Molecular Psychiatry (2011) 16, 273–281

www.nature.com/mp

ORIGINAL ARTICLE

Social network determinants of depression

JN Rosenquist[1,2], JH Fowler[3] and NA Christakis[4,5]

[1]Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA; [2]Department of Health Care Policy, Harvard Medical School, Boston, MA, USA; [3]Department of Political Science, University of California, San Diego, CA, USA; [4]Department of Health Care Policy, Harvard Medical School, Cambridge, MA, USA and [5]Department of Sociology, Harvard University, Cambridge, MA, USA

The etiology of depression has long been thought to include social environmental factors. To quantitatively explore the novel possibility of person-to-person spread and network-level determination of depressive symptoms, analyses were performed on a densely interconnected social network of 12 067 people assessed repeatedly over 32 years as part of the Framingham Heart Study. Longitudinal statistical models were used to examine whether depressive symptoms in one person were associated with similar scores in friends, co-workers, siblings, spouses and neighbors. Depressive symptoms were assessed using CES-D scores that were available for subjects in three waves measured between 1983 and 2001. Results showed both low and high CES-D scores (and classification as being depressed) in a given period were strongly correlated with such scores in one's friends and neighbors. This association extended up to three degrees of separation (to one's friends' friends' friends). Female friends appear to be especially influential in the spread of depression from one person to another. The results are robust to multiple network simulation and estimation methods, suggesting that network phenomena appear relevant to the epidemiology of depression and would benefit from further study.
Molecular Psychiatry (2011) **16,** 273–281; doi:10.1038/mp.2010.13; published online 16 March 2010

Keywords: depression; social networks; sociology; social norms; mood

Introduction

Depression is a significant cause of worldwide morbidity and mortality. Current estimates suggest a lifetime incidence of between 13.3 and 17.1% in the United States and a yearly cross-sectional prevalence ranging from 2.3–4.9%.[1] Using any number of metrics, the cost of depression is enormous. For example, disability, morbidity and mortality resulting from depression was estimated to cost $86 billion in the year 2000 alone.[2]

The etiology of depression as an illness has been conceptualized to have a number of interacting biological, psychological and social components.[3] This idea that social forces may impact mood symptoms was first hypothesized over 100 years ago in the context of suicide by the sociologist Emile Durkheim. He noted that suicide rates stayed the same across time and across groups even though the individual members of those groups came and went.[4] Durkheim's conclusion was that whether people took their own lives depended in part on the kind of society they inhabited. He noted that although depression and suicide were seen as entirely individualistic, they may be partly driven by social forces. More recent work on the social influences on depression find a significant correlation between social factors such as child abuse, disruptions in family functioning, stressful life events and neighborhood characteristics.[5–8]

The literature on social determinants of disease has been augmented in recent years by a growing literature focused on understanding the role of social network structure on individual outcomes. Recent work has yielded results suggesting that traits such as obesity, smoking behavior, happiness and loneliness may spread along social networks over time.[9–15] A person's structural position within a network, such as their transitivity (whether their friends are friends with each other) and centrality (whether they are located in the middle or edge of the network) have been found to affect the development of traits and behaviors. For example, Bearman and Moody found that social isolation and (among women) having friends who were not friends with each other were two factors predictive for suicidal ideation, suggesting the structural components of a person's network impacted their behavior.[16]

In addition to such structural effects of network position, there may also be influence effects, whereby depression might spread among friends, family members, co-workers and neighbors. While such influence effects may have an intuitive appeal (most people can no doubt think of instances where they found themselves influenced by a family member or friend), it is crucial to distinguish among three

Correspondence: Dr JN Rosenquist, Health Care Policy, Harvard Medical School, 180 Longwood Avenue, Boston, MA 02115, USA. E-mail: jrosenqu@gmail.com
Received 29 May 2009; revised 26 December 2009; accepted 27 December 2009; published online 16 March 2010

processes: (1) induction, whereby depression in one person actually causes the depression of others; (2) homophily, whereby depressed individuals choose one another as friends and become connected (that is, the tendency of like to attract like);[17] or (3) confounding, whereby connected individuals jointly experience contemporaneous exposures (such as an economic downturn or co-residence in a poor neighborhood[13]). To distinguish among these effects requires repeated measures of depression,[18] longitudinal information about network ties and (ideally) information about the nature or direction of the ties (for example, who nominated whom as a friend).

This paper tests the hypothesis that depressive symptoms may spread from person to person to person in social networks. Also tested is the hypothesis that the structure of social networks may influence, and/or may be influenced by, changes in the CES-D scores of its members over time. Finally, the analyses consider induction, homophily and confounding as possible explanations for these effects. A unique, longitudinal data set that contains rich social network data as well as measures of depressive symptoms is used for our analyses.

