Open Access: free full text available online
For the full versions of these articles see bmj.com

RESEARCH

Dynamic spread of happiness in a large social network: longitudinal analysis of the Framingham Heart Study social network

James H Fowler,[1] Nicholas A Christakis[2]

EDITORIAL by Steptoe and Diez Roux
RESEARCH p28

[1]Department of Political Science, University of California, San Diego, CA, USA
[2]Department of Health Care Policy, Harvard Medical School, and Department of Sociology, Harvard University, Cambridge, MA, USA
Correspondence to: N A Christakis christak@hcp.med.harvard.edu

Cite this as: *BMJ* 2008;337:a2338
doi:10.1136/bmj.a2338

ABSTRACT

Objectives To evaluate whether happiness can spread from person to person and whether niches of happiness form within social networks.

Design Longitudinal social network analysis.

Setting Framingham Heart Study social network.

Participants 4739 individuals followed from 1983 to 2003.

Main outcome measures Happiness measured with validated four item scale; broad array of attributes of social networks and diverse social ties.

Results Clusters of happy and unhappy people are visible in the network, and the relationship between people's happiness extends up to three degrees of separation (for example, to the friends of one's friends' friends). People who are surrounded by many happy people and those who are central in the network are more likely to become happy in the future. Longitudinal statistical models suggest that clusters of happiness result from the spread of happiness and not just a tendency for people to associate with similar individuals. A friend who lives within a mile (about 1.6 km) and who becomes happy increases the probability that a person is happy by 25% (95% confidence interval 1% to 57%). Similar effects are seen in coresident spouses (8%, 0.2% to 16%), siblings who live within a mile (14%, 1% to 28%), and next door neighbours (34%, 7% to 70%). Effects are not seen between coworkers. The effect decays with time and with geographical separation.

Conclusions People's happiness depends on the happiness of others with whom they are connected. This provides further justification for seeing happiness, like health, as a collective phenomenon.

INTRODUCTION

Emotional states can be transferred directly from one individual to another by mimicry and "emotional contagion,"[1] perhaps by the copying of emotionally relevant actions, particularly facial expressions. We were interested in the impact of social network structure on happiness, in the spread of happiness between connected individuals, and in whether there are geographical or temporal constraints on such spread.

METHODS

Participants

The Framingham Heart Study was initiated in 1948, when 5209 people in Framingham, Massachusetts, were enrolled into the "original cohort." In 1971, the "offspring cohort," composed of most of the children of the original cohort, and their spouses, was enrolled. This cohort of 5124 people has had almost no loss to follow-up other than death (only 10 people dropped out). Enrolment of the so called "third generation cohort," consisting of 4095 children of the offspring cohort, began in 2002. At regular intervals participants in all these cohorts come to a central facility for detailed examinations and collection of survey data.

Network ascertainment

We used the offspring cohort as the source of 5124 key individuals to study—whom we term "egos." Each ego in this cohort is connected to other people via friendship, family, spousal, neighbour, and coworker relationships. Each relationship is a "social tie." Each person who has a relationship with an ego was called an "alter." We wanted to know how each of these alters influences an ego. Many of the alters also happened to be members of a studied cohort in Framingham, which means that we had access to detailed information about both the focal group (the "egos") and the people to whom they were connected (the "alters"). Overall, within the entire Framingham Heart Study social network, there were 12 067 individuals who were connected at some point in 1971-2003.

We created the network dataset from administrative tracking sheets used since 1971 to identify people close to participants for the purpose of follow-up. Participants were asked to identify their relatives, "close friends," place of residence, and place of work to ensure they could be contacted every two to four years for follow-up.

The study recorded complete information about all first order relatives (parents, spouses, siblings, children), whether alive or dead, and at least one close friend at up to seven examinations from 1971 to 2003. Home address was also coded to determine neighbour relationships. Specific information about place of

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a2338

employment allowed us to identify ties to coworkers within the network. For any given ego, a particular alter can be in only one mutually exclusive category—that is, spouse, sibling, friend, coworker, or neighbour. There were 53 228 observed social ties between the 5124 egos and any other alters in any of the Framingham Heart Study cohorts, yielding an average of 10.4 ties to family, friends, and coworkers over the course of follow-up.

Given the compact nature of the Framingham social network in the period 1971-2007, many of the nominated contacts were also participants of one or another Framingham Heart Study cohort[2 3] so we have detailed survey and physical examination information about both the ego and the alter. Importantly, 45% of the 5124 egos were connected via friendship to another person in the study; there were 3604 unique observed friendships for an average of 0.7 friendship ties per ego. We can study three different types. An "ego perceived friend" means the ego nominates an alter as a friend, but the nomination is not reciprocated. An "alter perceived friend" means that an alter nominates the ego as a friend but not vice versa. Finally, a "mutual friend" is one in which the nomination is reciprocal.

