

## SPECIAL ARTICLE

# 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**

From the Department of Health Care Policy, Harvard Medical School, Boston (N.A.C.); the Department of Medicine, Mt. Auburn Hospital, Cambridge, MA (N.A.C.); the Department of Sociology, Harvard University, Cambridge, MA (N.A.C.); and the Department of Political Science, University of California, San Diego, San Diego (J.H.F.). Address reprint requests to Dr. Christakis at the Department of Health Care Policy, Harvard Medical School, 180 Longwood Ave., Boston, MA 02115, or at christakis@hcp.med.harvard.edu.

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 (body-mass index [the weight in kilograms divided by the square of the height in meters],  $\geq 30$ ) 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 than 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.

N Engl J Med 2007;357:370-9.

Copyright © 2007 Massachusetts Medical Society.

THE PREVALENCE OF OBESITY HAS INCREASED from 23% to 31% over the recent past in the United States, and 66% of adults are overweight.<sup>1,2</sup> Proposed explanations for the obesity epidemic include societal changes that promote both inactivity and food consumption.<sup>3</sup> The fact that the increase in obesity during this period cannot be explained by genetics<sup>4,5</sup> and has occurred among all socioeconomic groups<sup>1</sup> provides support for a broad set of social and environmental explanations. Since diverse phenomena can spread within social networks,<sup>6-10</sup> 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.<sup>11,12</sup> 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 obese social contacts might change a person's tolerance for being obese 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.<sup>13</sup> Even infectious causes of obesity are conceivable.<sup>14,15</sup>

We evaluated a network of 12,067 people who underwent repeated measurements over a period of 32 years. We examined several aspects of the spread of obesity, including the existence of clusters of obese persons within the network, the association between one person's weight gain and weight gain among his or her social contacts, the dependence of this association on the nature of the social ties (e.g., ties between friends of different kinds, siblings, spouses, and neighbors), and the influence of sex, smoking behavior, and geographic distance between the domiciles of persons in the social network.

## METHODS

### SOURCE DATA

The Framingham Heart Study was initiated in 1948, when 5209 people were enrolled in the orig-

inal cohort.<sup>16</sup> 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.<sup>17</sup> 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 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

### Glossary

**Ego:** The person whose behavior is being analyzed.

**Alter:** A person connected to the ego who may influence the behavior of the ego.

**Node:** 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.

**Tie:** A connection between two nodes that can be either one-way (directed) or two-way (bilateral). In this study, all family ties (e.g., between siblings and parents) as well as marital ties 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.

**Degree of separation:** 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.

**Homophily:** The tendency for people to choose relationships with people who have similar attributes.

**Induction:** The spread of a behavior or trait from one person to another.

**Cluster:** A group of nodes, each of which is connected to at least one other node in the group.

**Connected component:** 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 of the network.

to the study participants 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 periods centered in 1973, 1981, 1985, 1989, 1992, 1997, and 1999. Detailed home addresses were also recorded at each time point; we used this information to calculate the geographic distance between people.

Many of the named alters on these sheets also were members of Framingham Heart Study cohorts. 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 an ego identifies an alter as a friend; an “alter-perceived friendship,” in which an alter identifies an ego as a friend; and a “mutual friendship,” in which the identification 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, fol-

lowed 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  $\geq 17$  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<sup>18</sup> algorithm in Pajek software.<sup>19</sup> We generated videos of the network by means of the Social Network Image Animator (known as SoNIA).<sup>20</sup> We examined whether our data conformed to theoretical network models such as the small-world,<sup>19</sup> scale-free,<sup>21</sup> and hierarchical types<sup>22</sup> (see the Supplementary Appendix, available with the full text of this article at [www.nejm.org](http://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”).<sup>21,23,24</sup> 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).<sup>25</sup>

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 point ( $t$ ); and most pertinent, the alter's obesity status at times  $t$  and  $t+1$ .<sup>25</sup> We used generalized estimating equations to account for multiple observations of the same ego across examinations and across ego–alter pairs.<sup>26</sup> We assumed an independent working correlation structure for the clusters.<sup>26,27</sup>

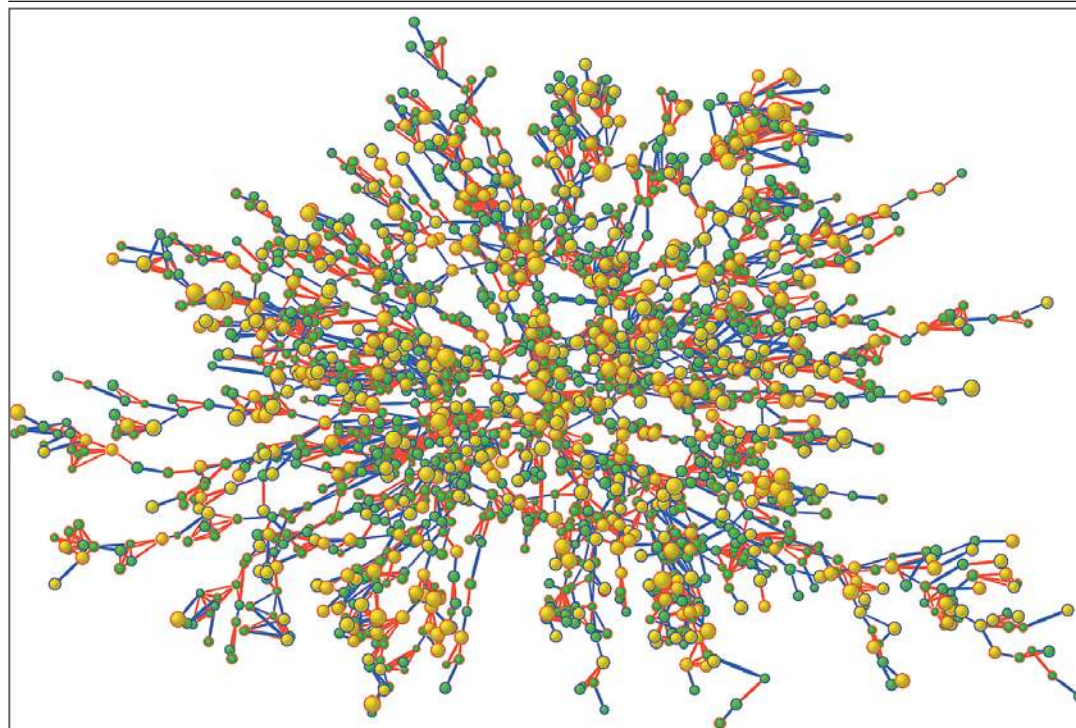
The use of a time-lagged dependent variable (lagged to the previous examination) eliminated serial correlation in the errors (evaluated with a Lagrange multiplier test<sup>28</sup>) and also substantially controlled for the ego's genetic endowment and any intrinsic, stable predisposition to obesity. The use of a lagged independent variable for an alter's weight status controlled for homophily.<sup>25</sup> 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 affect-

ing 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.

