EPJ Data Science
a SpringerOpen Journal

REGULAR ARTICLE

Open Access

The dynamics of health behavior sentiments on a large online social network

Marcel Salathé[1,2,3]*, Duy Q Vu[4], Shashank Khandelwal[1,2] and David R Hunter[1,4]

*Correspondence: salathe@psu.edu
[1]Center for Infectious Disease Dynamics, Penn State University, University Park, PA, USA
[2]Department of Biology, Penn State University, University Park, PA, USA
Full list of author information is available at the end of the article

Abstract

Modifiable health behaviors, a leading cause of illness and death in many countries, are often driven by individual beliefs and sentiments about health and disease. Individual behaviors affecting health outcomes are increasingly modulated by social networks, for example through the associations of like-minded individuals - homophily - or through peer influence effects. Using a statistical approach to measure the individual temporal effects of a large number of variables pertaining to social network statistics, we investigate the spread of a health sentiment towards a new vaccine on Twitter, a large online social network. We find that the effects of neighborhood size and exposure intensity are qualitatively very different depending on the type of sentiment. Generally, we find that larger numbers of opinionated neighbors inhibit the expression of sentiments. We also find that exposure to negative sentiment is contagious - by which we merely mean predictive of future negative sentiment expression - while exposure to positive sentiments is generally not. In fact, exposure to positive sentiments can even predict increased negative sentiment expression. Our results suggest that the effects of peer influence and social contagion on the dynamics of behavioral spread on social networks are strongly content-dependent.
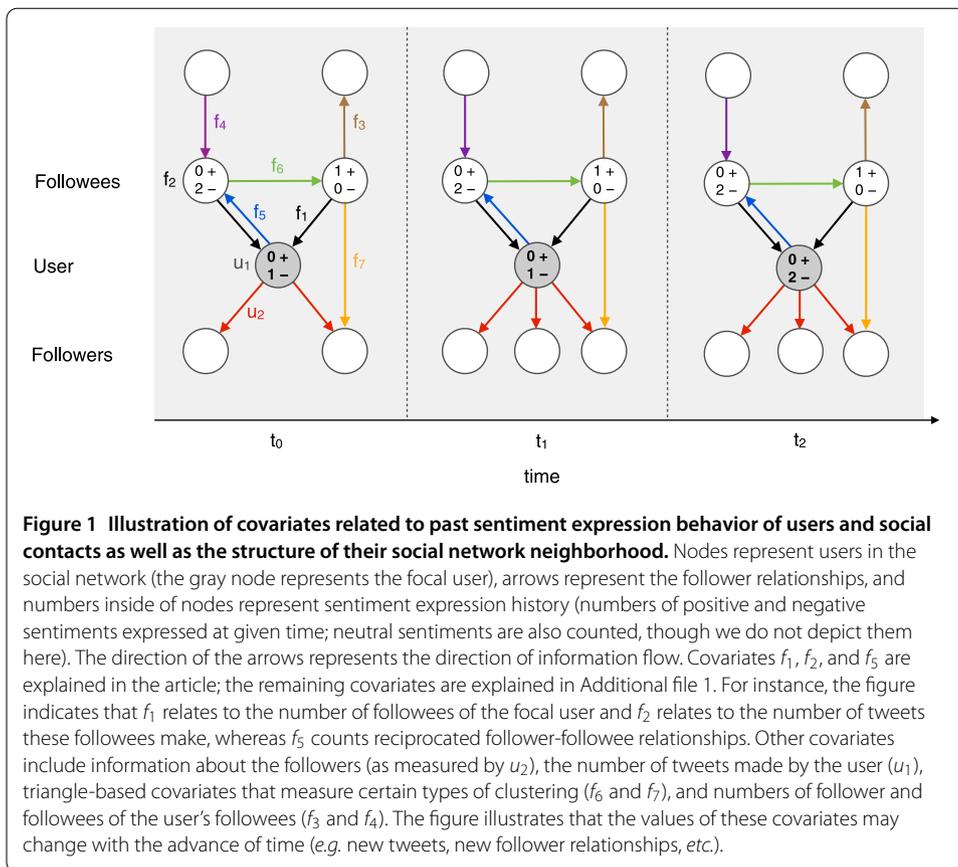
Keywords: social media; social network; diffusion; health behavior; contagion

Social networks play an important role in affecting the dynamics of health behaviors and the associated diseases [1–3], but identifying the main drivers of health behavior spread in social networks has been challenging. The observation that health behavior dynamics follow the patterns of social contacts - *e.g.* that behaviors are often clustered [4, 5] and positively assorted at the dyadic level [6, 7] - can be explained by multiple processes, the two most prominent being homophily and social influence. The homophily hypothesis posits that social contacts are a product of likemindedness, whereas the social influence hypothesis posits that likemindedness is a product of social contacts. Measuring and distinguishing between the effects of homophily and social influence can be difficult in observational studies [6, 8, 9], but is important for the development of health behavior intervention strategies. Vaccination behavior is a prime example of a health behavior shaping disease dynamics: outbreaks of vaccine preventable disease are more likely if overall vaccination rates decline [10], or if vaccination refusal is clustered in local communities [11, 12]. The continuously evolving public concern about vaccines despite the overwhelming scientific evidence on the safety of vaccines reflect the need for an increased understanding on how such sentiments spread over time [13].

