Polio vaccine hesitancy in the networks and neighborhoods of Malegaon, India

Jukka-Pekka Onnela [a, *], Bruce E. Landon [b], Anna-Lea Kahn [c], Danish Ahmed [d], Harish Verma [c], A. James O'Malley [e], Sunil Bahl [d], Roland W. Sutter [c], Nicholas A. Christakis [f]

[a] Department of Biostatistics, Harvard T.H. Chan School of Public Health, USA
[b] Department of Health Care Policy, Harvard Medical School, Division of General Medicine and Primary Care, Beth Israel Deaconess Medical Center, USA
[c] World Health Organization, Geneva, Switzerland
[d] National Polio Surveillance Project, WHO, India
[e] The Department of Biomedical Data Science, The Dartmouth Institute for Health Policy and Clinical Practice, Geisel School of Medicine at Dartmouth, USA
[f] Yale Institute for Network Science, USA

ARTICLE INFO

ABSTRACT

Objectives: Eradication and control of childhood diseases through immunization can only work if parents allow their children to be vaccinated. To learn about social network factors associated with polio vaccine hesitancy, we investigated social and spatial clustering of households by their vaccine acceptance status in Malegaon, India, an area known for vaccine refusal and repeated detection of polio cases.
Methods: We interviewed family heads from 2462 households in 25 neighborhoods in July 2012 and constructed social networks based on advice seeking from other households. We restricted our main analyses to surveyed households for which we also had data on whether they accepted the polio vaccine for their eligible children or not.
Results: Data from 2452 households was retained and these households made 2012 nominations to 830 households. Vaccine-refusing households had fewer outgoing ties than vaccine-accepting households. After excluding 24 isolated households, vaccine-refusing households had 189% more nominations to other vaccine-refusing households (93% more in the largest component of the network) compared to vaccine-accepting households, revealing that vaccine-refusing households cluster in the social network. Since roughly half of all ties connect households within neighborhoods, vaccine-refusing clusters lie in spatially localized "pockets".
Conclusions: The social (and spatial) clustering of vaccine-refusing households could be leveraged to tailor communication strategies to improve vaccine acceptance and community perceptions of immunization programs for polio and other vaccine-preventable diseases.

© 2016 Published by Elsevier Ltd.

1. Introduction

The Global Polio Eradication Initiative (GPEI), a partnership between national governments and five core agencies, including the World Health Organization (WHO), Rotary International, the US Centers for Disease Control and Prevention (CDC), the United Nations Children's Fund (UNICEF), and the Bill and Melinda Gates Foundation (BMGF) is striving to eradicate polio worldwide. Polio incidence in the developing world, especially among under-served and hard-to-reach populations, has been reduced through increased rates of vaccination supported by sustained media campaigns and by mobilizing community leaders (Obregón et al., 2009). Since its launch, the number of polio cases has been reduced from an estimated 350,000 in 1988 to 358 in 2014 (Data in WHO HQ as of 27 January 2015). In addition, the number of countries with endemic polio has been reduced from 125 to three, namely Pakistan, Afghanistan and Nigeria. India had its last case in 2011 and the whole South-East Asia region was certified polio-free in 2014.

In spite of these achievements, polio remains an international

* Corresponding author. Harvard T.H. Chan School of Public Health, 655 Huntington Avenue, Room 2-423, Boston, MA 02115, USA.
E-mail address: onnela@hsph.harvard.edu (J.-P. Onnela).

concern. In May 2014, renewed outbreaks led the WHO's International Health Regulations Emergency Committee to declare the situation a Public Health Emergency of International Concern (WHO). In particular, segments of susceptible populations continue to resist having their children vaccinated for a variety of reasons. Rumors and misinformation are likely to propagate through social connections, and they might amplify the rates of vaccine refusal. GPEI supported the present study in order to understand better vaccine-refusing behavior.

That vaccine acceptance might be related to social connections should not be surprising. People are connected, and so their health is connected (Smith and Christakis, 2008). Social networks play an important role in human health and disease, through mechanisms such as provision of social support, social engagement, and access to resources, as well as through the dissemination of information and behaviors (Valente, 2010; Berkman et al., 2000). We would expect the structure of social networks to play a critical role in funneling both information and misinformation, including that related to vaccines (Christakis and Fowler, 2010).

Our goal was to investigate whether the resistance of households to having their young children vaccinated against polio might be related to similar resistance in households to which they are socially connected. If children of unvaccinated households interact frequently, they might reinforce inaccurate beliefs and be substantially more likely to contract and propagate the disease, as herd immunity is reduced within these clusters. Learning about the possible existence of such clusters is therefore critical for eradicating infectious diseases like polio in developing countries by identifying "social pockets" of vaccine hesitancy. Moreover, to the extent that such clusters can be identified, public health officials might be able to use more targeted approaches to overcoming vaccine resistance. We also investigated differences in vaccine acceptance across the surveyed neighborhoods and the geographic distribution of nominations across neighborhoods. This enabled us to distinguish between network and neighborhood effects; in short, does your behavior with respect to polio vaccines depend on whom you know, where you live, or both?

