

HHS Public Access

Author manuscript *Health Info Libr J.* Author manuscript; available in PMC 2020 June 29.

Published in final edited form as:

Health Info Libr J. 2020 March ; 37(1): 26-34. doi:10.1111/hir.12274.

NETWORK ANALYSIS OF INTRA-HOSPITAL TRANSFERS AND HOSPITAL-ONSET CLOSTRIDIUM DIFFICILE INFECTION

Megan McHaney-Lindstrom, MA¹, Courtney Hebert, MD, MS², Harvey Miller, PhD¹, Susan Moffatt-Bruce, MD, PhD, MBA², Elisabeth Root, PhD¹

¹Department of Geography, Ohio State University, Columbus, OH 43210

²Department of Biomedical Informatics, Ohio State University, Columbus, OH 43210

Abstract

Objectives—To explore how SNA can be used to analyze intra-hospital patient networks of individuals with a HAI for further analysis in a GIS environment.

Methods—A case and control study design was used to select 2,008 patients. We retrieved locational data for the patients, which was then translated into a network with the SNA software and then GIS software. Overall metrics were calculated for the SNA based on three datasets and further analyzed with a GIS.

Results—The SNA analysis compared cases to controls indicating significant differences in the overall structure of the networks. A GIS visual representation of these metrics was developed, showing spatial variation across the example hospital floor.

Discussion—This study confirmed the importance that intra-hospital patient networks play in the transmission of HAIs, highlighting opportunities for interventions utilizing these data. Due to spatial variation differences, further research is necessary to confirm this is not a localized phenomenon, but instead a common situation occurring within many hospitals.

Conclusion—Utilizing SNA and GIS analysis in conjunction with one another provided a data rich environment in which the risk inherent in intra-hospital transfer networks was quantified, visualized, and interpreted for potential interventions.

Keywords

Communicable Diseases; Information Science; Information Systems; Public Health

INTRODUCTION

Hospitals are unique, complex, and constantly changing, constructed environments. They contain flows of people and objects which potentially contribute to the transmission of infectious pathogens within the hospital environment. New technologies, such as Real Time Locating Systems can successfully track material objects within hospital environments (Kamel Boulos & Berry, 2012). Likewise, data systems such as electronic health records

Conflict of Interest Statement: There are no conflicts of interest to declare.

(EHRs) or Radio Frequency Identification (RFID) can be used to virtually track individuals as they move throughout the hospital (Cao, Jones, & Sheng, 2014). Despite these advances, proactively identifying the mechanisms by which hospital acquired infections (HAI) spread remains a challenge.

A quarter of hospital patients will contract a HAI during their stay, making control of these pathogens a top concern for hospitals (Yokoe et al., 2014). *Clostridium difficile* is a sporeforming gram-positive anaerobic bacillus associated with diarrhea, which can vary from mild to life-threatening (Cohen et al., 2010). *C. difficile* infection (CDI) is associated with prolonged hospital stays and antibiotic usage and spreads through direct or indirect contact with the bacteria (Cohen et al., 2010). An important aspect of *C. difficile* is the pathogen's ability to live on surfaces for up to five months, suggesting that structures within the hospital environment are important risk factors for transmission (Cohen et al., 2010; Kaatz et al., 1988). As patients move between rooms during a hospital stay they are exposed to a greater number of environments, and in theory could encounter more surfaces harboring bacteria. Prior research from our group shows the number of intra-hospital transfers is positively associated with hospital-onset CDI (HO-CDI), indicating the importance of understanding patient movements within the hospital environment in HO-CDI acquisition (Author Citation Withheld).

From a spatial perspective, patient intra-hospital transfers connect hospital spaces through both associated individuals and equipment. The innovative application of information technology and statistical analysis such as social network analysis (SNA) and geographical information systems (GIS) makes it possible to visualize potential vehicles of HO-CDI in new ways. The current study leverages these techniques/methodologies to understand which hospital rooms are associated with infection acquisition. We suggest this information is crucial to the institution of impactful interventions attempting to reduce HO-CDI, because it provides empirical evidence on which rooms should be targeted for intensive cleaning.