Springer

Studying the dynamics of health behaviors on social networks can also be resource-intensive because social network data must often be inferred indirectly, and many health behaviors are complex and thus difficult to quantify. In recent years, online social media services have emerged as novel data sources where short messages are publicly shared, allowing for a detailed picture of the flow of information from person to person in large-scale networks. We have conducted a study to investigate the temporal dynamics of a readily quantifiable health sentiment - the intent to get vaccinated against a novel pandemic virus - on an online social network involving more than 100,000 people, and more than 4 million directed relationships among them. The health sentiment dynamics captured on this network are given by time-stamped messages published by the online social network users, retrospectively classified as expressing positive, neutral or negative sentiments about the intent to get immunized with pandemic influenza H1N1 vaccine [7]. Although not directly measuring the health behavior, the data were shown to explain a large fraction of the spatial variance in CDC-estimated influenza A H1N1 vaccination rates. Insofar as the dynamics of these sentiment have shaped the dynamics of the health behavior, we are interested in the factors affecting the spread of health sentiments in the social network.

The data were collected from the online social networking service Twitter (www.twitter.com), where users post short messages (so-called 'tweets') of up to 140 characters that are then broadcast to their followers. Follower relationships are directional - if user A chooses to follow user B, user A will receive messages from user B, but user B will not receive messages from user A. In this case, we call user A a follower of user B, and user B a followee of user A (although followees are sometimes referred to as 'friends' in the media, we prefer the term followee because it more clearly conveys the direction of the relationship). Nevertheless, user B may also choose to follow user A, in which case a bidirectional relationship is established, and both users will receive messages from each other. An application programming interface (API) provided by Twitter allows for the collection of tweets matching a given set of requirements (*e.g.*, containing a keyword), as well as the collection of follower and followee relationships among users. After data collection, machine learning algorithms were employed to label tweets as negative, positive or neutral with respect to the intent of getting vaccinated against influenza H1N1. Of the 477,768 collected tweets, 318,379 were classified as relevant to the influenza A(H1N1) vaccine. Of those, 255,828 were classified as neutral, 26,667 as negative, and 35,884 as positive. As our data collection efforts were whitelisted by Twitter (a practice that Twitter has now officially discontinued), we are confident that our data set represents the entirety of relevant content. We used an ensemble method combining a naive Bayes and a maximum entropy classifier with an accuracy of 84.29%. The full methodology is described in Salathé and Khandelwal 2011 [7].

In order to identify significant contributors to the likelihood that a user in the social network will express an opinionated (*i.e.*, positive or negative) sentiment in the future, we use an approach that estimates the individual effects of numerous covariates related to the past sentiment expression behavior of users and social contacts as well as the structure of their social network neighborhood (Figure 1). We associate two counting processes, $N_i^+(t)$ and $N_i^-(t)$, with each user i to count the number of positive and negative messages that the user has sent by time t [14]. This results in multivariate counting processes $N^+(t) = [N_i^+(t), \ldots, N_n^+(t)]$ and $N^-(t) = [N_i^-(t), \ldots, N_n^-(t)]$, where n is the number of users in the network. By a mathematical result called the Doob-Meyer theorem [14], each of these

Figure 1 Illustration of covariates related to past sentiment expression behavior of users and social contacts as well as the structure of their social network neighborhood. Nodes represent users in the social network (the gray node represents the focal user), arrows represent the follower relationships, and numbers inside of nodes represent sentiment expression history (numbers of positive and negative sentiments expressed at given time; neutral sentiments are also counted, though we do not depict them here). The direction of the arrows represents the direction of information flow. Covariates f_1, f_2, and f_5 are explained in the article; the remaining covariates are explained in Additional file 1. For instance, the figure indicates that f_1 relates to the number of followees of the focal user and f_2 relates to the number of tweets these followees make, whereas f_5 counts reciprocated follower-followee relationships. Other covariates include information about the followers (as measured by u_2), the number of tweets made by the user (u_1), triangle-based covariates that measure certain types of clustering (f_6 and f_7), and numbers of follower and followees of the user's followees (f_3 and f_4). The figure illustrates that the values of these covariates may change with the advance of time (*e.g.* new tweets, new follower relationships, *etc.*).

(random) counting processes can be decomposed into an integrated conditional intensity process (the signal) and a random process called a martingale (the noise). We denote the conditional intensity functions for positive and negative tweeting events by user i as $\lambda^+(i, t|\beta^+, H_{t-})$ and $\lambda^-(i, t|\beta^-, H_{t-})$, where H_{t-} is the network right before time t, and β^+ and β^- are vectors of parameters.

