HOSTED BY
ELSEVIER
Contents lists available at ScienceDirect
Asian Pacific Journal of Tropical Medicine
journal homepage: http://ees.elsevier.com/apjtm

Original Research http://dx.doi.org/10.1016/j.apjtm.2016.07.005

Social network analysis and modeling of cellphone-based syndromic surveillance data for Ebola in Sierra Leone

Jia B. Kangbai*

Department of Epidemiology, College of Public Health University of Kentucky, Lexington, KY 40536-0003, USA

ARTICLE INFO

Article history:
Received 17 May 2016
Received in revised form 16 Jun 2016
Accepted 1 Jul 2016
Available online 26 Jul 2016

Keywords:
Ebola
Syndromic
Surveillance
Cellphone
Outdegree
Centrality

ABSTRACT

Objective: To explore and visualize the connectivity of suspected Ebola cases and surveillance callers who used cellphone technology in Moyamba District in Sierra Leone for Ebola surveillance, and to examine the demographic differences and characteristics of Ebola surveillance callers who make more calls as well as those callers who are more likely to make at least one positive Ebola call.

Methods: Surveillance data for 393 suspected Ebola cases (192 males, 201 females) were collected from October 23, 2014 to June 28, 2015 using cellphone technology. UCINET and NetDraw software were used to explore and visualize the social connectivity between callers and suspected Ebola cases. Poisson and logistic regression analyses were used to do multivariable analysis.

Results: The entire social network was comprised of 393 ties and 745 nodes. Women (*AOR* = 0.33, 95% *CI* [0.14, 0.81]) were associated with decreased odds of making at least one positive Ebola surveillance call compared to men. Women (*IR* = 0.63, 95% *CI* [0.49, 0.82]) were also associated with making fewer Ebola surveillance calls compared to men.

Conclusion: Social network visualization can analyze syndromic surveillance data for Ebola collected by cellphone technology with unique insights.

1. Introduction

Ebola virus disease (EVD) was first identified in 1976 in Zaire but there has been more cases of the disease in other countries in recent time [1]. EVD is characterised by febrile illnesses and is naturally transmitted by animal and vector hosts [2]. The disease is caused by one of the four distinct hemorrhagic fever viruses (HFVs) family: *Flaviviradae, Arenaviradae, Bunyaviradae, and Filoviradae*. *Ebola virus* of the family *Filoviradae* which was responsible for the 2014 West African Ebola outbreak has 5 different virus strains: Sudan virus, Tai Forest virus, Reston virus, Ebola virus, and Bundibugyo virus. The 2014 outbreak is the largest so far and recorded more than 20000 cases and 10000 deaths [3].

Sierra Leone was one of the countries greatly affected by the 2014 West Africa Ebola outbreak [4]. Sierra Leone's first Ebola case was confirmed on 27th May 2014. Prior to 2014 Sierra Leone's preparedness for an EVD outbreak was lacking. The country's EVD response involved the formation of an Ebola technical task force that was responsible for EVD surveillance, case identification, case tracking and monitoring.

Contact tracing, Ebola case identification, treatment, management, and the effective response to both patients and the community have been shown to be effective for EVD surveillance in the past [5]. Correct coordination between Ebola case isolation and treatment, contact tracing and follow-up of contact for 21 days after exposure was very effective in controlling the spread of Ebola outbreak in West Africa in 2014 [6]. Sierra Leone embarked on using cellphone technology to assist with its EVD surveillance in 2014.

Cellphone technology has revolutionized disease surveillance by serving as a channel through which people reveal their public health concerns, locations, and movements. In Haiti, public health experts successfully predicted the spatial evolution of the 2010 cholera outbreak using cellphone calls and SMS messages obtained from more than 2 million mobile phone SIM cards [7].

*Corresponding author: Jia B. Kangbai, MPH, MSc, CTCM&H, Department of Epidemiology, College of Public Health University of Kentucky, Lexington, KY 40536-0003, USA.
E-mail: jiakangbai@hotmail.com
Peer review under responsibility of Hainan Medical College.

