

Investigating cholera transmission using social networks and space in rural Bangladesh

Introduction

Research has shown that social networks can function as a means of disease transmission and thus influence health outcomes. Within such “transmission networks,” a parasitic agent, infection, or other physical component is transferred via the connections that exist amongst the nodes in a given network; examples include the diffusion of sexually transmitted infections or communicable diseases such as tuberculosis. This study will examine the transmission of cholera as a result of existing social networks amongst household units in rural Bangladesh. Furthermore, it will incorporate geographic distance into the network in order to analyze the effects of space on transmission dynamics.

This research is of interest and contributes to existing literature and discussion for a number of reasons. Firstly, while social network analysis is used in disease transmission research, the focus has mainly been on sexually transmitted infections such as HIV/AIDS (Jaffe et al, 1983; Auerbach et al, 1984; Bell et al, 1999). Tuberculosis transmission as a result of social networks has also undergone study (Fitzpatrick et al, 2001; Klovdahl et al, 2001), but there is a significant lack of research on the diffusion of other non-sexually transmitted infectious diseases via social linkages. Secondly, given that there is a relationship between cholera incidence and an individual’s social network, this research will examine the effects of geography on that relationship, primarily by spatial distance between the actors in the network. Specifically, the question is whether distance modifies the effects of social linkages on cholera diffusion, which may prompt further research into what environmental variables affect this social-spatial relationship. This will thus contribute to the existing literature on geography, social networks, and health. Finally, this research adds to discussion of cholera disease ecology and the dominant form of transmission, specifically the difference between primary (contracting the disease from environmental reservoirs) and secondary (contracting the disease from an infected individual) diffusion.

Background and Significance

Cholera is a global public health threat, affecting people in the developing world who live with sub-standard sanitary conditions, including Bangladesh (Faruque et al, 1995; Albert et al, 1993). The disease is currently considered endemic in many countries, with two serogroups of the pathogen *Vibrio cholerae* (O1 and O139) causing outbreaks, most often in vulnerable populations with limited access to clean water and sanitation (Faruque et al, 1995). In Bangladesh, diarrheal diseases including cholera are a significant cause of death in the population (Black et al, 1981, 1982). Such illness causes not only unnecessary suffering, but also decreases economic productivity. Investigating and understanding the causes and patterns of cholera transmission remains an important step in improving the quality of life of people in Bangladesh and other cholera-endemic areas, as this can inform intervention practices and reduce the incidence of the disease.

Cholera transmission is a debated subject; several means of contracting the disease are possible. This has brought about discussion as to whether direct infection through the local environment plays a greater role in transmission than diffusion through person-to-person contact. Miller, Feachem and Drasar (1985) have proposed a model in which cholera infects populations in endemic areas by two different means. The first, known as primary transmission, originates in local estuarine environments where *V. cholerae* is able to survive, spreading to the individual by means of contaminated drinking water or the consumption of contaminated shellfish or aquatic plants. Secondary transmission is, in contrast, a result of cholera diffusion amongst infected individuals. The former is expected to show scattered cholera cases around environmental reservoirs where individuals have come into contact with the pathogen, while the latter is more likely to occur in spatial clusters where individuals have facilitated the spread of the disease to others with whom they interact (Miller, Feacham, and Drasar, 1985). Recent studies in both Peru and Bangladesh support this model (Franco et al, 1997; Ruiz-Moreno et al, in prep), but debate still exists as to whether primary or secondary constitutes the dominant form of transmission.

Additionally, little is known about the diffusion of cholera as a result of social networks, though a great deal of research has examined sexually transmitted infections and other communicable diseases as a result of social and sexual networks and mixing patterns. Social network analysis methods are used to map and measure relationships between social entities, or nodes in a network (Wasserman & Faust, 1994; Hanneman, 2001). They are also utilized in various health-related fields to measure the social relationships that influence disease outcomes and diffusion or health interventions (Morris, 2004). While cholera is able to infect an individual directly through exposure to the pathogen in the environment, it is also spread through consumption of contaminated food or drinking water, and person-to-person contact. These factors are all potentially related to daily interactions one has with other individuals and households. Furthermore, disease diffusion in social networks is affected by geographic space and the local environment.