Specifically, our models for the intensity functions $\lambda^+(i, t|\beta^+, H_{t-})$ and $\lambda^-(i, t|\beta^-, H_{t-})$ are Cox proportional hazards models [15], taking the form

$$\lambda^+\left(i, t|\beta^+, H_{t-}\right) = \lambda_0^+(t) \exp\left[\beta^+ \cdot s(i, H_{t-})\right] \tag{1}$$
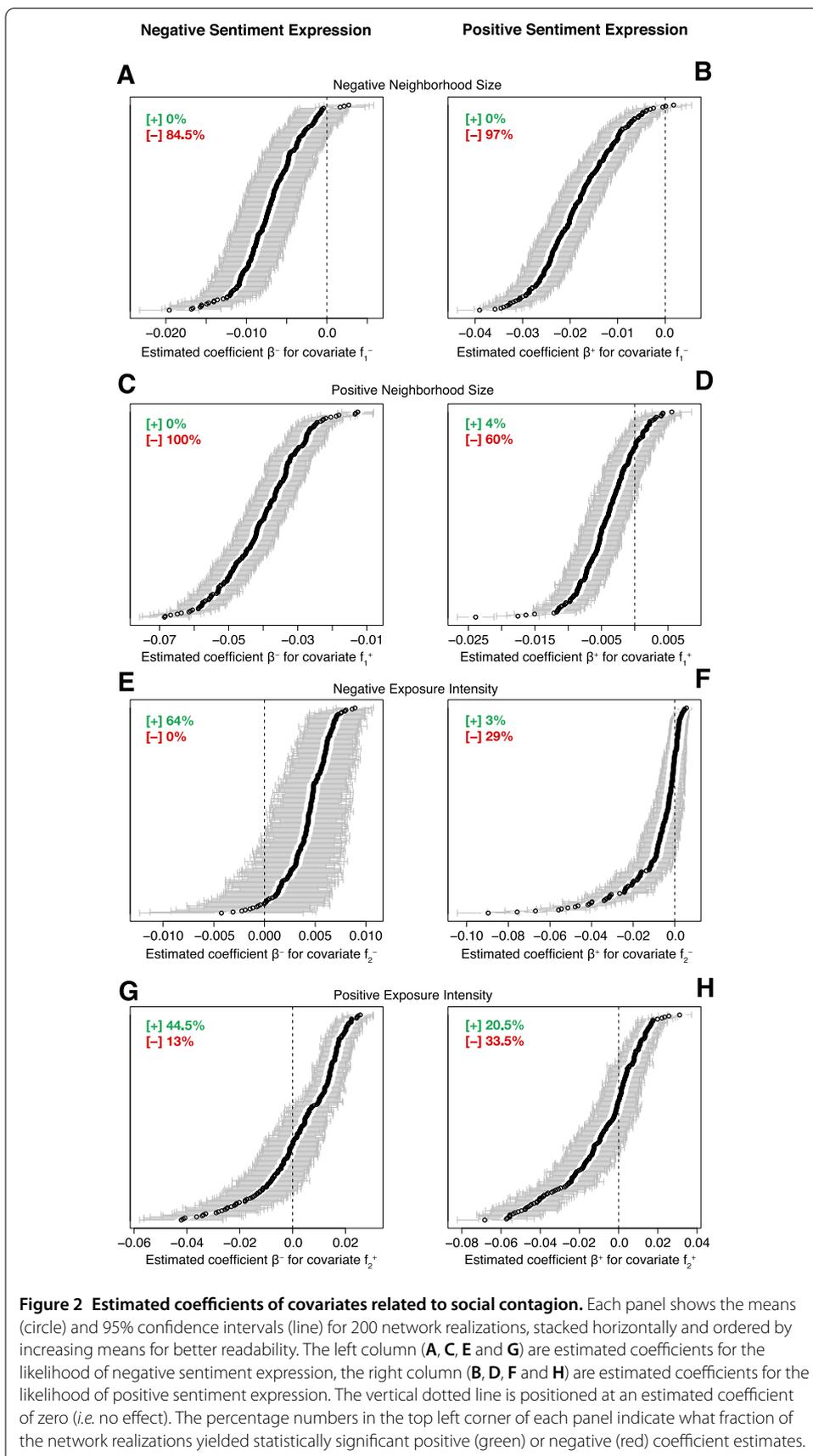
(similarly for λ^-). Here, $s(i, H_{t-})$ is a vector of model-specific covariates, such as node degree and other network statistics deemed appropriate for explaining the intensity of events, which may depend on both the particular node i and the network history H_{t-} up to time t. In our model, each of the network covariates is multiplied by a corresponding element of one of the beta vectors, much like covariates in a regression model are multiplied by regression coefficients. Hence, the statistical significance of the estimated beta coefficients and their signs tell us how the corresponding covariates predict sentiment expression after correcting for all other covariate effects. Notably, the covariate vectors are not constant in time; this fact, in addition to the multivariate counting process response, distinguishes our approach from that of other studies of Twitter data such as that of Golder and Macy [16], who model multivariate continuous (not counting process) responses as functions of fixed predictor variables. Our choice of the Cox proportional hazards model
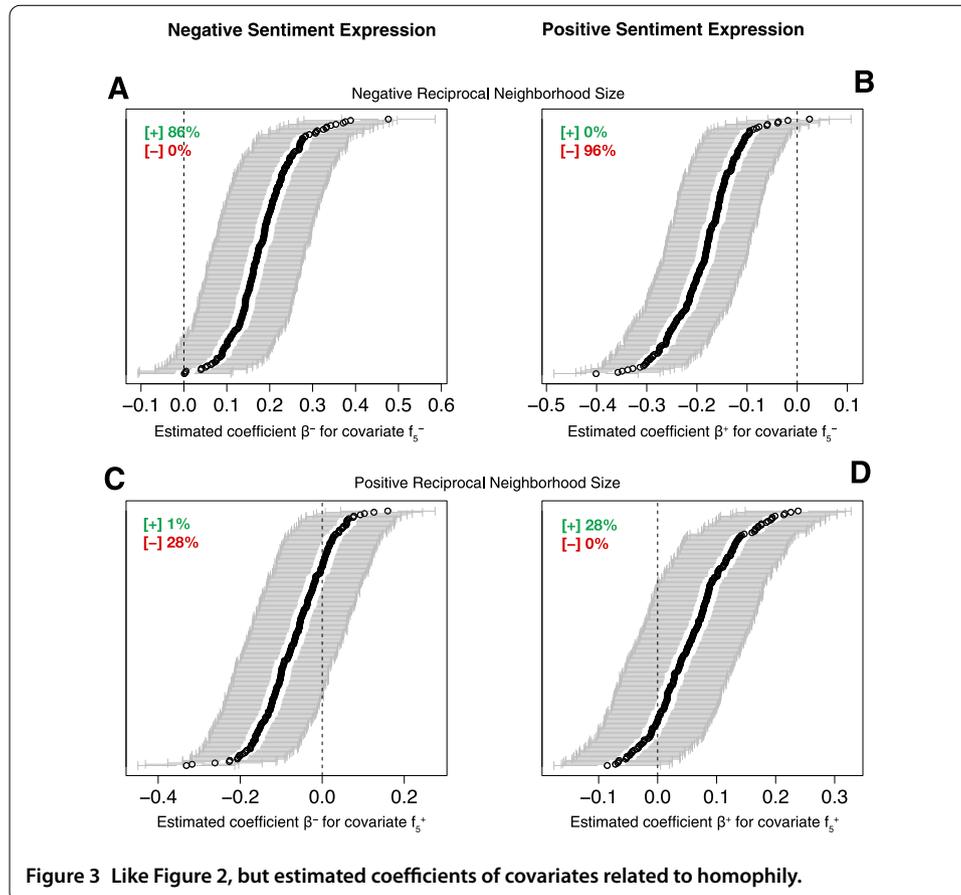
in equation (1) is largely due to the wide use of this model not only in the case of independently sampled survival-time data for which it was originally developed but, more recently, in the counting process context where observations are not necessarily independent. This choice entails an assumption that the coefficients do not change over time and that the covariates influence the intensity function multiplicatively; alternatives such as the Aalen additive model, discussed below, use different assumptions.