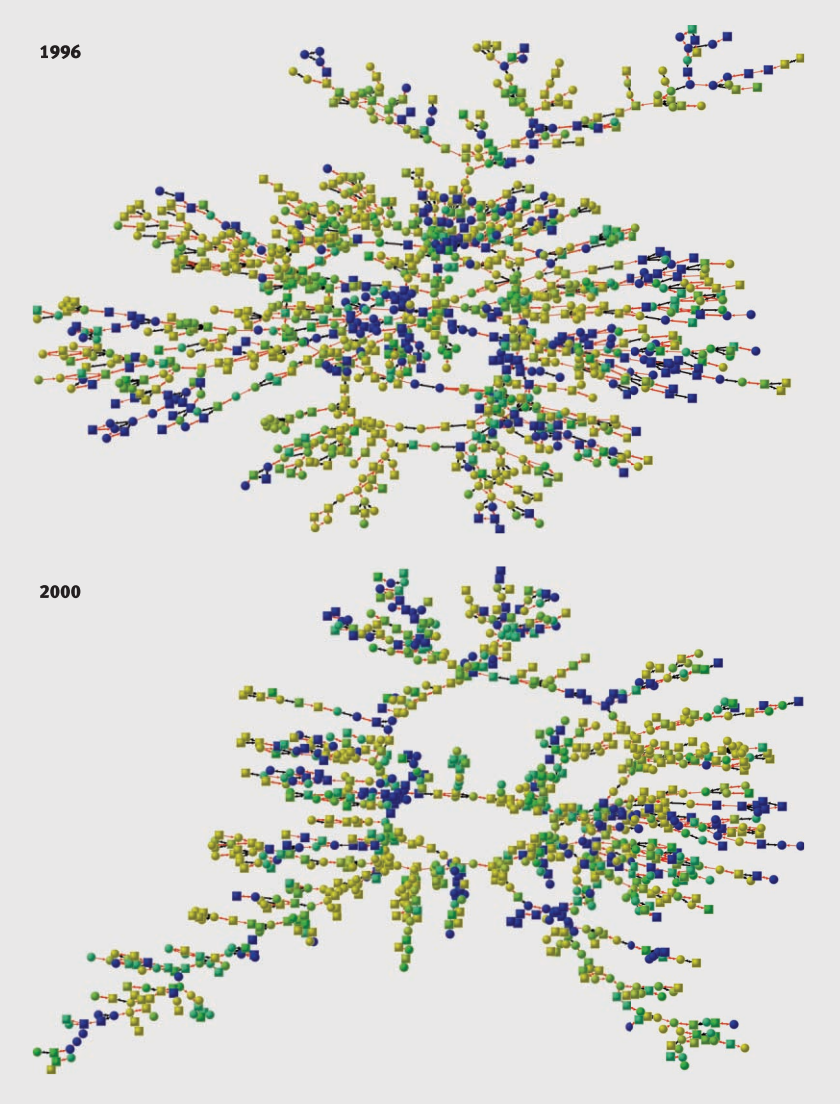


Fig 1 | Happiness clusters in the Framingham social network. Graphs show largest component of friends, spouses, and siblings at exam 6 (centred on year 1996, showing 1181 individuals) and exam 7 (year 2000, showing 1020 individuals). Each node represents one person (circles are female, squares are male). Lines between nodes indicate relationship (black for siblings, red for friends and spouses). Node colour denotes mean happiness of ego and all directly connected (distance 1) alters, with blue shades indicating least happy and yellow shades indicating most happy (shades of green are intermediate)

We hypothesised that the influence a friend has on an ego would be affected by the type of friendship, with the strongest effects occurring between mutual friends, followed by ego perceived friendships, followed by alter perceived friendships.

At inception, 53% of the egos were women; mean age was 38 years (range 21-70); and mean education was 1.6 years of college (range 0-≥17 years of education). Measures of occupational prestige for each ego at each wave were also available (see appendix on bmj.com). We studied 4739 of the 5124 egos who were alive in 1983 (the first time happiness was measured in the Framingham study). All participants were followed until 2003 as were any ties to alters noted during the time period 1983-2003.

Measures

The median year of examination for the offspring cohort was 1986 for exam 5, 1996 for exam 6, and 2000 for exam 7. We assessed happiness with the Center for Epidemiological Studies depression scale (CES-D). See bmj.com for details.

We were interested not just in whether individuals were happy or not but also in changes in their happiness over time. We used the previous wave as a baseline measure and evaluated the probability of an ego being happy at a succeeding wave. At follow-up, the prevalence of happiness was 61% in exam 6 and 59% in exam 7. The mean index score was 10.7 in exam 6 and 10.6 in exam 7. Between exams 6 and 7, for example, 16% of individuals became happy, 13% became unhappy, 49% remained happy, and 22% remained unhappy.

Network analysis

Social networks consist of two elements: individuals (nodes) and the relationships (social ties) between them. Once all the nodes and ties are known, one can draw pictures of the network and discern every person's position within it. Within a network, one can speak of the "distance" between two people (also known as the "degree of separation"), which is the shortest path in the network from one person to another.

Once a full set of individuals and ties is observed, there is only one network per se. This network, however, can be analysed or drawn in various ways. We used the Kamada-Kawai algorithm to prepare images of networks (fig 1). This algorithm is a visualisation tool that iteratively repositions nodes to reduce the number of ties that cross each other.

To test whether clustering of happy and unhappy people is due to chance, we compared the observed clustering to the clustering in 1000 randomly generated

networks in which we preserved the network topology and the overall prevalence of happiness but in which we randomly shuffled the assignment of the happiness value to each node.[4] If clustering is occurring, then the probability that an alter is happy given that an ego is happy should be higher in the observed network than in the random networks. This procedure also allowed us to generate confidence intervals and measure how far, in terms of social distance, the correlation in happiness between ego and alter reaches.

Measures of centrality in networks capture the extent to which a node connects, or lies between, other nodes, and hence its tendency to be positioned near the centre of his or her local network. See bmj.com for details and description of eigenvector centrality used here.

Statistical analysis

The association between the happiness of individuals connected to each other, and the clustering within the network, could be attributed to at least three processes: induction, whereby happiness in one person causes the happiness of others; homophily, whereby happy individuals choose one another as friends and become connected[5]; or confounding, whereby connected individuals jointly experience contemporaneous exposures. To distinguish between these effects requires repeated measures of happiness,[6 7] longitudinal information about network ties, and information about the nature or direction of the ties (for example, who nominated whom as a friend).

We evaluated regression models of ego happiness as a function of ego's age, sex, education, and happiness in the previous exam, and of the happiness of an alter in the current and previous exam. Inclusion of ego happiness in the previous exam helps to eliminate serial correlation in the errors and also substantially controls for ego's genetic endowment and any intrinsic stable predilection to be happy. Alter's happiness in the previous exam helps to control for homophily.[6 7] We evaluated the possibility of omitted variables or contemporaneous events or exposures in explaining the associations by examining how the type or direction of the social relationship between ego and alter affects the association between them. We also examined the possible role of exposure to neighbourhood factors by examining maps (see appendix on bmj.com).

The main coefficient of interest in these regression models is the one related to contemporaneous happiness in alters—that is, the extent to which an alter's present happiness, net of the alter's previous happiness, is associated with an ego's present happiness, net of the ego's previous happiness.[6 7] We used generalised estimating equation procedures to account for multiple observations of the same ego across waves and across ego-alter pairings.[8] We assumed an independent working correlation structure for the clusters.[9] See bmj.com.