This study will examine cholera diffusion as a result of social networks and spatial distance; in this sense, it will contribute to the ongoing discussion about transmission as well as the current knowledge of cholera disease ecology. It will take into consideration the social networks of individuals and identify whether or not these networks facilitate

cholera transmission. Because cholera often spreads as a result of behavioral practices—specifically those relating to water, food, and person-to-person contact—and because such behaviors may take place outside of the home, understanding the social network of the study population may be important to identify disease diffusion patterns, as well as points of transmission and the ways in which geographical distance and other environmental variables modify the effects of the social network. This research will further inform understanding of transmission dynamics as they relate to social relationships, environment, and space.

If a significant relationship is indeed found between number of connections to people with cholera and individual cholera occurrence, and if the distribution of cholera has a spatial pattern in terms of concentration and transmission points, then further investigation of transmission means and locations will be warranted in order to effectively intervene and isolate individuals and households at high risk of both contracting and diffusing cholera. Understanding the role that geographic distance plays in the process of disease diffusion will also inform intervention techniques for most effective use of vaccines and other public health measures. While this specific research deals exclusively with cholera, the models developed for this study may be applicable to other diarrheal diseases with similar modes of transmission, as well as non-diarrheal diseases that also diffuse through social networks and are affected by geographic distance.

Theoretical Framework

This research will examine cholera, space, and social networks using the theoretical framework of disease ecology, which is situated within the subdiscipline of medical geography. From this perspective, disease does not exist independently of a host and an environment; in order to fully understand the illnesses and ailments that manifest themselves in a population, one must address and examine the external social, economic, behavioral, cultural, environmental and biological factors in addition to the characteristics of the illness itself. May (1958) and Dubos (1965) introduced the concept of disease ecology, broadening previously held perceptions of human-environment interaction by identifying the numerous aspects involved in the relationship between pathogen and host. In *The Ecology of Human Disease* (1958), May stated that disease does not occur in a human being without the presence of two factors: an environmental stimulus (which may be organic, inorganic, or socio-cultural) and a factor that conditions the response of the body to this stimulus. Stimuli will vary by environment, geographical location, and cultural realm. The response of the individual to these stimuli, then, in whatever form and combination it occurs in, is dictated by genetics and phenotypical traits and varies by individual. The effects of previous stimuli may leave one with a better natural protection or immunity, or it may work to make an individual increasingly susceptible to those stimuli in the future. Each individual also has a point at which a stimulus will cause a response, termed a threshold (May, 1958). Taking into consideration these numerous potential stimuli, disease ecology is concerned with “the ways human behavior, in its cultural and socioeconomic context, interacts with environmental conditions to produce or prevent disease (Meade and Earickson, 2005).” While pathogens and their specific characteristics affect the experience of disease in an individual, it is also of critical importance to consider a person’s lifestyle, encompassing his or her choices, habits and daily environmental interactions. Fully understanding a disease pattern necessitates the study of the population and the various aspects of their lives (Dubos, 1965).

An additional model of human-environmental interaction has been proposed by Meade (1977), known as the “triangle of human ecology.” The three vertices of this triangle represent culture, environment and population, with the interactions between the three providing a foundation for integrating and analyzing the factors that contribute to ecologies of disease. Culture in this framework, for example, represents observable behavioral traits, such as dietary preferences and house type, but also the population’s perceptions of reality and understanding. The environment vertex refers to a comprehensive understanding of the daily and local habitat, rather than a complete understanding of the entire surrounding world. Population represents characteristics of people as biological organisms (age, sex and genetics) rather than as social beings, which directly impacts their interaction with culture and environment. This holistic and integrated perspective presented in Meade’s triangle of human ecology is essential to medical geography and disease ecology, as the interdependent elements all play a vital role in the relationship between population, health, and place.

This research will apply the concepts of disease ecology and the triangle of human ecology to investigate cholera transmission dynamics. In this model, population will include factors such as age and sex. Culture is represented by the behavior and customs of the population that research has linked to an increased risk of developing cholera, e.g. direct consumption of contaminated water or the use of contaminated water bodies for bathing or other purposes (Fuduka et al, 1995; Tauxe et al, 1988; Spira et al, 1980; Swerdlow et al, 1992; Birmingham et al, 1997); consumption of contaminated food (Rabbani et al, 1999; Estrada-Garcia et al, 1996); as well as practices within the household related to food and drink preparation, e.g. preventative measures such as boiling water and cooking food thoroughly (Mintz et al, 1994) or hand-washing. Related to these factors is the social network component, also considered a sociocultural/behavioral aspect. An individual has a network of other individuals with whom he or she may visit and

interact with on a regular basis. In these instances, contaminated food or drink may be consumed within the household, or nearby surface water containing the cholera pathogen used for bathing or other purposes. The connections that an individual has to others are thus a cultural factor related to risk encountered in everyday interactions. Finally, the local environment in this model is represented by the location of the population in relation to environmental factors that affect cholera risk, such as living within a flood-controlled area, proximity to surface water, population density, proximity to latrines, and additional variables that manifest themselves in the surrounding local environment (Emch, 1999). Additionally, spatial distance between *individuals* and households is considered an environmental variable here, as it represents housing density and can also be examined for its effects on cholera outcomes.