2. Methods

2.1. Study design

We carried out a population-based study of 2462 households in contiguous high-risk (by polio planning definitions) neighborhoods in Malegaon Municipal Corporation in the Nashik district of Maharashtra state in the western region of India, about 280 km northeast of Mumbai. The study protocol was reviewed and approved by Research Ethics Review Committee of WHO Geneva and the Haffekine Research Institute Mumbai.

Twenty-five teams each consisting of a professional surveyor (recruited through Nielsen Private Limited) and a female volunteer student from a local medical school (Mohammadia Tibbia College) collected data through in-person household surveys in all of the 25 neighborhoods. The interviews took place over a three-week period in July 2012, and the targeted respondents were household heads. On completing the paper questionnaires, survey teams handed them over for linking with information on oral polio vaccine (OPV) vaccination status for applicable children in each household and a quality crosscheck. The OPV vaccine status data were available from WHO that had concluded a supplementary polio immunization campaign just prior to our study. As part of this campaign, the organization classified a household as vaccine-accepting if the household, after confirming eligibility, accepted the vaccine for eligible children in the household; a household was classified as vaccine-reluctant if it accepted the vaccine on the second attempt;

and finally a household was classified as vaccine-refusing if it continued to refuse to have their vaccine-eligible children vaccinated even after the second attempt. Once the OPV vaccination status and quality crosscheck had been completed, the questionnaires were transmitted to the market research firm for coding and anonymous entry into the study database.

Using suitable "name generators" (Shakya et al., 2015), household heads were asked to identify up to four contacts with whom they discuss general issues (e.g., business, sports, personal matters, or issues that affect their community) and up to four contacts with whom they discuss health-related issues. More specifically, household heads were asked to nominate households by naming either household heads or other household members in response to the name generator questions. If the nominated person was not a household head, we implicitly re-directed the nomination to the household head of the corresponding household (see Fig. 1). Although the surveyed households were located within a defined geographic region, there were no restrictions on whom the respondents could name. Consequently, many of the identified contacts were outside the surveyed area with unknown vaccination status (as they were not included in the supplementary immunization campaign used to collect vaccine statuses), and these contacts are omitted from most of our analyses below. Such limitations are unavoidable in social network studies (Perkins et al., 2014).

In addition, respondents were asked their name and occupation, the highest level of education attained, and the number of people (both adults and children) living in the household. Finally, socioeconomic status was measured by a series of questions asking about the number of rooms in the house, whether there was a separate toilet in the home, and whether they owned a television, mobile phone, or cooking gas cylinder. See Supplemental Material for survey details.

2.2. Network measures and concepts

We summarize some of the basic network concepts relevant to our study here; the interested reader can find more details in the growing literature on networks (Valente, 2010; Newman, 2010; Borgatti et al., 2013; Wasserman and Faust, 1994). When the responses to name generator questions are represented as a network, *nodes* (or *vertices*) correspond to individuals and *ties* (or *edges*) correspond to nominations between any two individuals. We call this network the *nomination network*. A pair of nodes, connected or not, is called a *dyad*. The person doing the nomination is often called the *ego* and the person(s) nominated the *alter*(s). In this case, the responses of subjects to name generator questions gives rise to a directed network, where, by convention, the direction of the edges is from the person doing the nomination to the person(s) nominated. The number of out-going edges is called that person's *out-degree;* in this case it is bound by the study design to be less than or equal to eight, the maximum number of allowed nominations in the two categories combined. The number of in-coming edges is called that person's *in-degree;* the size of network minus 1 is the natural upper limit on its value given that any person in the study could be nominated by any number of other persons. *Network neighbors* of a node refer to the set of nodes to which any given node is connected by ties where the direction of ties is usually ignored (unless stated otherwise). Nodes with a disproportionately high in-degree (large number of nominations) are called *hubs*. There is no universally accepted definition of what constitutes a hub, and here we settled for a definition that a hub is any node with an in-degree that is four standard deviations above the mean in-degree. A *connected component* refers to a piece of the network that has no ties whatsoever with the rest of the network. Most empirical networks consist of several components, but even then it
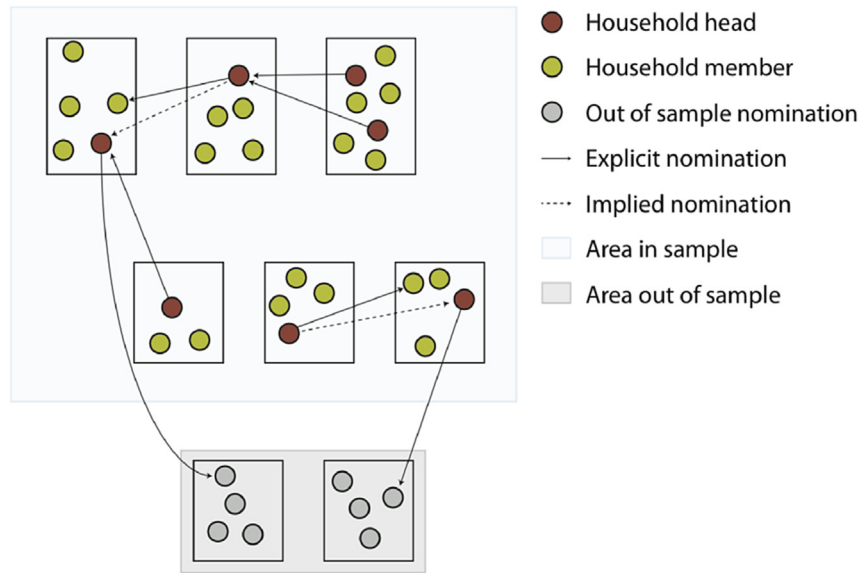
Fig. 1. Schematic of sampling design. The rectangles correspond to households and the circles inside the rectangles to household members. One of the household members is the household head. In this study, household heads were asked to nominate households by naming either household heads or other household members in response to the name generator questions. If the nominated person was not a household head, we implicitly re-directed the nomination to the household head of the corresponding household. This means that all households in the network can be treated the same way; in particular, we do not distinguish between households based on whether the corresponding household head or household member was the target of the nomination.