Social Network Analysis

Social network analysis (SNA) is traditionally a methodology used to describe, visualize, and quantify social connections between people. Public health uses of SNA include topics ranging from the spread of sexually transmitted infections through sexual networks to examining collaboration and mentorship between health professionals (Doherty et al., 2012; Petrescu-Prahova et al., 2015). Most relevant to the current research are studies conducted on patient sharing and inter-hospital transfers. These studies have indicated that hospitals which are more geographically central to the populations they serve and those with more inter-hospital transfers increase the spread of hospital acquired infections (Iwashyna, Christie, Moody, Kahn, & Asch, 2009; Lee et al., 2011). However, to our knowledge no previous studies have examined the network structure of *intra*-hospital patient transfers and how this network might contribute to the spread of HAIs within a hospital.

Networks used in SNA are composed of two main parts, the nodes and edges. In the visualization of social networks, the nodes are represented as points. They often represent individuals, or in the case of this research, the rooms that individuals occupy during a hospital stay. The edges are shown as the lines, or relationships, that connect two nodes.

Here, edges represent the physical movement of a patient from one node (hospital room) to another, thereby linking the starting and ending nodes.

SNA produces a number of overall metrics which can be used to examine the qualities of a network. Here we use vertices (number of nodes), total edges (number of intra-hospital transfers), average geodesic distance (average length of paths taken) and eigenvector centrality (calculated relative 'importance' of node to entire network). We used SNA to compare the movement networks of patients diagnosed with HO-CDI to similar patients without HO-CDI to examine whether the network structure of HO-CDI patients differed from non-HO-CDI patients.

Geographic Information Systems

Information technology is now a critical component of public health practice and management. Geographic Information Systems (GIS) have been used in public health informatics as a descriptive and analytic technique that allows visualization and analysis of geographic patterns and spatial associations at various different scales (Fletcher-Larety & Caprarelli, 2016; Ricketts, 2003). Within a GIS, the spatial location of an object (person, room, hospital) is linked to the attributes or characteristics of that object. Different spatial objects can also be linked using their spatial location; for example, a patient can be linked to a hospital room and assigned the attributes of that room. Within a hospital GIS, rooms and spaces could be given attributes - such as the number of beds, presence of sinks for hand washing, and in-room bathrooms - and spatial relationships related to disease transmission such as whether rooms are next to each other or linked by flows of patients or staff. While GIS has been widely applied in public health, there are few examples specific to hospitals, whereby the layout of hospital buildings, floors, and rooms are integrated into a GIS environment. This means that very little spatial analysis has been conducted to examine how spatial relationships between rooms or floor influence disease transmission within the hospital setting.

OBJECTIVES

By applying SNA to electronic health records from a large regional health center, this study: 1) investigated whether the intra-hospital transfer network structures of patients with HO-CDI differed from the networks of similar patients without HO-CDI and 2) explored whether network structures could be transformed into a visualization that could be utilized by hospital epidemiologists for intervention purposes. To address the first question, we examined whether network graphs varied between patients that acquired HO-CDI and those that did not. We expected the patient rooms associated with HO-CDI to have higher eigenvector centrality, indicating those rooms have a higher connectivity within the network due to more patients being transferred into or out of them. To address the second question, we translated the SNA analysis into a GIS to identify high-risk rooms that were spatially correlated. We expected to see high-risk rooms exhibit a local (floor-level) clustering pattern indicative of the spread of CDI pathogen by people and objects utilized on a floor. Results from this study enabled us to recommend ways the hospital administration can utilize SNA

and GIS analysis of hospital EHR data to potentially implement new interventions to mitigate HO-CDI spread within the hospital environment.

METHODS

Data Acquisition and Preparation

We obtained patient data from The Ohio State University Wexner Medical Center Information Warehouse. The data was kept on a secure server behind the OSUWMC firewall to protect patient health information. We identified 502 cases of HO-CDI spanning from late 2013 through 2015 and 1506 controls were selected from the hospital population. Three controls per case were selected using exact matching on birth year and antibiotic usage, which are both predictors of HO-CDI (Fekety et al., 1981). We also applied nearest neighbor matching for the admittance department of the patients to ensure a similar distribution and patient type (Rubin, 1973). The R software package 'Matchit' was used to conduct the matching (Ho, Imai, King, & Stuart, 2011). In order to construct patient-level room transfer network, we extracted admission-discharge-transfer (ADT) data which included every unique room a patient was associated with throughout their hospital stay.