Data and Methods

Matlab, the study area, is located in rural Bangladesh approximately 53 kilometers south of Dhaka, at the confluence of the Meghna and Ganges rivers. The population is over 200,000 persons, with a population density of ~1000 persons per square kilometer. The population of Matlab resides in clusters of patrilineally-related households called *baris*, with an average of approximately six households per household and thirty people per *bari* (Ali et al, 2002). The study area has poor sanitary conditions, and much of the population uses surface water for daily purposes such as drinking, cooking, and cleaning. Due to a lack of a public sewage system, the population constructs its own latrines, often near river banks, bodies of water, or ditches in the ground, the latter which may overflow during the rainy season and contaminate surface water (Ali et al, 2002).

A demographic surveillance system has been in place in Matlab since 1966, recording vital demographic events in the study area population (D’Souza, 1981). Individuals in the database who live in the study area are identified by a registration number received upon entry into the surveillance system, whether at birth or through in-migration. Individuals are also assigned a current registration number that corresponds to their current residence in the study area. Data on cholera cases was obtained from hospital records from the local area treatment center run by the International Center for Diarrheal Disease Research in Bangladesh, which is also the only treatment center in the surveillance area. In order to perform spatial analyses of population and health related events in Matlab a vector GIS of the study area was created in 1994 using aerial photography (Ali et al, 2002). This spatial database includes geographic features such as *baris*, a flood-control embankment, rivers, villages, and health facilities.

The specific data used in this study is obtained through the ICDDR,B's demographic and health surveillance system (DHSS), as described above. The dataset that will be used for this study is a subset of a larger dataset containing information on every individual that has ever resided within the study area from 1983 to 2003, including their date of birth, sex, and *bari* of residence. Each *bari* that an individual has lived in is recorded, along with the specific time period. Each person is identified by an RID (Registration Identification Number). Additionally, this study will use data on individuals diagnosed with cholera during the same time period, which also includes RID, gender, residence, and date of diagnosis (for an example of rates of cholera by year, see Table 1). For spatial analysis purposes, the Matlab GIS will be used in conjunction with information on the *baris* of residence of individuals in the study population. The spatial data on *baris* within the GIS allows for identifying the geographic location of a person in addition to the other available data.

Year	1993	1994	1995	1996	1997	1998	1999	2000	2001
Cases	795	827	422	304	580	662	280	145	117
Total	207332	207848	208459	210297	211517	213073	214344	214370	203922
Rate	0.0038	0.004	0.002	0.0014	0.0027	0.0031	0.0013	0.001	0.001

Table 1: Number of cholera cases, total population, and rate of cholera cases in the Matlab study area, 1993-2001

Using the available data on all *baris* of residence for a given individual, a number of social networks will be created for analysis in conjunction with cholera data. Table 2 shows an example of the data describing three individuals, represented by the RID, migrating between *baris* during similar time periods:

RID	<i>Bari</i> at time 1	Entrance Date	Exit Date	<i>Bari</i> at time 2	Entrance Date	Exit Date
A101	10	1/1/1994	3/5/1996	8	3/6/1996	10/4/2000
A102	19	7/18/1995	4/17/1996	34	4/18/1996	1/15/2002
A103	45	10/12/1991	3/29/1996	22	3/30/1996	8/3/1998

Table 2: Example of demographic data

Using this information, it is possible to assume the existence of social linkages between a person's past and present *baris*, based on the date of migration. Figure 1 shows a diagram of these linkages, illustrated as lines connecting nodes, which here represent *baris*:

