The Spread of Obesity in a Large Social Network Over 32 Years.

The Harvard community has made this article openly available. **Please share** how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th>Citation</th>
<th>Christakis, Nicholas A. and James H. Fowler. 2007. The spread of obesity in a large social network over 32 years. The New England Journal of Medicine 357, no. 4: 370-379.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Published Version</td>
<td><a href="http://dx.doi.org/10.1056/NEJMs066082">http://dx.doi.org/10.1056/NEJMs066082</a></td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:3710802">http://nrs.harvard.edu/urn-3:HUL.InstRepos:3710802</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
The Spread of Obesity in a Large Social Network Over 32 Years

Nicholas A. Christakis, M.D., Ph.D., M.P.H., and James H. Fowler, Ph.D.

ABSTRACT

BACKGROUND
The prevalence of obesity has increased substantially over the past 30 years. We performed a quantitative analysis of the nature and extent of the person-to-person spread of obesity as a possible factor contributing to the obesity epidemic.

METHODS
We evaluated a densely interconnected social network of 12,067 people assessed repeatedly from 1971 to 2003 as part of the Framingham Heart Study. The body-mass index was available for all subjects. We used longitudinal statistical models to examine whether weight gain in one person was associated with weight gain in his or her friends, siblings, spouse, and neighbors.

RESULTS
Discernible clusters of obese persons were present in the network at all time points, and the clusters extended to three degrees of separation. These clusters did not appear to be solely attributable to the selective formation of social ties among obese persons. A person’s chances of becoming obese increased by 57% (95% confidence interval [CI], 6 to 123) if he or she had a friend who became obese in a given interval. Among pairs of adult siblings, if one sibling became obese, the chance that the other would become obese increased by 40% (95% CI, 21 to 60). If one spouse became obese, the likelihood that the other spouse would become obese increased by 37% (95% CI, 7 to 73). These effects were not seen among neighbors in the immediate geographic location. Persons of the same sex had relatively greater influence on each other as compared with those of the opposite sex. The spread of smoking cessation did not account for the spread of obesity in the network.

CONCLUSIONS
Network phenomena appear to be relevant to the biologic and behavioral trait of obesity, and obesity appears to spread through social ties. These findings have implications for clinical and public health interventions.
THE PREVALENCE OF OBESITY HAS INCREASED FROM 23% TO 31% OVER THE RECENT PAST IN THE UNITED STATES, AND 66% OF ADULTS ARE NOW OVERWEIGHT.\textsuperscript{1,2} PROPOSED EXPLANATIONS FOR THE OBESITY EPIDEMIC INCLUDE SOCIETAL CHANGES THAT PROMOTE BOTH INACTIVITY AND FOOD CONSUMPTION.\textsuperscript{3} THE FACT THAT THE INCREASE IN OBESITY DURING THIS PERIOD CANNOT BE EXPLAINED BY GENETICS\textsuperscript{4,5} AND HAS OCCURRED AMONG ALL SOCIOECONOMIC GROUPS\textsuperscript{1} PROVIDES SUPPORT FOR A BROAD SET OF SOCIAL AND ENVIRONMENTAL EXPLANATIONS. SINCE DIVERSE PHENOMENA CAN SPREAD WITHIN SOCIAL NETWORKS,\textsuperscript{6-10} WE CONDUCTED A STUDY TO DETERMINE WHETHER OBESITY MIGHT ALSO SPREAD FROM PERSON TO PERSON, POSSIBLY CONTRIBUTING TO THE EPIDEMIC, AND IF SO, HOW THE SPREAD MIGHT OCCUR.