is typical that one component is much larger than the others. This component is called the *largest connected component* (LCC) and its size is measured as the number of nodes that it contains. Finally, an *induced subgraph* refers to a subset of the network that consists of a specified set of nodes and all edges that fall between these nodes in the underlying network.

2.3. Construction of the polio vaccine network

Connected individuals within a social network tend to resemble each other (McPherson et al., 2001). There are at least three different mechanisms that can produce clustering of individuals by their attributes: influence, selection, and confounding. Briefly, influence refers to a person inducing a behavior in another person, thereby making their attributes similar; selection refers to the tendency of people to seek out others who are like them but implies no influence; and confounding refers to an external factor that has an effect on the attribute(s) of both individuals, typically such that they resemble each other even if influence and selection are absent. Cross-sectional data like ours do not allow for disentangling these potential mechanisms behind the observed similarity of connected individuals, but it is nevertheless important from the perspective of targeted interventions to learn about potential clustering of nodes by their attributes.

The *nomination network*, after excluding 10 households with incomplete data and a further 24 households that neither made nor received any nominations, consisted of 804 components with the LCC containing 6113 of the 11,828 (51.7%) network nodes and 6647 of the 11,655 (57.0%) network ties. All of the other components were substantially smaller; for example, the 2nd largest component had only 59 nodes. All network nodes had at least one tie (incoming or outgoing); 1600 nodes had no incoming ties and 9400 had no outgoing ties. We constructed the *vaccine network* as the subgraph of the nomination network that is induced by the set of nodes with known polio vaccination statuses. The vaccine network had 2428 nodes and 1355 edges, whereas the LCC of the vaccine network, had 710 nodes and 813 edges. See Table 1 for numerical summaries of

Table 1
We report here the number of nodes and the number of edges in the nomination network, in the largest connected component (LCC) of the nomination network, the vaccine network (subset of the nomination network with nodes having known vaccination status), and the LCC of the vaccine network.

Network	Number of nodes	Number of edges
Nomination network	11,828	11,655
Nomination network LCC	6113	6647
Vaccine network	2428	1355
Vaccine network LCC	710	813

the nomination network and the vaccine network.

Although study subjects could nominate anyone in the region, the vaccine status of households was known for only those people in the surveyed neighborhoods. The percentage of nominations into the sampled neighborhoods was 14.2% for vaccine-accepting households, 16.2% for vaccine-reluctant households, 12.7% for vaccine-refusing households, and 15.0% for households with no vaccine-eligible children. These sample proportions were not statistically significantly different from one another ($\chi^2 = 5.11$, $p = 0.16$), which supported the extraction of the vaccine network from the larger underlying nomination network without potentially introducing a bias across households by their vaccine status.

2.4. Statistical analyses: tie prediction

We constructed a simple statistical model to predict the existence of a tie in the vaccine network between two households based on the observed household attributes in order to detect which attributes are most highly predictive of ties. We modeled the binary status of each dyad using logistic regression where differences of nodal attributes across the dyad were used as predictors. In other words, we formed the predictors by considering all node pairs in the network, connected or not, and regressed them on the differences in nodal attributes across all pairs. The observed outcome for each node pair was set to 0 if the nodes were not connected and to 1 if they were connected. We adopted this

modeling strategy due to its simplicity and computational performance, but other modeling approaches, such as more general exponential random graph models (ERGMs) (Lusher et al., 2013) or latent-space models (Hoff et al., 2002) could be considered here, subject to computational limitations.

The nodal attributes we considered were vaccine status, neighborhood, sum and difference of degree (taken as undirected, combining general and health categories), education, and mobile phone ownership (ownership of other items, such as cooking cylinder, were dropped because of their limited predictive power). When analyzing vaccine acceptance, we combined the reluctant and refuser groups and then set the vaccine status indicator to one if the vaccine statuses of the two nodes were identical. The pooling is justified from the substantive perspective in that both reluctant and refusing households exhibit some level of vaccine non-acceptance. For neighborhood, a bivariate indicator predictor was set to one if the two neighborhoods were the same; otherwise it was set to zero.

