A network approach to understanding obesogenic environments for children in Pennsylvania

Emily A. Knapp*, Usama Bilal, Bridget T. Burke, Geoff B. Dougherty and Thomas A. Glass

Johns Hopkins School of Public Health. 615 N. Wolfe Street, Baltimore, MD 21205.

*E-mail: eknapp2@jhu.edu.

This article was edited by Eric Quintane.

Abstract

Network methods have been applied to obesity to map connections between obesity-related genes, model biological feedback mechanisms and potential interventions, and to understand the spread of obesity through social networks. However, network methods have not been applied to understanding the obesogenic environment. Here, we created a network of 32 features of communities hypothesized to be related to obesity. Data from an existing study of determinants of obesity among 1,288 communities in Pennsylvania were used. Spearman correlation coefficients were used to describe the bivariate association between each pair of features. These correlations were used to create a network in which the nodes are community features and weighted edges are the strength of the correlations among those nodes. Modules of clustered features were identified using the walktrap method. This network was plotted, and then examined separately for communities stratified by quartiles of child obesity prevalence. We also examined the relationship between measures of network centrality and child obesity prevalence. The overall structure of the network suggests that environmental features geographically co-occur, and features of the environment that were more highly correlated with body mass index were more central to the network. Three clusters were identified: a crime-related cluster, a food-environment and land use-related cluster, and a physical activity-related cluster. The structure of connections between features of the environment differed between communities with the highest and lowest burden of childhood obesity, and a higher degree of average correlation was observed in the heaviest communities. Network methods may help to explicate the concept of the obesogenic environment, and ultimately to illuminate features of the environment that may serve as levers of community-level intervention.

Keywords

Obesity environment networks.

Networks are everywhere (Barabasi, 2007, 2012, 2009, 2013). However, in public health, network science has only now begun to make significant in-roads. To date, network science has made contributions in diverse areas of biomedical research including cellular communication in cancer (Stites et al., 2007; Berger et al., 2012; Gill et al., 2014; Mutation

Consequences and Pathway Analysis working group of the International Cancer Genome Consortium, 2015), protein–protein interactions (Jeong et al., 2001), and complex disease interactions (Barabasi, 2007; Goh et al., 2007; Hidalgo et al., 2009; Zhou et al., 2014). Common features link these diverse applications, including high dimensional data and emer-

gent patterns not easily visible in bivariate space. Networks depict relationships among objects in a system, and network methods help identify structures that influence system behavior.

Obesity is a challenge for traditional public health research because we do not at present have a robust explanation for the temporal and spatial patterns of the obesity epidemic (Galea et al., 2010). This has led obesity researchers to seek alternative, systems science-oriented methods and approaches (Burke and Heiland, 2007; Huang et al., 2009; Finegood, 2011). Network science has made important contributions in obesity research along several dimensions. First, network methods have been used to identify complex linkages among obesity-related genes in animal models (Chen and Zhang, 2013). Second, researchers have conceptualized the 'stress-response network' to understand how feedback within biological systems leads to exacerbation and habituation that results in obesogenic growth (Dallman et al., 2003, 2006). Network approaches have been used to study interactions among organizations and components of obesity interventions (Leroux et al., 2013; Marks et al., 2013), and applied to causal loop diagrams to identify leverage points for intervention (McGlashan et al., 2016). Several studies have focused on how obesity and physical activity spread through populations like infection (Crandall, 1988; Christakis and Fowler, 2007; Blanchflower et al., 2009; Hammond, 2010; Hill et al., 2010; Ali et al., 2012; El-Sayed et al., 2012; Gesell et al., 2012; Simpkins et al., 2013; Hammond and Ornstein, 2014). Others have examined how obesity impacts social relationships (Brewis et al., 2011; de la Haye et al., 2011; Ali et al., 2012). Despite these advances, most network studies in obesity have focused on the structure of linkages between individuals connected through social ties. We are aware of no studies to date that focus on the structure of linkages between features of the environment, thought to be the main drivers of the obesity epidemic.

The concept of the 'obesogenic environment' was first proposed in the 1990s as a model for evaluating the contribution of environmental factors to obesity (Hill and Peters, 1998; Swinburn et al., 1999). The obesogenic environment assumes a pattern of spatially co-occurring features that jointly influence obesity risk. There is little doubt that aspects of the food and physical activity environment are important, but the question of how to identify patterns of features within the obesity environment remains unanswered. Tools to examine the connections between features of the obesogenic environment are needed. Network analysis can be used to describe relationships (edges) between objects (nodes), allowing for the characterization of network-level features that are otherwise hidden. Network methods also allow us to visualize these connections, facilitating understanding of a very complex epidemic and potentially prioritizing areas for intervention. In this study, we characterize the obesogenic environment with community features as nodes, and correlations between those features as edges. A version of this methodology has been used in neurologic and genetic research and is commonly referred to as 'Weighted Correlation Network Analysis' (Fox et al., 2005; Zhang and Horvath, 2005). Our approach examines the structure of the relationships among multiple community features, instead of examining each community feature as an independent cause of obesity.