WHEREAS OBESITY HAS BEEN STIGMATIZED IN THE PAST, ATTITUDES MAY BE CHANGING.\textsuperscript{11,12} TO THE EXTENT THAT OBESITY IS A PRODUCT OF VOLUNTARY CHOICES OR BEHAVIORS, THE FACT THAT PEOPLE ARE EMBEDDED IN SOCIAL NETWORKS AND ARE INFLUENCED BY THE EVIDENT APPEARANCE AND BEHAVIORS OF THOSE AROUND THEM SUGGESTS THAT WEIGHT GAIN IN ONE PERSON MIGHT INFLUENCE WEIGHT GAIN IN OTHERS. HAVING OBSESE SOCIAL CONTACTS MIGHT CHANGE A PERSON’S TOLERANCE FOR BEING OBSESE OR MIGHT INFLUENCE HIS OR HER ADOPTION OF SPECIFIC BEHAVIORS (E.G., SMOKING, EATING, AND EXERCISING). IN ADDITION TO SUCH STRICTLY SOCIAL MECHANISMS, IT IS PLAUSIBLE THAT PHYSIOLOGICAL IMITATION MIGHT OCCUR; AREAS OF THE BRAIN THAT CORRESPOND TO ACTIONS SUCH AS EATING FOOD MAY BE STIMULATED IF THESE ACTIONS ARE OBSERVED IN OTHERS.\textsuperscript{13} EVEN INFECTIOUS CAUSES OF OBESITY ARE CONCEIVABLE.\textsuperscript{14,15}


METHODS

SOURCE DATA

The Framingham Heart Study was initiated in 1948, when 5209 people were enrolled in the original cohort.\textsuperscript{16} The Framingham Offspring Study began in 1971, when most of the children of members of the original cohort and their spouses were enrolled in the offspring cohort.\textsuperscript{17} There has been almost no loss to follow-up other than death in this cohort of 5124 people; only 10 people left the study. In 2002, the third-generation cohort, consisting of 4095 children of the offspring cohort, was initiated. All participants undergo physical examinations (including measurements of height and weight), and they complete written questionnaires at regular intervals.

NETWORK ASCERTAINMENT

For our study, we used the offspring cohort as the source of 5124 key subjects, or “egos,” as they are called in social-network analysis. Any persons to whom the egos are linked — in any of the Framingham Heart Study cohorts — can, however, serve as “alters.” Overall, 12,067 living egos and alters were connected at some point during the study period (1971 to 2003).

To create the network data set, we entered information about the offspring cohort into a computer. This information was derived from archived, handwritten administrative tracking sheets that had been used since 1971 to identify people close

<table>
<thead>
<tr>
<th>Glossary</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ego:</strong> The person whose behavior is being analyzed.</td>
</tr>
<tr>
<td><strong>Alter:</strong> A person connected to the ego who may influence the behavior of the ego.</td>
</tr>
<tr>
<td><strong>Node:</strong> An object that may or may not be connected to other objects in a network. In this study, nodes represent people in the Framingham Heart Study cohorts.</td>
</tr>
<tr>
<td><strong>Tie:</strong> A connection between two nodes that can be either one way (directed) or two ways (bilateral). In this study, all family ties (e.g., between siblings, parents, and spouses) are bilateral, but friendship ties are directional since a subject may identify someone as a friend who does not necessarily identify that person as a friend in return.</td>
</tr>
<tr>
<td><strong>Degree of separation:</strong> The social distance between two people as measured by the smallest number of intermediaries between an ego and other members of the network. For a given ego, alters are degree 1, since they are directly connected to the ego. Nodes that are connected to the alters but not to the ego are degree 2 (alters’ alters). Nodes that are connected to the alters’ alters but not to the ego are degree 3, and so on.</td>
</tr>
<tr>
<td><strong>Homophily:</strong> The tendency for people to choose relationships with people who have similar attributes.</td>
</tr>
<tr>
<td><strong>Induction:</strong> The spread of a behavior or trait from one person to another.</td>
</tr>
<tr>
<td><strong>Cluster:</strong> A group of nodes, each of which is connected to at least one other node in the group.</td>
</tr>
<tr>
<td><strong>Connected component:</strong> Part of a social network in which all persons have a social tie to at least one other person and no person is connected to a member of any other component.</td>
</tr>
</tbody>
</table>
to the study participants in order to facilitate follow-up. These sheets contain valuable, previously unused social-network information because they systematically and comprehensively identify relatives and friends named by the ego. The tracking sheets provide complete information about all first-order relatives (parents, spouses, siblings, and children), whether they are alive or dead, and at least one “close friend” at each of seven examinations between 1971 and 2003. The examinations took place during 3-year intervals centered in 1973, 1981, 1985, 1989, 1992, 1997, and 1999.