Materials and methods

Source data

Our study uses data obtained from participants in the Framingham Heart Study (FHS). The FHS is a population-based, longitudinal, observational cohort study that was initiated in 1948 to prospectively investigate risk factors for cardiovascular disease. Since then, it has come to be composed of four separate but related cohort populations: (1) the 'Original Cohort' enrolled in 1948 ($N=5209$); (2) the 'Offspring Cohort' (the children of the Original Cohort and spouses of the children) enrolled in 1971 ($N=5124$); (3) the 'Omni Cohort' enrolled in 1994 ($N=508$, designed to increase ethnic diversity of participants); and (4) the 'Generation 3 Cohort' (the grandchildren of the Original Cohort) enrolled beginning in 2002 ($N=4095$). The Original Cohort captured the majority of the adult residents of Framingham in 1948 whereas the Offspring Cohort included the great majority of the living offspring of the Original Cohort in 1971, and their spouses. Published reports describe these cohorts in more detail.[19–21]

Continuous surveillance and serial examinations of these cohorts are the source of our longitudinal data. Participant data includes physical exam, laboratory, battery testing (such as the Mini-Mental status exam), questionnaire results and basic demographic information. The Offspring study data is drawn from exams completed roughly every 4 years over a 32-year period (1971–2003), whereas the Original Cohort has data available for approximately every 2 years over a 60-year period. Within all cohorts, there is minimal (<1%) loss to follow-up because of out-migration. For the purposes of the analyses reported here, exam waves for the Original cohort were aligned with those of the Offspring cohort, so that all subjects were treated as having been examined at just seven waves (in the same time windows as the Offspring, as detailed in Supplementary Table S1).

The Offspring Cohort comprises the main source of subjects as it is the source of 'egos' (the focal individuals in the network). However, other FHS participants are included when listed as social contacts by the egos (known as 'alters'). Therefore, whereas egos come only from the Offspring Cohort, alters are drawn from the entire set of FHS cohorts (including also the Offspring Cohort itself). This explains why the total number of individuals in the FHS social network is 12 067, a number that includes individuals from multiple cohorts (5124 from the offspring cohort, 3403 from the original cohort and 3540 from other cohorts). Participant compliance with examinations is excellent, with each wave having a participation rate of about 80% (detailed in Supplementary Table S1).

To ascertain the network ties, a separate data set was created that linked individuals through self-described social ties. Specifically, information from archived, handwritten documents that had been used by FHS staff members to help keep track of individuals was computerized. These sheets record the answers when all 5124 of the egos were asked to comprehensively identify friends, neighbors (based on address), coworkers (based on place of employment) and relatives. Although these tracking sheets were used as a way to optimize participant follow-up, they also implicitly contain valuable social network information. Another unique feature of these administrative records that makes them valuable for social network research relates to the compact nature of the Framingham population in the period from 1971 to 2007. This feature meant that many of the nominated contacts were themselves also participants of one or another FHS cohort, thus allowing for detailed data on them as well.

Through these self-described ties, we developed network links from FHS Offspring participants to other participants in any of the four FHS cohorts. Thus, for example, it is possible to know which participants have a relationship (for example, spouse, sibling, friend, coworker, neighbor) with other participants. It is interesting to note that each link between two people might be identified by either party identifying the other; this observation is most relevant to the 'friend' link, as we can make this link either when A nominates B as a friend, or when B nominates A (and, as discussed below, this directionality is also methodologically important). People in any of the FHS cohorts may marry or befriend or work with or live next to each other. Figure 1 is an illustration of these network ties, and shows the largest connected part of the network (known as a 'component'), of friends, spouses, and siblings to illustrate the clustering of moderately depressed (green nodes) and very depressed (blue nodes) people in 2000.

Finally, complete records of participants' (and their contacts') addresses since 1971 were used in our analyses. Because of the high accuracy of addresses in the FHS data (even though people spread out across the USA over time), and the wealth of information available about each subject's residential history, we have been able to correctly assign addresses to virtually all subjects. Through address mapping technologies, it was possible to determine (1) who is whose neighbor, and (2) what the geographical distance between individuals was.[22]

Outcome measures

Depression is measured using the Center for Epidemiological Studies Depression Scale (CES-D). The CES-D is well established as a screening method for depression with good reliability and validity.[23–26] The scale consists of a 20-item questionnaire where subjects are asked how often during the previous week they experienced a particular feeling that is associated with depression, with four possible answers, 0–1 days, 1–2 days, 3–4 days and 5–7 days. Scoring yields a scale from 0 (least depressed) to 60 (most depressed), with a score of 16 or above used to identify individuals with depressive illness. The CES-D was administered between 1983 and 2001 at times corresponding to the 5th, 6th, and 7th examinations of the Offspring Cohort. The median year of examination for these individuals was 1986 for exam 5, 1996 for exam 6 and 2000 for exam 7. Table 1 shows summary statistics for the network, including mean CES-D scores as well as social ties and demographic information. The score means and incidence of depression over time (CES-D over >16) track with other national estimates.[1]

Figure 1 Depression Clusters in the Framingham Social Network. This graph shows the largest component of friends, spouses and siblings at exam 7 (centered on the year 2000). There are 957 individuals shown. Each node represents a subject and its shape denotes gender (circles are male, squares are female). Lines between nodes indicate relationship (red for siblings, black for friends and spouses). Node color denotes the percentile score of the mean level of depression in ego and all directly connected (distance 1) alters, with yellow being below the 80th percentile, shades of green being the 80th to 95th percentile, and blue being above the 95th percentile (the most depressed).