2.5. Statistical analyses: clustering of households by their polio vaccination status

To investigate the statistical significance of the clustering of households in the LCC of the vaccine network, we developed a simple non-parametric resampling procedure to generate the distribution of the test statistic under the null. We first identified all nodes in a given category (say, accepting) and then examined them one at a time. For each such node, we identified its network neighbors but instead of using their observed vaccine statuses, we sampled those statuses with replacement from the observed distribution of all vaccine statuses in the vaccine network. We applied this procedure to every node in each of the four categories, thus covering all nodes in the vaccine network, and computed the proportion of neighboring accepting, reluctant, refusing, and vaccine-ineligible households for them. We repeated this procedure 10,000 times. The observed value of the test static was obtained by using the actual vaccine status of each node.

3. Results

A total of 2462 household respondents were surveyed, and after excluding 10 respondents due to incomplete data, 2452 (or 99.6%) of the households approached for the interview were included in the study. Of these, 1355 included a child between the ages of 0 and 5 eligible for the vaccine for whom we had data on vaccine status; the ineligible households reported not having children in this age category and were excluded. Among eligible households, 1074 accepted the vaccine, 137 were reluctant, and 144 refused the vaccine. The 2452 respondents identified a total of 10,228 unique alters (some of whom were also egos) who were nominated a total of 13,819 times (some alters being nominated multiple times). Many of the nominations fell outside the sampled neighborhoods: a total of 2012 of the nominations (edges), or 14.6%, were to 830 households located within the sampled neighborhoods, and these sampled neighborhoods contained 2452 of the 11,828, or 20.7%, of the nodes in the network. In constructing the vaccine network, we excluded 24 households that had neither incoming nor outgoing ties, resulting in the vaccine network having 2428 nodes.

Households varied in terms of their possessions and level of education. From less common to more common items: 519 (21.6%) had a cooking cylinder, 710 (29.6%) a toilet, 1080 (45.0%) a TV, and 1752 (72.9%) a mobile phone. Education was coded as a categorical variable, and the categories and the number of people in them were as follows: 1 = no school (459), 2 = primary school (1220), 3 = middle school (369), 4 = high school (173), 5 = intermediate or

post high school diploma (63), 6 = graduate or post graduate (42), 7 = professional or honors (22), 8 = religious (Islamic) education (54). See Table 2 for a summary of household attributes and network nominations.

3.1. Nominations

Fig. 2 shows the estimated mean out-degree and associated standard errors for general nominations, health nominations, and both in the nomination network. Combining general and health nominations, vaccine-accepting households made on average 4.93 nominations, vaccine-reluctant households 4.92 nominations, and vaccine-refusing households 4.41 nominations. The differences of the means were statistically significant (ANOVA $F = 5.97$, $p < 0.001$). The mean out-degree of 4.77 for households without an eligible child was somewhat lower than that for accepting households (t-test $t = -2.40$, $p = 0.02$). There was also a statistically significant difference in the mean out-degrees for general ties (ANOVA $F = 4.60$, $p < 0.001$) and health ties (ANOVA $F = 3.19$, $p = 0.02$) between vaccine-accepting and vaccine-refusing households. There were no significant differences (ANOVA $F = 0.22$, $p = 0.88$) among the mean in-degrees combining general and health categories (0.60 for vaccine-accepting, 0.64 for vaccine-reluctant, 0.57 for vaccine-refusing, and 0.64 for vaccine-ineligible households).

Using an in-degree cut-off value of 6 for hubs (see Methods), a total of 13 nodes (0.5%) in the vaccine network qualified as hubs. Hub households had 0.72 rooms more than non-hub households ($p = 0.01$), had 1.5 fewer people living per room ($p = 0.01$), and appeared to be more educated than non-hub households (0.95 units higher on the education scale, $p = 0.09$). In addition, 92% of hub households had a mobile phone compared to the 73% of non-hub households ($p = 0.03$) (The p-values for hubs are based on a two-tailed variant of the Welch 2-sample test). There was no evidence of different neighborhoods having different proportions of hub households ($\chi^2 = 20.4$, $d = 24$, $p = 0.68$). We found that 6 out of 13 hub households were vaccine-accepting, and the remaining 7 households had no vaccine eligible children. No hub was vaccine-

Table 2
Summary of nominations and household attributes for 2452 households after excluding 10 households due to incomplete data. Note that of these households, 24 had neither incoming nor outgoing edges and were therefore not included in the vaccine network.