Many of the named alters on these sheets also were members of Framingham Heart Study cohorts. Detailed home addresses were also recorded at each time point; we used this information to calculate the geographic distance between people. This newly computerized database thus identifies the network links among participants at each examination and longitudinally from one examination to the next. As a person’s family changed because of birth, death, marriage, or divorce, and as contacts changed because of residential moves or new friendships, this information was recorded. Furthermore, dates of birth and death were available from separate Framingham Heart Study files.

Overall, there were 38,611 observed social and family ties to the 5124 egos, yielding an average of 7.5 ties per ego (not including neighbors). For example, 83% of the spouses of egos were directly and repeatedly observed at the time of examination, and 87% percent of egos with siblings had at least one sibling in the network. For 10% of the egos, an immediate neighbor also participated in the study; more expansive definitions of neighbors yielded similar results.

A total of 45% of the 5124 egos were connected through friendship to another person in the network. There were 3604 unique, observed friendships, for an average of 0.7 friendship tie per ego. Because friendship identifications are directional, we studied three different kinds of friendships: an “ego-perceived friendship,” in which the relation is reciprocal. We hypothesized that a friend’s social influence on an ego would be affected by the type of friendship, with the strongest effects occurring in mutual friendships, followed by ego-perceived friendships, followed by alter-perceived friendships. Our reasoning was that the person making the identification esteems the other person and may wish to emulate him or her.

We included only persons older than 21 years of age at any observation point and subsequently. At the inception of the study, 53% of the egos were women, the mean age of the egos was 38 years (range, 21 to 70), and their mean educational level was 13.6 years (range, no education to ≥17 years of education).

The study data are available from the Framingham Heart Study. The study was approved by the institutional review board at Harvard Medical School; all subjects provided written informed consent.

STATISTICAL ANALYSIS

We graphed the network with the use of the Kamada-Kawai algorithm in Pajek software. We generated videos of the network by means of the Social Network Image Animator. We examined whether our data conformed to theoretical network models such as the small-world, scale-free, and hierarchical types (see the Supplementary Appendix, available with the full text of this article at www.nejm.org).

We defined obesity as a body-mass index (the weight in kilograms divided by the square of the height in meters) of 30 or more. Analyses in which the body-mass index was a continuous variable did not yield different results.

We considered three explanations for the clustering of obese people. First, egos might choose to associate with like alters (“homophily”). Second, egos and alters might share attributes or jointly experience unobserved contemporaneous events that cause their weight to vary at the same time (confounding). Third, alters might exert social influence or peer effects on egos (“induction”). Distinguishing the interpersonal induction of obesity from homophily requires dynamic, longitudinal network information about the emergence of ties between people (“nodes”) in a network and also about the attributes of nodes (i.e., repeated measures of the body-mass index).

The basic statistical analysis involved the specification of longitudinal logistic-regression models in which the ego’s obesity status at any given examination or time point (t+1) was a function of various attributes, such as the ego’s age, sex, and
educational level; the ego’s obesity status at the previous time (t); and most pertinent, the alter’s obesity status at times t and t+1. We used generalized estimating equations to account for multiple observations of the same ego across examinations and across ego–alter pairs. We assumed an independent working correlation structure for the clusters.