RESULTS

Examination of the social network indicates that happy people tend to be connected to one another. Figure 1 shows the largest connected network component in 1996 and 2000 based on a restricted set of ties among siblings, spouses, and friends (coworker and neighbours are excluded to simplify the image).

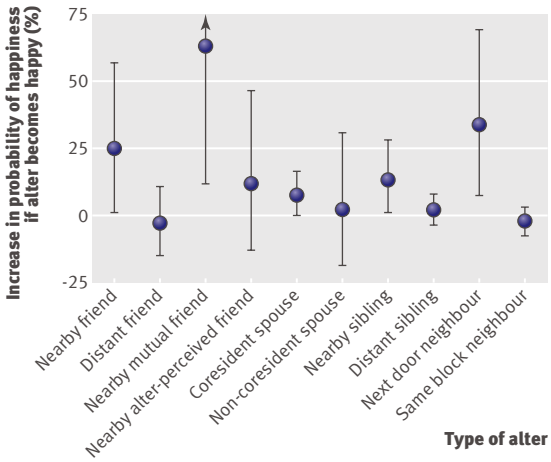


Fig 2 | Alter type and happiness in the Framingham social network. Friends, spouses, siblings, and neighbours significantly influence happiness, but only if they live close to ego. Effects estimated with generalised estimating equation logit models of happiness on several different subsamples of network (see table S6 in appendix on bmj.com)

The clusters of happy and unhappy people seen in the network are significantly larger than expected by chance. We can calculate the relationship of ego and alter happiness at various degrees of separation by measuring the probability that an ego is happy when an alter is happy and comparing it to the same probability in a simulated network in which we retain the observed network ties and prevalence of happiness, but randomly shuffle the observed happiness between nodes. We found the association between ego and alter happiness is significant up to three degrees of separation. A person is 15.3% (95% confidence interval 12.2% to 18.8%) more likely to be happy if a directly connected alter (distance one) is happy. The effect for distance two alters is 9.8% (7.0% to 12.9%) and for distance three alters is 5.6% (2.4% to 9.0%). See bmj.com.

Figure 1 also suggests a relation between network centrality and happiness: people at the core of their local networks seem more likely to be happy, while those on the periphery seem more likely to be unhappy. We tested this by computing eigenvector centrality measures for each subject. Generalised estimating equation regressions show that ego centrality is significantly associated with improved future happiness: a 2 SD increase in centrality (from low to medium or medium to high) increases the probability of being happy at the next examination by 14% (1% to 29%, P=0.03). Moreover, the relation between centrality and future happiness remained significant even when we controlled for age, education, and the total number of family and non-family alters. Thus, it is not only the number of direct ties but also the number of indirect ties that influence future happiness. The better connected are one's friends and family, the more likely one will attain happiness in the future. Conversely, happiness itself does not increase a person's centrality at subsequent time points (see

WHAT IS ALREADY KNOWN ON THIS TOPIC

Previous work on happiness and wellbeing has focused on socioeconomic and genetic factors

Research on emotional contagion has shown that one person's mood might fleetingly determine the mood of others

Whether happiness spreads broadly and more permanently across social networks is unknown

WHAT THIS STUDY ADDS

Happiness is a network phenomenon, clustering in groups of people that extend up to three degrees of separation (for example, to one's friends' friends' friends)

Happiness spreads broadly in social networks

Network characteristics independently predict which individuals will be happy years into the future

appendix on bmj.com). That is, network centrality leads to happiness rather than the other way around.

We specified generalised estimating equation regression models of ego happiness with the number of happy and unhappy alters in the previous exam as key predictors. The relation is highly significant, with each happy alter increasing the probability the ego is happy by about 9% (P=0.001), and each unhappy alter decreasing it by 7% (P=0.004). We also evaluated the simultaneous effect of total number of alters (whether happy or unhappy) and the fraction of alters who are happy. These models show that happy alters consistently influence ego happiness more than unhappy alters, and only the total number of happy alters remains significant in all specifications (see appendix on bmj.com). Thus, the social network effect of happiness is multiplicative and asymmetric. Each additional happy alter increases the likelihood of happiness, but each additional unhappy alter has little or no effect. The emotional state of a person's social relationships is more important to one's own emotional state than the total number of those relationships.

The principal determinant of a person's happiness was their previous happiness; individuals who were happy at one wave were about three times more likely than unhappy people to be happy at the subsequent observation, depending on what class of alters were included in the model. Age, sex, and education had effects consistent with previous research, with women being less happy then men and educated people being slightly happier (see appendix on bmj.com).

Our main interest was the impact on an ego of the happiness of others. Figure 2 shows the results of generalised estimating equation models that distinguish effects for friends, spouses, siblings, coworkers, and neighbours. We can use these results to estimate what would happen to the happiness of the ego if the alter were "switched" from being unhappy to being happy. "Nearby" friends (who live within a mile (1.6 km)) and who become happy increase the probability ego is happy by 25% (1% to 57%). "Distant" friends (who live more than a mile away) have no significant effect on ego. Nearby mutual friends have a stronger effect than nearby ordinary friends; when they become happy it increases the probability ego will be happy by 63% (12% to 148%). In contrast, the influence of nearby alter perceived friends is much weaker and not significant (12%, −13% to 47%). If the associations in the social network were merely caused by confounding, these effect sizes for different types of friendships should be more similar. That is, if some third factor were explaining both ego and alter happiness, it should not respect the directionality of the tie.

To further explore whether distance affects the spread of happiness, we varied the cut off for nearby friends. We found an ego is 42% (6% to 95%) more likely to be happy if a friend who lives less than half a mile (0.8 km) away becomes happy (net of controls, including ego's baseline happiness). In contrast, the effect is only 22% (2% to 45%) for friends who live less than two miles (3.2 km) away, and it declines and ceases to be significant at greater distances.

Past research also suggests that changes in happiness are temporary and that people get used to good or bad fortune after some time. An ego is 45% (4% to 122%) more likely to be happy if a friend who was examined in the past half year becomes happy. In contrast, the effect is only 35% (6% to 77%) for friends who were examined within the past year, and it declines and ceases to be significant at greater periods of time.