Primary Analysis

To apply social network analysis techniques, ADT data were transformed using NodeXL which calculates the order of room visitation and disease status. Rooms visited by patients during their stay were 'nodes', while their movement between the rooms created edges. This was done separately for both the cases and controls; ultimately resulting in two SNA datasets. Two additional datasets were created for cases only: one for only rooms attributed to HO-CDI (room of 'onset' status or ROS) and a second for rooms prior to onset leading to the rooms of onset (room prior to room of onset or RPROS) (Figure 1). These two additional datasets provided a focused view of the path to disease onset in the case population and allowed for movements between rooms that led to the onset of HO-CDI multiple times to be identified more clearly. Beyond the initial SNA, these datasets isolated movements closer to the onset of HO-CDI, where we conducted SNA analysis examining overall network metrics, most commonly used edges, and nodes most associated with HO-CDI.

To further visualize the SNA data, we created a hospital GIS by integrating shapefiles, a GIS data file that allows storage of an object's locational, shape, and attribute characteristics, and joined SNA output metrics to these shapefiles using hospital room number. We conducted a case study using one floor of the hospital to examine how geostatistical methods could be used to further analyze and visualize the output derived from a SNA. First, a simple map was created that identified nodes (rooms) that had the highest numbers of HO-CDI associated with them. Next, the Local Indicator of Spatial Autocorrelation (LISA) statistical tests were used to identify locations where significantly high or low values are located near one another, identifying the exact locations (rooms) of these extreme values, thereby visualizing hot spots (groups of high values) or cold spots (groups of low values) (Anselin, 1995). Neighboring rooms were defined using the queen's contiguity network (neighboring rooms as those that either share a side or corner with the room of interest).

RESULTS

Social Network Analysis

The SNA provided a unique visualization of the case and control networks. Figure 2 shows edges, or patient pathways, taken more than one time in each network, with darker lines denoting more highly used pathways. The notable difference between the two is the centralized pattern of the control group (Figure 2A) and the seemingly more dispersed pattern of the case group (Figure 2B). This indicates that patients with HO-CDI are visiting a set of rooms more frequently than the control group, which could be an area harboring the HO-CDI bacteria or rooms associated with higher risk such as ICUs, warranting further investigation.

Table 1 shows the overall SNA metrics for both the controls and the cases. For nearly every measure, the SNA for cases had higher values than the controls. The vertices count, or count of number of rooms visited, for the case dataset was higher than the control, indicating that infected patients visited a substantially higher number of rooms. Additionally, the edge count was higher for the cases, suggesting that more intra-hospital transfers occurred in the case network. Edgeweight, or the number of times a specific edge was traveled by patients, results indicated that the same pathways were traveled more often by the case group than the control; Further, nearly all of the control groups' 16 highest edgeweight pathways led to discharge from the hospital, while none of 16 highest edgeweight pathways in the case led to discharge out-degree (amount of times an edge ended at a particular room) and average out-degree (amount of time an edge left a particular room) were higher for the case population than the controls, showing the increased level of movement the case group had over the control. Finally, the eigenvector centrality measurement, or how well connected a room was to the rest of network, reflected the results that 'exit' was a highly central node for the control group- resulting in a higher average.

T-tests were performed to examine differences in edgeweight and eigenvector centrality between case and control networks. Cases had statistically significantly higher values of edgeweight (P < 0.001); while, controls had statistically significantly higher values for eigenvector centrality (P < 0.001).

SNA analysis was also performed on the ROS and RPROS datasets. SNA visualization of results suggest the presence of concentric rings of risk with the rooms with the greatest number of HO-CDI cases in the innermost ring (Figure 3). The central node that all of the edges lead to in Figure 3 represents the 'onset' status of HO-CDI, therefore the nodes surrounding it represent the rooms attributed with the onset of disease. The inner most ring could be interpreted as the rooms that may need the most thorough additional cleaning or additional investigation into potential room-based risk factors, as they lead most frequently to the disease status, while both the intermediate and distal ring are likely of less concern.