We use exactly the same covariates in both models even though the coefficient vectors are different. The network covariates as summarized in Figure 1 capture a number of important aspects of network history H_{t-} thought to be relevant for the dynamics of sentiment expression. A detailed description of all the covariates, along with a full list of the corresponding coefficient estimates and their p-values, is given in Additional file 1. Although we do not discuss them in the current paper, alternative methods for modeling $\lambda^+(i, t)$ exist. For instance, Vu *et al.* [17] discuss the so-called Aalen additive model for a similar situation, in which the effects of the covariates $s(i, H_{t-})$ are additive, rather than multiplicative, and the coefficients β^+ and β^- may be assumed to change over time.

The coefficient vector β^+ in model given by equation (1), along with the vector β^- corresponding to the analogous model for negative tweeting intensity, is estimated using maximum partial likelihood. This is standard practice for Cox proportional hazards models, whose partial likelihood functions do not suffer from multi-modality due to the fact that the log-partial-likelihood is concave [18] with maximizers known to have desirable statistical properties [18, 19]. However, the computations are difficult in the present case because of the size of the dataset. Thus, we employ the computational innovations for caching the time series of network statistic updates outlined in Section 3.2 of Vu *et al.* [20]. Using standard statistical theory for the counting process approach to the Cox model [14, 19], we may also obtain confidence intervals for each coefficient. These confidence intervals do not take into account the error introduced by the possible misclassification of the sentiment expressed in each tweet by the automatic classifier we employ. Therefore, we do not base our statistical inferences on the single set of confidence intervals, but instead employ a series of random reclassifications of each tweet (the four categories being positive, negative, neutral, or unrelated to vaccination), based on a smaller set of test tweets used for calibration and using a method we detail in Additional file 1. In all, 200 different random reclassifications of every tweet are employed, and each such reclassification leads to a new realization of the network to which we apply our statistical estimation method. The resulting profile of 200 95% confidence intervals for every individual coefficient allows us to examine, in aggregate, the direction of each covariate's effect as well as its robustness against the misclassifications inherent in the automatic classification process. Examples of these sets of confidence intervals are presented in Figures 2 and 3 (with more given in Additional file 1).