Happiness spreads significantly more through same sex relations than opposite sex relations (P=0.02, see appendix on bmj.com), possibly helping to explain why friends and next door neighbours exhibit stronger effects than spouses (who in our sample were all opposite sex). It should be noted, however, that the difference in effect size for friends and spouses is not significant (see appendix on bmj.com).

Finally, similarity in socioeconomic status probably cannot explain the clustering of happy people as next door neighbours have a much stronger influence than neighbours who live a few doors down in the same neighbourhood (and who consequently have similar housing, wealth, and environmental exposures). Moreover, the geographical distribution of happiness is not systematically related to local levels of either income or education (see maps in appendix on bmj.com). Both of these factors suggest that contextual effects are probably not driving our results.

DISCUSSION

While there are many determinants of happiness, whether an individual is happy also depends on whether others in the individual's social network are happy. Happy people tend to be located in the centre of their local social networks and in large clusters of other happy people. Happiness of an individual is associated with the happiness of people up to three degrees removed in the social network. Happiness, in other words, is not merely a function of individual experience or individual choice but is also a property of groups of people. Indeed, changes in individual happiness can ripple through social networks and generate large scale structure in the network, giving rise to clusters of happy and unhappy individuals. These results are even more remarkable considering that happiness requires close physical proximity to spread and that the effect decays over time.

Our results are consistent with previous work on the evolutionary basis of human emotions and with work focusing on the fleeting direct spread of emotions. In addition to their internal and psychological relevance,[10] emotions play a specifically social role: when humans experience emotions, they do not generally keep them to themselves but tend to show them. Like laughter and smiling, the emotion of happiness might serve the evolutionarily adaptive purpose of enhancing social bonds.[11]

The spread of happiness seems to reach up to three degrees of separation, just like the spread of obesity[2] and smoking behaviour,[3] suggesting a "three degrees rule" that might apply to many phenomena across many human social networks.

Our findings also have relevance for public health. To the extent that clinical or policy manoeuvres increase the happiness of one person, they might have cascade effects on others, thereby enhancing the efficacy and cost effectiveness of the intervention.[12] People are embedded in social networks and the health and wellbeing of one person affects the health and wellbeing of others. This fundamental fact of existence provides a conceptual justification for the specialty of public health. Human happiness is not merely the province of isolated individuals.

We thank Laurie Meneades, Rebecca Joyce, Molly Collins, Marian Bellwood, and Karen Mutalik for the expert assistance required to build the analytic data. We thank Chris Dawes, Dan Gilbert, Tom Keegan, Erez Lieberman, Andrew Oswald, Mark Pachucki, and Holly Shakya for helpful suggestions regarding the manuscript.

Contributors: See bmj.com.
Funding: This work was supported by the National Institute on Aging (R-01 AG24448, P-01 AG031093) and a Pioneer Award from the Robert Wood Johnson Foundation; NHLBI's Framingham Heart Study is supported by contract number N01-HC-25195. Neither author has a dependent relationship with any of the funding agencies.
Competing interests: None declared.
Ethical approval: This work was approved by the Harvard institutional review board; the parent Framingham Heart Study has separate IRB approval. All participants gave informed consent.
Provenance and peer review: Not commissioned; externally peer reviewed.

1 Hatfield E, Cacioppo JT, Rapson RL. *Emotional contagion*. New York: Cambridge University Press, 1994.
2 Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. *N Engl J Med* 2007;357:370-9.
3 Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network. *N Engl J Med* 2008;358:2249-58.
4 Szabo G, Barabasi AL. *Network effects in service usage*. http://lanl.arxiv.org/abs/physics/0611177.
5 McPherson M, Smith-Lovin L, Cook JM. Birds of a feather: homophily in social networks. *Ann Rev Sociol* 2001;27:415-44.
6 Fowler JH, Christakis NA. Estimating peer effects on health in social networks. *J Health Econ* 2008;27:1400-5.
7 Carrington PJ, Scott J, Wasserman S. *Models and methods in social network analysis*. Cambridge: Cambridge University Press, 2005.
8 Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986;73:13-22.
9 Schildcrout JS. Regression analysis of longitudinal binary data with time-dependent environmental covariates: bias and efficiency. *Biostatistics* 2005;6:633-52.
10 Gilbert DT, Wilson TD. Prospection: experiencing the future. *Science* 2007;317:1351-4.
11 Gervais M, Wilson DS. The evolution and functions of laughter and humor: a synthetic approach. *Q Rev Biol* 2005;80:395-430.
12 Christakis NA. Social networks and collateral health effects. *BMJ* 2004;329:184-5.

Accepted: 10 September 2008

Commentary: Understanding social network analysis

Peter Sainsbury

Sydney South West Area Health Service, Locked Bag 7008, Liverpool, NSW 1871, Australia
sainsburyp@email.cs.nsw.gov.au

Cite this as: *BMJ* 2008;337:a1957
doi:10.1136/bmj.a1957

In the linked study, Fowler and Christakis investigated the new and intriguing hypothesis that people's happiness is influenced by, among other things, the happiness of their acquaintances, particularly first degree relatives, close friends, neighbours, and coworkers.[1] The authors cleverly use the Framingham heart study's existing database that includes, fortuitously rather than by design, information that can be used for social network analysis. Their results broadly confirm this hypothesis, but many readers will be unfamiliar with social network analysis, confused by the analytical techniques, and unsure about the validity of the findings.

Humans are unavoidably social beings. Consequently, not only does society exist, but its existence is inevitable, and each person is influenced in many ways by society at large and individuals and groups within it. It follows that to understand the attributes of individuals (for instance their behaviour and health) the research toolkit must include methods that explore the social relationships between people. Social network analysis is one such method.

Put simply, by asking study participants to list the people they know, and which acquaintances know each other, social network analysis researchers seek to represent visually and analyse quantitatively the web of relationships around and among people. Of course, in reality, it is more complex than this. For instance, researchers may focus on the relationships around each individual or they may aggregate these to construct the more complex web of relationships within a community (for instance a business organisation or a town); or researchers may focus on everyone known to each study participant or, more commonly, on a particular group of their acquaintances (for instance their family or the people they see daily). Depending on their specific aims, researchers carefully phrase their questions to participants (the "name generators") to identify the types of acquaintances they are interested in.