Page 5

Geographic Information Systems

Figure 4 shows a visualization of the SNA metrics utilizing the spatially referenced hospital floor. The eigenvector of the nodes (rooms) in RPROS is shown on the left, while a LISA map of the same area identifies clusters of similar eigenvector values. Eigenvectors were chosen because it can identify the most well-connected rooms through patient movements; highly connected rooms see higher patient traffic which may act as vectors for disease transmission. The geovisualization of the SNA metrics for this hospital floor show that the SNA eigenvector values vary over space. This spatial heterogeneity in the data indicates that there are rooms that have a higher risk associated with HO-CDI. The LISA map highlights rooms that have a high value of eigenvector and are surrounded by other high values (highhigh in pink), rooms with high values surrounded by rooms with low values (high-low in red), rooms with low values surrounded by rooms with high values (low-high in dark blue) and rooms with low values surrounded by rooms with low values (low-low in light blue). In our results, the LISA analysis only identified high-low and low-high clusters. This indicates that there are rooms that are key to the intra-hospital transfer network. The spatial correlation of these rooms is noteworthy since HO-CDI can survive for long periods of time on surfaces and is easily spread through both direct and indirect contact. Areas highlighted by the LISA map could reveal rooms that have not received adequate sanitization and require further intervention by hospital staff. These are important points of intervention that could lead to a reduction in HO-CDI diagnosis.

DISCUSSION

Social Network Analysis

This study explored whether the network structures (as defined by intra-hospital room transfers) of patients with HO-CDI differed from similar patients without HO-CDI and whether this could be translated into information for hospital epidemiologist's prevention efforts. The SNA visualization of the entire patient intra-hospital network showed that the network for HO-CDI cases had several key nodes; while, in comparison, the control network had only one main node, highlighting that patients with HO-CDI seem to be visiting the same set of rooms more frequently. If the key nodes, or rooms, of the case network are identified, they would be ideal for supplementary cleaning---an important step for control of hospital acquired infections. Furthermore, the case network had higher values for nearly every overall metric, and most importantly a significantly higher edgeweight value. Higher edgeweight values indicate that more similar pathways are taken by HO-CDI cases relative to controls, providing important information on sets of rooms that CDI patients are transferred between; these rooms should receive focus for more thorough or frequent cleaning to remove the C. difficile spores. Also of interest is that for patients that served as controls, the highest edgeweight values led to hospital discharge, meaning these patients were not consistently transferred to similar sets of rooms. Rather, most of these patients were admitted, had few transfers within the hospital during their stay, and then left the hospital when treatment was over. By comparison, none of the highest edgeweights for the CDI cases led to exit. This suggests that CDI patients were admitted, had significantly more transfers within the hospital, and many visited a similar set of rooms during their stay. These represent very different hospital stay experiences between patients with HO-CDI and those

without. The comparison of these two networks confirmed that the intra-hospital transfer network structure of patients with HO-CDI and those without is different. This is applicable beyond the hospital network at this hospital system and for HO-CDI research in general, due to the pervasive nature of HAIs. The higher rate of intra-hospital transfers that patients who are diagnosed with CDI could be the cause of the infection, as they are exposed to more environments during the process; however, further research needs to be conducted to draw a conclusion.

The most important result from the ROS dataset was the identification of high-risk rooms. Figure 3 was arranged with rooms having the highest edgeweight, or times that room was attributed with HO-CDI, located in the center, therefore identifying 'risk rings'. The centermost ring would be the most opportune location for intervention as they are the rooms that lead most often to the onset of HO-CDI in the case dataset. A recommended intervention might be targeted cleaning via chlorine releasing disinfectants and with UV light cleaning equipment (Fawley et al., 2007; Levin, Riley, Parrish, English, & Ahn, 2013). A secondary possible recommendation could be to initiate outbreak investigation protocols using the patients who contracted CDI in the high-risk rooms.

The edgeweight and eigenvector values for the RPROS dataset also help identify particularly high-risk rooms. High edgeweight values identify the pathways most taken by patients in the hospital, while eigenvector centrality identifies the nodes most central to the network. Similarly to the 'risk rings', those rooms identified as having both high edgeweight and eigenvector centrality should be targeted for additional cleaning, because they are both highly connected to other rooms and lie along heavily used pathways. Both characteristics facilitate the spread of infectious disease. Targeting rooms with high edgeweight and eigenvector centrality could have a significant impact on reducing the spread of HO-CDI.