The literature demonstrates a strong relationship between environmental features that impact diet and physical activity. However, existing studies have focused on single obesity-related features in isolation, most often evaluated for their linear associations with obesity. There has been scant attention to the interdependence of these environmental features and how relationships between obesogenic features of the environment are structured and may create qualitatively different risk environments for obesity. We draw from network analysis tools to study these interrelations between features of the environment and explore how they relate to spatial patterns of obesity prevalence. We are guided by the view that transportation systems, cultural variation, markets, and other system dynamics create clusters of obesity-related features that may have synergistic and aggregative effects on population behavior. Market forces lead to clusters of restaurants, stores and activity spaces in the built environment (Hidalgo and Castañer, 2015). This clustering may potentiate the effect of any one facility by increasing the joint effect of a built and social environment designed to deliver excess calories with maximal efficiency. Therefore, the clustering of features and the existence of central bridging nodes that link disparate clusters may point toward novel targets for research and intervention.

Our primary aim is to explore the utility of network analysis methods to characterize the linkages among a set of 32 spatially patterned features of the obesogenic environment. We created a weighted network of community features from 1,288 communities in Pennsylvania, and examined the relationship between centrality and clustering measures and a commonly used metric of childhood overweight and

obesity (percentage of children with body mass index (BMI) percentile≥85th).

Methods

Our goal was to model the network of hypothesized obesity-related features of local environments to better understand how network and node centrality and clustering provide insights about the role of environments in child and adolescent obesity.

Data

Our study was based on data from a study of children from 1,288 communities in central and northeastern Pennsylvania served by Geisinger Health System. From the system's electronic medical records system (EMR), we previously received data on all patients between 3 and 18 years old who visited a Geisinger primary care physician from 2001 to 2012. The sample included 163,473 children and 523,674 visits. The sample is representative of the youth population in the region (Schwartz et al., 2011). This study was approved by institutional review boards at Geisinger Health System and Johns Hopkins School of Public Health.

Children were previously assigned to one of 1,288 communities based on their geocoded home address. Communities consisted of census tracts within cities, and minor civil divisions (townships and boroughs) outside of cities (Schwartz et al., 2011). From the Geisinger EMR we obtained longitudinal height and weight measurements for children. Implausible BMI values, defined as five standard deviations above and below the median, were assumed to be mismeasurement or data entry errors and were deleted using the standard CDC SAS Program (Schwartz et al., 2011). We calculated z-scores for individual BMI, estimated community-level mean BMI z-score, and percent of children who are overweight or obese (BMI greater than or equal to the 85th percentile for age and sex). We then categorized communities according to quartiles of the percent of children with BMI at or above the 85th percentile.

To characterize obesity-related features of the environment we assembled a corpus of 32 variables hypothesized to be related to obesity based on existing research. These variables include demographic, economic, and geographic information from publicly available datasets, including those published by the U.S. Census, the Federal Bureau of Investigation, and two commercial data vendors, InfoUSA and Dun & Bradstreet, that provided registries of commercial food and physical activity establishments categorized using standard industry codes. Table 1 describes the community features we studied. This list was selected based on attributes that are well accepted in the literature, have acceptable measurement properties, and span a range of content domains and relations with some being related to physical activity and some related to diet. This set of variables has been used in our previous research to characterize diverse aspects of the obesity-related environment in communities (Nau et al., 2015). We categorized all variables in quintile or quartile rank scores to preserve the rank position of variables that are often poorly distributed. After reviewing variable distributions and Spearman, Pearson, and Phi correlation matrices, we chose Spearman correlations as the best representation of variable distributions and the strength of connections.

Network methods

Given the complex nature of obesogenic environments, we looked for a way to best characterize the relationships between all 32 community features. We needed to honor both the pairwise (bivariate) correlations between variables and the structures that emerge from these pairwise correlations. We used a method analogous to Weighted Correlation Network Analysis (Zhang and Horvath, 2005). We generated a data array of covariates (32 obesity-related community features) that we treat as nodes in a network of interconnected environmental attributes. Edges were operationalized as the strength of the bivariate correlation between each pair of attributes. Bivariate correlations were estimated using pairwise Spearman correlation coefficients between the community variables. Because we were primarily interested in the strength of linkages among nodes and there is controversy over the direction of the relationships between some of these variables and obesity, we chose to use the absolute value of the correlation between variables.

All 992 pairwise correlations were then converted to a weighted undirected adjacency matrix where each cell was the correlation between any two variables. We created a weighted graph from this adjacency matrix using the R package iGraph (version 1.0.1) (Csardi and Nepusz, 2006), specifying pairwise correlations as the edge weights. From this graph, we obtained five sets of results.

First, we plotted an overall network graph using all 1,288 communities. The coordinates of each node were computed using a force-based algorithm, the Fruchterman-Reingold algorithm (Fruchterman and Reingold, 1991), where the attraction between nodes is

Table 1. Obesity-related community features included in network analysis.