1995-7645/Copyright © 2016 Hainan Medical College. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Following the 2014 Ebola outbreak, researchers in West Africa called for the use of routine syndromic surveillance systems that rely on data supplied via cellphones. Researchers also recommended the use of cellphones for relaying surveillance data from communities affected by an EVD outbreak to peripheral health centers [8]. In Guinea, smartphones were used to communicate real-time surveillance data for contact tracing and case identification during the 2014 West Africa EVD outbreak [9]. Given the surge in global cellphone usage and the increasing popularity of cellphone-based epidemiological surveillance more research is needed to explore its efficiency and community uptake. Furthermore, few studies have taken advantage of the network-like data generated from cellphone-based surveillance in which callers and cases are interconnected through a web of calls.

This study used social network analysis to evaluate Ebola surveillance data collected in Sierra Leone by cellphone technology. This study specifically used network analysis to explore and visualize the connectivity between suspected Ebola cases and callers who used cellphone technology for Ebola surveillance; to examine demographic differences in surveillance callers who used the cellphone surveillance system; and to determine the characteristics of efficient Ebola surveillance callers (i.e., Ebola surveillance callers whose calls were subsequently confirmed to be Ebola positive).

2. Material and methods

2.1. Study setting and participants

Surveillance data were collected from Moyamba District in southern Sierra Leone from October 2014 to June 2015 using cellphones. The district is approximately 6 902 sq/km with a population of 258 506 [10]. In 2014 the population of Sierra Leone was estimated to be 6 315 627 of which approximately 4 215 000 (70%) had cellphones [11]. The Moyamba District Health Manager Team (MDHMT) Ebola Taskforce, in collaboration with Action Contra la famine (ACF), started cellphone-based syndromic surveillance for Ebola in October 2014. Specifically, community members in Moyamba District were encouraged to call the MDHMT Ebola surveillance hotline to report suspected Ebola cases (dead or alive). Community members who called the center provided their names, telephone number, and the village in which they were residing. The caller also provided the name of the person that they suspected to have Ebola and the person's sex, age and the village of the person that they suspected to have Ebola. The caller could call about any person, regardless of their age, gender, or ethnicity. Within 24 h of receiving a call, the MDHMT Ebola Taskforce would dispatch a Community Health Officer and Ebola contact tracers to the location of the person suspected to have Ebola and transfer him/her to an Ebola treatment center for diagnosis. If the suspected Ebola case was deceased his/her body was transferred to the local mortuary for safe burial. Deceased persons however were not tested for Ebola.

This study is an analysis of Ebola surveillance data collected by the MDHMT Ebola Taskforce from October 2014 through May 2015. The dataset for analysis included callers' names, their telephone numbers, and their village of residence, as well as the cases' name, sex, and location (village), whether the person was sick or dead at the time the surveillance call was made, and if there were sick or dead people at the residence of the suspected Ebola case at the time the call was made to the MDHMT call center. The dataset also included the Ebola lab result for suspected Ebola cases. There were 353 surveillance callers and 393 suspected Ebola cases, including one caller who subsequently became a suspected Ebola case. By cross-referencing names and demographic characteristics of callers and cases, a whole network covering the entire surveillance period was constructed and the connectivity of callers and cases was determined. UCINET software [12] was used to analyze the network of callers and suspected Ebola cases. Specifically, UCINET was used to compute degree centrality. NetDraw [12] was used to visualize the network components in order to depict the social ties among callers and suspected Ebola cases. The analysis was also used to estimate the number of calls made by each caller; this estimate served as the outcome in the multivariate analysis.

2.2. Analysis

SAS 9.2 version [13] was used for descriptive and summary statistical analysis of the characteristics of the callers and suspected Ebola cases. Poisson and logistic regressions were used to determine the factors associated with the number of Ebola surveillance calls made and the likelihood of making at least one positive Ebola surveillance call respectively. Specifically, Poisson regression was used to determine the gender difference in the number of Ebola surveillance calls made controlling in the model for the week in which the surveillance call was made, and the Ebola prevalence of the village in which calls were made. Logistic regression was used to determine the gender difference associated with making at least one positive Ebola surveillance call, controlling in the model for the week in which Ebola surveillance calls were made, whether the person for whom the Ebola surveillance call was made was sick or dead, and the Ebola prevalence of the village in which calls were made.