Analytic methods

To evaluate the association of an ego's social network with an ego's depressive symptoms, various factors were included in our models, ranging from the prospective effect of alters' symptoms, social network variables and other control variables. Ego CES-D scores were regressed on ego age, gender, education and depression in the previous exam, as well as alter age, gender and depression in the current and previous exam. Ego depression at the previous exam was included to eliminate serial correlation in the errors and also control ego's genetic endowment and any intrinsic, stable tendency to be depressed. Including the alter's depression at the previous exam helps control for homophily[25] as shown by monte carlo simulations in previous work.[27]

The key coefficient in the model that measures the effect of induction is the variable for contemporaneous alter depression. Generalized estimating equation (GEE) procedures were used to account for multiple observations of the same ego across waves and across ego-alter pairings.[28] An independent working correlation structure was assumed for the clusters.[29] These analyses underlie the results presented in Figure 4. The GEE regression models provide parameter estimates in the form of β-coefficients, whereas the results reported in the text and in Figures 4 and 5 are in the form of risk ratios, which are related to the exponential coefficients. Mean effect

Table 1 Summary statistics for egos and alters (across all waves)

Variable	*Mean*	*s.d.*	*Min.*	*Max*	*Observed cases*
CES-D score	5.84	7.32	0	54	7603
Depression status (CES-D score 16+)	0.10	0.30	0	1	7603
Number of close friends	0.90	0.90	0	6	8309
Number of family members	2.81	3.07	0	23	8309
Network centrality (times 1000)	0.87	12.90	0	235.35	8309
Female	0.55	0.50	0	1	8309
Years of education	13.57	2.41	2	17	7159
Age	63.79	11.84	29.67	101.28	8309

Abbreviation: CES-D, Center for Epidemiological Studies Depression Scale.

sizes and 95% confidence intervals are calculated by simulating first difference in alter contemporaneous depression status (changing from 0–1) using 1000 randomly drawn sets of estimates from the coefficient covariance matrix and assuming all other variables are held at their means.[30] Results were checked using a linear specification on the raw CES-D score and none of these models changed the significance of any reported result.

The models include exam fixed effects, which, combined with age at baseline, account for the aging of the population. The sample size is shown for each model, reflecting the total number of all relevant ties, with multiple observations for each tie if it was observed in more than one exam, and allowing for the possibility that a given person can have multiple ties. Such multiple ties were handled with GEE procedures, clustering on ego.

To test for the possibility of omitted variables or contemporaneous events explaining the associations, we used longitudinal models and also examined how the type or direction of the social relationship between ego and alter affects the association between ego and alter. If unobserved factors drive the association between ego and alter, then directionality of friendship should not be relevant. Depression in the ego and the alter will move up and down together in response to the unobserved factors. In contrast, if an ego names an alter as a friend but the alter does not reciprocate, then we assume a causal relationship would indicate that the alter would significantly affect the ego, but the ego would not necessarily affect the alter.

The sensitivity of our results to model specification was tested by conducting numerous other analyses (not shown here) each of which had various strengths and limitations, but none of which yielded substantially different results than those presented. For example, although only a single friend was identified for most of the egos at any given time, the question of how multiple observations on some egos (who had more than one friend on one or more waves) affect the s.e. of our models was considered. Huber–White sandwich estimates with clustering on the egos yielded very similar results. In another case, the presence of serial correlation in the GEE models was tested using a Lagrange multiplier test, which found none after including the lagged dependent variable.[31] To check for multicollinearity, we measured the variance inflation factor for all variables in each regression reported here. All variance inflation factor values were 1.2 or lower, far below the value of 2.5 that typically warrants concern.

The Kamada–Kawai algorithm was used to prepare images of the networks, such as that in Figure 1.[32] The algorithm is a visualization tool that iteratively repositions nodes to reduce the number of ties that cross each other. The fundamental pattern of ties in a social network (known as the 'topology') is fixed, but how this pattern is visually rendered depends on the analyst's objectives.

To be sure any clustering of depressed people shown in Figure 1 is not simply due to chance, the following permutation test was implemented: the observed network was compared with 1000 randomly generated networks in which we preserved the network topology and the overall prevalence of depression but in which the assignment of the depression value was randomly shuffled to each node.[33] For this test, depression was dichotomized to be 1 if the respondent had a CES-D score of 16 or greater, and 0 otherwise. If clustering in the social network is occurring, then the probability that an alter is depressed given that an ego is depressed should be higher in the observed network than in the random networks. This procedure also allows us to generate confidence intervals and measure how far, in terms of social distance, the correlation in depression between ego and alter reaches.

Results

Figure 1 shows the largest connected component of friends, spouses and siblings and illustrates the clustering of moderately depressed (green nodes) and very depressed (blue nodes) people. In addition to illustrating clustering, this graph also visually suggests a relationship between depressive symptoms and being socially peripheral (being located on the edge of the network).

Figure 2 shows the correlation of alters depression scores over time based on their degrees of separation from an ego. The results suggest that there is a significant relationship between ego and alter depression, and this relationship extends up to three degrees of separation. In other words, a person's depression depends not just on his friend's depression, but also extends to his friend's friend and his friend's friend's friend. The full network shows that subjects are 93% (95% CI 59–135%) more likely to be depressed if a person they are directly connected to (at one degree of separation) is depressed. The size of the effect for people at two degrees of separation (the friend of a friend) is 43% (95% CI 21–70%) and for people at three degrees of separation (the friend of a friend of a friend) is 37% (95% CI 16–60%). At four degrees of separation the effect disappears (–2%, 95% CI –15–11%)., a result that is in line with other results that have shown similar drop-offs after three degrees of separation, including obesity, smoking, happiness and loneliness.[9–12]