Geographic Information Systems

Maps are a visual language that are familiar and widely understood, and the GIS visualization allowed the data to be transferred from the realm of connections simply based upon patient movements and into a visual representation of the high-risk rooms. Mapping the eigenvector centrality metric identified the rooms most crucial to the intra-hospital HO-CDI patient transfer network. The LISA statistic then identified high eigenvector rooms on a hospital floor that were surrounded by low eigenvector rooms for the RPROS dataset. This migration of the data from a SNA framework into a GIS revealed key information in the form of rooms that were central to the spread of HO-CDI. This provides two important pieces of information. First, by highlighting particular high-risk rooms it provides a starting point for further investigation into the mechanisms or structures which increase risk. These rooms could have characteristics within them, or in the areas around them, that harbor or propagate HO-CDI since they are crucial to the intra-hospital transfer network. It could also indicate a diffusion of the diseasebetween spaces through potential vectors such as patients, employees, or equipment. Second, it provides hospital administration with visualized data to make decisions regarding interventions. Since these rooms are most associated with the intra-hospital transfer network it could identify rooms that are not receiving adequate

sanitization and make help to delegate resources more efficiently, ultimately cutting off these key rooms from the larger network.

Limitations

The data was collected over several years, while HO-CDI spores have been known to live up to five months on surfaces. Therefore, an appropriate time lag would need to be created if the goal was the creation of a real-time application to account for how long bacteria could live in a room post HO-CDI patient occupancy. However, because HO-CDI is a rare outcome this could prove difficult. A limitation in the GIS analysis is having data grouped by floor. Preliminary results indicated that between floor spatial correlation of eigenvector and edgeweight values were insignificant for the hospital studied, however this could be highly dependent on the space and hospital being analyzed, and configuration of air ducts and other non-obvious connections which could link high-risk rooms. Additional limitations include the fact that a significant proportion of HO-CDI is likely due to patients who are colonized by the bacteria prior to admission and receive in-hospital antibiotics, causing active disease. For these cases additional room cleaning would have little impact.

CONCLUSION

This research shows the feasibility of a real-time implementation of both SNA and GIS analysis to help inform hospital staff on which rooms are highly associated to HO-CDI patients by using retrospective data. Improved cleaning practices, or more intensive regimens, could be instituted for rooms identified as being highly associated to HO-CDI patients. These rooms were first found by using SNA, then further delineated by the GIS analysis. Of highest concern should be the rooms that the LISA map showed as high valued surrounded by high values. Cleaning that has been identified for HO-CDI bacteria includes chlorine releasing products, bleach based products, and UV light sanitization (Fawley et al., 2007; Levin et al., 2013). These are all modalities currently used at our institution, however these types of analyzes would allow us to better target cleaning. More than just identifying the most high-risk rooms, the GIS allowed identification of these rooms in their geographic reality. This is key for the allocation of cleaning staff or assignment of patients to certain wards. Further, this GIS could be integrated into applications for infection preventionists to identify the rooms and patient pathways of high-risk in real time to help with decision making as outbreaks unfold. The next step for the combination SNA and GIS analysis is the implementation of real-time applications to analyze and visualize patient data to halt outbreaks before they begin.

Acknowledgments

Funding Statement: This project was supported by the Institute for the Design of Environments Aligned for Patient Safety (IDEA4PS) at The Ohio State University which is sponsored by the Agency for Healthcare Research & Quality (P30HS024379). The authors' views do not necessarily represent the views of AHRQ.

References

Anselin L (1995). Local Indicators of Spatial Association—LISA. Geographical Analysis, 27(2), 93– 115.