Variable identifier	Feature
C-1	Violent crime per 100,000 population
C-2	Crimes against person per 100,000 pop
C-3	Crimes against property per 100,000 pop
F-1	Grocery stores and supermarkets per square mile
F-2	Gas stations and convenience stores per square mile
F-3	Snack stores (donuts, pretzels, ice cream) per square mile
F-4	All food establishments per square mile
F-5	Fast food chain restaurants, count
F-6	All retail food establishments per square mile
F-7	All food service establishments per square mile
F-8	Diversity of food establishments in 9 categories
F-9	Limited service restaurants per capita
F-10	Full service restaurants per capita
F-11	Bars and taverns per capita
F-12	Health food and gourmet stores per capita
F-13	Fruit and vegetable stores and stalls per sqare mile
F-14	Discount stores per square mile
L-1	Average block length
L-2	Household density
L-3	Road intersection density
L-4	Road segment length diversity
P-1	Diversity of physical activity establishments in 6 categories
P-2	Indoor recreational centers per square mile
P-3	Outdoor recreational centers per square mile
P-4	Public outdoor parks and recreational spaces per street mile
P-5	All physical activity establishments per square mile
P-6	Indoor fitness and recreational facilities per street miles
P-7	Outdoor fitness & recreational facilities per capita
P-8	Indoor recreational clubs and organizations per square mile
P-9	Outdoor recreational clubs and organizations per square mile
T-1	Vehicle miles traveled on main roads (total)
T-2	Vehicle miles traveled on main roads per capita

Notes: Data are from combined Info USA and D&B databases, 2000; the Federal Bureau of Investigation Uniform Crime Reporting System; U.S. Census 2000; and Pennsylvania State Government.

proportional to the strength of the correlation between environmental features (nodes). We implemented the version of the algorithm in the R package qgraph (version 1.3.2) (Epskamp et al., 2012). The second set of results represents the same graph stratified by community obesity burden (quartiles). For ease of interpretation we show plots corresponding to quartile 1 and 4 (the thinnest and heaviest communities, respectively (see Fig. 1 (overall network) and Fig. 2 (stratified network)).
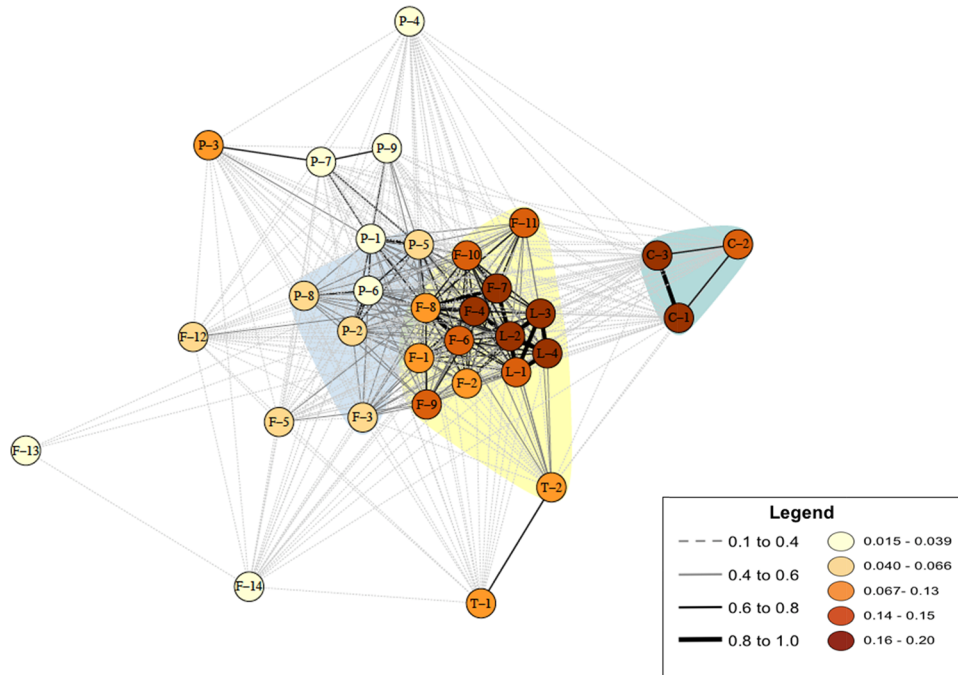
Figure 1: **Network graph for 1,288 communities in Pennsylvania**. This shows a graph of the network of connections between attributes of communities in 1,288 communities in Pennsylvania. Each node in the network represents one feature of the communities, and the edges in the network are absolute values of Spearman correlation coefficients. The bivariate correlation between each variable and average body mass index (BMI) z-score is shown by the shading of each node, with darker colors representing stronger absolute correlation with average community BMI z-score. The strength of absolute correlation between two nodes is represented by the darkness and thickness of the lines connecting the variables. A thick, dark line may represent either a strong positive or a strong negative correlation. Modules of highly connected variables were created using the walktrap method.

Third, in order to better understand the relationships between the components of the obesogenic environment, we sought to obtain a measure of clustering and community structure that allowed us to evaluate if network structure was different across communities classified by prevalence of childhood obesity. We conducted a module detection analysis using the walktrap method (Pons and Latapy, 2005) that performs a series of 'random walks' between nodes. The probability of 'walking' from one node to the next is proportional to the weight of the edge between the nodes, meaning that a walk is more likely to occur between two highly correlated nodes. Each node is restricted to membership in one module. This creates modules of variables that are highly connected to each other. We then calculated the normalized network modularity score (Newman, 2006), which quantifies the strength of the connections within and between modules. A higher modularity score indicates a network with high within-module connectivity and low between-module connectivity. We calculated

the modularity score for the overall network graph and each of the four graphs based on community strata of burden of childhood obesity (see Table 2). We used the pairwise correlation between variables (nodes) as a weight in the computation of modularity. Fourth, we calculated an overall measure of network centrality by calculating the average network degree (Barrat et al., 2004). In a weighted undirected network like ours, average network degree is the mean of all pairwise correlations (Barrat et al., 2004). A high average network degree represents a network that has an overall tighter correlation between all nodes. We calculated average network degree for the overall network graph and each of the four network graphs based on obesity prevalence (see Table 2).