Variable	Mean	SD	Median	Range
Out-degree	4.83	1.49	5	0–8
Out-degree (general)	3.04	0.87	3	0–4
Out-degree (health)	2.39	0.90	2	0–4
In-degree	0.62	1.35	0	0–37
In-degree (general)	0.48	1.08	0	0–27
In-degree (health)	0.29	0.79	0	0–16
Number of people in household	6.34	3.16	6	1–37
Number of rooms in household	1.44	0.85	1	1–9
Education	2.44	1.41	2	1–8

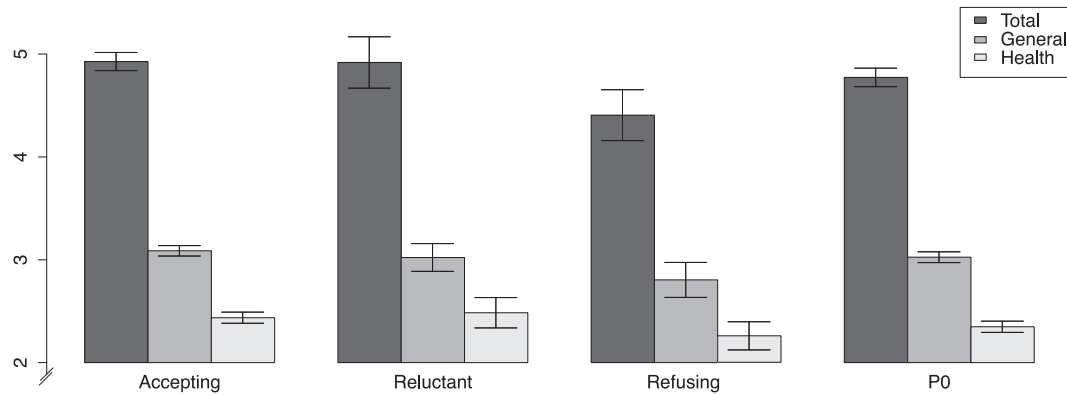| Variable	N	%
Cooking cylinder	519	21.6%
Toilet	710	29.6%
TV	1080	45.0%
Mobile phone	1752	72.9%
Education		
1: No school	459	19.1%
2: Primary school	1220	50.8%
3: Middle school	369	15.4%
4: High school	173	7.2%
5: Intermediate diploma	63	2.6%
6: Graduate or post graduate	42	1.7%
7: Professional	22	0.9%
8: Islamic education	54	2.2%

Fig. 2. Mean out-degrees (total, general, health), and the corresponding 95% confidence intervals, stratified by vaccine status of the recipient. Note that the vertical axis starts at 2.

reluctant or vaccine-refusing.

3.2. Neighborhoods

The neighborhoods varied considerably in terms of their vaccine status. The mean percentage of (i) vaccine-accepting households was 44.0% (range 33.3%–54.4%), (ii) vaccine-reluctant households was 5.6% (range 0%–12.7%), and (iii) vaccine-refusing households was 5.8% (range 0%–12.1%). The vaccine status composition of neighborhoods varied more than chance alone (χ^2-test $p < 0.001$).

We also investigated the proportion of nominations to households located in the same neighborhood as the nominating household, and we aggregated these household-level results to the neighborhood level. Note that we only considered nominations within and across the sampled neighborhoods, not those outside the surveyed neighborhoods. Taken over all neighborhoods, 48.9% (range 23.8%–66.7%) of nominations were to households in the same neighborhood. The fact that about half of all ties were located within neighborhoods suggests strong spatial clustering of network ties (by which we mean clustering with respect to defined boundaries of neighborhoods, not clustering with respect to physical distance).

3.3. Clustering of nodes by vaccine status

The LCC of the vaccine network had mean out-degree of 2.34 for accepting households, 2.13 for reluctant households, and 2.06 for refusing households, thus retaining the trend for out-degrees by vaccine status observed in the full nomination network. An undirected version of the LCC of the vaccine network (obtained by ignoring edge directions) is shown in Fig. 3. The visualization suggests that nodes of a given vaccine status might cluster together. Using the resampling approach described earlier, we found that reluctant households had more reluctant alters and more refusing alters than accepting households did. Furthermore, the proportion of refusing alters was higher for refusing households than for reluctant households. These findings can be summarized in two ways (see Table 3). First, the percentage of neighboring households that were either *reluctant or refusing* was 17.1% for refusing households but only 10.8% for accepting households, resulting in a difference of 17.1% − 10.8% = 6.3% [*p*-value: 0.08; 95% CI: −6.4%, 7.2%]. Second, and more important, the observed difference between refusing and accepting households in the percentage of *neighboring refusing households* was 11.4% − 5.9% = 5.5% [*p*-value: 0.05; 95% CI: −5.0%, 5.5%]. Phrased differently, vaccine-refusing households had 93% more ties to vaccine-refusing households than vaccine-accepting households did (Fig. 4).

The results on clustering of nodes by vaccine status so far

pertained to the LCC of the vaccine network. When we considered all components of the vaccine network, not just the LCC, the results were qualitatively similar, but the extent of clustering was stronger still: the percentage of neighboring *reluctant or refusing* households was 19.7% for refusing households and 9.6% for accepting households, whereas the percentage of neighboring *refusing* households was 13.6% for refusing households and 4.7% for accepting households. Vaccine-refusing households now had 189% more ties to vaccine-refusing households than vaccine-accepting households did.