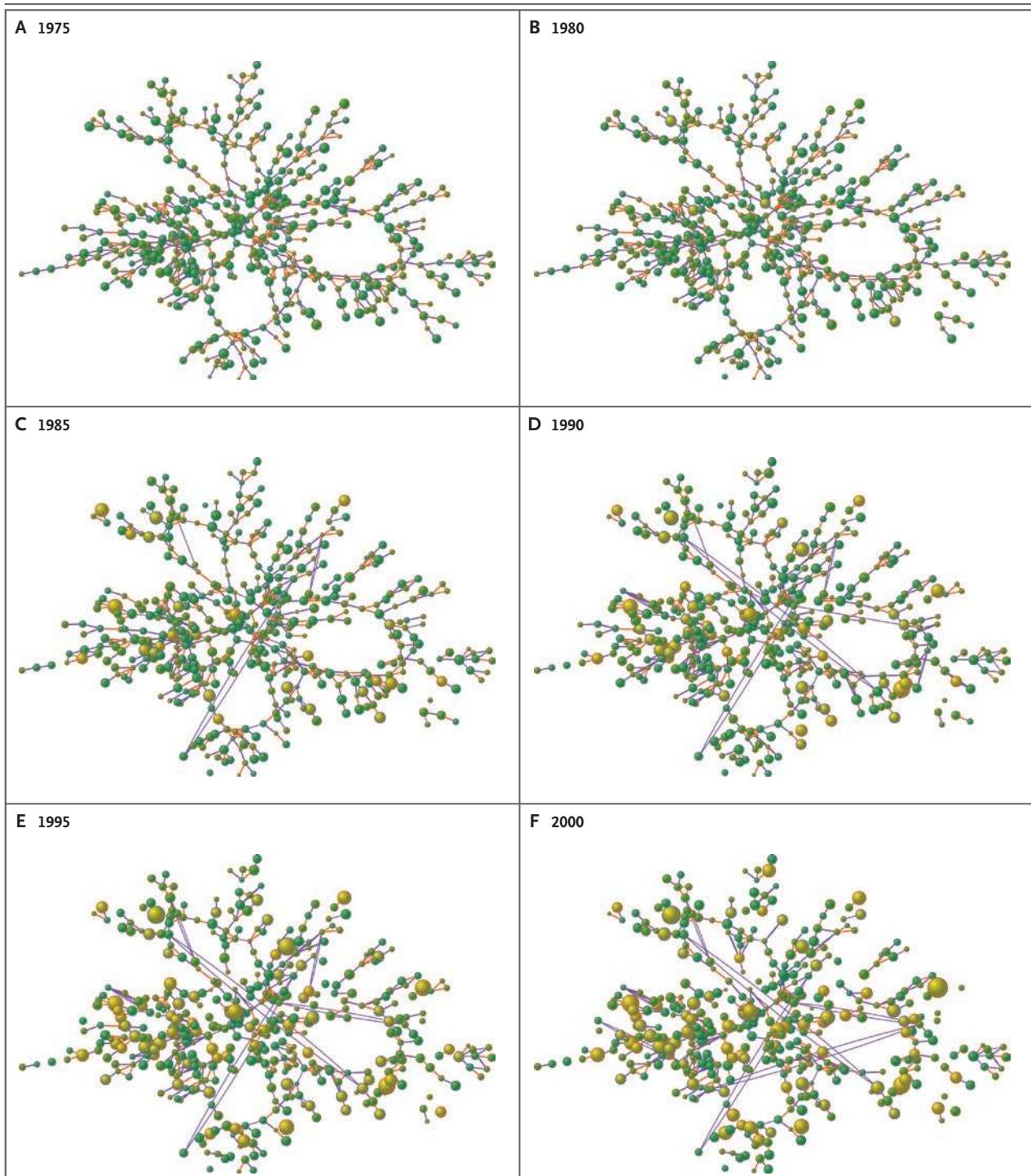
We calculated 95% confidence intervals by simulating the first difference in the alter's contem-



**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 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,  $\geq 30$ ) and green denotes a nonobese person. 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 According to 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,  $\geq 30$ ) and green denotes a nonobese person. The colors of the ties between the circles indicate the relationship between them: purple denotes a friendship or a marital tie and orange denotes a familial tie. The disappearance of a circle from one year to another 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](http://www.nejm.org).

**Figure 3. Effect of Social and Geographic Distance from Obese Alters on the Probability of an Ego's Obesity in the Social Network of the Framingham Heart Study.**