Fifth, we examined the association between the centrality of a node and its correlation with the prevalence of childhood obesity. For this, we plotted the degree centrality of each node against that node's correlation with the prevalence of childhood obesity (Fig. 3).
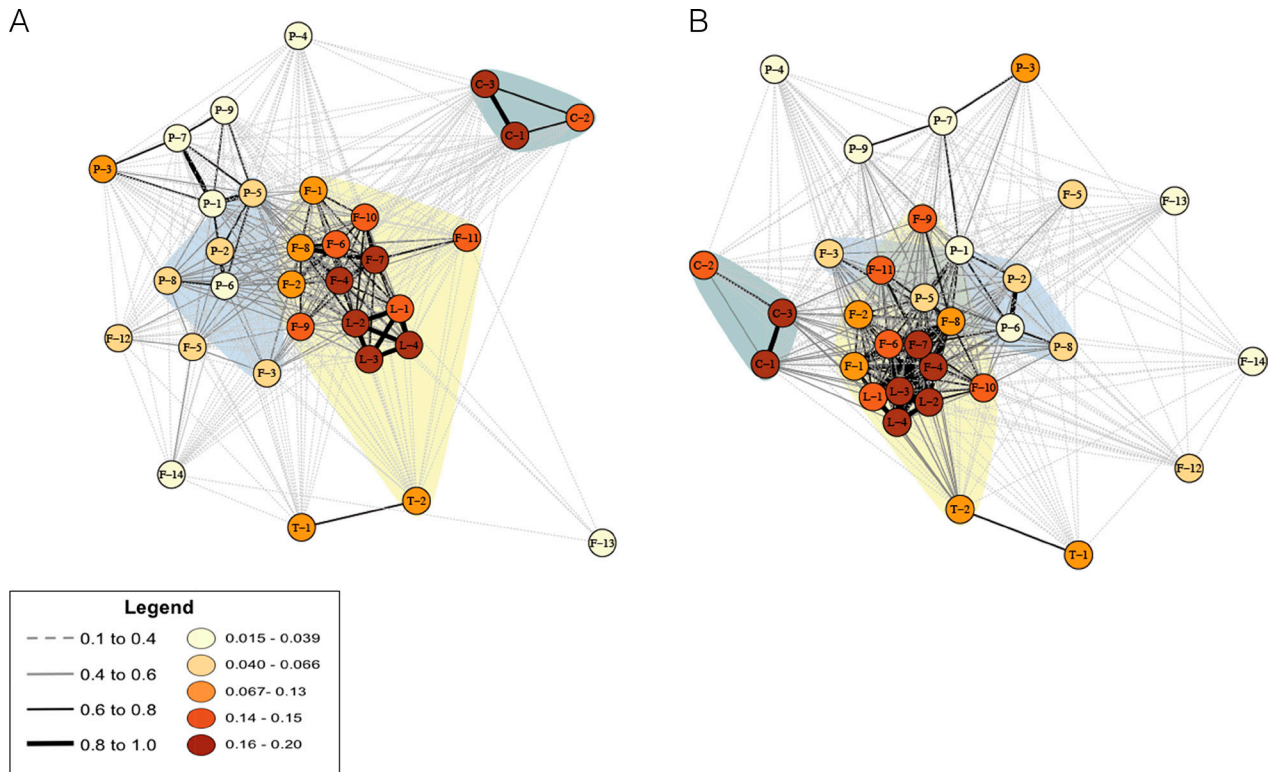
Figure 2: **Network Graphs for 1288 Communities in Pennsylvania, by Quartile of Percent of Children at or Above the 85th Percentile of BMIz**. In communities in the lowest quartile of percent of children who are overweight or obese (A: left), community features appear to be less tightly clustered, i.e., co-occur less often, than in communities in the highest quartile of community BMIz (B: right).

Results

The purpose of this analysis was to apply network methodology to characterize patterns of linkages and interactions among obesity-related environ-mental features among communities in Pennsylvania. Figure 1 is a graph of the network of connections (pairwise correlations) between nodes (obesity-related features) in 1,288 Pennsylvania communities.

Table 2. Network modularity and average network degree in the overall network and by quartile of prevalence of childhood obesity.