2.3. Ethics and privacy

The University of Kentucky Institutional Review Board reviewed the protocol for the secondary analyses described in this capstone and determined that it met federal criteria to be exempt.

3. Results

3.1. Descriptive characteristics of suspected Ebola cases

Surveillance data for 393 suspected Ebola cases (192 males, 201 females) were collected from October 23, 2014 to June 28, 2015 using cellphone technology. The descriptive characteristics of suspected Ebola cases, callers, status of Ebola suspected cases, lab results and type of Ebola surveillance call made are presented in Table 1. The average age of the suspected Ebola cases was 23.5 years (standard deviation = 29.5). Three hundred and twenty-four (82.4%) of the suspected Ebola cases were reported sick at the time data was collected while 69 (17.6%) were deceased. Two hundred and twenty (68%) of the sick suspected Ebola cases were females while 104 (32%) were males. Twenty-

Table 1

Descriptive characteristics and status of suspected Ebola cases and callers, lab result and type of call (Suspected cases, n = 393).

Characteristics	*N* (%)
Sex	
Male	192 (48.9)
Female	201 (51.2)
Status	
Sick	324 (82.4)
Dead	69 (17.6)
Lab result (among 324 sick suspected Ebola cases)	
Positive	25 (8.0)
Negative	295 (91.0)
Unknown	4 (1.0)
Village Ebola prevalence – median (interquartile range)	0 (0.0)
Callers (n=353)	
Sex	
Male	301 (85.3)
Female	52 (14.7)
Number of calls per person – median (interquartile range)	1 (0.0)
One call	339 (93.9)
Two or more calls	54 (6.1)

five (8%) of the sick suspected Ebola cases were later confirmed as infected with Ebola, 295 (91%) were tested negative for Ebola while 4 (1%) had unknown infections.

3.2. Descriptive characteristics of surveillance callers

Three hundred and fifty-three callers reported syndromic surveillance data for Ebola for the period under review. Three hundred and one (85.3%) of the callers were males while fifty-two (14.7%) were females. Three hundred and thirty-six (95.2%) callers made one call while seventeen (4.8%) made multiple calls. Of the multiple callers; 14 made two calls, one made 3 calls, one made 10 calls, and another made 13 calls.

Of the 353 callers; twenty-two (2 females and 20 males) reported suspected Ebola cases with 100% efficiency (*i.e.*, all the calls made were later confirmed to be Ebola positive). Three male callers made two calls each with one call for each resulting in one confirmed Ebola case (*i.e.* 50% efficiency). All other callers had 0% efficiency (*i.e.*, every suspected Ebola case about whom they called tested negative).

3.3. Descriptive characteristics of calls

Three hundred and ninety three calls were made by 353 callers (male = 341, females = 52) using cellphone technology for Ebola syndromic surveillance from October 23, 2014 to June 28, 2015 (Table 2). Men made fewer (16.7%) surveillance calls about dead suspected Ebola cases compared to women (17.3%). Majority (83.3%) of the calls made by men were about sick suspected Ebola cases. The percentage of men making at least one positive Ebola surveillance call was nearly twice (6.8%) as high than that for women (3.7%).

3.4. Social network analysis

UCINET software was used to construct a network showing the connectivity between Ebola surveillance callers and Ebola suspected cases for the entire surveillance period. The entire network was comprised of 745 nodes with 393 ties covering 253 villages. Dyads (i.e., single calls) composed the largest proportion of the network components. For clarity, a smaller network with no dyads was constructed for optimal visualization (Figure 1). Figure 1 shows the network comprising 45 nodes and 33 ties (i.e. calls) between Ebola surveillance callers and positive Ebola cases. The largest component in the smaller network has a size of 13 and the majority of its other components are triads (19 triplets).

Table 2

Distribution of status of suspected cases, lab results, number and percentages of positive and negative calls by gender of callers.