We base our estimates on only the final 45 days of the data collection time period in order to ensure that they are based on a maximally accurate network representation. Network relationships could only be captured once a user had been identified as messaging about H1N1 vaccination, so cumulative network information improves toward the end of the data collection period. In particular, we have had to make the simplifying assumption that all users are in the network for the entire period on which estimates are to be based - as the exact time a user begins to follow other users' tweets is not observed - and we found this assumption to be suspect beyond 45 days from the end of the data collection period.

Figure 2 Estimated coefficients of covariates related to social contagion. Each panel shows the means (circle) and 95% confidence intervals (line) for 200 network realizations, stacked horizontally and ordered by increasing means for better readability. The left column (**A**, **C**, **E** and **G**) are estimated coefficients for the likelihood of negative sentiment expression, the right column (**B**, **D**, **F** and **H**) are estimated coefficients for the likelihood of positive sentiment expression. The vertical dotted line is positioned at an estimated coefficient of zero (*i.e.* no effect). The percentage numbers in the top left corner of each panel indicate what fraction of the network realizations yielded statistically significant positive (green) or negative (red) coefficient estimates.

Negative Sentiment Expression **Positive Sentiment Expression**

A Negative Reciprocal Neighborhood Size **B**

[+] 86%   [+] 0%
[−] 0%   [−] 96%

Estimated coefficient β^- for covariate f_5^-

Estimated coefficient β^+ for covariate f_5^-

C Positive Reciprocal Neighborhood Size **D**

[+] 1%   [+] 28%
[−] 28%   [−] 0%

Estimated coefficient β^- for covariate f_5^+

Estimated coefficient β^+ for covariate f_5^+

Figure 3 Like Figure 2, but estimated coefficients of covariates related to homophily.

For the time period of 45 days, 98,235 out of 101,853 nodes (96.4%), and 4,209,361 out of 4,858,985 edges (86.6%) are already observed. Therefore, our choice of 45 days represents a balance between the desire to minimize bias due to violations of our simplifying assumption and the desire to use as much data as possible to improve the precision of our estimates. We have verified that results based on a 30-day window were not qualitatively different (in terms of sign and statistical significance) than those based on a 45-day window, whereas a 60-day window appears to introduce bias due to violations of the simplifying assumption.

Because our main interest is in assessing the effects of homophily and social contagion on the health sentiment dynamics in the network, we would like to measure the effects of both how many opinionated people a user is connected to, as well as how many opinionated messages a user is exposed to. These two effects are often confounded because on average, the more people a user is connected to, the more messages a user is exposed to. We therefore define covariates that separate these two effects as much as possible. A further important consideration is that users cannot simply be classified as positive or negative in their overall opinions because over the course of time they might have expressed different sentiments in numerous tweets. To address this issue, each followee is weighted by the fraction of opinionated (positive or negative) tweets he or she makes. The following paragraph gives precise definitions of these three positive-sentiment covariates as employed by the vector $s(i, H_{t-})$ of the model given by equation (1). The three corresponding negative-

sentiment covariates are defined similarly. The full set of covariates, of which there are 24 in our full model, is explained in Additional file 1.

In order to measure the extent to which a user is connected to people expressing positive or negative sentiments, we define the opinionated neighborhood size of a user to be the number of followees. The corresponding covariate, $f_1^+(i,t)$ as indicated in Figure 1, is defined as

$$f_1^+(i,t) = \sum_{j \in F(i,t)} \frac{N^+(j,t)}{N^a(j,t)}, \tag{2}$$

where $F(i,t)$ is the set of followees of i at time t and $N^+(j,t)$ and $N^a(j,t)$ are, respectively, the number of positive tweets and the total number of vaccination-related tweets (positive, negative, or neutral, but excluding any tweets not related to H1N1 vaccination) made by j before time t. We take the opinionated reciprocal neighborhood fraction of a user to be the proportion of followees that are reciprocal (*i.e.*, who are also followers), weighted by the positivity fraction. The corresponding covariate, $f_5^+(i,t)$ in Figure 1, is defined as

$$f_5^+(i,t) = \frac{1}{f_1^+(i,t)} \sum_{j \in F(i,t)} \frac{N^+(j,t)}{N^a(j,t)} Y_{ji}(t), \tag{3}$$

where $Y(t)$ is the adjacency matrix of the network at time t and thus $Y_{ji}(t)$ is the indicator that j follows i at time t. Finally, we define the average opinionated exposure intensity to be the weighted number of opinionated tweets by followees, normalized by the sum of the weights (to minimize the confounding with $f_1^+(i,t)$ as mentioned above). The corresponding covariate is

$$f_2^+(i,t) = \frac{1}{f_1^+(i,t)} \sum_{j \in F(i,t)} \frac{N^+(j,t)}{N^a(j,t)} N^+(j,t). \tag{4}$$