	Quartiles of prevalence of childhood obesity				
	Quartile 1	Quartile 2	Quartile 3	Quartile 4	Overall network
Network modularity	0.1891	0.2681	0.1181	0.0982	0.1496
Average network degree	0.3318	0.3533	0.3584	0.3623	0.3507

Note: Quartile 1 represents communities with the lowest prevalence of child overweight and obesity, and quartile 4 represents communities with the highest prevalence of child overweight and obesity. Higher modularity indicates a network with high within-module connectivity and low between-module connectivity. Average network degree is the mean of all pairwise correlations.
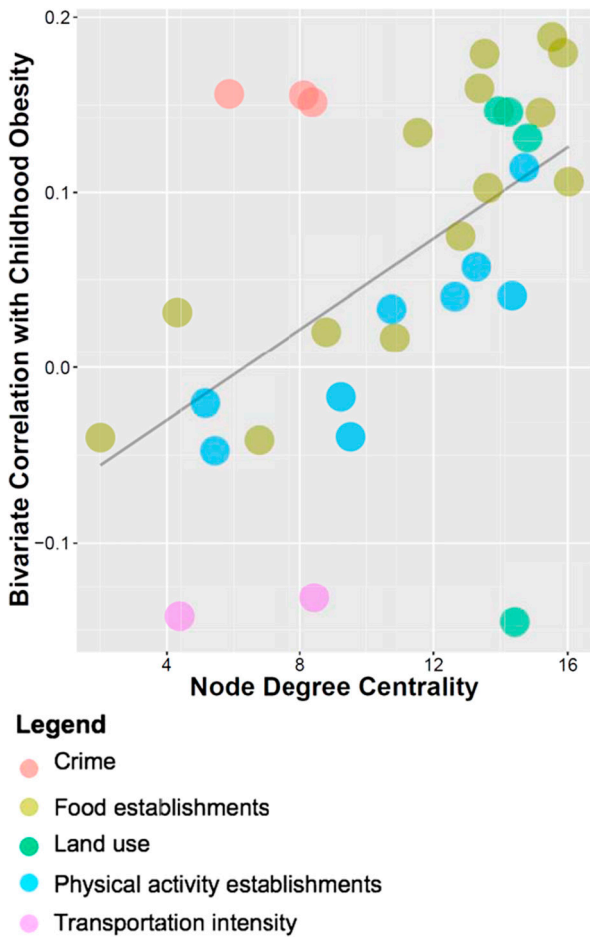
Figure 3: **Association of degree centrality of each community feature with prevalence of overweight and obesity among children**. Correlation between community features and body mass index is stronger for more central variables of the obesity-related network features ($R = 0.51$).

The graph illustrates three important network characteristics. First, three clusters of tightly-connected variables were identified using the walktrap method. A cluster of the three crime-related variables (rates per 100,000 persons of crimes against property, crimes against persons, and all Part I offenses) can be seen (green shading), and is weakly linked to the main network. This suggests that communities with high rates of violent crime (i.e., assault) also have high rates of crime against property (i.e., arson). Crime rates appear to be moderately correlated with obesity rates as indicated by the dark color of the crime-related nodes. A second cluster is identified consisting of features representing land use patterns, transportation, and food

establishment density (yellow shading). We believe this represents the spatial clustering that occurs in the context of suburban sprawl with co-location of establishments in high-volume transportation corridors. The nodes at the heart of this cluster include household density (per square mile) and all food establishments per square mile. This second cluster appears to be the tightest. Eleven of the 14 nodes have an above average absolute correlation with obesity. The model identified a third cluster (pale blue shading) consisting mostly of features that describe the physical activity environment. These include diversity of physical activity establishments, outdoor recreational facilities per square mile, snack stores (e.g., donuts, pretzels, ice cream) per square mile, indoor recreation centers per square mile, all physical activity establishments per square mile, indoor fitness and recreational facilities per street mile, and indoor recreational clubs and facilities per square mile.

In both the second and third clusters, the nodes that are most highly correlated with obesity (indicated by darker node color), are more central in the network overall, as well as within each cluster. Not all of the food or physical activity nodes are clustered. At the edge of the graph we see several food or physical activity nodes that are not as tightly coupled (including parks and big box stores). Finally, the overall structure of the network suggests that elements of these communities are geographically clustered and are not randomly dispersed across communities, especially features of the physical, food, and land use environments.

Next, we were interested in whether the structure of this network of features varied across strata of community obesity burden. Figure 2 shows the result of running a similar network model separately by quartile of percent of children at or above the 85th percentile on BMI-z, a threshold widely considered to be indicative of overweight and obesity burden. Among communities in the lowest quartile of obesity prevalence (Fig. 2A), community features appear to be less tightly connected than in communities in the highest quartile of obesity prevalence (Fig. 2B). This is also described by the higher modularity in Table 2. For example, among lower obesity prevalence communities, crime is weakly linked to the land use-food- physical activity cluster; but in higher obesity prevalence communities, crime is more tightly linked to this cluster. It is not just that quantities of these features are larger in heavier communities, but that the connections between features are also altered: communities that give rise to higher rates of childhood obesity are structured differently than those with less child obesity.

Table 2 shows the results of the network structure analysis, overall, and stratified by quartile of obesity. The overall network has a positive modularity of 0.15, indicating that the nodes (environmental features) show a degree of clustering (as compared to a random distribution of nodes with no clustering). In the analysis stratified by prevalence of childhood obesity, communities in the 1st and 2nd quartile (thinnest communities) show a higher modularity compared to communities in the 3rd and 4th quartile (heaviest communities) (modularity of 0.19 and 0.27 *vs*. 0.12 and 0.09, respectively). This means that the modules of variables in thinner communities are either more clustered within each module or have weaker connections to variables in other modules, and that in the heaviest communities variables (nodes) exhibit a lower degree of clustering in modules (as can be seen in Fig. 2). For example, a comparison of the two panels in Figure 2 demonstrates that the crime-related cluster shown in green has fewer strong ties (shown by darker lines) to the center of the network in the thinner communities on the left panel compared to the heavier communities in the right panel. Similarly, the average network degree is higher in the heaviest communities (degree = 0.362) compared with the thinnest communities (degree = 0.332), representing higher average correlation (i.e., stronger connections), between variables in communities with higher prevalence of childhood obesity.