Characteristics	Male	Female
Number of calls	341 (86.77)	52 (13.23)
Number of positive calls	23 (0.59)	2 (0.05)
Call Status of suspected cases – n(%)		
Sick	284 (83.30)	43 (82.70)
Dead	57 (16.70)	9 (17.30)
Lab result (for 327 sick suspected Ebola reported calls)		
Positive	23 (8.10)	2 (4.70)
Negative	261 (91.90)	41 (95.30)
Number of dead suspected cases not tested	57 (16.70)	9 (17.30)
Number who made at least one positive call	23 (6.80)	2 (3.70)

3.5. Factors associated with number of calls

Multivariable analysis was done using Poisson regression to determine the factors associated with the number of Ebola surveillance calls made and the results are presented in Table 3. There was a gender difference in the number of Ebola surveillance calls made. Men made more Ebola surveillance calls than women ($P < 0.01$). In multivariable analysis, adjusting for other variables in the model, female gender ($IR = 0.63$, 95% CI [0.49, 0.82]) was associated with decreased incidence of making Ebola surveillance calls than male gender. The goodness-of-fit chi-squared test for the Poisson regression was 420.690 ($df = 389$, $P = 0.13$) which indicates that the regression model fits the study data well.

3.6. Factors associated with making at least one positive call

Logistic regression was used to do multivariable analysis to examine the likelihood of an Ebola surveillance call yielding a positive Ebola laboratory result and the results are presented in Table 4. There was a gender difference in the likelihood of making at least one positive Ebola surveillance call. Women were less likely to make at least one Ebola positive call compared to men ($P = 0.01$). Holding other covariates in the model constant, the odds ratio estimate for the gender difference (reference group: males) is 0.33 (95% CI [0.14, 0.81]). This indicates that women were significantly less likely to make at least one positive Ebola surveillance call than men holding other covariates in the model constant (95% CI [0.14, 0.81]).

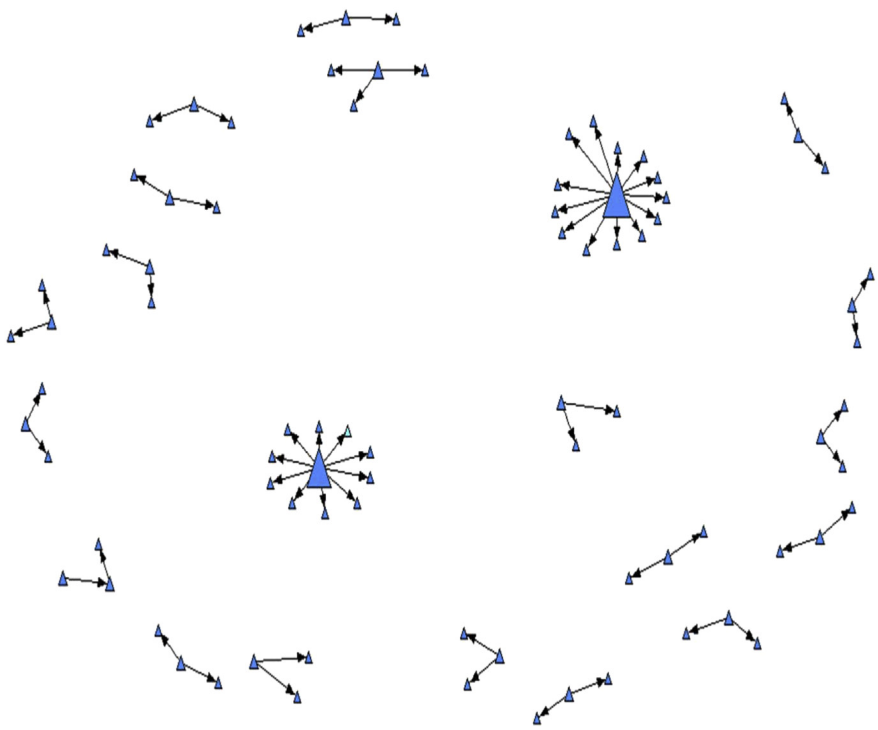

Figure 1. Network showing the connectivity between Ebola surveillance callers and Ebola suspected cases for the entire surveillance period.