The use of a time-lagged dependent variable (lagged to the previous examination) eliminated serial correlation in the errors (evaluated with a Lagrange multiplier test) and also substantially controlled for the ego’s genetic endowment and any intrinsic, stable predisposition to obesity. The use of a lagged independent variable for an alter’s weight status controlled for homophily. The key variable of interest was an alter’s obesity at time t+1. A significant coefficient for this variable would suggest either that an alter’s weight affected an ego’s weight or that an ego and an alter experienced contemporaneous events affecting both their weights. We estimated these models in varied ego–alter pair types.

To evaluate the possibility that omitted variables or unobserved events might explain the associations, we examined how the type or direction of the social relationship between the ego and the alter affected the association between the ego’s obesity and the alter’s obesity. For example, if unobserved factors drove the association between the ego’s obesity and the alter’s obesity, then the directionality of friendship should not have been relevant.

We evaluated the role of a possible spread in smoking-cessation behavior as a contributor to the spread of obesity by adding variables for the smoking status of egos and alters at times t and t+1 to the foregoing models. We also analyzed the role of geographic distance between egos and alters by adding such a variable.

We calculated 95% confidence intervals by simulating the first difference in the alter’s contem-

---

**Figure 1.** Largest Connected Subcomponent of the Social Network in the Framingham Heart Study in the Year 2000. Each circle (node) represents one person in the data set. There are 2200 persons in this subcomponent of the social network. Circles with red borders denote women and circles with blue borders denote men. The interior color of the circles indicates the person’s obesity status: yellow denotes an obese person (body-mass index, >30) and green denotes a nonobese person. The size of each circle is proportional to the person's body-mass index. The colors of the ties between the nodes indicate the relationship between them: purple denotes a friendship or marital tie and orange denotes a familial tie.
Figure 2. Part of the Social Network from the Framingham Heart Study with Information about Body-Mass Index, by Year.

Each circle (node) represents one person in the data set. Circles with red borders denote women and circles with blue borders denote men. The size of each circle is proportional to the person’s body-mass index. The interior color of the circles indicates the person’s obesity status: yellow denotes an obese person (body-mass index, >30) and green denotes a nonobese person. The colors of the ties between the circles indicate the relationship between them: purple denotes a friendship or marital tie and orange denotes a familial tie. The disappearance of a circle indicates the person’s death, and the disappearance of a tie between the circles indicates that the relationship between the two persons no longer exists. The largest connected subcomponent of the whole network and the change in obesity over the 32-year study period, are shown in an animation that is available with the full text of this article at www.nejm.org.
poraneous obesity (changing from 0 to 1), using 1000 randomly drawn sets of estimates from the coefficient-covariance matrix and assuming mean values for all other variables. All tests were two-tailed. The sensitivity of the results was assessed with multiple additional analyses (see the Supplementary Appendix).

**RESULTS**

Figure 1 depicts the largest connected subcomponent of the social network in the year 2000. This network is sufficiently dense to obscure much of the underlying structure, although regions of the network with clusters of obese or nonobese persons can be seen. Figure 2 illustrates the spread of obesity between adjoining nodes in a part of the network over time. A video (available with the full text of this article at www.nejm.org) depicts the evolution of the largest component of the network and shows the progress of the obesity epidemic over the 32-year study period.