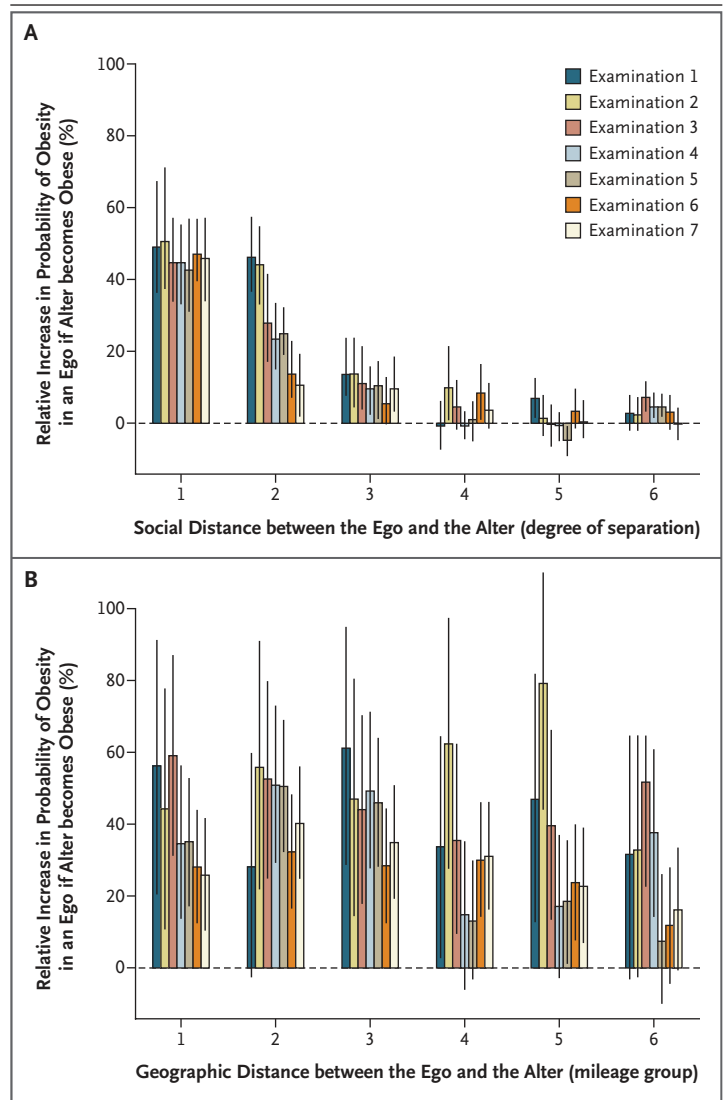
Panel A shows the mean effect of an ego's social proximity to an obese alter; this effect is derived by comparing the conditional probability of obesity in the observed network with the probability of obesity in identical networks (with topology preserved) in which the same number of obese persons is randomly distributed. The social distance between the alter and the ego is represented by degrees of separation (1 denotes one degree of separation from the ego, 2 denotes two degrees of separation from the ego, and so forth). The examination took place at seven time points. Panel B shows the mean effect of an ego's geographic proximity to an obese alter. We ranked all geographic distances (derived from geocoding) between the homes of directly connected egos and alters (i.e., those pairs at one degree of separation) and created six groups of equal size. This figure shows the effects observed for the six mileage groups (based on their average distance): 1 denotes 0 miles (i.e., closest to the alter's home), 2 denotes 0.26 mile, 3 denotes 1.5 miles, 4 denotes 3.4 miles, 5 denotes 9.3 miles, and 6 denotes 471 miles (i.e., farthest from the alter's home). There is no trend in geographic distance. I bars for both panels show 95% confidence intervals based on 1000 simulations. To convert miles to kilometers, multiply by 1.6.

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.<sup>29</sup> All tests were two-tailed. The sensitivity of the results was assessed with multiple additional analyses (see the Supplementary Appendix).

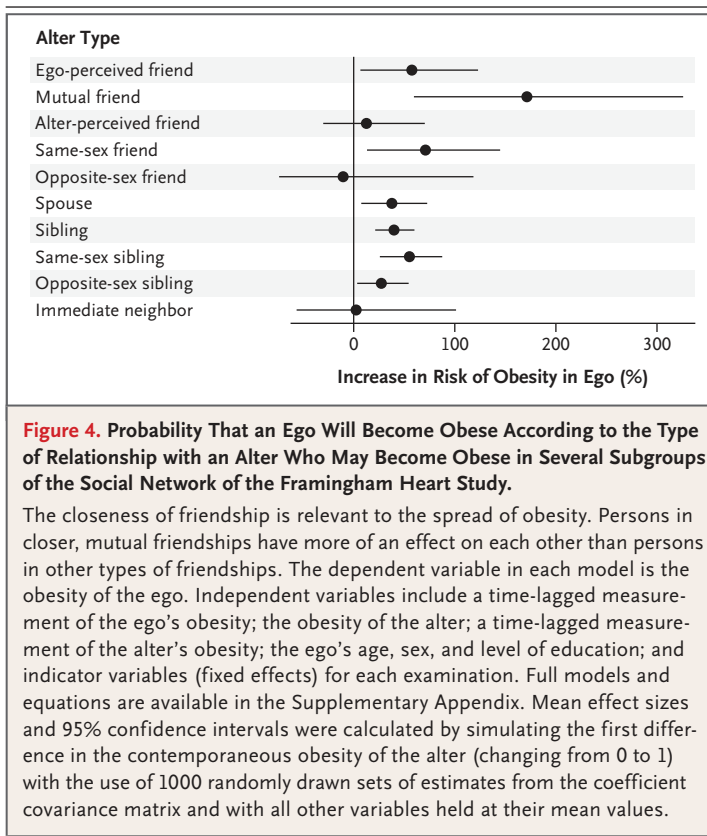
## RESULTS

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](http://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 preva-



lence 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 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 was 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 was 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 was virtually unchanged; smoking behavior does not appear to be instrumental to the spread of obesity. Models that included the geographic distance between the ego and alter corroborated the result shown in Figure 3B: geographic distance did not modify 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).<sup>30</sup> In addition, the control in our models for an alter's previous weight status accounts for a possible tendency of obese people to form ties among themselves. Finally, the findings regarding the directional nature of the effects of friendships are especially important with regard to the interpersonal induction of obesity because they suggest that friends do not simultaneously become obese as a result of contemporaneous exposures to unobserved factors. If the friends did become obese at the same time,

any such exposures should have an equally strong influence regardless of the directionality of friendship. This observation also points to the specifically social nature of these associations, since the asymmetry in the process may arise from the fact that the person who identifies another person as a friend esteems the other person.

Finally, pairs of friends and siblings of the same sex appeared to have more influence on the weight gain of each other than did pairs of friends and siblings of the opposite sex. This finding also provides support for the social nature of any induction of obesity, since it seems likely that people are influenced more by those they resemble than by those they do not. Conversely, spouses, who share much of their physical environment, may not affect each other's weight gain as much as mutual friends do; in the case of spouses, the opposite-sex effects and friendship effects may counteract each another.

Obesity in alters might influence obesity in egos by diverse psychosocial means, such as changing the ego's norms about the acceptability of being overweight, more directly influencing the ego's behaviors (e.g., affecting food consumption), or both. Other mechanisms are also possible. Unfortunately, our data do not permit a detailed examination. However, some insight into possible mechanisms can be gained from a consideration of the roles of smoking and geographic distance in obesity. The tendency of persons to gain weight when they stop smoking is well known,<sup>31</sup> and the coincidence of a decrease in smoking and an increase in obesity in the overall population has been noted.<sup>32</sup> 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,<sup>33</sup> 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,<sup>34-36</sup> in part because people's perceptions of their own risk of illness may depend on the people around them.<sup>37</sup> 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.<sup>34,35,38,39</sup> People are connected, and so their health is connected.<sup>40,41</sup> Consequently, medical and public health interventions might be more cost-effective than initially supposed, since health improvements in one person might spread to others.<sup>42</sup> 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 Emelia Benjamin, Virginia Chang, Scott Desposato, Felix Elwert, Peter Marsden, Joanne Murabito, James O'Malley, Barbara McNeil, Mark Pachucki, Mark Pletcher, Mason A. Porter, Darren Schreiber, Richard Suzman, and Alan Zaslavsky for helpful comments.