Table 3
Multivariate analysis of factors associated with the number of Ebola surveillance calls.

Characteristic	*IR*	95% *CI*	*P*-value
Female versus male gender	0.63	(0.49, 0.82)	<0.01*
Village Ebola prevalence	0.95	(0.88, 1.02)	0.17
Week	0.90	(0.89, 0.91)	<0.0001*

IR = Incidence ratio; CI = Confidence interval, $^{*}P < 0.05$.

Table 4
Multivariate analysis of factors associated with making at least one positive Ebola surveillance call.

Characteristic	Adjusted *OR*	95% *CI*	*P*-value
Female versus male gender	0.33	(0.14, 0.81)	0.01*
Week (1 unit change in week)	1.05	(0.96, 1.15)	0.31
Sick versus dead status	0.52	(0.17, 1.54)	0.24

OR = Odds ratio; *CI* = Confidence interval, $^{*}P < 0.05$.

4. Discussion

Ebola syndromic surveillance data collected by cellphone technology in Sierra Leone was analyzed by visualizing the connectivity of 353 callers and 393 suspected Ebola cases using network analysis. Adjusting the other covariates in the model, women made fewer Ebola surveillance calls than men (*IR* = 0.63, 95% *CI* [0.49, 0.82]). From multivariable analysis, women were significantly less likely to make at least one positive Ebola surveillance call than men holding other covariates in the model constant (*AOR* = 0.33, 95% *CI* [0.14, 0.81]).

No study has investigated the association between gender, number of Ebola surveillance calls made during an Ebola outbreak, and the likelihood of making at least one positive surveillance call. The reasons for men in Sierra Leone making more surveillance calls than women may be attributed to the higher gender gap in literacy. Forty-five percent of men in Sierra Leone are educated compared to 26% for women [10]. This higher gender literacy gap makes men more likely to be employed than women. Sierra Leonean women like their counterparts in Guinea and Liberia experience significant gender-based discrimination in terms of employment and education [14]. This gender-based difference in terms of employment and education also means high property-ownership gap between men and women in Sierra Leone. Sierra Leonean women are 43% less likely than their men to own house, television, computers and a cellphone [15].

Men in Sierra Leone were more likely to make at least one positive surveillance call than women probably due to their increased ability to accurately identified signs and symptoms of Ebola than women. This advantage by men may have been acquired from years of academic and practical clinical training. There are almost twice as many educated men than women in Sierra Leone [10]. Studies have shown that experience obtained from years of academic training and clinical practice is associated with post predictive value accuracy in syndromic surveillance [16]. In syndromic surveillance, post predictive value or positive predictive value is the probability of disease given a positive syndromic surveillance test.

This study used few variables relating to the people in the network because surveillance officers were not provided sufficient, or complete information by callers. It may also be possible that MDHMT Ebola surveillance team officers may have incorrectly recorded the names of the callers, suspected Ebola cases or both. This limitations implies that few variables were available for use in the multivariate analysis using Poisson and logistic regressions. It could also be possible that there were some confounders that were unmeasured and hence unaccounted for in the final multivariable models. The presence of such unmeasured confounders may have also interfered in the result obtained from the models.

In spite of these assumed limitations, health policy and decision makers should not be dissuaded from employing the use of social network analysis to assist with outbreak surveillance and investigation during infectious disease outbreaks. Findings from this study show that men are more active in making Ebola surveillance calls using cellphone technology than women, and that men were also more likely to make at least one positive Ebola surveillance call than women. This study also shows that visualization of simple network of surveillance data can provide a better understanding of the distribution of suspected Ebola cases during outbreak of the disease. Social network analysis could help outbreak investigators and disease intervention specialists to control and managed disease outbreak by connecting suspected Ebola cases to each other and to surveillance officers. Such analysis could also help to identify those people in a community that are at high risk. Also, the combination of cellphone technology with user-friendly and open-source social network software will provide an important adjunct to the traditional measures of epidemiologic surveillance. The ability to analyze the social network of suspected cases during disease outbreak in near or real time may greatly help to decide the best use of disease control resources. This study underscores the need to develop, improve and utilize available tools for outbreak investigation and control.