3.4. Predicting ties based on network and neighborhood effects

The statistically significant predictors of dyadic similarity were bivariate degree sum, bivariate degree difference, and residence in the same neighborhood (Table 4). Because a relatively high fraction of ties were within neighborhoods, knowing that two households were located in the same neighborhood substantially increased the log-odds for them to be connected. All else equal, high-degree individuals were more likely to be connected to one another than predicted based on their individual degrees. Note that difference in degree decreased the probability of the two households being connected. The degree-difference predictor informs us that the probability for two households to be connected is higher in the former case, even though in both cases the degree sum is the same. This finding is consistent with the notion of dyadic similarity, i.e., households tend to connect to other households that are similar to them, and, in this case, that similarity manifests itself by the households being close in degree.

4. Discussion

Using a population-based, household-level survey in Malegaon, India, we find that vaccine-refusing households have fewer outgoing ties than vaccine-accepting households and that they had fewer health ties than vaccine-accepting households. Moreover, the surveyed neighborhoods varied significantly by their vaccine status composition, and, on average, 49% of network nominations were to households in the same neighborhood. This suggests that social ties are spatially localized. Finally, vaccine-refusing households nominated on average 93% more vaccine-refusing households in the LCC of the vaccine network than vaccine-accepting households did, meaning that vaccine refusers clustered in the social network and not just within neighborhoods. (This number was 189% when we considered all components, not just the LCC.) Thus, there was both social and geographic clustering. These findings parallel recent findings regarding the clustering of latrine use in India (Shakya et al., 2014a, 2014b).
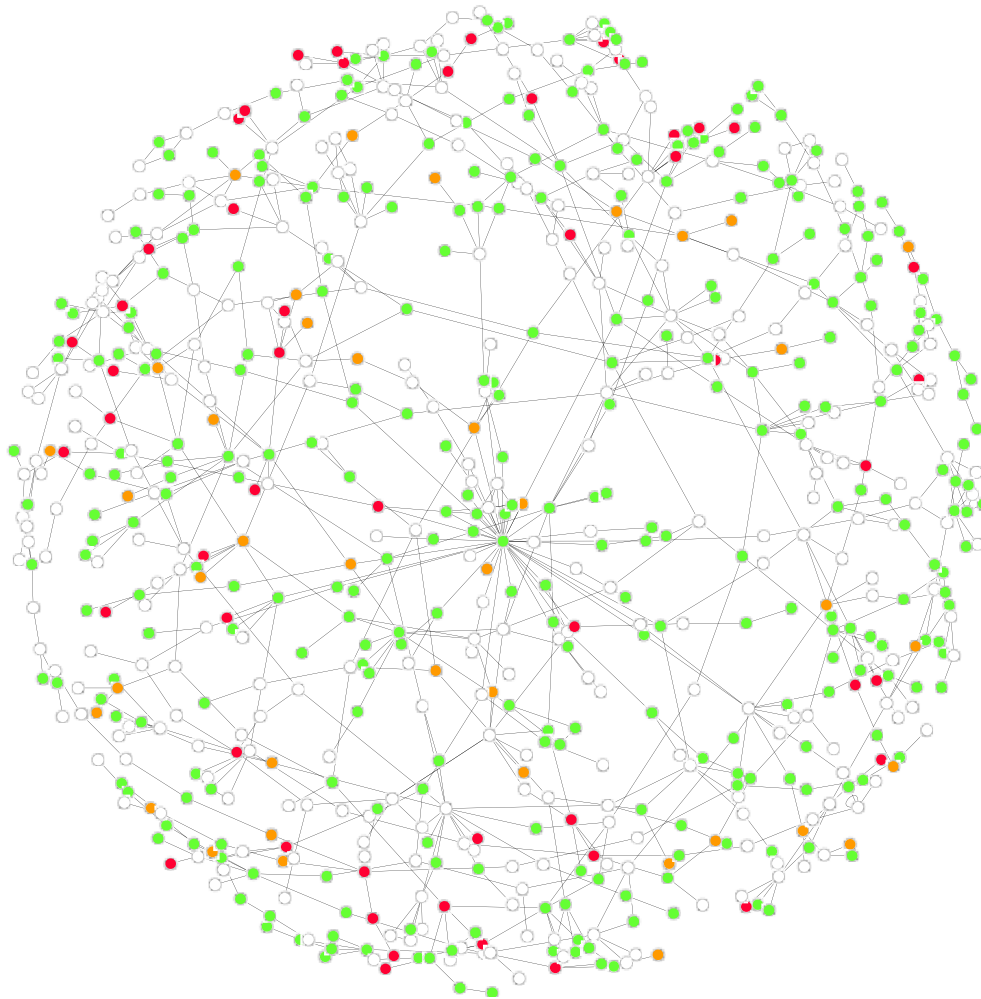
Fig. 3. Largest connected component (LCC) of the vaccine network. Node colors: green = accepting (1), orange = reluctant (2), red = refusing (3), and white = ineligible household (4). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3
Mean fraction of accepting, reluctant, refusing, and ineligible alter nodes stratified by the vaccine status of the ego nodes (rows) in the given category in the vaccine network and the LCC of the vaccine network. Households were deemed ineligible if they reported not having any children between the ages of 0 and 5. To convert from percentages to counts, multiply the former by the number of nodes N.