We focus our attention on the six coefficients corresponding to the covariates described above, *i.e.*, $f_1^+(i,t), f_2^+(i,t)$, and $f_5^+(i,t)$ and their corresponding negative-sentiment covariates. We do not study the remaining 18 coefficients in the model with the same level of detail, both for the sake of simplicity and because our interest lies primarily in those effects that relate directly to social contagion. However, it is important that the other statistics, all of which are explained in Additional file 1, are included in the model, since this means that the six coefficients we discuss are estimated after accounting for the effects of all of the other statistics. For instance, we account for possible triangle-based clustering effects by including terms for average number of shared followers (of followees) and average number of shared followees (of followees); as we mention below, these terms control for some types of homophily. Readers interested in statistics used in different applications might compare the statistics used in the citation network examples of Vu *et al.* [20] or the social network and email examples of Vu *et al.* [17] and Perry and Wolfe [18].

The results are summarized in Figures 2 and 3, which simultaneously account for two different types of uncertainty. Error due to selecting a random sample of individuals from a hypothetical infinite population of potential Twitter users, as represented by the model, is expressed by the 95% confidence intervals, whereas error due to misclassifying sentiments is captured by the 200 randomly reclassified samples. The percentages in green

and red are therefore the proportion of times we could expect our dataset to result in rejecting the null hypothesis of no effect and concluding that a positive (green) or negative (red) effect exists; we may therefore understand these values as bootstrapped probabilities that our dataset will produce these two statistical results. Generally, larger opinionated neighborhood sizes have an inhibitory effect on the expression of opinionated sentiments (Figure 2A-D): While both larger positive and larger negative neighborhood sizes have the expected inhibitory effect on the expression of the opposite sentiments (*i.e.*, negative and positive, respectively), they also predict diminished expression of that same sentiment. If we look at the opinionated reciprocal neighborhood size (Figure 3), we see that the effects are content-dependent, *i.e.*, the effects are different for negative and positive sentiments. On one hand, larger positive reciprocal neighborhood sizes do not generally have a significant predictive effect on the rate of expressing opinionated sentiments. On the other hand, increasing negative reciprocal neighborhood size has the expected effect of increasing the likelihood of expressing a negative sentiment, and decreasing the likelihood of expressing a positive sentiment. Finally, the predictive effects of opinionated exposure intensity are also content-dependent (Figure 2E-H). While a range of outcomes are observed in the 200 network realizations obtained *via* reclassifying each tweet's sentiment (as explained earlier), there is a sizable fraction of outcomes that show unexpected effects. In particular, in a substantial fraction of cases, being exposed to an increased intensity of positive tweets is predictive of increased intensity of negative sentiment (Figure 2G), as well as decreased intensity of positive sentiment (Figure 2H). Finally, the past expression of a sentiment by an individual predicts an increased propensity for that individual to express that same sentiment again, a finding that is very consistent across all 200 network realizations (see Figure S1 in Additional file 1).

It is worthwhile to consider these results in the context of what the statistics are expected to measure. Our main interest is in identifying the extent to which social contagion and homophily drive sentiment dynamics within the social network. In an observational study like the present study, causality cannot be established. Furthermore, disentangling effects of homophily and contagion is notoriously hard [8] because they are often confounded. Our approach tries to minimize these issues as much as possible. We use the term social contagion to mean the extent to which exposure to a given sentiment is predictive of future expression of that sentiment. Previous studies have focused on binary outcomes such as the adoption (*vs.* non-adoption) of a service [6, 21], and have measured exposure as the number of social contacts that have adopted the service previously. Our methodology allows us to consider more complex measures of exposure: For instance, in the present analysis we measure both the number of social contacts expressing a given sentiment as well as the intensity with which the sentiment is expressed. Thus, both the opinionated neighborhood size as well as the average opinionated exposure intensity relate to social contagion as defined above. Homophily, on the other hand, is assessed by the opinionated reciprocal neighborhood size of a user, *i.e.*, the weighted number of reciprocal followees, or followees who are also followers of that user.