- Cao Q, Jones DR, & Sheng H (2014). Contained nomadic information environments: Technology, organization, and environment influences on adoption of hospital RFID patient tracking. Information and Management, 51(2), 225–239.
- Cohen SH, Gerding DN, Johnson S, Kelly CP, Loo VG, McDonald LC, Pepin J, et al. (2010). Clinical practice guidelines for Clostridium difficile infection in adults: 2010 update by the society for healthcare epidemiology of America (SHEA) and the infectious diseases society of America (IDSA). Infection control and hospital epidemiology: the official journal of the Society of Hospital Epidemiologists of America, 31(5), 431–55.
- Doherty IA, Serre ML, Gesink D, Adimora AA, Muth SQ, Leone PA, & Miller WC (2012). Sexual networks, surveillance, and geographical space during Syphilis Outbreaks in Rural North Carolina. Epidemiology, 23(6), 845. [PubMed: 23007041]
- Fawley WN, Underwood S, Freeman J, Baines SD, Saxton K, Stephenson K, Owens RC, et al. (2007). Efficacy of Hospital Cleaning Agents and Germicides Against Epidemic Clostridium difficile Strains. Infection Control & Hospital Epidemiology, 28(08), 920–925. [PubMed: 17620238]
- Fekety R, Kim KH, Brown D, Batts DH, Cudmore M, & Silva J (1981). Epidemiology of antibioticassociated colitis. Isolation of clostridium difficile from the hospital environment. The American Journal of Medicine, 70(4), 906–908. [PubMed: 7211925]
- Fletcher-Larety SM, & Caprarelli G (2016). Application of GIS technology in public health: successes and challenges. Parasitology, 143(4), 401–415. [PubMed: 26831619]
- Ho DE, Imai K, King G, & Stuart EA (2011). MatchIt : Nonparametric Preprocessing for Parametric Causal Inference. Journal of Statistical Software, 42(8), 1–28.
- Iwashyna TJ, Christie JD, Moody J, Kahn JM, & Asch DA (2009). The structure of critical care transfer networks. Medical Care, 47(7), 787–793. [PubMed: 19536030]
- Kaatz GW, Gitlin SD, Schaberg DR, Wilson KH, Kauffman C. a , Seo SM, & Fekety R (1988). Acquisition of Clostridium difficile from the hospital environment. American journal of epidemiology, 127(6), 1289–1294. [PubMed: 2835900]
- Kamel Boulos MN, & Berry G (2012). Real-time locating systems (RTLS) in healthcare: A condensed primer. International Journal of Health Geographics, 11, 25. [PubMed: 22741760]
- Lee BY, McGlone SM, Song Y, Avery TR, Eubank S, Chang CC, Bailey RR, et al. (2011). Social network analysis of patient sharing among hospitals in Orange County, California. American journal of public health, 101(4), 707–713. [PubMed: 21330578]
- Levin J, Riley LS, Parrish C, English D, & Ahn S (2013). The effect of portable pulsed xenon ultraviolet light after terminal cleaning on hospital-associated Clostridium difficile infection in a community hospital. American Journal of Infection Control, 41(8), 746–748. [PubMed: 23685092]
- Petrescu-Prahova M, Belza B, Leith K, Allen P, Coe NB, & Anderson LA (2015). Using Social Network Analysis to Assess Mentorship and Collaboration in a Public Health Network. Preventing Chronic Disease, 12(E130), 1–10.
- Ricketts TC (2003). Geographic Information Systems and Public Health. Annual Review of Public Health, 24(1), 1–6.
- Rubin DB. (1973). Matching to Remove Bias in Observational Studies. Biometrics, 29(1), 159–183.
- Yokoe DS, Anderson DJ, Berenholtz SM, Calfee DP, Dubberke ER, Ellingson KD, Gerding DN, et al. (2014). A compendium of strategies to prevent healthcare-associated infections in acute care hospitals: 2014 Updates. American Journal of Infection Control, 42(8), 820–828. [PubMed: 25087135]

Key Messages:

- **1.** SNA highlights how patients with HAIs are moved differently within the hospital.
- **2.** GIS can be used to further investigate potentially infectious rooms identified through analysis with SNA.
- **3.** GIS and SNA can help better inform hospitals on cleaning efforts to help potentially reduce HAI occurrence.





Rooms of 'onset' status (ROS) and the room prior to room of 'onset' status (RPROS)



Figure 2.

Visualization of SNA results for A) HO-CDI Controls and B) Cases. Points represent nodes (rooms) and lines represent edges (connections between rooms through patient transfer). Darker lines indicate frequently traveled pathways.



Figure 3. ROS 'Onset' Risk Rings



Figure 4. RPROS Eigenvector and LISA Map

Table 1.

SNA Overall Metrics.

Graph Metric	CASE	CONTROL		
Graph Type	Directed	Directed	Difference	P-value
Vertices	770	692	Case +	-
Total Edges	2466	1477	Case +	-
Average Edgeweight	1.6434	1.3697	Case +	0.001
Average Eigenvector Centrality	0.0013	0.0014	Control +	< 0.001