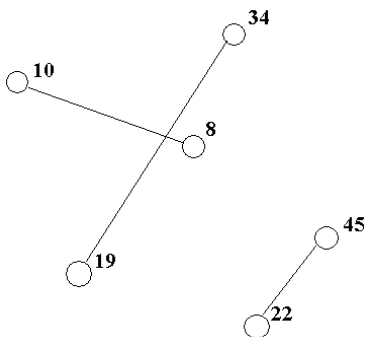


Figure 1: Example of links between nodes in the network

Using comprehensive migration data from 1993 to 2003, various social networks can be constructed based on the movement of individuals between *baris* and the assumption that a linkage between two *baris* is created via the person who has migrated during the previous time period, from time x to time y . Once a network is created representing the structure at time y , each individual is assigned a value based on whether or not cholera was contracted in the period between time y and time z . In addition, each individual is assigned a value representing the number of people with cholera they are connected to via the linkages between *baris* as well as the number of people overall. Using this information it is possible to test whether or not the number of social connections an individual has to others with cholera is a predictor of individual cholera outcome.¹

In order to integrate the geographic component, this study will use the Matlab GIS data in addition to the social network data. Geographic distance between all connected *baris* will be measured using the Hawth's Tools point-distance extension in ArcMap. The spatial distance between any two connected *baris* will then be added as a value into the network data. Assuming that there is a relationship between number of connections to people with cholera and individual cholera outcome, these connections will be weighed by Euclidian distance in order to assess whether space has an effect on the diffusion of cholera across networks.

Expected Results

The main hypothesis is that social networks are a possible means by which cholera transmission occurs, and thus individuals with higher numbers of social connections to others with cholera are at higher risk of developing the disease. Because cholera is spread not only through direct contact with the pathogen in the environment, but also through person-to-person contact, social networks may be an accurate way of representing these transmission patterns. Furthermore, it is expected that the effects of the network are modified by distance. Social networks do not occur in a vacuum; various environmental variables may affect the strength of social relationships (i.e., less frequent visits to *baris* that are at a greater distance). As an alternate and additional means of looking at network connections and cholera, it is planned to test whether or not the number of individuals with cholera who migrate into a certain *bari* within a time period is related to the number of cholera cases in that *bari*.

This research has important implications for public health, as it will add to knowledge of potential prevention of cholera in the study area and other cholera-endemic areas; furthermore, it will help in understanding the dominant form of cholera transmission in this particular setting, which is still a debated subject amongst researchers. Additionally, it will contribute to knowledge of how social networks function as disease transmission pathways in the case of non-sexually transmitted infections. By integrating geographic information, it will also present a realistic view of the strength of these networks in a real-life setting and describe the relationship between distance, social linkages, and disease risk. If the initial findings confirm these hypotheses, there is also potential for further tests based on the wealth of local environmental data available in order to further understand social network-environment interaction.

¹ In order to ensure that the results are robust, I will also test whether or not the number of cholera in-degrees (people with cholera migrating into a *bari*) between time x and time y is related to the number of individuals in that *bari* that develop cholera between time y and time z .