Fowler and Christakis's study has several strengths. Firstly, when the information was collected it was not intended that it would be used to measure happiness, analyse social networks, or explore this hypothesis. Consequently, the original data collection was not

biased by the researchers' desire to confirm this hypothesis or by the participants' wishes to give socially desirable answers. Secondly, although social network analysis is complex and unfamiliar to many, this research method is commonly used by sociologists, community psychologists, and others. Thirdly, despite the sometimes large and overlapping confidence intervals, the results are internally consistent and robust to sensitivity analyses.

We should be cautious, however, for several reasons. Firstly, a single community and a single database that was not designed to tackle this hypothesis was studied—perhaps Framingham is unique in some way; perhaps the data collection incorporated an unknown systematic bias that produced these results. Secondly, the findings concerning friends must be viewed cautiously because the name generator used seems unlikely to have encouraged respondents with several close friends to name more than one. From a social network analysis viewpoint it would have been preferable to ask respondents to name all their close friends. This would have generated more complete networks and made it more likely that mutual friends would have been identified. Also, the size of the influence of distant friends (friends of friends' friends; 5.6%) seems overly large when the influence of a happy friend is only 14%. Thirdly, the measure of happiness is well validated as a measure of "positive affect," but it will be interesting to see if similar results are produced with different measures of happiness. Fourthly, happiness is not everything; unhappy acquaintances may contribute something other than happiness to our lives.

In summary, Fowler and Christakis have produced valuable, exciting, and reasonably robust results that will stimulate new and productive lines of enquiry in happiness studies. However, we must not expect all the details of their findings to be confirmed in subsequent work. Don't drop your unhappy friends yet.

Competing interests: None declared, but PS donated the £100 (€126; $178) fee for writing this article to Amnesty International.
Provenance and peer review: Commissioned; not externally peer reviewed.

1 Fowler JH, Christakis NA. Dynamic spread of happiness in a large social network: longitudinal analysis of the Framingham Heart Study social network. *BMJ* 2008;337:a2338.

Detecting implausible social network effects in acne, height, and headaches: longitudinal analysis

Ethan Cohen-Cole,[1] Jason M Fletcher[2]

EDITORIAL by Steptoe and Diez Roux
RESEARCH p23

[1]Federal Reserve Bank of Boston, 600 Atlantic Avenue, Boston, MA 02210
[2]Yale University, School of Public Health, 60 College Street, New Haven, CT 06510
Correspondence to: J Fletcher
jason.fletcher@yale.edu

Cite this as: *BMJ* 2008;337:a2533
doi:10.1136/bmj.a2533

ABSTRACT

Objective To investigate whether "network effects" can be detected for health outcomes that are unlikely to be subject to network phenomena.
Design Statistical analysis common in network studies, such as logistic regression analysis, controlled for own and friend's lagged health status. Analyses controlled for environmental confounders.
Setting Subsamples of the National Longitudinal Study of Adolescent Health (Add Health).
Participants 4300 to 5400 male and female adolescents who nominated a friend in the dataset and who were both longitudinally surveyed.
Measurements Health outcomes, including headache severity, acne severity, and height self reported by respondents in 1994-5, 1995-6, and 2000-1.
Results Significant network effects were observed in the acquisition of acne, headaches, and height. A friend's acne problems increased an individual's odds of acne problems (odds ratio 1.47, 95% confidence interval 0.93 to 2.33). The likelihood that an individual had headaches also increased with the presence of a friend with headaches (1.62, 0.91 to 2.89); and an individual's height increased by 20% of his/her friends' height (0.15 to 0.26). Each of these results was estimated by using standard methods found in several publications. After adjustment for environmental confounders, however, the results become uniformly smaller and insignificant.
Conclusions Researchers should be cautious in attributing correlations in health outcomes of close friends to social network effects, especially when environmental confounders are not adequately controlled for in the analysis.

INTRODUCTION

Providing credible estimates of the effects of social networks in choices and outcomes in health is important for suggesting policies that could improve health via social networks. For example, Christakis and Fowler have presented evidence of the person to person spread of obesity and quitting smoking among friends.[1 2] Raspe et al proposed that back pain might be a "communicable disease."[3]

Many methods used to estimate social network effects are subject to potentially large biases that result in the increased likelihood of detecting social network effects where none exists. For example, the use of standard econometric methods on peer effects substantially reduces evidence of social network effects in obesity.[4] Previous work that claimed to find social contagion in the diffusion of prescription drugs was confounded by marketing effects.[5]

Using standard methods we examined whether one can "find" network effects using common methods even in health outcomes that are unlikely to be transmitted socially: acne, headaches, and height.

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a2533

Table 1 | Association between own health status and friend's health status (skin problems and headache) at baseline and extended specifications. Figures are odds ratios with robust standard errors

Outcome	Unadjusted	Adjusted*
Skin problems		
Control for environmental confounding	No	Yes
Lagged own skin problem	6.48† (1.09)	4.90† (0.85)
Friend's skin problem	1.62‡ (0.48)	1.23 (0.35)
Lagged friend's skin problem	0.89 (0.29)	0.75 (0.24)
Wave	1.78 (1.06)	1.56 (1.06)
Fixed effects	None	School
No of observations	4540	3856
Headache		
Control for environmental confounding	No	Yes
Lagged own headache problem	5.20† (0.86)	3.85† (0.62)
Friend's headache problem	1.47‡ (0.34)	1.14 (0.27)
Lagged friend's headache problem	1.13 (0.26)	1.09 (0.25)
Wave	0.67 (0.36)	0.77 (0.43)
Fixed effects	None	School
No of observations	5292	4750

*Sex, age, race, maternal education, family income, grade level in wave 1, indicator for missing family information.
†P<0.001.
‡P<0.1.

METHODS

We focused on two main empirical difficulties in estimating social network effects within reference groups: firstly, that friendship selection is non-random and, secondly, that confounding factors affect all members of the reference group. The first creates correlations in health outcomes because individuals in good (or bad) health tend to associate with other individuals in good (or bad) health. This non-random pattern of association across individuals can lead to correlations in health outcomes between friends that are not caused by direct social network effects.