Figure 3A characterizes clusters within the entire network more formally. To quantify these clusters, we compared the whole observed network with simulated networks with the same network topology and the same overall prevalence of obesity as the observed network, but with the incidence of obesity randomly distributed among the nodes (in what we call “random body-mass–index networks”). If clustering is occurring, then the probability that an alter will be obese, given that an ego is known to be obese, should be higher in the observed network than in the random body-mass–index networks. What we call the “reach” of the clusters is the point, in terms of an alter’s degree of separation from any given ego, at which the probability of an alter’s obesity is no longer related to whether the ego is obese. In all of the examinations (from 1971 through 2003), the risk of obesity among alters with simulated networks with the same network topology and the same overall prevalence of obesity as the observed network, but with the incidence of obesity randomly distributed among the nodes (in what we call “random body-mass–index networks”). If clustering is occurring, then the probability that an alter will be obese, given that an ego is known to be obese, should be higher in the observed network than in the random body-mass–index networks. What we call the “reach” of the clusters is the point, in terms of an alter’s degree of separation from any given ego, at which the probability of an alter’s obesity is no longer related to whether the ego is obese. In all of the examinations (from 1971 through 2003), the risk of obesity among alters
The closeness of friendship is relevant to the spread of obesity. Persons in closer, mutual friendships have more of an effect on each other than persons in other types of friendships. The dependent variable in each model is the obesity of the ego. Independent variables include a time-lagged measurement of the ego's obesity; the obesity of the alter; a time-lagged measurement of the alter's obesity; the ego's age, sex, and level of education; and indicator variables (fixed effects) for each examination. Full models and equations are available in the Supplementary Appendix. Mean effect sizes and 95% confidence intervals were calculated by simulating the first difference in the contemporaneous obesity of the alter (changing from 0 to 1) with the use of 1000 randomly drawn sets of estimates from the coefficient-covariance matrix and with all other variables held at their mean values.

who were connected to an obese ego (at one degree of separation) was about 45% higher in the observed network than in a random network. The risk of obesity is also about 20% higher for alters' alters (at two degrees of separation) and about 10% higher for alters' alters' alters (at three degrees of separation). By the fourth degree of separation, there was no excess relationship between an ego's obesity and the alter's obesity. Hence, the reach of the obesity clusters was three degrees.

Figure 3B indicates that the effect of geographic distance is different from the effect of social distance. Whereas increasing social distance appeared to decrease the effect of an alter on an ego, increasing geographic distance did not. The obesity of the most geographically distant alters correlated as strongly with an ego's obesity as did the obesity of the geographically closest alters. These results suggest that social distance plays a stronger role than geographic distance in the spread of behaviors or norms associated with obesity.

We evaluated the extent of interpersonal association in obesity with the use of regression analysis. Our models account for homophily by including a time-lagged measurement of the alter's obesity. We evaluated the possible role of unobserved contemporaneous events by separately analyzing models of subgroups of the data involving various ego–alter pairings. Figure 4 summarizes the associations.

If an ego stated that an alter was his or her friend, the ego's chances of becoming obese appeared to increase by 57% (95% confidence interval [CI], 6 to 123) if the alter became obese. However, the type of friendship appeared to be important. Between mutual friends, the ego's risk of obesity increased by 171% (95% CI, 59 to 326) if an alter became obese. In contrast, there was no statistically meaningful relationship when the friendship was perceived by the alter but not the ego (P=0.70). Thus, influence in friendship ties appeared to be directional.

The sex of the ego and alter also appeared to be important. When the sample was restricted to same-sex friendships (87% of the total), the probability of obesity in an ego increased by 71% (95% CI, 13 to 145) if the alter became obese. For friends of the opposite sex, however, there was no significant association (P=0.64). Among friends of the same sex, a man had a 100% (95% CI, 26 to 197) increase in the chance of becoming obese if his male friend became obese, whereas the female-to-female spread of obesity was not significant (38% increased chance; 95% CI, –39 to 161).

Among pairs of adult siblings, one sibling's chance of becoming obese increased by 40% (95% CI, 21 to 60) if the other sibling became obese. This phenomenon appeared to be more marked among siblings of the same sex (55%; 95% CI, 26 to 88) than among siblings of the opposite sex (27%; 95% CI, 3 to 54), although the difference was not significant (P=0.16). Among brothers, an ego's chance of becoming obese increased by 44% (95% CI, 6 to 91) if his alter became obese, and among sisters, an ego's chance of becoming obese increased by 67% (95% CI, 27 to 114) if her alter became obese. Obesity in a sibling of the opposite sex did not affect the chance that the other sibling would become obese.