To evaluate the impact of one aspect of network structure, our analyses, shown in Table 2, suggest that people with more friends and social connections are less likely to experience depressive symptoms in the future (Model 1), in keeping with past work. Each extra connection reduces the CES-D score by about 0.3 points. Interestingly, the same model shows that the number of family members has no effect at all ($P = 0.32$). The results in Model 2 suggest that people who feel depressed are likely to have significantly fewer friends in the future. In fact, compared with

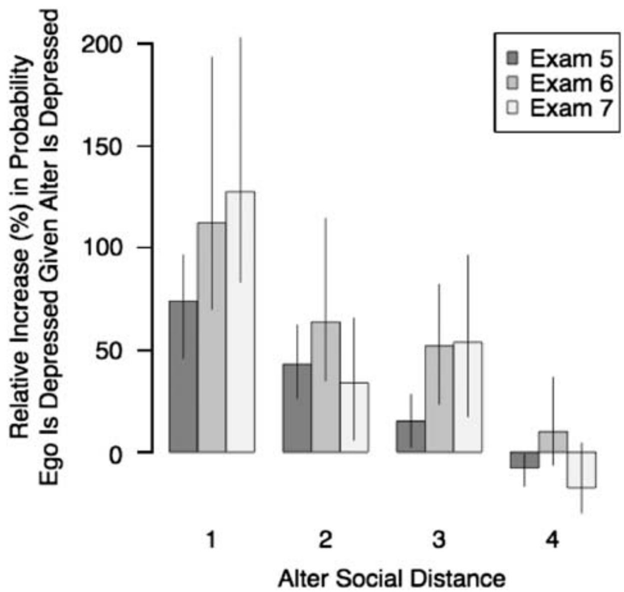

Figure 2 Social Distance and Depression in the Framingham Social Network. This figure shows for each exam the percentage increase in the likelihood a given ego is depressed if a friend or family member at a certain social distance is depressed (where depressed is defined as a score greater than 16 or greater on the CES-D). Values are derived by comparing the conditional probability of being depressed in the observed network with an identical network (with topology and incidence of depression preserved) in which the same number of depressed subjects are randomly distributed. Alter social distance refers to closest social distance between the alter and ego (alter = distance 1, alter's alter = distance 2, etc.). Error bars show 95% confidence intervals. CES-D, Center for Epidemiological Studies Depression Scale.

people who show no depressive symptoms, they will lose about 6% of their friends on average over a roughly 4-year period. For comparison, Model 3 shows that depression has no effect on the future number of family members a person has. These results are symmetric to both incoming and outgoing ties (not shown—available on request); that is, depressed people tend to receive fewer friendship nominations, but they also tend to name fewer people as friends as well. These results suggest that a person's depression may shape their social network as well as be shaped by it (further evidence of this effect is included in the online appendix).

In Table 3, we show how depression is influenced by an additional measure called 'eigenvector centrality' which indicates how central a person is in the whole network (see appendix for formal definition). The larger this value, the better connected a person is to all people in the network either directly via friends and family or indirectly via the friends and family of their friends and family. The model shows that network centrality in the previous exam significantly decreases the likelihood of depression, and it does so even when we control for the number of direct ties to friends and family. This suggests that a person's relationship to the whole network is important, above and beyond how well connected a person is to immediate friends and family.

Figure 3 shows the smoothed bivariate relationship between the fraction of a person's friends and family who are depressed at one exam, and the likelihood they will be depressed at the following exam. Here, the issue is not how many contacts a person has or how central a person is, but whether a large or small

Table 2 Prospective influence of friends and family on depression and vice versa

	Dependent variable								
	Current CES-D score			*Current number of friends*			*Current number of family*		
	Co-eff	*s.e.*	P-*value*	*Co-eff*	*s.e.*	P-*value*	*Co-eff*	*s.e.*	P-*value*
Previous CES-D score	0.456	0.019	0.000	−0.002	0.001	0.008	−0.001	0.001	0.142
Previous number of friends	−0.289	0.093	0.002	0.901	0.007	0.000	0.933	0.003	0.000
Previous number of family	0.024	0.025	0.323	−0.003	0.002	0.042	−0.029	0.007	0.000
Age	0.022	0.008	0.007	−0.002	0.000	0.000	0.001	0.001	0.004
Years of education	−0.169	0.037	0.000	0.002	0.002	0.206	−0.006	0.003	0.020
Female	1.106	0.159	0.000	−0.015	0.009	0.091	0.015	0.012	0.213
Exam 7	1.159	0.169	0.000	0.005	0.009	0.543	0.039	0.012	0.001
Constant	−4.373	1.450	0.003	0.113	0.076	0.135	−0.250	0.093	0.007
Deviance	244229			723			1291		
Null deviance	327768			4893			57482		
N	6113			6113			6113		

Abbreviation: CES-D, Center for Epidemiological Studies Depression Scale.
Results for linear regression of ego's CES-D score, number of friends, and number of family members at current exam on previous CES-D score, number of friends, and number of family plus other covariates. Models were estimated using a general estimating equation (GEE) with clustering on the ego and an independent working covariance structure. Models with an exchangeable correlation structure yielded poorer fit. Fit statistics show the sum of squared deviance between predicted and observed values for the model and a null model with no covariates.