The second difficulty, environmental confounding, can occur when a feature of the shared environment affects all individuals in the same reference group. For example, a fast food restaurant, or gym opening near a school could simultaneously affect the weight of all individuals in networks within the school. Importantly, the presence of (often unmeasured) shared surroundings can lead to erroneously implicating social network effects in individual outcomes where none exists.

The problem of selection has been addressed through the use of fixed effects or influencing random assignments of reference groups. Other authors include lagged variables in the empirical model, though in general this practice produces biased results.[4] New research combines multiple strategies to address the multiple difficulties.[6]

The problem of confounding has generally been addressed by controlling for a rich set of individual, family, and environmental characteristics or using fixed effects at the group level.[7] The problem, of course, is that social groups are often faced with similar environmental characteristics. If these are neglected, one can improperly interpret the results to imply that true "network effects" exist.

We argue that the test statistics drawn from what we call the "standard" approach are incorrect. In particular, because research has not accounted for the problems above, the standard errors from the simple models will be unreasonably small. If standard errors are too small, a researcher will be more likely to reject, incorrectly, any given null hypothesis.

It is simple in empirical work to assume that whatever information is available in the dataset is the same information that describes the social environment in which people live. The problem is that the datasets used were rarely, if ever, constructed with this type of analysis in mind. Inclusion of the individual's race, income, etc, might be reasonable proxies for some studies but cannot distinguish two otherwise similar groups that have different environments. Estimating a regression of any type without this salient information might show a "network effect" if one school, for example, is next to a fast food restaurant and another is not.

Though there are different approaches to estimating social network effects, we focused on the approach used in Christakis and Fowler,[1 2] which addresses selection issues by controlling for lagged health outcomes. Unfortunately, unless selection is conditioned only on this variable, these methods might lead to spurious results. For example, if friendships are formed based on characteristics like self esteem, and if self esteem affects both current weight and future weight in differing ways, then adjustment for current peer weight status will not capture the self selection of friends based on self esteem that also affects future weight.[8] Finally, in the presence of social network effects, the use of lagged variables can lead to bias in estimation apart from the issues of self selection.[9]

The second issue with these methods—confounding—hinges on whether appropriate variables are included in the regression analysis. Common environmental exposures might produce the appearance of "social network effects" if not controlled for in the empirical models, particularly if only the type and direction of the friendship networks are used for adjustment.

Data source

We use the Add Health dataset to examine social network effects in three health outcomes for a national sample of adolescents. A full description of the sample design, data, and documentation is available at www.cpc.unc.edu/addhealth.

As we intended to investigate potential biases in previous methods, we looked at three health outcomes that could not credibly be subject to social network effects and were available in all three waves of the data: self reports of skin problems, self reports of headaches, and height over time. See bmj.com for details of measurements.

We had information on friends for over 5000 individuals, about 2000-3000 of whom were followed over time along with at least one same sex friend, depending on the health outcome. Nearly two thirds of the individuals in our sample were matched to only one friend's health because of the sample design. We selected

Table 2 | Association between own height and friend's height difficulties with multicollinearity and confounding*. Figures are odds ratios with robust standard errors

Variable	Col 1	Col 2	Col 3	Col 4
Lagged own height	0.89† (0.01)	—	0.88† (0.01)	0.881† (0.01)
Friend's height	0.21† (0.03)	0.18† (0.022)	0.01‡ (0.01)	0.009 (0.01)
Lagged friend's height	−0.21† (0.03)	—	—	—
Wave	1.37† (0.08)	0.11 (0.21)	1.678† (0.09)	1.580† (0.11)
Fixed effects				
Observations	4284	4284	4284	4284
R^2	0.91	0.53	0.91	0.91

*Additional controls: sex, age, race, maternal education, family income, grade level in wave 1, indicator for missing family information. Column 1 recreates previous specifications by controlling both own and friend's lagged height. Column 2 repeats column 1 except lagged own height and lagged friend's height are not controlled. Column 3 repeats column 2 with the addition of lagged own height. Column 4 repeats column 3 with the addition of controls for environmental confounding in the form of school level fixed effects.
†P<0.001.
‡P<0.1.

only one friend to be consistent with previous research. These sample sizes gave us about 4000 person year observations for each analysis. See bmj.com for summary statistics.

Though there are several important differences between the Add Health and the Framingham Heart Study used in previous research,[1 2] the two datasets are sufficiently similar to use to evaluate the role of transmission mechanisms.[4]

RESULTS

Table 1 shows baseline estimates for self reported skin problems, with unadjusted data and data adjusted for environmental confounding through the use of school fixed effects. Our baseline unadjusted results suggest that having a friend with skin problems increases the respondent's chances of skin problems (odds ratio 1.62, 95% confidence interval 0.91 to 2.89). While the unadjusted result is significant only at the 10% level, our intention is not to state that the effects are relevant but to show that the findings are fragile. The 5% standard, typically drawn from Fisher's recommended threshold, was never intended to be a strict one, but rather reasonable guidance as a threshold for rejection of the null.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Recent research has shown that individuals who are socially connected also engage in similar health behaviours and have similar health outcomes

Socially connected individuals might have similar outcomes because they share similar environments, because they purposefully select their connections, or because their connections causally influence their behaviours and outcomes

These competing hypotheses are difficult to distinguish using many current empirical models, which might lead to the detection of causal "social network effects" where none exists

WHAT THIS STUDY ADDS

Current empirical methods used to estimate causal social network effects might detect implausible network effects, including "contagion" in headaches, skin problems, and height between adolescent friends

Caution is needed in attributing causality in empirical studies of social network effects; empirical models are needed that can distinguish causal and non-causal channels of social influence

This empirical model generates such large results that we would reject a null at the 10% level even when the true contagion effect is zero. The magnitude of our result (relative risk of 1.58) is similar to the 57% increase in risk of becoming obese when a friend is obese, as reported elsewhere,[1] and larger than the 36% increase in quitting smoking when a friend quits.[2]

For self reported headache problems our unadjusted baseline results suggest that having a friend with headache problems increases the respondent's chances of headache problems. Again, this result is marginally significant but quite large in magnitude.