Figure 3 shows the relationship between the degree centrality of each community feature (node) with the bivariate correlation of that feature with childhood overweight and obesity prevalence (percent of kids above the 85th BMI percentile). Each dot represents one of the 32 community features. The correlation between the degree of each feature and its correlation with the prevalence of childhood obesity is positive ($r = 0.51$), indicating that more 'central' variables have a stronger association with the outcome. For example, fresh fruit and vegetable stands per square mile has a low correlation with community obesity, and can be seen in Figure 1 as a variable far from the center of the network and with only a few weak ties into the rest of the network.

Discussion

We applied network methodology in order to describe linkages between community features associated with obesity. We used network analysis to characterize the obesogenic environment: instead of treating individual features of communities in isolation, this method honors the interactions and spatial co-occurrence that make up this landscape of obesity risk.

This work suggests that (i) there are identifiable clusters of environmental features; (ii) that the level of connectivity and structure of features in the network may be informative; and (iii) that features more highly associated with obesity are more likely to be central in the network of community features. Three clusters were identified in the overall network: a cluster of crime-related variables that was weakly linked into the main network, and food and land use and physical activity clusters, respectively. In communities stratified by prevalence of childhood obesity, the structure and overall connectivity of the network appeared to differ by level of obesity. Not only are the values of these attributes different in the heaviest and thinnest communities, but also the patterns of connections are different. We also found that centrality alone, as measured by degree, is correlated with obesity. Obesity-related features are therefore more tightly geographically clustered. This may be evidence of synergy between features of the obesogenic environment, of non-independent features of communities that join forces to shape obesity risk.

Understanding and intervening on the drivers of the obesity epidemic is a challenge for obesity researchers and policy makers. Obesity is complex and has multiple drivers at the individual, community, and state and national levels (Huang et al., 2009). Traditional methods such as regression models fail to account for interaction between multiple factors at multiple scales, the complexity and importance of contextual factors, and feedback loops and other dynamic processes (Hammond, 2009). Although our work is preliminary, it suggests that systems approaches to obesity may be useful for characterizing linkages among features of the environment. Despite the recognition that environmental features of communities play a strong role in the obesity epidemic, network methods to characterize linkages between attributes of communities has been underutilized. The structure and strength of these linkages may provide evidence for geographic areas or types of clusters of features that would be most efficient for intervention.

Network methods, especially graphical methods, could be used to help set priorities for obesity-related interventions in communities. For example, food establishments exhibited both high centrality (as measured by degree) in our network and high correlation with childhood obesity (Fig. 3). Using these network graphs (e.g., Fig. 2), we can narrow in on features such as these that may have far reaching effects, if intervened upon. This is consistent with the

literature on 'food swamps' and 'food deserts' – but helps to prioritize interventions in this area because these features are more central. This could point to the effectiveness of intervening on such variables that are highly central in the network, and thus may have more far-reaching effects than intervention on less central variables. Network methods may help identify such synergistic actors that could have large effects on obesity due to their connections to other variables.

In particular, our work points toward possible interventions regarding community-zoning policies. Our network graphs show tight clusters of food-related (e.g., grocery and convenience stores, fast food and full-service restaurants) and land use (e.g., road block length, household density) features that are strongly correlated with obesity. Restructuring the community environment may be a promising avenue for obesity prevention. Considering communities to be complex systems where multiple interrelated phenomena act together to create an obesogenic environment, these methods also push us to consider intervening on not just the environmental features themselves, but also on the linkages between features. This is a new way to approach the obesity epidemic – by looking for factors that may be linking features, or that can be manipulated to disrupt harmful connections. For example, the crime-related cluster is more tightly linked to the network among communities with more childhood obesity. Further research into the underlying causes of this linkage (and why it differs in communities stratified by childhood obesity prevalence) may illuminate important drivers of the obesity epidemic.

This work also has methodological implications for obesity research. Future work should explore the mechanisms for how these clusters are associated with increased obesity prevalence, and whether interventions on features in this network change the network structure itself. This future research should consider the relationships, or clustering, of these features. Evaluating independent associations between any single feature and obesity rates would ignore the complex inter-relations this work has highlighted. Other methods that acknowledge these clusters of features, such as latent variable methods (Nau et al., 2015), may be more appropriate to honor the way that environmental features cluster together and to uncover unobserved sources of the correlation observed in this network.

We have data from a large and diverse geographic area that includes urban, rural, and suburban communities. However, this analysis is exploratory. We are not able to rule out the possibility that popula-

tion density and development may be a common cause of many of the variables we selected. This is potentially a source of bias or a possible explanation for the clustering of features of the environment on which our study is based. It is widely recognized that obesity-related characteristics of communities are geographically correlated. The reasons for those correlations are not well understood. Our results, we believe, support the utility of network methods for the study of environments that are not formed randomly, but which are shaped by diverse market and demographic forces that may be important in driving spatial variation in obesity rates.