The finding that the opinionated neighborhood size generally has an inhibitory effect on the likelihood of expressing any opinionated sentiment (Figure 2A-D) is difficult to interpret in the context of a standard contagion framework, because contagion is normally associated with spread, rather than inhibition. For example, it makes intuitive sense that a larger number of negative followees should lead to a reduction in the expression of pos-

itive sentiments. The finding that it also leads to a reduction in the expression of negative sentiments is harder to interpret, but nevertheless agrees with the general pattern of inhibition. When looking at the average opinionated exposure intensity (Figure 2E-H), a different picture emerges. The results are rather sensitive to misclassification of the messages, but the most stable result (64% of all network realizations, Figure 2E) is that increased average negative exposure intensity does predict increased negative sentiment expression, in line with the expectation of social contagion. Surprisingly, the second most stable result (44.5% of all network realizations, Figure 2G) is that the average positive exposure intensity does also predict increased negative sentiment expression. Equally surprisingly, the third most stable result (33.5% of network realizations, Figure 2H) is that higher average positive exposure intensity predicts decreased positive sentiment expression. Taken together, the results suggest that exposure to negative sentiment is contagious - by which we merely mean predictive of future negative sentiment expression - while exposure to positive sentiments is generally not. They also suggest that exposure to increased intensity of opinionated sentiments has on balance led to increased negative sentiment expression and decreased positive sentiment expression, overall favoring the spread of negative vaccination sentiments.

The lack of detailed information about the users prohibits us from assessing manifest homophily, and our analysis is thus subject to the problem of latent homophily which is generally confounded with contagion [8]. We assess homophily with the opinionated reciprocal neighborhood size of a user, which is the weighted number of reciprocal followees (*i.e.*, followees who are also followers of that user). Bidirectional follower relationships mean that two users are interested in receiving messages from each other, which we assume to indicate that the users may share similar interests, which in turn suggests homophily. To further reduce the confounding effects of homophily and contagion, our model contains covariates for the number of shared followees and followers. These covariates are expected to control for latent homophily to a certain extent, since homophily is known to manifest itself in network clustering [8, 22]. Our findings suggest that the effects of homophily, insofar as we can measure it, are content-dependent: the positive reciprocal neighborhood size does generally not have significant effects (Figure 3C and D), while increasing negative reciprocal neighborhood size has the expected effects of predicting decreased positive and increased negative sentiment expression (Figure 3A and B). This finding further contributes to favoring the spread of negative vaccination sentiments.

Overall, the finding that the effects of various network covariates are strongly content-dependent suggests that a standard contagion framework might be too constrained to understand the health sentiment dynamics occurring on this network. By standard contagion framework, we mean the conceptual idea that increased exposure to any given agent (whether biological or social) will lead to an increased transmission - and predict an increased adoption - of that agent. In such a framework, the expectation is that there is a positive relationship between exposure and the consequent adoption of whatever it is individuals are exposed to. In our data, the only effect that corresponds to this pattern is that increased negative exposure intensity does predict increased negative sentiment expression. All the other results suggest that increased exposure predicts either a decrease of the same sentiment expression or an increase of the opposite sentiment expression.

From a public health perspective, the results raise some questions about the design of health behavior communication strategies. In particular, the notion that increased posi-

tive exposure intensity predicts increased negative sentiments could be of great concern if this turns out to be a consistent finding in future studies, since it would indicate that the level of positive messaging needs to be assessed carefully. Equally worrisome is the notion that the identified effects overall seem to favor the spread of negative sentiments, but not the spread of positive sentiments. This suggests that increased attention should be given to the prevention and control of negative sentiments (particularly if based on rumors, misinformation, misunderstandings, *etc.*). A recent study [23] has found that the popularity of documents shared on Twitter decreased significantly faster if the documents contained more words related to negative emotion, rather than to positive emotion. In general, the ability to measure the dynamics of sentiments on online networks generates opportunities to dramatically reduce the time lag between communication strategies and the assessment of the effects of those strategies.

The study framework has a number of limitations that need to be taken into account when assessing its applicability. First, our study design has been set up to catch expression of sentiments only (rather than actual vaccination behavior), but users might have been affected by exposure to sentiments from social contacts without ever expressing these sentiments themselves. For example, a user exposed to many negative messages may have been influenced and adopted a negative stance on H1N1 vaccination, but the user might not consequently have expressed that opinion in the network. Thus, a substantial fraction of actual contagion may have gone unnoticed. Conversely, peer pressure effects may have driven users to express a certain sentiment online even though they personally hold a different opinion (and behave differently from what one would expect based on the expressed sentiment), leading to false positives. Future research should address the question to what extent health sentiments expressed online overlap with actual health behaviors. Moreover, our study design ignores the possibility that follower relationships may have been established because users already share the same opinion on vaccination. While it is not unlikely that vaccination sentiments can be a contributor to establishing follower relationships, we believe that overall it had a small effect in the short period of time on which our analysis is based. Finally, the content of short messages like the ones studied here is subjective and open to interpretation by the reader of the message. Given the sometimes strong dependency of the effect on network realizations, this is an important problem that needs to be addressed in the future.