Conflict of interest statement

There is no conflict of interest attached to this manuscript.

Acknowledgement

Many thanks to Dr. April Young, Dr. Erin Abner, Dr. W. Jay Christian and Prof. Wayne T. Sanderson-of the Department of Epidemiology in the College of Public Health at the University of Kentucky for proof reading, and providing guidance during this study. Thanks goes also to the Mr. Mohamed Kallon of Action Contra la famin in Moyamba District in Sierra Leone for providing the Ebola surveillance data that was used in this study.

References

[1] Sarwar UN, Sitar S, Ledgerwood JE. Filovirus emergence and vaccine development: a perspective for health care practitioners in travel medicine. *Travel Med Infect Dis* 2011; **9**(3): 126-134.

[2] Hong JE, Hong KJ, Choi WY, Lee WJ, Choi YH, Jeong CH, et al. Ebola hemorrhagic fever and the current state of vaccine development. *Osong Public Health Res Perspect* 2014; **5**(6): 378-382.

[3] Zawilinska B, Kosz-Vnenchak M. General introduction into the Ebola virus biology and disease. *Folia Med Cracov* 2014; **54**(3): 57-65.

[4] Dallatomasina S. Ebola outbreak in rural West Africa: epidemiology, clinical features and outcomes. *Trop Med Int Health* 2015; **4**(2015): 448-454.

[5] Tambo E, Xiao-Nong Z. Acquired immunity and asymptomatic reservoir impact on frontline and airport ebola outbreak syndromic surveillance and response. *Infect Dis Poverty* 2014; **3**: 41.

[6] Pandey A, Atkins KE, Medlock J, Wenzel N, Townsend JP, Childs JE, et al. Strategies for containing Ebola in West Africa. *Science* 2014; **346**(6212): 991-995.

[7] Bengtsson L, Gaudart J, Lu X, Moore S, Wetter E, Sallah K, et al. Using mobile phone data to predict the spatial spread of cholera. *Sci Rep* 2015; **5**: 8923.

[8] Ansumana R, Bonwitt J, Stenger DA, Jacobsen KH. Ebola in Sierra Leone: a call for action. *Lancet* 2014; **384**(9940): 303.

[9] Sacks JA, Zehe E, Redick C, Bah A, Cowger K, Camara M, et al. Introduction of mobile health tools to support Ebola surveillance and contact tracing in Guinea. *Glob Health Sci Pract* 2015; **3**(4): 646-659.

[10] Statistics Sierra Leone. *Sierra Leone demographic and health survey 2013*. 4th ed. Sierra Leone: Government of Sierra Leone; 2013.

[11] GlobalEDGE Team. *Global resources and information directory*. 2015 [Online] Avaliable at: http://globaledge.msu.edu/global-resources [Accessed on 22nd June, 2016]

[12] Borgatti SP, Everett MG, Freeman LC. *Ucinet for Windows: software for social network analysis*. Harvard: Analytic Technologies; 2002.

[13] SAS Institute. *SAS/STAT 9.2 User's guide*. Cary, NC: Author; 2004.

[14] Manivanna A. Gender inequalities in access to information about Ebola as gender-based violence. *Harv Hum Rts J* 2015; **28**(6). Avaliable at: http://harvardhrj.com/2015/06/gender-inequalities-in-access-to-information-about-ebola-as-gender-based-violence/ [Accessed on 22nd June, 2015]

[15] Bhandari M. Using technology to close the gender gap in Sierra Leone. The Guardian [Online]. Available at: https://www.theguardian.com/global-development/poverty-matters/2011/aug/29/closing-gender-technology-gap-sierra-leone [Accessed on 30th August, 2011].

[16] Cadieux G, Buckeridge DL, Jacques A, Libman M, Dendukuri N, Tamblyn R. Patient, physician, encounter, and billing characteristics predict the accuracy of syndromic surveillance case definitions. *BMC Public Health* 2012; **12**: 166.