## REFERENCES

1. Chang VW, Lauderdale DS. Income disparities in body mass index and obesity in the United States, 1971-2002. *Arch Intern Med* 2005;165:2122-8.
2. Hedley AA, Ogden CL, Johnson CL, Carroll MD, Curtin LR, Flegal KM. Prevalence of overweight and obesity among US children, adolescents, and adults, 1999-2002. *JAMA* 2004;291:2847-50.
3. Hill JO, Peters JC. Environmental contributions to the obesity epidemic. *Science* 1998;280:1371-4.
4. Stunkard AJ, Sorensen TI, Hanis C, et al. An adoption study of human obesity. *N Engl J Med* 1986;314:193-8.
5. Stunkard AJ, Harris JR, Pedersen NL, McClearn GE. The body-mass index of twins who have been reared apart. *N Engl J Med* 1990;322:1483-7.
6. Newman MEJ. The structure and function of complex networks. *SIAM Review* 2003;45:167-256.
7. Nowak MA, Sigmund K. Evolution of indirect reciprocity. *Nature* 2005;437:1291-8.
8. Bearman PS, Moody J, Stovel K. Chains of affection: the structure of adolescent romantic and sexual networks. *Am J Sociol* 2004;110:44-91.
9. Liljeros F, Edling CR, Nunes Amaral LA. Sexual networks: implications for the transmission of sexually transmitted infections. *Microbes Infect* 2003;5:189-96.
10. Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. *Nature* 1998;393:440-2.
11. Sobal J. The size acceptance movement and the social construction of body weight. In: Sobal J, Maurer D, eds. *Weighty issues: fatness and thinness as social problems*. New York: Aldine de Gruyter, 1999:231-49.
12. Chang VW, Christakis NA. Medical modeling of obesity: a transition from action to experience in a 20th century American medical textbook. *Social Health Illness* 2002;24:151-77.
13. Fogassi L, Ferrari PF, Gesierich B, Rozzi S, Chersi F, Rizzolatti G. Parietal lobe: from action organization to intention understanding. *Science* 2005;308:662-7.
14. Whigham LD, Israel BA, Atkinson RL. Adipogenic potential of multiple human adenoviruses in vivo and in vitro in animals. *Am J Physiol Regul Integr Comp Physiol* 2006;290:R190-R194.
15. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 2006;444:1027-31.
16. Dawber TR. *The Framingham Study: the epidemiology of atherosclerotic disease*. Cambridge, MA: Harvard University Press, 1980.
17. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham Offspring Study: design and preliminary data. *Prev Med* 1975;4:518-25.
18. Kamada T, Kawai S. An algorithm for drawing general undirected graphs. *Information Processing Letters* 1989;31:7-15.
19. Batagelj V, Mrvar A. Pajek — analysis and visualization of large networks. In: Junger M, Mutzel P, eds. *Graph drawing software*. Berlin: Springer, 2003:77-103.
20. Moody J, McFarland DA, Bender-deMoll S. Visualizing network dynamics. *Am J Sociol* 2005;110:1206-41.
21. Barabasi AL, Albert R. Emergence of scaling in random networks. *Science* 1999;286:509-12.
22. Ravasz E, Barabasi AL. Hierarchical organization in complex networks. *Phys Rev E Stat Nonlin Soft matter Phys* 2003;67:026112.
23. McPherson M, Smith-Lovin L, Cook JM. Birds of a feather: homophily in social networks. *Ann Rev Sociol* 2001;27:415-44.
24. Sackett DL, Anderson GD, Milner R, Feinleib M, Kannel WB. Concordance for coronary risk factors among spouses. *Circulation* 1975;52:589-95.
25. Carrington PJ, Scott J, Wasserman S. *Models and methods in social network analysis*. New York: Cambridge University Press, 2005.
26. Liang K-Y, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986;73:13-22.
27. Schildcrout JS, Heagerty PJ. Regression analysis of longitudinal binary data with time-dependent environmental covariates: bias and efficiency. *Biostatistics* 2005;6:633-52.
28. Beck N. Time-series-cross-section data: what have we learned in the past few years? *Annu Rev Polit Sci* 2001;4:271-93.

29. King G, Tomz M, Wittenberg J. Making the most of statistical analyses: improving interpretation and presentation. *Am J Polit Sci* 2000;44:341-55.
30. Atwood LD, Heard-Costa NL, Fox CS, Jaquish CE, Cupples LA. Sex and age specific effects of chromosomal regions linked to body mass index in the Framingham Study. *BMC Genet* 2006;7:7.
31. Eisenberg D, Quinn BC. Estimating the effect of smoking cessation on weight gain: an instrumental variable approach. *Health Serv Res* 2006;41:2255-66.
32. Philipson TJ, Posner RA. The long-run growth in obesity as a function of technological change. *Perspect Biol Med* 2003;46:Suppl 3:S87-S107.
33. Andrews JA, Tildesley E, Hops H, Li F. The influence of peers on young adult substance use. *Health Psychol* 2002;21:349-57.
34. Wing RR, Jeffery RW. Benefits of recruiting participants with friends and increasing social support for weight loss and maintenance. *J Consult Clin Psychol* 1999;67:132-8.
35. Malchodi CS, Oncken C, Dornelas EA, Caramanica L, Gregonis E, Curry SL. The effects of peer counseling on smoking cessation and reduction. *Obstet Gynecol* 2003;101:504-10.
36. Bruckner H, Bearman PS. After the promise: the STD consequences of adolescent virginity pledges. *J Adolesc Health* 2005;36:271-8.
37. Montgomery GH, Erbllich J, DiLorenzo R, Bovbjerg DH. Family and friends with disease: their impact on perceived risk. *Prev Med* 2003;37:242-9.
38. McKnight AJ, McPherson K. Evaluation of peer intervention training for high school alcohol safety education. *Accid Anal Prev* 1986;18:339-47.
39. Wechsler H, Moeykens B, Davenport A, Castillo S, Hansen J. The adverse impact of heavy episodic drinkers on other college students. *J Stud Alcohol* 1995;56:628-34.
40. Christakis NA, Allison PD. Mortality after the hospitalization of a spouse. *N Engl J Med* 2006;354:719-30.
41. Granovetter MS. The strength of weak ties. *Am J Sociol* 1973;78:1360-80.
42. Christakis NA. Social networks and collateral health effects. *BMJ* 2004;329:184-5.

Copyright © 2007 Massachusetts Medical Society.