	Accepting alter	Reluctant alter	Refusing alter	Ineligible alter	Reluctant or refusing alter
Vaccine Network ($N = 2428$)					
Accepting ego	47.0%	4.9%	4.7%	43.4%	9.6%
Reluctant ego	37.0%	10.4%	5.8%	46.8%	16.2%
Refusing ego	36.7%	6.1%	13.6%	43.5%	19.7%
Ineligible ego	40.3%	5.8%	5.1%	48.8%	10.9%
Vaccine Network LCC ($N = 710$)					
Accepting ego	48.6%	4.9%	5.9%	40.6%	10.8%
Reluctant ego	43.4%	7.2%	7.2%	42.2%	14.5%
Refusing ego	41.9%	5.7%	11.4%	41.0%	17.1%
Ineligible ego	43.2%	5.0%	6.2%	45.5%	11.2%

Influence, selection, and confounding are difficult to disentangle in observational studies, especially if there are pertinent unmeasured factors (Christakis and Fowler, 2007, 2008; Shalizi and Thomas, 2010; Christakis and Fowler, 2013). However, if longitudinal data were available, it would be possible to use a model-based approach to get a sense of the extent to which each mechanism might be driving the behavior. Here, given the cross-sectional nature of the study, these types of approaches are not feasible. In addition, other methods, such as those involving the directionality of ties (Christakis and Fowler, 2007) or instrumental variables (O'Malley et al., 2014) are not possible here either.

The present study has a number of further limitations. First, a large number of households nominated households outside the surveyed neighborhoods and vaccine status was only available for the households in the surveyed neighborhoods. Second, the network model used to predict the existence vs. non-existence of ties assumes dyadic independence, which is likely to be violated. However, because the networks studied here were relatively
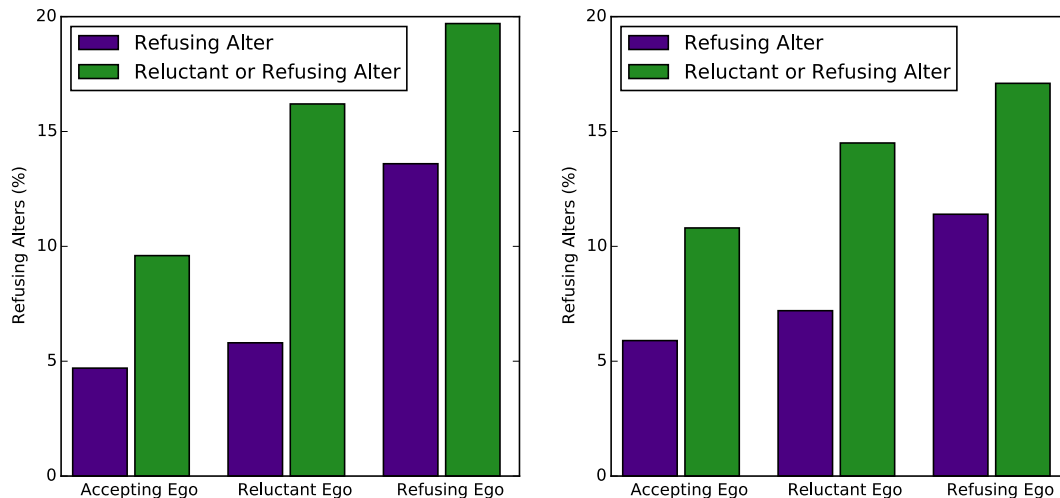
Fig. 4. Mean percentage of refusing alters (purple) and reluctant or refusing alters (green) stratified by the vaccine status of the ego nodes in the vaccine network (left panel) and in the largest connected component (LCC) of the vaccine network (right panel). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4
Estimated regression coefficients for a logistic regression model, where tie status is regressed on the difference in the attributes for the corresponding pair of individuals (dyad). The model was estimated on the largest connected component (LCC) of the vaccine network.

Variable	Coefficient	SE	z-score	p-value
Intercept	−7.316	0.066	−110.693	0.000
Vaccine status equal	0.007	0.052	0.131	0.896
Neighborhood ID equal	2.711	0.053	51.392	0.000
Degree sum	0.265	0.010	25.804	0.000
Degree difference	−0.179	0.011	−15.917	0.000
Education difference	0.010	0.017	0.574	0.566
SES difference	−0.039	0.053	−0.730	0.466

sparse, we believe this approach was reasonable.