References

- Albert, M. J., A.K. Siddique, M.S. Islam, A.S.G. Faruque, M.Ansaruzzaman, S.M. Faruque, and R.B. Sack. 1993. A large outbreak of clinical cholera due to *Vibrio cholerae* non-O1 in Bangladesh. *Lancet*, 341:704.
- Ali, M., Emch, M., Donnay, J. P., Yunus, M., & Sack, R. B. 2002a. Identifying environmental risk factors for endemic cholera: A raster GIS approach. *Health & Place*, 8(3), 201-210.
- Ali, M., Emch, M., Donnay, J. P., Yunus, M., & Sack, R. B. 2002b. The spatial epidemiology of cholera in an endemic area of Bangladesh. *Social Science & Medicine*, 55(6), 1015-1024.
- Auerbach, D.M., W.W. Darrow, H.W. Jaffe, and J.W. Curran. 1984. Cluster of Cases of the Acquired Immune Deficiency Syndrome: Patients Linked by Sexual Contact. *The American Journal of Medicine*, Volume 76(3):487-92.
- Bell, D.C., J.S. Atkinson and J.W. Carlson. 1999. Centrality measures for disease transmission networks. *Social Networks*, Volume 21, Issue 1, 1-21.
- Birmingham et al. 1997. Epidemic cholera in Burundi: patterns of transmission in the Great Rift Valley Lake region. *The Lancet*, 349: 981–85.
- Black, R.E., Brown, K.H., Becker, S., Alim, A.R.M.A., Huq, I., 1982. Incidence of diarrhea and association with known pathogens. *American Journal of Epidemiology*, 115 (3), 315±324.
- Black, R.E., Merson, M.H., Huq, I., Alim, A.R.M.A., Yunus, M.D., 1981. Incidence and severity of rotavirus and *Escherichia coli* diarrhoea in rural Bangladesh. *The Lancet* 1, 141±143.
- D'Souza, S., 1981. A population laboratory for studying disease processes and mortality: the demographic surveillance system in Matlab, Bangladesh. Special Publication No. 13, ICDDR,B, Dhaka, Bangladesh.
- Dubos, R. J. (1965). *Man adapting*. New Haven: Yale University Press.
- Emch, M. 1999. Diarrheal disease risk in Matlab, Bangladesh. *Social Science & Medicine*, 49(4), 519-530.
- Estrada-Garcia, T, and Mintz, ED. 1996. Cholera: Foodborne transmission and its prevention. *European Journal of Epidemiology*, Vol. 12, no. 5, 461-469.
- Faruque, S. M., S.K. Roy, A.R. Alim, A.K. Siddique, and M.J. Albert. Molecular epidemiology of toxigenic *Vibrio cholerae* in Bangladesh studied by numerical analysis of rRNA gene restriction patterns. *Journal of Clinical Microbiology*, 33(11): 2833–2838.
- Fitzpatrick, LK et al. 2001. A Preventable Outbreak of Tuberculosis Investigated through an Intricate Social Network. *Clinical Infectious Diseases*, Vol. 33, no. 11, 1801-1806.
- Fuduka, J.M, Augusta, Y., L. Chaparro, M. Campos, and E. Chea. 1995. Clinical characteristics and risk factors for *Vibrio cholerae* infection in children. *The Journal of Pediatrics*, Vol. 126, No 6., 882-886.
- Hanneman, R.A. (2001) *Introduction to social network methods*. Riverside, California: Department of Sociology, University of California at Riverside.
- Jaffe, H.W., et al. 1983. National case-control study of Kaposi's sarcoma and *Pneumocystis carinii* pneumonia in homosexual men: part 1, epidemiologic results. *Annals of Internal Medicine*, 99(2):145-51.
- Klov Dahl, A.S. 2001. Networks and tuberculosis: an undetected community outbreak involving public places. *Social Science & Medicine* Volume 52, Issue 5, 681-694.

- May, Jacques M. 1958. *The Ecology of Human Disease*. MD Publications, Inc., New York.
- Meade, M. S. 1977. Medical geography as human ecology: The dimension of population movement. *The Geographical Review*, 67(4), 379-393.
- Meade, Melinda S. and Earickson, Robert J. 2005. *Medical Geography*. Guilford Press.
- Pyle, Gerald S. 1979. *Applied Medical Geography*. V.H. Winston and Sons, Washington, D.C.
- Mintz, E.D., T. Popovic, and P.A. Blake. 1994. Transmission of *Vibrio Cholerae* 01. In *Vibrio Cholerae and Cholera: Molecular to Global Perspectives*. (Wachsmuth, K.I., Blake, P.A. and Olsik, O., Eds.), pp. 321-344. American Society for Microbiology Press, Washington, DC.
- Morris, M. (2004) *Network Epidemiology: A Handbook for Survey Design and Data Collection*. Oxford: Oxford University Press.
- Rabbani, GH, and Greenough WB 3d. 1999. Food as a Vehicle of Transmission of Cholera. *Journal of Diarrhoeal Diseases Research*, 1999 Mar;17(1):1-9.
- Spira et al. 1980. Microbiological surveillance of intra-neighborhood El Tor cholera transmission in rural Bangladesh. *Bulletin of the World Health Organization (WHO)*, 33(5):627-36.
- Swerdlow et al. 1992. Waterborne transmission of epidemic cholera in Trujillo, Peru: lessons for a continent at risk. *The Lancet*, Jul 4;340(8810):28-33.
- Tauxe et al. 1988. Epidemic Cholera in Mali: High Mortality and Multiple Routes of Transmission in a Famine Area. *Epidemiology and Infection*, Vol. 100, No. 2, 279-289.
- Wasserman, S. and Faust, K. (1994) *Social network analysis: Methods and applications*. Cambridge: Cambridge University Press.
- Wylie, J.L., Cabral, T., and Jolly, A.M. 2005. Identification of Networks of Sexually Transmitted Infection: A Molecular, Geographic, and Social Network Analysis. *Journal of Infectious Diseases*, Vol. 191, no. 6, 899-906.
- Wylie, J.L., Shah, L., and Jolly, A. 2007. Incorporating geographic settings into a social network analysis of injection drug use and bloodborne pathogen prevalence. *Health and Place*. 2007 Sep;13(3):617-28.