#### FULL TEXT OF ALL JOURNAL ARTICLES ON THE WORLD WIDE WEB

Access to the complete text of the *Journal* on the Internet is free to all subscribers. To use this Web site, subscribers should go to the *Journal*'s home page ([www.nejm.org](http://www.nejm.org)) and register by entering their names and subscriber numbers as they appear on their mailing labels. After this one-time registration, subscribers can use their passwords to log on for electronic access to the entire *Journal* from any computer that is connected to the Internet. Features include a library of all issues since January 1993 and abstracts since January 1975, a full-text search capacity, and a personal archive for saving articles and search results of interest. All articles can be printed in a format that is virtually identical to that of the typeset pages. Beginning 6 months after publication, the full text of all Original Articles and Special Articles is available free to nonsubscribers who have completed a brief registration.



# Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic<sup>☆,☆☆</sup>

Ethan Cohen-Cole<sup>a,1</sup>, Jason M. Fletcher<sup>b,\*</sup>

<sup>a</sup> Federal Reserve Bank of Boston, 600 Atlantic Avenue, Boston, MA 02210, United States

<sup>b</sup> Yale University, School of Public Health, Division of Health Policy and Administration, 60 College Street, New Haven, CT 06510, United States

## ARTICLE INFO

### Article history:

Received 12 December 2007

Received in revised form 16 April 2008

Accepted 25 April 2008

Available online 9 May 2008

### JEL classification:

D10

D71

I19

J11

Z13

### Keywords:

Obesity

Peer effects

Social networks

## ABSTRACT

This note's aim is to investigate the sensitivity of Christakis and Fowler's claim [Christakis, N., Fowler, J., 2007. The spread of obesity in a large social network over 32 years. The New England Journal of Medicine 357, 370–379] that obesity has spread through social networks. It is well known in the economics literature that failure to include contextual effects can lead to spurious inference on “social network effects.” We replicate the NEJM results using their specification and a complementary dataset. We find that point estimates of the “social network effect” are reduced and become statistically indistinguishable from zero once standard econometric techniques are implemented. We further note the presence of estimation bias resulting from use of an incorrectly specified dynamic model.

© 2008 Elsevier B.V. All rights reserved.

## 1. Introduction

The United States has experienced a startling increase in average weight and in obesity over the past few decades (Flegal et al., 2002; Hedley et al., 2004). Though this phenomenon is by now well known and has been widely discussed and debated, there is still little consensus on its causes. One proposed explanation for the increase in obesity is long run technological changes that have impacted food prices as well as the propensity to exercise (Philipson and Posner, 2003; Cutler et al., 2003). Though some observers include genetic variation as a potential explanation for the rise of obesity because of the large estimates of heritability of obesity (Stunkard et al., 1990; Coady et al., 2002), most researchers acknowledge that genetic explanations are unlikely to explain the rapid increase in obesity over a relatively short period of time.

<sup>☆</sup> The views in this paper are solely those of the authors and do not reflect official positions of the Federal Reserve Bank of Boston or the Federal Reserve System.

<sup>☆☆</sup> Further exchanges between the authors may appear on their home pages. Jason Fletcher <http://www.med.yale.edu/eph/faculty/fletcher.html>; Ethan Cohen-Cole <http://www.bos.frb.org/economic/econbios/cohen-cole.htm>; James Fowler <http://jh.fowler.ucsd.edu/>; Nicholas Christakis <http://christakis.med.harvard.edu/>.

\* Corresponding author. Tel.: +1 203 785 5760; fax: +1 203 785 6287.

E-mail address: [jason.fletcher@yale.edu](mailto:jason.fletcher@yale.edu) (J.M. Fletcher).

<sup>1</sup> Tel.: +1 617 973 3294; fax: +1 617 573 5494.

One particularly interesting hypothesis recently explored by Nicholas Christakis and James Fowler (henceforth, CF) in the *New England Journal of Medicine* is that obesity may spread through “social networks effects.”<sup>2</sup> In fact, CF report that their findings suggest that social networks indeed facilitate the spread of obesity.<sup>3</sup> This provocative finding was detailed in many media sources, including the front page of the *New York Times*.<sup>4</sup> *USA Today* coverage indicated that “Obesity is contagious” and “. . . pick your friends carefully. . .” (Hellmich, 2007).<sup>5</sup> CF suggest some potential mechanisms by which this may occur, including that having obese peers may change a person’s tolerance for being obese or may influence weight-related behaviors such as eating habits, smoking, or exercise. Additional mechanisms suggested by CF include infectious causes of obesity or physiological imitation.

However, as is well known in the economics literature, there are alternative hypotheses that also potentially explain the empirical finding that friends’ weight is correlated across time that do not require the presence of social network effects. As CF identify in their study, there are at least three reasons why the weight status of individuals could be clustered within reference groups.<sup>6</sup> The first is that individuals could choose their friends based on factors associated with weight or weight trajectories. In economics, this is typically referred to as selection (CF as homophily). Thus, friendship selection could directly lead to the correlation between friends’ weight or weight gain without an individual’s weight causally affecting his friend’s weight through a social network effect. Second, individuals may adjust behavior because of exposure to common influences. These effects are typically referred to as *contextual* influences (CF as confounding). For example, the opening of a fast food restaurant, convenience store, gym, etc. near a school could simultaneously affect the weight of all friends in a school’s social network. Importantly, the presence of (often unmeasured) shared surroundings can lead to erroneously implicating social network effects in individual outcomes where none exist.<sup>7</sup> Finally, individuals may alter their behavior as others in their group change theirs. Economists are now generally labeling this an *endogenous* social effect (CF as social network effects).

We point to three problems with the CF method. First, CF do not include a sufficiently broad set of contextual effects to account for a range of hypothesized causes of the epidemic. Second, the CF method of controlling for selection is much too narrow in scope. Third, the CF dynamic model as estimated produced coefficients with large degrees of bias (Liu et al., 2006).

Once the first two errors are corrected, evidence for endogenous causes of obesity is thin. We find that the CF results are not robust. In fact, the econometric evidence points strongly to shared environmental factors as the principle operative social mechanism underlying the positive correlation in weight status within reference groups.<sup>8</sup> We find this remarkable given the preponderance of contexts in which endogenous effects appear present and the fact that this class of empirical models appears to generate the appearance of effects quite easily (Krauth, 2006). Our findings point to the difficulty in labeling the source of social effects, particularly in contexts with a direct policy reference. The public health implications given endogenous versus contextual drivers of obesity are quite different.

