M.T., et al. 2012. "HPV vaccine recommendations." *Pediatrics* 129 (3):602-605.

Brewer, N.T., et al. 2007. "Predictors of HPV vaccine acceptability: a theory-informed, systematic review." *Preventive medicine* 45 (2):107-114.

Chow, J.C.-C., et al. 2003. "Racial/ethnic disparities in the use of mental health services in poverty areas." *American Journal of Public Health* 93 (5):792-797.

Chunara, R., et al. 2013. "Assessing the Online Social Environment for Surveillance of Obesity Prevalence." *PLOS ONE* 8 (4):e61373.

Darden, P.M., et al. 2013. "Reasons for Not Vaccinating Adolescents: National Immunization Survey of Teens, 2008–2010." *Pediatrics* 131 (4):645-651.

De Choudhury, M., et al. 2013. "Predicting Depression via Social Media." ICWSM.

Gostin, L.O. 2011. "Mandatory HPV vaccination and political debate." *Jama* 306 (15):1699-1700.

Joachims, T. 1998. *Text categorization with support vector machines: Learning with many relevant features*: Springer.

Lamb, A., et al. 2013. "Separating Fact from Fear: Tracking Flu Infections on Twitter." HLT-NAACL.

Salathé, M., et al. 2011. "Assessing vaccination sentiments with online social media: implications for infectious disease dynamics and control." *PLoS computational biology* 7 (10):e1002199.

Salathé, M., et al. 2013. "The dynamics of health behavior sentiments on a large online social network." *EPJ Data Science* 2 (1):1-12.

Washington Post. 2012. "House passes bill to repeal HPV vaccine for girls." Accessed Oct. 8, 2015. http://goo.gl/fEIB7o.

Wei, F., et al. 2013. "Geographic variability in human papillomavirus vaccination among US young women." *American journal of preventive medicine* 44 (2):154-157.