The dynamics of sentiments and behaviors on social networks is of great importance in many fields concerning human affairs [24], and particularly also in the health domain. There is an increased understanding that modifiable health behaviors are a key contributor to health outcomes [25], and that health behavior modification might be a key strategy to control major public health issues, both from the perspective of prevention (vaccination, smoking cessation, diet modification, *etc.*) and treatment (adherence to treatment plans, antibiotic overuse, *etc.*) strategies. The rapid worldwide adoption of online social network services means that an increasing fraction of (mis-)information diffusion is occurring on these networks. The methods and findings presented here are a small step towards an increased understanding of these dynamics, demonstrating both the promise and the challenges associated with these large and often unstructured data sets. In addition to online experiments [26, 27], analysis of large-scale, high-resolution observational data will provide a much better picture of the dynamics of health behavior diffusion on social networks.

Additional material

> **Additional file 1: Supplementary Materials for "The Dynamics of Health Behavior Sentiments on a Large Online Social Network"**

Author details
[1]Center for Infectious Disease Dynamics, Penn State University, University Park, PA, USA. [2]Department of Biology, Penn State University, University Park, PA, USA. [3]Department of Computer Sciences and Engineering, Penn State University, University Park, PA, USA. [4]Department of Statistics, Penn State University, University Park, PA, USA.

References
1. Smith KP, Christakis NA (2008) Social networks and health. Annu Rev Sociol 34:405-429
2. Christakis NA, Fowler JH (2008) The collective dynamics of smoking in a large social network. N Engl J Med 358:2249-2258
3. Valente TW (2010) Social networks and health. Oxford University Press, Oxford
4. Christakis NA, Fowler JH (2007) The spread of obesity in a large social network over 32 years. N Engl J Med 357:370-379
5. Schuit AJ, van Loon AJM, Tijhuis M, Ocké M (2002) Clustering of lifestyle risk factors in a general adult population. Prev Med 35:219-224
6. Aral S, Muchnik L, Sundararajan A (2009) Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. Proc Natl Acad Sci USA 106:21544-21549
7. Salathé M, Khandelwal S (2011) Assessing vaccination sentiments with online social media: implications for infectious disease dynamics and control. PLoS Comput Biol 7:e1002199
8. Shalizi CR, Thomas AC (2011) Homophily and contagion are generically confounded in observational social network studies. Sociol Methods Res 40:211-239
9. An W (2011) In: Scott J, Carrington PJ (eds) The SAGE handbook of social network analysis. Thousand Oaks, Sage, pp 514-532
10. Jansen VAA et al (2003) Measles outbreaks in a population with declining vaccine uptake. Science 301:804
11. Salathé M, Bonhoeffer S (2008) The effect of opinion clustering on disease outbreaks. J R Soc Interface 5:1505-1508
12. Omer SB, Enger KS, Moulton LH (2008) Geographic clustering of nonmedical exemptions to school immunization requirements and associations with geographic clustering of pertussis. Am J Epidemiol 168:1389-1396
13. Omer SB, Salmon DA, Orenstein WA, deHart MP, Halsey N (2009) Vaccine refusal, mandatory immunization, and the risks of vaccine-preventable diseases. N Engl J Med 360:1981-1988
14. Aalen OO, Borgan Ø, Gjessing HK (2008) Survival and event history analysis. Springer, Berlin
15. Cox DR (1972) Regression models and life-tables (with discussion). J R Stat Soc, Ser B, Stat Methodol 34:187-220
16. Golder SA, Macy MW (2011) Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. Science 333:1878-1881
17. Vu DQ, Asuncion AU, Hunter DR, Smyth P (2011) In: Proceedings of the 24th international conference on neural information processing systems (NIPS 2011), vol 24, pp 2492-2500
18. Perry PO, Wolfe PJ (2010) Point process modeling for directed interaction networks. arXiv:1011.1703
19. Andersen PK, Gill RD (1982) Cox's regression model for counting processes: a large sample study. Ann Stat 10:1100-1120
20. Vu DQ, Asuncion AU, Hunter DR, Smyth P (2011) In: Proceedings of 28th international conference on machine learning (ICML 2011), pp 857-864
21. Ugander J, Backstrom L, Marlow C, Kleinberg J (2012) Structural diversity in social contagion. Proc Natl Acad Sci USA 109:5962-5966
22. McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a feather: homophily in social networks. Annu Rev Sociol 27:415-444
23. Wu S, Tan C, Kleinberg J, Macy M (2011) Does bad news go away faster? In: Proc. 5th international AAAI conference on weblogs and social media
24. Rogers EM (2003) Diffusion of innovations, 5th edn. Free Press, New York
25. Mokdad AH, Marks JS, Stroup DF, Gerberding JL (2004) Actual causes of death in the United States, 2000. JAMA J Am Med Assoc 291:1238-1245
26. Centola D (2010) The spread of behavior in an online social network experiment. Science 329:1194-1197
27. Centola D (2011) An experimental study of homophily in the adoption of health behavior. Science 334:1269-1272