## 2. Data

We use the Add Health dataset to examine whether there are social network effects in weight outcomes for a national sample of adolescents who transition into early adulthood.<sup>9</sup> Importantly, we have information on friends for approximately 5000 individuals, nearly 2000 of whom are followed over time along with at least one same-sex friend.<sup>10</sup> This sample size gives us nearly 4000 person-year observations, slightly larger than the 3000 used by CF. Summary statistics for our sample are presented in Table 1.<sup>11</sup>

Though there are several important differences between the Add Health and the Framingham Heart Study (FHS) used in CF, the two data sets are sufficiently similar to use to evaluate the role of transmission mechanisms. One important advantage of the Add Health is that it is a national sample of 7–12th graders in 1994/1995 instead of being confined to a smaller population with a wider age range. Another advantage is that individuals in our data are in a setting (high school) that is, in principle, more social than the varied lives of individuals in the Framingham study. Finally, by focusing on a national sample

<sup>2</sup> In the language of the Christakis and Fowler article, social network effects are equivalent to *endogenous* social effects (Manski, 1993). We use the terms interchangeably.

<sup>3</sup> They state, “. . . our observations suggest an important role for a process involving the induction and person-to-person spread of obesity” (p. 377), “The spread of obesity in social networks appears to be a factor in the obesity epidemic” (p. 378), and “. . . obesity appears to spread through social ties” in the summary of their conclusions (p. 370).

<sup>4</sup> See for example, Kolata (2007) for the *New York Times*, Hellmich (2007) for *USA Today*, and Dembner (2007) for the *Boston Globe*.

<sup>5</sup> An official with the National Institute on Aging is quoted in Dembner (2007) as calling the study, “one of the most exciting studies in medical sociology . . . in decades”.

<sup>6</sup> Manski (1993) provided the seminal contribution in economics in describing the alternative mechanisms that can lead to correlated outcomes within reference groups.

<sup>7</sup> See Fletcher (2007) and Lee (2007) on the importance of unobserved group-level characteristics in estimating peer effects.

<sup>8</sup> While many empirical studies label contextual effects as observable group-level phenomena, our main result is to show that the social network effects estimated by CF can be accounted for by unobserved group-level controls. Of course, in Manski’s nomenclature, these are also contextual effects.

<sup>9</sup> A full description of the sample design, data, and documentation is available at: <http://www.cpc.unc.edu/addhealth>.

<sup>10</sup> For individuals for whom more than one friend’s BMI information is available, we select the friend with the highest nomination (1st–5th). Nearly two-thirds of the individuals in our sample are only matched to one friend’s BMI due to the sample design. We select only one friend to be consistent with CF.

<sup>11</sup> Table 1A in the appendix of our working paper (Cohen-Cole and Fletcher, 2008) shows that our analysis sample has similar characteristics as the full Add Health sample.



**Table 1**Summary statistics Add Health, analysis sample  $N = 1988$ ; two observations per person

Variable	Wave	Mean	Std. Dev.
BMI	2	22.66	4.38
BMI	3	26.29	5.97
Friend BMI	2	22.50	4.10
Friend BMI	3	26.03	5.58
Obese	2	0.07	0.26
Obese	3	0.21	0.41
Friend obese	2	0.06	0.24
Friend obese	3	0.20	0.40
Male	All	0.44	0.50
Age	2	16.24	1.51
Age	3	21.70	1.50
Black	All	0.16	0.37
Hispanic	All	0.15	0.36
Maternal education <sup>a</sup>	All	13.34	2.34
Family income <sup>a</sup>	All	49.72	46.98
Missing family information	All	0.28	0.45

<sup>a</sup> Imputed missing data.

of individuals in an active social setting,<sup>12</sup> we are largely able to rule out that our findings may be due to the nature of our dataset. Thus, while one may believe that the finding of network effects in the FHS suggests an age or social setting difference in effects, it would be particularly difficult to make a generalized claim using the CF results in the absence of similar results from a national sample.

Similarities between the data sets include the time lags between interviews (approximately 3 years on average) and the type of information that was collected (see Table 4A in [Cohen-Cole and Fletcher, 2008](#)). It will be important for our baseline results to closely mirror those found by CF in order to be able to compare the results from our extended analyses. For brevity and to maximize the overlapping information contained in each dataset, we focus on one of the most interesting and robust results in CF—the positive relationship between the weight status of individuals and their same-sex friends.

### 3. Social network effects vs. shared experience

Central to our discussion is the distinction between *endogenous* effects, also labeled “induction” or social network effects by CF, and *contextual* effects. In the case of obesity, one can think of endogenous effects as describing the propensity to become obese because of the direct interaction with another individual. One may decide to eat more (or higher caloric foods) because their friend, spouse, neighbor does so. Because the two individuals are directly connected, they may influence each other for a variety of reasons. Contextual influences, on the other hand, reflect the shared surroundings of the members of a group that could lead to similar weight outcomes.<sup>13</sup> Without detailed information on an individual's characteristics, choices, preferences, and environment, it is difficult to discern whether two friends' simultaneous weight gain is attributable to their friendship or to an exposure of a common environmental factor ([Manski, 1993](#); [Durlauf, 2004](#)).

## 4. Empirical methods

### 4.1. CF specification and replication

CF use data on obesity status for an individual (in their terminology, an “Ego”) at a given point in time and estimate its relationship to the obesity status of a friend, spouse or relative (an “Alter”) as well as its relationship to the Ego's age, gender, educational level, and past obesity status. The CF specification uses the BMI of an Ego ( $i$ ) who lives in community ( $c$ ) at time ( $t + 1$ )<sup>14</sup> as a function of the individual's previous BMI, the individual's measured characteristics ( $X$ ), Alter's ( $j$ ) current and previous BMI, and an unobserved error term:

$$BMI_{ic,t+1}^{ego} = \theta_1 BMI_{ic,t+1}^{alter} + \theta_2 BMI_{jc,t}^{alter} + \theta_3 BMI_{ic,t}^{ego} + \beta X_{ic,t+1}^{ego} + \varepsilon_{ic,t+1}^{ego} \quad (1)$$

This type of model has three features that can, and regularly do, impact inference. First, if unmeasured community-level variables, which we will denote  $c_{c,t+1}$ , are positively correlated with individuals' BMIs, this will bias the estimate of the social network effect. Critically, the absence of a relevant contextual variable can lead to spurious inference on the endogenous

<sup>12</sup> Studies of peer effects often focus on school and college settings because of the frequent and repeated contact amongst individuals.

<sup>13</sup> Consider the case of two friends: each friend may simultaneously become obese due to the recent introduction of a high caloric restaurant near his place of residence. Though both soon have a higher BMI, this would not have been attributable to their friendship. The difficulty, of course, is distinguishing between the two potential mechanisms.

<sup>14</sup> Recall that in the Add Health data, there are three waves of surveys for which we have BMI data.