Table 3 Prospective influence of centrality on depression

	Dependent variable: current CES-D score		
	Co-eff	*s.e.*	P-*value*
Previous network centrality	**−8.285**	**4.142**	**0.045**
Previous number of friends	−0.278	0.094	0.003
Previous number of family	0.456	0.019	0.000
Previous CES-D score	0.036	0.028	0.195
Age	0.022	0.008	0.006
Years of education	−0.168	0.037	0.000
Female	1.110	0.159	0.000
Exam 7	1.161	0.169	0.000
Constant	−4.469	1.454	0.002
Deviance	244061		
Null deviance	327588		
N	6113		

Abbreviation: CES-D, Center for Epidemiological Studies Depression Scale.
Model of ego's CES-D score at current exam regressed on measures from the previous exam including ego's network centrality, number of friends, number of family, plus other covariates. The model was estimated using a general estimating equation (GEE) with clustering on the ego and an independent working covariance structure. Models with an exchangeable correlation structure yielded poorer fit. Fit statistics show sum of squared deviance between predicted and observed values for the model and a null model with no covariates. The main results (coefficients in bold) show that network centrality is associated with a decrease in future depressive symptoms, even controlling for the number of friends and family. This suggests that connection to more socially distant alters (for example, friends of friends) also influences depression.

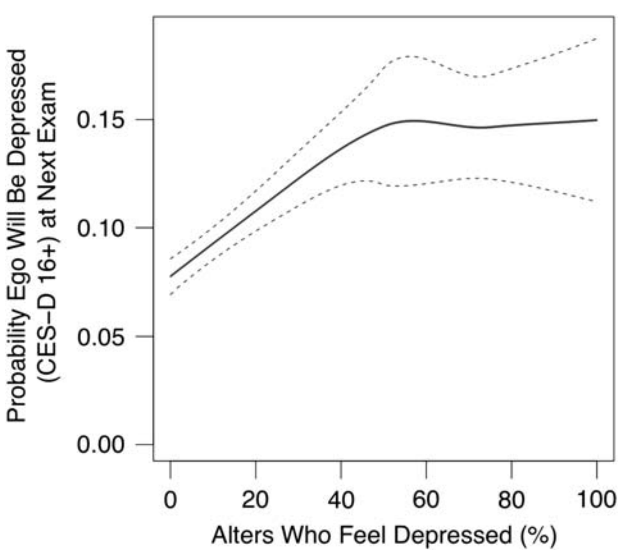

Figure 3 Depressed Alters in the Framingham Social Network. This plot shows that the probability of being depressed (CES-D score of 16 or greater) in exams 6 and 7 is positively associated with the fraction of their friends and family in the previous exam who are depressed. Blue line shows smoothed relationship based on bivariate LOESS regression, and dotted lines indicate 95% confidence intervals. CES-D, Center for Epidemiological Studies Depression Scale.

Table 4 Influence of number of depressed alters on ego depression

	Dependent variable: CES-D score		
	Co-eff	*s.e.*	P-*value*
Previous number of depressed alters	0.762	0.190	0.000
Previous number of non-depressed alters	−0.128	0.058	0.026
Previous CES-D score	0.431	0.021	0.000
Age	−0.011	0.010	0.265
Years of education	−0.138	0.041	0.001
Female	1.129	0.172	0.000
Exam 7	1.457	0.191	0.000
Constant	−4.926	1.611	0.002
Deviance	190439		
Null deviance	251500		
N	4913		

Abbreviation: CES-D, Center for Epidemiological Studies Depression Scale.
Results for linear regression of ego's depression on previous depression, number of depressed friends and family (16 or greater on CES-D), number of non-depressed friends and family, and other covariates. Models were estimated using a general estimating equation (GEE) with clustering on the ego and an independent working covariance structure. Models with an exchangeable correlation structure yielded poorer fit. Fit statistics show sum of squared deviance between predicted and observed values for the model and a null model with no covariates.

fraction of those contacts is also depressed. The relationship is significant and nearly doubles the likelihood of depression for the average person who is surrounded by other depressed people compared with those who are not connected to anyone who is depressed.

Analyses testing the impact of depressed alters over time are presented in Table 4. The results show that each additional depressed alter significantly increases the number of days a subject feels depressed ($P<0.001$). Conversely, each additional non-depressed alter significantly decreases the number of days a subject feels depressed ($P=0.026$). But these effects are asymmetric: depressed alters are about six times more influential than non-depressed alters, and the difference in these effect sizes is itself significant ($P=0.0006$). Therefore, the feeling of depression seems to spread more easily than its absence.

Results of GEE models that estimate specific effects for friends, spouses, co-workers, siblings and neighbors are presented in Figure 4. If a friend is depressed, this increases the probability ego and is depressed by 118% (95% CI 6–290%). Among friends, it is possible to distinguish additional possibilities. As each person was asked to name a friend, and not all of these nominations were reciprocated, we have ego-perceived friends (denoted 'friends'), 'alter-perceived friends' (alter named ego as a friend, but not vice versa) and 'mutual friends' (ego and alter nominated each other). If mutual friends are depressed, it increases the likelihood of depression for the ego by 359% (95% CI: 25–1095%). In contrast, the influence