In the adjusted results, we show that adding simple controls for environmental confounding reduces the "social network effect" by over 50% and renders the results indistinguishable from zero. We controlled for school level fixed effects in our empirical models to control for all environmental conditions shared by students in the same school. Previous research controlled only for a limited set of individual level variables and time effects but no other shared environmental factors. Inclusion of school level fixed effects does not provide precise information on network specific confounders, but even a relatively blunt measure is effective at correcting spurious results.

Finally, we examined the "social network effects" of height between friends. We used ordinary least squares regression analysis because of the continuous nature of the outcome (table 2). Our baseline findings in column 1 suggest strong contagion in height between friends over time. This finding for height is driven by a different specification error that is particularly acute for height but might be more generally applicable to other health outcomes that do not change often over time. For the case of height, the correlation between waves is greater than 0.95. This high multicollinearity probably generates the "social network effects" in a case (height) where we would expect no true contagious effects. The association between friend's current height and the respondent's current height was 0.18, but much of this correlation can be explained with adjustment for lagged respondent height, but friend's height is still associated with own height with a P value <0.10 and a small magnitude. Finally, we controlled for school fixed effects, which reduced the magnitude of the "social network effect" of height.

DISCUSSION

The methods of detecting "social network effects" of health outcomes commonly found in the recent medical literature might produce effects where none exists. The presence of network effects in three health outcomes—headaches, skin problems, and height—disappeared after we controlled for environmental confounders. These methods might produce fragile results and consequently can produce premature claims of social network effects in health outcomes. Lack of controlling for confounding factors is not a solution in itself, but any individual study needs to fully articulate the necessary assumptions and explain how common identification issues apply to the study.

Strengths and weaknesses

We used common empirical methods and sets of control variables to show that the evidence of social network effects can largely be eliminated after adjustment for environmental confounders. Weaknesses of the study include our inability to test additional implausible health outcomes within this sample and the marginal levels of significance for two of our health outcomes. See bmj.com.

There is a need for caution when attributing causality to correlations in health outcomes between friends using non-experimental data. While it will probably not be harmful for policy makers and clinicians to attempt to use social networks to spread the benefits of health interventions and information, the current evidence is not yet strong enough to suggest clear evidence based recommendations.

We thank David Paltiel and the reviewers and editorial board for helpful comments that substantially improved the paper.
Contributors: See bmj.com.
Funding: No external funding was used to support this research.
Competing interests: None declared.
Ethical approval: The Yale HIC exempted this research.
Provenance and peer review: Not commissioned; externally peer reviewed.

1 Christakis N, Fowler J. The spread of obesity in a large social network over 32 years. *N Engl J Med* 2007;357:370-9.
2 Christakis N, Fowler J. The collective dynamics of smoking in a large social network. *N Engl J Med* 2008;358:249-58.
3 Raspe H, Hueppe A, Neuhauser H. Back pain, a communicable disease? *Int J Epidemiol* 2008;37:69-74.
4 Cohen-Cole E, Fletcher JM. Is obesity contagious? Social network vs environmental factors in the obesity epidemic. *J Health Econ* 2008;27:1382-7.
5 Van den Bulte C, Lilien GL. Medical innovation revisited: social contagion versus marketing effort. *Am J Sociol* 2001;106:1409-35.
6 Fletcher JM. *Social interactions and smoking: an IV/FE approach*. Yale University Working Paper. 2008. http://search.ssrn.com/sol3/papers.cfm?abstract_id=1069912
7 Arcidiacono P, Nicholson S. Peer effects in medical school. *J Public Econ* 2005;89:327-50.
8 Wertheim EH, Paxton SJ, Maude D, Szmukler GI, Gibbons K, Hiller L. Psychosocial predictors of weight loss behaviors and binge eating in adolescent girls and boys. *Int J Eat Disord* 1992;12:151-60.
9 Liu X, Lee LF, Kagel J. *Dynamic discrete choice models with lagged social interactions: with an application to a signaling game experiment*. Ohio State: Mimeo, 2006.

Accepted: 3 November 2008

Elbow extension test to rule out elbow fracture: multicentre, prospective validation and observational study of diagnostic accuracy in adults and children

A Appelboam,[1] A D Reuben,[1] J R Benger,[2] F Beech,[3] J Dutson,[4] S Haig,[3] I Higginson,[2] J A Klein,[5] S Le Roux,[6] S S M Saranga,[6] R Taylor,[1] J Vickery,[1] R J Powell,[7] G Lloyd[1]

EDITORIAL by Mackway-Jones

[1]Emergency Department, Royal Devon and Exeter Foundation NHS Trust, Exeter EX2 5DW
[2]Emergency Department, United Bristol Healthcare NHS Trust, Bristol BS2 8HW
[3]Emergency Department, Bath Royal United Hospital NHS Trust, Bath BA1 3NG
[4]Emergency Department, Bristol Royal Infirmary, United Bristol Healthcare NHS Trust, Bristol BS1
[5]Emergency Department, Musgrove Park Hospital, Taunton and Somerset NHS Trust, Taunton TA1 5DA
[6]Emergency Department, Bristol Children's Hospital, United Bristol Healthcare NHS Trust, Bristol BS3 8BJ
[7]Research and Development Support Unit, Royal Devon and Exeter Foundation NHS Trust, Exeter EX2 5DW
Correspondence to: G Lloyd
gavin.lloyd@rdeft.nhs.uk