**Table 2**

Association between own-obesity status and friend's obesity status comparing results across studies and methods

Outcome: ego currently obese?			
Logit specification	CF	Baseline replication	School trends
Alter currently obese?	0.62*** (0.24) [1.71]	0.588*** (0.160) [1.80]	0.411** (0.162) [1.51]
Ego previously obese	4.38*** (0.19)	4.617*** (0.24)	4.634*** (0.27)

Note: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ . Robust standard errors in parentheses. Odds ratios in brackets. CF results in Table S1 of web appendix, "same-sex friend" column. We do not include fixed effects in this table as results for short panel logistic specifications are known to be biased.

variable. CF appear to approach confounding by examining the type and direction of the friendship networks. For example, if individual A declared himself a friend of individual B but not vice-versa, then a social network effect should appear for A but not B. While network structures can be useful for identification of social network effects, their presence does not rule out the possibility that confounding environmental effects overlap and influence the decisions of network members.<sup>15</sup> Since CF never directly control for environmental factors, we view their results using directionality of friendship nominations as suggestive rather than conclusive.

Second, similarly incorrect inference can result if an individual's error term ( $\varepsilon_{ic,t+1}^{\text{ego}}$ ) is correlated with his friend's BMI. For example, the (unmeasured) propensity to exercise or be involved in sports might increase the chances that two thin individuals are friends. Selection issues appear in this context. CF also provide an ad hoc solution for this issue. Their claim is that "the use of a lagged independent variable for an Alter's weight status controlled for homophily." Unless selection is conditioned *only* on this variable, this statement is spurious. As well, in the presence of social interactions, the use of lagged variables can lead to bias in estimation, a topic we turn to now.

Third, the presence of a lagged dependent (or independent) variable in a social interactions model can lead to substantial biases in estimation. Liu et al. (2006) find, using simulation evidence, that a model very similar to (1) above leads to significant bias in estimation relating to the dynamic role of social interactions.<sup>16</sup> Most importantly for our note is that misspecification of the model or error structure can lead to very large biases and thus incorrect inference. This essentially highlights the well known fact in this literature that inference errors are particularly hard to avoid. Inclusion of a lagged dependent variable appears to magnify the problem.<sup>17</sup>

Acknowledging these concerns, we take the CF model as is and replicate it using the Add Health data. Using Eq. (1) above, a logistic regression model as proposed by CF, we find the odds that an Ego becomes obese increase by 80% if his or her Alter is also obese. Broadly, this is in line with the CF finding of 71% for same-sex friends. Our basic results are presented in Table 2 as logit coefficients in order to compare with the table in CF.

#### 4.2. Extension

We extend the CF model as follows. Assume the model specification now appears as

$$\text{BMI}_{ic,t+1}^{\text{ego}} = \theta_1 \text{BMI}_{ic,t+1}^{\text{alter}} + \theta_2 \text{BMI}_{jc,t}^{\text{alter}} + \theta_3 \text{BMI}_{ic,t}^{\text{ego}} + \beta X_{ic,t+1}^{\text{ego}} + c_{c,t+1} + \varepsilon_{ic,t+1}^{\text{ego}} \quad (2)$$

where we have added a set of environmental confounders ( $c_{c,t+1}$ ).<sup>18</sup> This formalizes the notion of contextual effects. We use either a time-invariant measure as above or a time-dependent set of location-specific (in our case, schools) covariates. Econometrically, this can be expressed as using the following for the environmental confounders:  $c_{c,t+1} = tc_{c,t+1}$ , where  $t$  is a time variable. These represent a much richer set of controls to absorb average changes in social context experienced by all individuals in the sample. To explain further, these school-specific trends account for any environmental factors shared by individuals at the same school. CF control for year effects, but their specification does not capture any shared confounders that also vary across geographic space. For example, CF can control for the fact that the density of fast food restaurants has increased over time but not the fact that the number of fast food restaurants has grown faster in some areas than other areas. For example, suppose that the number of fast food restaurants has grown faster in Boston, Massachusetts than in western Massachusetts. Controlling for year effects (which controls for the growth in the number of fast food restaurants across the states in a given year) is not as appealing as controlling for the number of fast food restaurant in an individual's local area.

<sup>15</sup> The argument for identification of social network effects using network architecture has been formalized in Bramoullé et al. (2007). The methodology is intriguing and may be sufficient to control for confounding in the CF case, however, CF do not employ it, and it is not clear whether the single-Alter structure of most of their data permit identification in this setup in any case.

<sup>16</sup> The size and direction of bias depends on the type of variable, sample size, panel length and type of misspecification. However, bias appears in essentially every case. That is, correct inference on this type of model requires some type of bias correction.

<sup>17</sup> Of course, simply omitting a lagged dependent variable where it should belong does not alleviate the inference problem!

<sup>18</sup> As mentioned previously, we measure an individual's community by the high school they attend. This equation represents our capture of school-level confounders as a fixed effect.

**Table 3**

Association between own-body mass index and friend's body mass index comparing results across studies and methods

Outcome: ego current BMI				
OLS specification	CF	Baseline replication	School trends	Trends and FE
Alter current BMI	0.05** (0.02)	0.054** (0.023)	0.037 (0.023)	0.033 (0.039)
CI	(.01, .09)	(.01, .10)	(−.01, .08)	(−.043, .109)

Note: \*\* $p < 0.05$ . Robust standard errors in parentheses. Baseline replication are results that are identical in specification to CF with the exception of data used. School Trends column updates CF specification by adding school-specific time trends. Trends and FE columns includes both school-specific trends and individual level fixed effects.

Without accounting for the trends, clustering of obesity in social networks that changed over time would incorrectly be absorbed in estimation by the endogenous variable. Though one, in principle, would want many more controls to account for additional contextual effects, we will note shortly that the endogenous effect vanishes even with this relatively simple characterization.

We add to our evaluation by accounting for self-selection of friends (homophily). This is accomplished by looking only at the change in BMI from the time of declaration of friendship until the subsequent weight measurement. Note the distinction between this method and the lagged independent variable used in CF. Our method allows us to distinguish between the desire to become friends based on similarity in weight, which would appear based on the simultaneous measurement of friendship and weight, and the friendship effect of weight gain.

We find that inclusion of a more complete set of controls for school-level environmental confounders leads to a large drop in the coefficient of interest. The odds fall to 50% in column 3 of Table 2 (logit coefficient reported). As well, Liu et al. (2006) report that an increase in the number of contextual variables reduces bias in most cases.