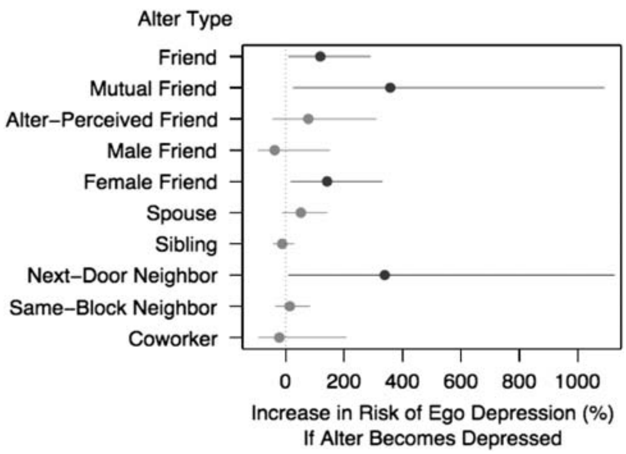

Figure 4 Alter Type and Depression in the Framingham Social Network. This graph shows the change in likelihood of depression given that an alter is depressed. Estimates are based on generalized estimating equation logit models of depression on several different sub-samples of the Framingham Heart Study Social Network. The dependent variable in each model is ego depression status and independent variables include lagged ego depression status, alter depression status, lagged alter depression status, ego age, gender, and education and fixed effects for each wave. Full models and equations are available in the appendix. Mean effect sizes and 95% confidence intervals were calculated by simulating first difference in alter contemporaneous depression status (changing from 0–1) using 1000 randomly drawn sets of estimates from coefficient covariance matrix and assuming all other variables are held at their means.

of alter-perceived friends is not significant ($P = 0.38$). These results suggest that the associations in the social network were not merely because of confounding, for, in the null hypothesis, one would assume that the significance and effect sizes for different types of friendships should be the same. For example, if some third factor were explaining both ego and alter depression, there should be no observed influences from directionality or strength of the social tie in question. Further analyses of person-to-person effects can be found in the online appendix. In particular, we show that the main results remain significant even when we control for dynamic changes in the number of network ties to each person, longitudinal attrition, occupational prestige and marital status. These results can also be found in the Supplementary online appendix.

An intriguing difference between inter-personal effects in depression and previous work with respect to happiness[11] is shown in gender effects. When friends are divided up by gender in Table 4, depression is found to spread much more easily from women than from men. When a female friend becomes depressed, it increases the probability that the ego is depressed by 142% (CI 18–331%). In contrast, when a male friend becomes depressed, on average it has no significant effect ($P = 0.34$). However, similar differences in receptivity were not found; men and women appear to be equally sensitive to their friends (Table 5).

Discussion

These results support the hypothesis that depressive symptoms as defined by CES-D scores can be observed to travel along social networks. They suggest that both decreases and increases in CES-D scores (and classifi-

Table 5 Association of friend's depression and ego depression

	Alter type				
	Friend	*Mutual friend*	*Alter-perceived friend*	*Male Friend*	*Female friend*
Alter currently depressed	0.78 (0.36)	1.43 (0.61)	0.49 (0.56)	−0.92 (0.96)	0.93 (0.38)
Alter previously depressed	0.68 (0.42)	2.23 (0.63)	0.71 (0.45)	0.43 (0.75)	0.77 (0.44)
Ego previously depressed	2.34 (0.31)	2.35 (0.72)	1.12 (0.53)	2.57 (0.71)	2.25 (0.33)
Exam 7	0.47 (0.27)	−0.75 (0.64)	0.22 (0.43)	−0.20 (0.50)	0.68 (0.32)
Ego's age	0.00 (0.02)	0.08 (0.04)	0.01 (0.02)	0.03 (0.04)	−0.02 (0.03)
Ego female	0.63 (0.40)	−0.35 (0.58)	1.09 (0.45)	1.42 (0.60)	0.30 (0.70)
Ego's years of education	−0.24 (0.09)	−0.42 (0.21)	0.03 (0.10)	−0.25 (0.13)	−0.25 (0.12)
Constant	−0.48 (2.13)	−3.24 (4.25)	−4.89 (2.41)	−2.48 (3.51)	0.79 (2.65)
Deviance	51	11	36	12	38
Null deviance	66	18	38	16	49
N	858	265	572	359	499

Coefficients and standard errors in parenthesis for linear logit regression of ego's depression status on covariates are shown. Observations for each model are restricted by type of relationship (for example, the leftmost model includes only observations in which the ego named the alter as a 'friend' in the previous and current period). Models were estimated using a general estimating equation with clustering on the ego and an independent working covariance structure. Models with an exchangeable correlation structure yielded poorer fit. Fit statistics show sum of squared deviance between predicted and observed values for the model and a null model with no covariates.

cation as being depressed) were strongly correlated with depression measures in one's friends and neighbors. The correlations decrease significantly the further away (in terms of degree of separation) an alter is in an ego's social network. This is, to our knowledge, the first such analysis performed looking specifically at depressive symptoms in such a network-based, longitudinal manner. These results are similar to previous work showing the spread of obesity, smoking and happiness within social networks.[9–11]

Another important finding related to the how CES-D scores and depression appeared to affect (and be affected by) the actual architecture of the network itself. The results suggest that not only may depressed mood spread across social ties, but also that depression depends on how connected individuals are and where they are located within social networks. Our work suggests that to understand someone's CES-D score, one must ask both where they are situated within a network (for example, how central they are to the network or how many friends they have), and how the people around them are actually feeling (quite apart from their topological location). An alternative interpretation of these results is that causality may be reversed, meaning that the onset of depression affects social networks themselves by inducing tie formation and/or dissolution; however, analyses of our data indicates that CES-D scores did not predict changes in ties, thus making this hypothesis less likely.[34]

A related topological finding comes from the analysis of individuals on the periphery of the network. These people have significantly worse CES-D scores and have fewer friends. Their isolation appears to be correlated not only with having fewer friends as baseline, but also the likelihood for them to cut any remaining ties that they have left. Given that these peripheral individuals' depression is correlated with future scores of those friends further within the network, it suggests that isolation as well as clustering may have an impact on the spread of depressive symptoms. This finding also suggests that selective targeting of more socially isolated individuals for interventions might be particularly cost effective from a societal standpoint, benefitting both them and others.