One lesson that was recently put forward to help tailor strategies to communities to counter vaccine resistance was to find out what is driving rumors and resistance by identifying key influencers of vaccine acceptance and refusal (Larson and Ghinai, 2011). In the US, discredited claims about the safety of the MMR vaccine have contributed to immunizations rates in some areas falling short of the targeted rate of 90% for children 19–35 months (Nyhan et al., 2014). A randomized trial concluded that pro-vaccine messages, however, do not always work as intended and, for some parents, may actually increase misperceptions or reduce vaccination intention (Nyhan et al., 2014).

While networks have the potential to propagate behavior change and interventions (Valente, 2012), convincing vaccine-refusing households of the benefits of the polio vaccine is likely to remain a difficult challenge. However, identification of these clusters is an important step in that it enables better targeting of subsequent interventions and, furthermore, establishes the relevant units of study, the clusters, for future investigations into different vaccine-adoption strategies. Moreover, taking into account people's social network position can enhance the effectiveness and efficiency of public health messaging and interventions: the effectiveness because people are more likely to change their perceptions and behavior if their friends do, and the efficiency because perceptions and behavioral changes can spread out from the primary targets of interventions, creating spillovers that may benefit whole populations.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.socscimed.2016.01.024

References

Berkman, L.F., Glass, T., Brissette, I., Seeman, T.E., 2000. From social integration to health: Durkheim in the new millennium. Soc. Sci. Med. 51 (6), 843–857.
Borgatti, S., Everett, M., Johnson, J., 2013. Analyzing Social Networks. Sage Publications.
Christakis, N.A., Fowler, J.H., 2007. The spread of obesity in a large social network over 32 years. N. Engl. J. Med. 357 (4), 370–379.
Christakis, N.A., Fowler, J.H., 2008. The collective dynamics of smoking in a large social network. N. Engl. J. Med. 358 (21), 2249–2258.
Christakis, N.A., Fowler, J.H., 2010. Social network sensors for early detection of contagious outbreaks. PLoS ONE 5 (9), e12948. http://dx.doi.org/10.1371/journal.pone.0012948.
Christakis, N.A., Fowler, J.H., 2013. Social contagion theory: examining dynamic social networks and human behavior. Stat. Med. 32 (4), 556–577.
Hoff, P.D., Raftery, A.E., Handcock, M.S., 2002. Latent space approaches to social network analysis. J. Am. Stat. Assoc. 97 (460), 1090–1098.
Larson, H.J., Ghinai, I., 2011. Lessons from polio eradication. Nature 473 (7348), 446–447.
Lusher, D., Koskinen, J., Robins, G., 2013. In: Lusher, D., Koskinen, J., Robins, G. (Eds.), Exponential Random Graph Models for Social Networks: Theory, Methods and Applications. Cambridge University Press.
McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: homophily in social networks. Annu. Rev. Sociol. 27, 415–444.
Newman, M.E.J., 2010. Networks: An Introduction. Oxford University Press, USA.
Nyhan, B., Reifler, J., Richey, S., Freed, G.L., 2014. Effective messages in vaccine promotion: a randomized trial. Pediatrics 133 (4), e835–e842.
Obregón, R., Chitnis, K., Morry, C., et al., 2009. Achieving polio eradication: a review of health communication evidence and lessons learned in India and Pakistan. Bull. World Health Organ. 87 (8), 624–630.
O'Malley, A.J., Elwert, F., Rosenquist, J.N., Zaslavsky, A.M., Christakis, N.A., 2014.

Estimating peer effects in longitudinal dyadic data using instrumental variables. Biometrics 70 (3), 506–515.

Perkins, J.M., Subramanian, S.V., Christakis, N.A., 2014. Social networks and health: a systematic review of sociocentric network studies in low- and middle-income countries. Soc. Sci. Med. 125, 60–78.

Shakya, H.B., Christakis, N.A., Fowler, J.H., 2014. Association between social network communities and health behavior: an observational sociocentric network study of latrine ownership in rural India. Am. J. Public Health 104 (5), 930–937.

Shakya, H.B., Christakis, N.A., Fowler, J.H., 2014. Social network predictors of latrine ownership. Soc. Sci. Med. 125, 129–138.

Shakya, H.B., Christakis, N.A., Fowler, J.H., 2016. An Exploratory Comparison of Name Generator Content: Data from Rural India (under review).

Shalizi, C.R., Thomas, A.C., 2010. Homophily and contagion are generically confounded in observational social network studies. Sociol. Methods Res. 40 (2), 27.

Smith, K.P., Christakis, N.A., 2008. Social networks and health. Annu. Rev. Sociol. 34 (1), 405–429.

Valente, T.W., 2010. Social Networks and Health: Models, Methods, and Applications. Oxford University Press, USA.

Valente, T.W., 2012. Network interventions. Science 337 (6090), 49–53. http://dx.doi.org/10.1126/science.1217330.

Wasserman, S., Faust, K., 1994. Social Network Analysis: Methods and Applications. In: Structural Analysis in the Social Sciences. Cambridge University Press.

WHO Statement on the Meeting of the International Health Regulations Emergency Committee Concerning the International Spread of Wild Poliovirus. http://www.who.int/mediacentre/news/statements/2014/polio-20140505/en/.