Among married couples, when an alter became obese, the spouse appeared to be 37% more likely (95% CI, 7 to 73) to become obese. Husbands and wives appeared to affect each other similarly.
(44% and 37%, respectively). Finally, we observed no effect on the risk that an ego would become obese if an immediate neighbor became obese.

We also investigated two factors that might mediate or modify the effect of an alter’s weight gain: his or her smoking behavior and geographic distance from the ego (see the Supplementary Appendix). We added measures of smoking behavior for the ego and the alter at both the current and previous examinations. The coefficient for the effect of the alter’s obesity on the intensity of the effect of the alter’s obesity on the ego.

**Discussion**

Our study suggests that obesity may spread in social networks in a quantifiable and discernable pattern that depends on the nature of social ties.

Moreover, social distance appears to be more important than geographic distance within these networks. Although connected persons might share an exposure to common environmental factors, the experience of simultaneous events, or other common features (e.g., genes) that cause them to gain or lose weight simultaneously, our observations suggest an important role for a process involving the induction and person-to-person spread of obesity.

Our findings that the weight gain of immediate neighbors did not affect the chance of weight gain in egos and that geographic distance did not modify the effect for other types of alters (e.g., friends or siblings) helps rule out common exposure to local environmental factors as an explanation for our observations. Our models also controlled for an ego’s previous weight status; this helps to account for sources of confounding that are stable over time (e.g., childhood experiences or genetic endowment). In addition, the control of smoking and an increase in obesity in the overall population has been noted. However, the present study indicates that regardless of whether smoking cessation causes weight gain in individual persons, and regardless of whether smoking-initiation or smoking-cessation behavior itself spreads from person to person, any spread in smoking behavior is not a significant factor in the spread of obesity. This finding indicates that smoking behavior does not mediate the interpersonal effect in the spread of obesity. However, in addition, it suggests that the psychosocial mechanisms of the spread of obesity may rely less on behavioral imitation than on a change in an ego’s general perception of the social norms regarding the acceptability of obesity. This point is further...
reinforced by the relevance of the directionality of friendship.

Hence, an ego may observe that an alter gains weight and then may accept weight gain in himself or herself. This weight gain in an ego might, in turn, be determined by various behaviors that an ego chooses to evince, and these behaviors need not be the same behaviors that an alter evinces. The observation that geographic distance does not modify the effect of an alter’s obesity also provides support for the concept that norms may be particularly relevant here. Behavioral effects might rely more on the frequency of contact (which one might reasonably expect to be attenuated with distance), whereas norms might not.

The spread of obesity in social networks appears to be a factor in the obesity epidemic. Yet the relevance of social influence also suggests that it may be possible to harness this same force to slow the spread of obesity. Network phenomena might be exploited to spread positive health behaviors, 34-36 in part because people’s perceptions of their own risk of illness may depend on the people around them.37 Smoking- and alcohol-cessation programs and weight-loss interventions that provide peer support — that is, that modify the person’s social network — are more successful than those that do not.34,35,38,39 People are connected, and so their health is connected.40,41 Consequently, medical and public health interventions might be more cost-effective than initially supposed, since health improvements in one person might spread to others.42 The observation that people are embedded in social networks suggests that both bad and good behaviors might spread over a range of social ties. This highlights the necessity of approaching obesity not only as a clinical problem but also as a public health problem.

Supported by a grant from the National Institutes of Health (NIH R-01 AG24448-01).

No potential conflict of interest relevant to this article was reported.

We thank Laurie Meneades, Rebecca Joyce, Molly Collins, Marian Bellwood, and Karen Mutalik for the expert assistance required to build the data set, and Emiliano M. Viterbo, James O’Malley, Barbara McNeil, Mark Pachucki, Mark Fletcher, Mason A. Porter, Darren Schreiber, Richard Suzanne, and Alan Zaslavsky for helpful comments.

REFERENCES
27. Schildcrout JS, Heagerty PJ. Regression analysis of longitudinal binary data with time-dependent environmental covariates;