There are some surprising findings with regards to types of alters and the observed effects. The gender of alters appears to be significant, with women appearing more influential than men over time. While recent work has challenged the idea that women are more 'emotional' than men, there remains a vast literature on the differences in emotional expression between men and women.[35,36] For example, women have been found to be more emotionally expressive than men, often using non-verbal cues to express their emotional state.[37–39] One conceivable hypothesis is that, based on these differences, women may be more effective at communicating certain mood states within dyads, a trait that may have an evolutionary origin.[40]

The results also would seem to suggest that spouses are significantly less influential than friends over the course of time. This finding, also found in previous work on happiness,[11] on the surface appears puzzling (and conceivably problematic). One would imagine that spouses would be particularly influential on each other. For example, work in cross-sectional data suggests a strong correlation between spouses with regards to mood symptoms.[41,42] However, further work by Siegel *et al*.[43] showed using that, when looking at longitudinal data on spousal mood, and the across-time correlation of spousal changes in mood symptoms, the magnitude of effects found were quite similar to those found in this paper. One possible explanation for these findings is that there is more homophily between couples than friends. Put another way, just as spouses are sometimes referred to as 'joined at the hip,' so might be their CES-D scores from one period to the next. Our models, which control for homophily on depression (that is, the tendency of spouses to choose each other based on a predilection to a certain mood), can thus uncover a possibly greater residual effect of induction between friends than between spouses. This finding would suggest that, while heterophily ('opposites attract') may be an important factor in mate selection for some traits, it might not be with regards to factors that influence CES-D scores.[17]

This work has a few notable limitations. It does not randomize individuals into social networks, thus leaving open the possibility that these results may in part reflect homophily-driven selection bias on the basis of unobserved phenotypes that influence the development and transmission of depressive symptoms over time. Our group is actively addressing this question through the use of quasi-experimental approaches and other methods in ongoing work. Another limitation is that the outcome measure (CES-D) is not a clinical tool, so it is not possible to make any specific conclusions about the spread of clinical depression in our sample. Also, the sample is somewhat homogenous and does not have a significant percentage of underrepresented minorities in it. Furthermore, our data set captures a limited number of close friendship ties, thus we cannot extend our analyses to more broad social network ties (such as online interactions). Finally, the FHS social network data set is unusual due to its longitudinal nature and relative completeness; it is our hope that the further collection of social network data in other settings will provide more opportunities to evaluate such hypotheses.

In conclusion, we consider these results to be an important step towards better understanding the impact of interpersonal relationships and social networks on the development of depressive symptoms. By identifying how the structure of networks may determine the spread of clinically relevant conditions (and vice versa), future policy may be able to target individuals within a network to maximize the impact of a policy. An example of this, known as 'seeding,' seeks to use well-connected individuals to spread information and is the subject of a number of ongoing research projects.[44] We hope that social networks and their effects will be considered in the design of future research studies and public policies seeking to address the important issue of depression in our society. People are connected, and so their mental health is connected.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

We thank Laurie Meneades, Rebecca Joyce, Molly Collins, Marian Bellwood and Karen Mutalik for the expert assistance required to build the analytical data. We thank Maurizio Fava for helpful comments regarding the article.