Since linear specifications more easily accommodate fixed effects and time trends, in Table 3 we estimate models predicting BMI. In our data the association between own-BMI and friend-BMI is 0.05, which is identical to the CF results.<sup>19</sup> After controlling for environmental confounding (through school-specific time trends), our estimate falls by over 30% to 0.037. Finally, we control for friendship selection by controlling for individual fixed effects.<sup>20</sup> With this inclusion, the coefficient of interest declines further and becomes statistically indistinguishable from zero. These results, using similar sample sizes and data as CF, suggest that the “social network effects” implicated in CF are not robust to standard economic critiques found in the peer effects literature. In fact, our results suggest that omitted group-level characteristics are the most likely mechanism responsible for correlated body weight within peer groups.<sup>21</sup>

## 5. Conclusion

Our evaluation suggests that the spread of obesity is related to the environment in which individuals live. Though we do not completely rule out the possibility of induction and person-to-person spread of obesity, our results suggest that shared environmental factors can cause the appearance of social network effects. While comparing results across datasets that are quite different in design and focus is usually fraught with difficulty, we were encouraged to be able to closely replicate findings of CF using similar methods; this provides suggestive evidence that the structure of obesity decisions is similar across the two samples. It is only when we incorporate more extensive controls than those found in CF that our results diverge. This divergence in results suggests caution in interpreting correlations in linked individuals' BMI as social network effects. In fact, we find evidence that community-level factors are able to explain a large share of the “social network effect” in our data.

We find results that suggest difficulty in distinguishing social network effects and environmental confounds of weight gain. And though we advise caution in interpreting the available evidence of a social contagion in weight, we concur with CF on the use of network phenomena, broadly writ, to help to ameliorate the epidemic. Some of the encouraging evidence mentioned in CF, in particular the fact that alcohol and smoking cessation programs are more effective when coupled with peer support, can be exploited independent of the degree of induction present. That is, one need not find evidence that obesity is spread via networks to draw on the logic of peer support to address important issues in public health such as smoking and obesity.<sup>22</sup> Our data show tightly interconnected networks of friends at the high school level and a high degree of coincidence of obesity within these networks; this can be exploited for interventions even if the cause of the obesity came from a joint

<sup>19</sup> OLS results are discussed in the CF article but not presented. The authors provided supplementary material to us upon request.

<sup>20</sup> In a supplemental appendix available upon request, we also show that our results here are robust to removing the lagged individual obesity status. Including the lagged individual status is akin to assuming that individuals follow an autoregressive process in obesity. Not only is this very difficult to verify, its incorporation makes inference on the object of interest (friend's influence) very difficult.

<sup>21</sup> The Liu et al. (2006) simulations also find a decrease in estimation error for an increase in the number of contextual effects suggesting that while our extensions may have bias, the finding that the endogenous effects is not significant is robust.

<sup>22</sup> In fact, none of the cited articles in CF illustrate a link between success in cessation programs and induction. This is a very interesting question and one that we encourage as a topic of future research. Until that point, the key relationship is simply the presence of a peer group for support—which can and does occur independent of the obesity transfer mechanism.

external source. Similar arguments can be made to support the notion that obesity can be approached from a public-health perspective rather than a clinical one.

## Acknowledgements

The authors thank Elizabeth Bradley, Paul Cleary, John Mullahy, David Paltiel, and Jody Sindelar for very helpful comments and Jonathan Morse for research assistance. This research uses data from Add Health, a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by grant P01-HD31921 from the National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. Special acknowledgment is due to Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Persons interested in obtaining data files from Add Health should contact Add Health, Carolina Population Center, 123 W. Franklin Street, Chapel Hill, NC 27516-2524 ([addhealth@unc.edu](mailto:addhealth@unc.edu)).

## References

- Bramoullé, Y., Djebbari, H., Fortin, B., 2007. Identification of Peer Effects through Social Networks. University of Laval Working Paper.
- Christakis, N., Fowler, J., 2007. The spread of obesity in a large social network over 32 years. *The New England Journal of Medicine* 357, 370–379.
- Coady, S.A., Jaquish, C.E., Fabsitz, R.R., Larson, M.G., Cupples, L.A., Myers, R.H., 2002. Genetic variability of adult body mass index: a longitudinal assessment in Framingham families. *Obesity* 10, 675–681.
- Cohen-Cole, Ethan, Fletcher, J.M., 2008. Is Obesity Contagious? Social Networks vs. Environmental Factors in the Obesity Epidemic. SSRN Working Paper.
- Cutler, D.M., Glaeser, E.L., Shapiro, J.M., 2003. Why have Americans become more obese? *Journal of Economic Perspectives* 17-3, 93–118.
- Dembner A., 2007. Obesity spreads to friends, study concludes. *Boston Globe*; July 26.
- Durlauf, S.N., 2004. Neighborhood effects. In: Henderson, J.V., Thisse, J.F. (Eds.), *Handbook of Regional and Urban Economics*, Vol. 4, 1 ed. Elsevier, Amsterdam, pp. 2173–2242.
- Flegal, K.M., Carroll, M.D., Ogden, C.L., Johnson, C.L., 2002. Prevalence and trends in obesity among US adults, 1999–2000. *Journal of the American Medical Association* 288, 1723–1727.
- Fletcher, J.M., 2007. Social Interactions and Smoking: An IV/FE Approach. Yale University Working Paper.
- Hedley, A.A., Ogden, C.L., Johnson, C.L., Carroll, M.D., Curtin, L.R., Flegal, K.M., 2004. Prevalence of overweight and obesity among US children, adolescents, and adults, 1999–2002. *Journal of the American Medical Association* 291, 2847–2850.
- Hellmich, N., 2007. Obesity Can Run in Social Circles. *USA Today*; July 26.
- Kolata, G., 2007. Find Yourself Packing It On? Blame Friends. *New York Times*; July 26.
- Krauth, B., 2006. Simulation based evidence of peer effects. *Journal of Econometrics* 133, 243–271.
- Lee, L.F., 2007. Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. *Journal of Econometrics* 140, 333–374.
- Liu, X., Lee, L.F., Kagel, J., 2006. Dynamic Discrete Choice Models with Lagged Social Interactions: With an Application to a Signaling Game Experiment. Ohio State; Mimeo.
- Manski, C.F., 1993. Identification of endogenous social effects: the reflection problem. *Review of Economic Studies* 60-3, 531–542.
- Philipson, T., Posner, R., 2003. The long-run growth in obesity as a function of technological change. *Perspectives in Biology and Medicine* 46-3 (Supplement), S87–S107.
- Stunkard, A.J., Harris, J.R., Pedersen, N.L., McClearn, G.E., 1990. The body-mass index of twins who have been reared apart. *New England Journal of Medicine* 322 (21), 1483–1487.