Conclusion

Network analysis may be a useful tool for evaluating obesogenic environments and other questions of interest in epidemiology. This preliminary analysis suggests that patterns of clustering and connections between features of the environment are important. Land use and food features are strongly linked (especially in 'heavier' communities), and features are more highly clustered in communities with higher average BMI. Network methods can illuminate patterns of linkages and key factors in obesogenic environments. Network position (centrality) is correlated with average BMI. Ultimately, the goal of this type of analysis would be to identify highly connected community features that can be used as levers of intervention to reduce population rates of obesity.

Acknowledgements

Literature Cited

Ali, M.M., Amialchuk, A. and Rizzo, J.A. 2012. The influence of body weight on social network ties among adolescents. *Economics and Human Biology* 10(1): 20–34.

Ali, M.M., Amialchuk, A., Gao, S. and Heiland, F. 2012. Adolescent weight gain and social networks: Is there a contagion effect?. *Applied Economics* 44(23): 2969–83.

Barabasi, A.L. 2007. Network medicine – from obesity to the 'diseasome'. *The New England Journal of Medicine* 357(4): 404–7, doi: 10.1056/NEJMe078114.

Barabasi, A.L. 2009. Scale-free networks: A decade and beyond. *Science* 325(5939): 412–3.

Barabasi, A.L. 2012. Network science: Luck or reason. *Nature* 489(7417): 507–8.

Barabasi, A.L. 2013. Network science. *Philosophical Transactions of the Royal Society A Mathematical Physicla and Engineering Science* 371(1987): 20120375.

Barrat, A., Barthélemy, M., Pastor-Satorras, R. and Vespignani, A. 2004. The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America* 101(11): 3747–52, doi: 10.1073/pnas.0400087101.

Berger, E., Vega, N., Vidal, H. and Geloen, A. 2012. Gene network analysis leads to functional validation of pathways linked to cancer cell growth and survival. *Biotechnology Journal* 7(11): 1395–404.

Blanchflower, D.G., Landeghem, B. and Oswald, A.J. 2009. Imitative obesity and relative utility. *Journal of the European Economic Association* 7(2–3): 528–38.

Brewis, A.A., Hruschka, D.J. and Wutich, A. 2011. Vulnerability to fat-stigma in women's everyday relationships. *Social Science and Medicine* 73(4): 491–7.

Burke, M.A. and Heiland, F. 2007. Social dynamics of obesity. *Economic Inquiry* 45(3): 571–91.

Chen, Z. and Zhang, W. 2013. Integrative analysis using module-guided random forests reveals correlated genetic factors related to mouse weight. *PLOS Computational Biology* 9(3): e1002956.

Christakis, N.A. and Fowler, J.H. 2007. The spread of obesity in a large social network over 32 years. *The New England Journal of Medicine* 357(4): 370–9.

Crandall, C.S. 1988. Social contagion of binge eating. *Journal of Personality and Social Psychology* 55(4): 588–98.

Csardi, G. and Nepusz, T. 2006. The igraph software package for complex network research. *Inter-Journal Complex Systems* 1695, 1–9.

Dallman, M.F., Pecoraro, N., Akana, S.F., La Fleur, S.E., Gomez, F., Houshyar, H., Bell, M.E., Bhatnagar, S., Laugero, K.D. and Manalo, S. 2003. Chronic stress and obesity: A new view of 'comfort food'. *Proceedings of Natlional Academy of Science of the United States of America* 100(20): 11696–701.

Dallman, M.F., Pecoraro, N.C., La Fleur, S.E., Warne, J.P., Ginsberg, A.B., Akana, S.F., Laugero, K.C., Houshyar, H., Strack, A.M., Bhatnagar, S. and Bell, M.E. 2006. Glucocorticoids, chronic stress, and obesity. *Progress in Brain Research* 153, 75–105.

de la Haye, K., Robins, G., Mohr, P. and Wilson, C. 2011. Homophily and contagion as explanations for weight similarities among adolescent friends. *Journal of Adolescent Health* 49(4): 421–7.

El-Sayed, A.M., Scarborough, P., Seemann, L. and Galea, S. 2012. Social network analysis and agent-based modeling in social epidemiology. *Epidemiologic Perspectives and Innovations* 9(1): 1.

Epskamp, S., Cramer, A.O.J., Waldorp, L.J., Schmittmann, V.D. and Borsboom, D. 2012. qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software* 48(4): 1–18.

Finegood, D.T. 2011. The complex systems science of obesity. in Cawley, J. (Ed.), *The Oxford Handbook of Social Science of Obesity*, Oxford University Press, New York, 208–36.

Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C. and Raichle, M.E. 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America* 102(27): 9673–8, doi: 10.1073/pnas.0504136102.

Fruchterman, T.M.J. and Reingold, E.M. 1991. Graph drawing by force-directed placement. *Software: Practice and Experience* 21(11): 1129–64, doi: 10.1002/spe.4380211102.

Galea, S., Riddle, M. and Kaplan, G.A. 2010. Causal thinking and complex system approaches in epidemiology. *International Journal of Epidemiology* 39(1): 97–106.

Gesell, S.B., Tesdahl, E. and Ruchman, E. 2012. The distribution of physical activity in an after-school friendship network. *Pediatrics* 129(6): 1064–71, doi: 10.1542/peds.2011-2567.