References

1 Fava M, Cassano P. Major depressive disorder and dysthymic disorder. In: Stern T, Rosenbaum J, Biederman J, Fava M, Rauch S, (eds). *The MGH Textbook of Comprehensive Clinical Psychiatry*. Mosby-Elsevier: Philadelphia, 2008.
2 Greenberg PE, Leong SA, Birnbuam HG. Cost of depression: current assessment and future directions. *Expert Rev Pharmacoecon Outcomes Res* 2001; **1**: 89–96.
3 Engel GL. The need for a new medical model. *Science* 1977; **196**: 129–136.
4 Durkheim E. *Suicide*. Translated by John A. Spaulding and George Simpson. Edited with an Introduction by George Simpson. The Free Press: Glencoe, IL, 1951, 405pp.
5 Vilhjalmsson R. Life stress, social support and clinical depression: a reanalysis of the literature. *Social Sci Med* 1993; **37**: 331–342.
6 Kim D. Blues from the neighborhood? neighborhood characteristics and depression. *Epidemiol Rev* 2008; **30**: 101–117.
7 Raphael B. Unmet need for prevention. In: Andrews G, Henderson S (eds). *Unmet Need in Psychiatry: Problems, Resources, Responses*. Cambridge University Press: New York, 1998, pp 138–139.
8 Subramanian SV, Kim D, Kawachi I. Covariation in the socioeconomic determinants of self rated health and happiness: a multivariate multilevel analysis of individuals and communities in the USA. *J Epidemiol Community Health* 2005; **59**: 664–669.
9 Christakis NA, Fowler JH. The Spread of obesity in a large social network over 32 years'. *New Engl J Med* 2007; **357**: 370–379.
10 Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network'. *New Engl J Med* 2008; **358**: 2249–2258.
11 Fowler JH, Christakis NA. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. *Br Med J* 2008; **337**: a2338.
12 Cacioppo JT, Fowler JH, Christakis NA. Alone in the crowd: the structure and spread of loneliness in a large social network. *J Pers Soc Psychol*, in press 2009; **97**: 977–991.
13 Cutler DM, Glaeser EL. Social interactions and smoking. *NBER Working Paper Series* 2007; W13477.
14 Trogdon JG, Nonnemaker J, Pais J. Peer effects in overweight adolescents. *J Health Econ* 2008; **25**: 1388–1399.
15 Hatfield E, Cacioppo JT, Rapson RL. *Emotional contagion*. Cambridge University Press: New York, 1994.
16 Bearman P, Moody J. Suicide and friendships among American adolescents. *Am J Public Health* 2004; **94**: 89–95.
17 McPherson M, Smith-Lovin L, Cook JM. Birds of a feather: homophily in social networks. *Ann Rev Sociol* 2001; **27**: 415–444.
18 Carrington PJ, Scott J, Wasserman S. *Models and Methods in Social Network Analysis*. Cambridge University Press: Cambridge, 2005.
19 Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families: the Framingham offspring study. *Am J Epidemiol* 1979; **110**: 281–290.
20 Cupples LA, D'Agostino RB. Survival following initial cardiovascular events: 30 year follow-up. In: Kannel WB, Wolf PA, Garrison RJ, (eds). *The Framingham Study: An epidemiological investigation of cardiovascular disease*. NHLBI, NIH: Bethesda, MD, 1988.
21 Quan SF, Howard BV, Iber C *et al*. The sleep heart health study: design, rationale, and methods. *Sleep* 1997; **20**: 1077–1085.
22 Fitzpatrick GL, Modlin ML. *Direct-Line Distances: International Edition*. The Scarecrow Press: Metuchen, NJ, 1986.
23 McDowell I, Newell C. *Measuring Health, a Guide to Rating Scales and Questionnaires*, 2nd edn. Oxford University Press: New York, 1996.
24 Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *J Appl Psychol Meas* 1977; **1**: 385–401.
25 Weissman MM, Sholomskas D, Pottenger M, Prusoff BA, Locke BZ. Assessing depressive symptoms in five psychiatric populations: a validation study. *Am J Epidemiol* 1977; **706**: 203–213.
26 Boyd JH, Weissman MM, Thompson WD, Meyers JK. Screening for depression in a community sample: understanding discrepancies between depression symptom and diagnostic scales. *Arch Gen Psychiatry* 1982; **39**: 1195–1200.
27 Fowler JH, Christakis NA. Estimating peer effects on health in social networks. *J Health Econ* 2008; **27**: 1400–1405.
28 Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986; **73**: 13–22.
29 Schildcrout JS. Regression analysis of longitudinal binary data with time-dependent environmental covariates: bias and efficiency. *Biostatistics* 2005; **6**: 633–652.
30 King G, Tomz M, Wittenberg J. Making the most of statistical analyses: improving interpretation and presentation. *Am J Pol Sci* 2000; **44**: 341–355.
31 Beck N. Time-series–cross-'tion data: what have we learned in the past few years? *Ann Rev Pol Sci* 2001; **4**: 271–293.
32 Kamada T, Kawai S. An algorithm for drawing general undirected graphs. *Inf Process Lett* 1989; **31**: 7–15.
33 Szabo G, Barabasi AL. Network effects in service usage. Available at http://lanl.arxiv.org/abs/physics/0611177 accessed 12 December 2007.
34 O'Malley AJ, Christakis NA. The role of health traits in the longitudinal formation and dissolution of friendship ties in a large social network over 32 years. 2009. *Working Paper*.
35 Fischer A (ed). *Gender and Emotion; social psychological perspectives*. Cambridge University Press: New York, 2000.
36 Canary DJ, Kathryn D. *Sex Differences and Similarities in Communication*. Lawrence Erlbaum Associates: Mahwah, New Jersey, 1998.
37 Eakins B, Eakins G. *Sex Differences in Human Communication* 1979, Houghton Mifflin Co.: Boston, pp 147–179.
38 Noller P. *Nonverbal Communication and Marital Interaction*. Pergamon Press: Oxford, 1984.
39 Hall JA, Carter JD, Horgan TG. Gender differences in non-verbal communication of emotion. In: Fischer AH (ed). *Gender and Emotion: Social Psychological Perspective*. Cambridge University Press: Cambridge England, 2000.
40 Geary DC. *Mxale, Female: The Evolution of Human Sex Differences*. American Psychological Association Press: Washington DC, 1998.
41 Christakis NA, Fowler JH. *Connected: The Suprising Power of Our Social Networks and How They Shape Our Lives*. Little Brown: New York, 2009.
42 Meyler D, Stimpson JP, Peek MK. Health concordance within couples: a systematic review. *Soc Sci Med* 2007; **64**: 2297–2310.
43 Siegel MJ, Bradley EH, Gallo WT, Kasl SV. The effect of spousal mental and physical health on husbands' and wives' depressive symptoms, among older adults. *J Aging Health* 2004; **16**: 398–425.
44 Watts DJ, Dodds PS. Networks, influence, and public opinion formation. *J Consum Res* 2007; **34**: 441–458.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (http://www.nature.com/mp)