Gill, R., Datta, S. and Datta, S. 2014. Differential network analysis in human cancer research. *Current Pharmaceutical Design* 20(1): 4–10.

Goh, K.I., Cusick, M.E., Valle, D., Childs, B., Vidal, M. and Barabasi, A.L. 2007. The human disease network. *Proceedings of Natlional Academy of Science of the United States of America* 104(21): 8685–90, doi: 10.1073/pnas.0701361104.

Hammond, R. 2009. Complex systems modeling for obesity research. *Preventing Chronic Disease* 6(3): 1–10.

Hammond, R.A. 2010. Social influence and obesity. *Current Opinion in Endocrinology, Diabetes and Obesity* 17(5): 467–71.

Hammond, R.A. and Ornstein, J.T. 2014. A model of social influence on body mass index. *Annals of the New York Academy of Science* 1331, 34–42.

Hidalgo, C.A. and Castañer, E.E. 2015. The amenity space and the evolution of neighborhoods. arXiv:1509.02868 [physics.soc-ph]

Hidalgo, C.A., Blumm, N., Barabasi, A.L. and Christakis, N.A. 2009. A dynamic network approach for the study of human phenotypes. *PLOS Computational Biology* 5(4): e1000353, doi: 10.1371/journal.pcbi.1000353.

Hill, A.L., Rand, D.G., Nowak, M.A. and Christakis, N.A. 2010. Infectious disease modeling of social contagion in networks. *PLOS Computational Biology* 6(11): e1000968.

Hill, J.O. and Peters, J.C. 1998. Environmental contributions to the obesity epidemic. *Science* 280(5368): 1371–4.

Huang, T.T., Drewnosksi, A., Kumanyika, S. and Glass, T.A. 2009. A systems-oriented multilevel framework for addressing obesity in the 21st century. *Preventing Chronic Disease* 6(3): A82.

Jeong, H., Mason, S.P., Barabasi, A.L. and Oltvai, Z.N. 2001. Lethality and centrality in protein networks. *Nature* 411(6833): 41–2, doi: 10.1038/35075138.

Leroux, J.S., Moore, S. and Dubé, L. 2013. Beyond the 'I' in the obesity epidemic: A review of social relational and network interventions on obesity. *Journal of Obesity* 2013, 348249.

McGlashan, J., Johnstone, M., Creighton, D., de la Haye, K. and Allender, S. 2016. Quantifying a systems map: Network analysis of a childhood obesity causal loop diagram. *PLOS ONE* 11(10): e0165459, doi: 10.1371/journal.pone.0165459.

Marks, J., Barnett, L.M., Foulkes, C., Hawe, P. and Allender, S. 2013. Using social network analysis to identify key child care center staff for obesity prevention interventions: A pilot study. *J Obes* 2013, 919287.

Mutation Consequences and Pathway Analysis working group of the International Cancer Genome Consortium 2015. Pathway and network analysis of cancer genomes. *Nature Methods* 12(7): 615–21.

Nau, C., Ellis, H., Huang, H., Schwartz, B.S., Hirsch, A., Bailey-Davis, L., Kress, A.M., Pollak, J. and Glass, T.A. 2015. Exploring the forest instead of the trees: An innovative method for defining obesogenic and obesoprotective environments. *Health Place* 35, 136–46, doi: 10.1016/j.healthplace.2015.08.002.

Newman, M.E.J. 2006. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America* 103(23): 8577–82, doi: 10.1073/pnas.0601602103.

Pons, P. and Latapy, M. 2005. Computing communities in large networks using random walks. in Yolum, P., Güngör, T., Gürgen, F. and Özturan, C. (Eds), *Computer and Information Sciences – ISCIS 2005: Proceedings of the 20th International Symposium, Istanbul, Turkey, October 26–28, 2005* Springer, Berlin, Heidelberg, 284–93.

Schwartz, B.S., Stewart, W.F., Godby, S., Pollak, J., Dewalle, J., Larson, S., Mercer, D.G. and Glass, T.A. 2011. Body mass index and the built and social environments in children and adolescents using electronic health records. *American Journal of Preventive Medicine* 41(4): e17–e28, doi: 10.1016/j.amepre.2011.06.038.

Simpkins, S.D., Schaefer, D.R., Price, C.D. and Vest, A.E. 2013. Adolescent friendships, BMI, and physical activity: Untangling selection and influence through longitudinal social network analysis. *Journal of Research Adolescence* 23(3): doi: 10.1111/j.1532-7795.2012.00836.x.

Stites, E.C., Trampont, P.C., Ma, Z. and Ravichandran, K.S. 2007. Network analysis of oncogenic Ras activation in cancer. *Science* 318(5849): 463–7.

Swinburn, B., Egger, G. and Raza, F. 1999. Dissecting obesogenic environments: The development and application of a framework for identifying and prioritizing environmental interventions for obesity. *Preventive Medicine* 29(6 Pt 1): 563–70, doi: 10.1006/pmed.1999.0585.

Zhang, B. and Horvath, S. 2005. A general framework for weighted gene co-expression network analysis. *Statistical Applications in Genetics and Molecular Biology* Epub 2005.

Zhou, X., Menche, J., Barabasi, A.L. and Sharma, A. 2014. Human symptoms-disease network. Nature Communication 5, 4212, doi: 10.1038/ncomms5212.