Cite this as: *BMJ* 2008;337:a2428
doi:10.1136/bmj.a2428

ABSTRACT

Objective To determine whether full elbow extension as assessed by the elbow extension test can be used in routine clinical practice to rule out bony injury in patients presenting with elbow injury.
Design Adults: multicentre prospective interventional validation study in secondary care. Children: multicentre prospective observational study in secondary care.
Setting Five emergency departments in southwest England.
Participants 2127 adults and children presenting to the emergency department with acute elbow injury.
Intervention Elbow extension test during routine care by clinical staff to determine the need for radiography in adults and to guide follow-up in children.
Main outcome measures Presence of elbow fracture on radiograph, or recovery with no indication for further review at 7-10 days.
Results Of 1740 eligible participants, 602 patients were able to fully extend their elbow; 17 of these patients had a fracture. Two adult patients with olecranon fractures needed a change in treatment. In the 1138 patients without full elbow extension, 521 fractures were identified. Overall, the test had sensitivity and specificity (95% confidence interval) for detecting elbow fracture of 96.8% (95.0 to 98.2) and 48.5% (45.6 to 51.4). Full elbow extension had a negative predictive value for fracture of 98.4% (96.3 to 99.5) in adults and 95.8% (92.6 to 97.8) in children. Negative likelihood ratios were 0.03 (0.01 to 0.08) in adults and 0.11 (0.06 to 0.19) in children.
Conclusion The elbow extension test can be used in routine practice to inform clinical decision making. Patients who cannot fully extend their elbow after injury should be referred for radiography, as they have a nearly 50% chance of fracture. For those able to fully extend their elbow, radiography can be deferred if the practitioner is confident that olecranon fracture is not present. Patients who do not undergo radiography should return if symptoms have not resolved within 7-10 days.

INTRODUCTION

Only a minority of patients with elbow injury have a fracture, and although clinical decision rules for other limb injuries are well recognised,[1 2] no guidelines have been established to indicate which patients with an elbow injury require radiography.

The elbow extension test has been proposed as a simple means of excluding the need for a radiograph, but has yet to be validated in routine practice and has not been well studied in children.[3-5]

Our objective was to determine whether the elbow extension test could be used in routine clinical practice to rule out bony injury in patients presenting with acute elbow injury.

Elbow extension test characteristics (95% confidence intervals shown in parentheses)

	Adults		Children		Combined	
	Fracture	Fracture or effusion	Fracture	Fracture or effusion	Fracture	Fracture or effusion
Sensitivity	98.4 (96.3 to 99.5)	97.3 (95.2 to 98.6)	94.6 (90.7 to 97.2)	93.7 (90.3 to 96.2)	96.8 (95.0 to 98.2)	95.8 (94.0 to 97.2)
Specificity	47.7 (43.7 to 51.6)	54.3 (50.1 to 58.6)	49.5 (45.2 to 53.7)	54.8 (50.3 to 59.2)	48.5 (45.6 to 51.4)	54.6 (51.5 to 57.6)
Negative predictive value	98.4 (96.3 to 99.5)	96.5 (93.8 to 98.2)	95.8 (92.6 to 97.8)	93.7 (90.1 to 96.2)	97.2 (95.5 to 98.3)	95.2 (93.1 to 96.7)
Positive predictive value	48.1 (44.2 to 52.0)	61.0 (57.2 to 64.8)	42.8 (38.4 to 47.3)	54.8 (50.3 to 59.2)	45.8 (42.9 to 48.7)	58.3 (55.4 to 61.2)
Positive likelihood ratio	1.88 (1.75 to 2.03)	2.13 (1.95 to 2.34)	1.87 (1.72 to 2.05)	2.07 (1.88 to 2.30)	1.88 (1.78 to 1.99)	2.11 (1.97 to 2.26)
Negative likelihood ratio	0.03 (0.01 to 0.08)	0.05 (0.03 to 0.09)	0.11 (0.06 to 0.19)	0.11 (0.07 to 0.18)	0.06 (0.04 to 0.10)	0.08 (0.05 to 0.11)

METHODS

Design and setting

We did a multicentre, prospective validation study in adults and an observational study in children who presented with acute elbow injury to five emergency departments. As the diagnostic accuracy of the test had not been assessed in children, we did not think an interventional study was justified in this group.

Participants

Adults (>15 years old) and children (3-15 years) presenting to the participating centres within 72 hours of elbow injury were consecutively recruited to the trials. For inclusion and exclusion criteria see bmj.com.

We judged that for the elbow extension test to be clinically acceptable as a single test for universal use to rule out elbow fracture, sensitivity needed to be greater than 99%. With the 3/n rule for zero numerators, 300 adults and 300 children with full elbow extension and no significant fracture would yield a test sensitivity of 100% for each group, with 95% confidence intervals between 99% and 100%.

Interventions

All patients with elbow injury were identified on arrival during normal registration and triage, and were given analgesia in accord with standard protocols. A doctor or emergency nurse practitioner then screened and recruited each patient during routine care. Recruitment rate was monitored and was constant between the centres.

The treating practitioner performed the standardised elbow extension test (box) as part of the examination. Adult patients with full extension did not undergo radiography and were discharged with analgesia and a sling as needed. Children underwent radiography at the discretion of the treating practitioner, regardless of the result of the elbow extension test. All patients who did not undergo radiography received a structured follow-up assessment by telephone at 7-10 days. Patients who met any of the recall criteria (inability to fully straighten their elbow, pain worsening or not improving, any functional problems, or any other concern) were recalled to the emergency department for radiography. Those not requiring recall were assumed not to have a clinically significant bony injury.

The reference standard was the final discharge diagnosis for patients followed up in an orthopaedic clinic, the formal report of a radiologist blinded to the result of the extension test for those not followed up in an orthopaedic clinic, and the result of the structured telephone interview at 7-10 days for those who did not undergo follow-up in an orthopaedic clinic or radiography.

We calculated test characteristics (sensitivity, specificity, predictive values and likelihood ratios) with 95% confidence intervals, and compared proportions by χ^2 test to obtain P values. See bmj.com.

RESULTS

We screened 2127 patients for eligibility over 21 months (July 2004-April 2006). Of these, 960 adults and 780 children were recruited to the study and underwent the elbow extension test. The age range of the adults was 16-94 (mean 38) years; 51% were male. Among the children, the age range was 3-15 (mean 10) years and 52% were male. The overall prevalence of fracture was 31% (538/1740). For recruitment and table showing results of the test see bmj.com.

Adults

Of the 958 adults included in the analysis, 313 (33%) were able to fully extend their elbow, and of these all but two were followed up. Five fractures were identified in those patients with full elbow extension, and of these two required operative intervention (both olecranon fractures).

Seven hundred and five adults (73%) underwent radiography at their first visit. Fifty eight protocol violations occurred, mostly by temporary staff (52 patients), but also in patients who underwent radiography for a potential foreign body (three) or at the request of their general practitioner (three).

Of the 647 adults who could not fully extend their injured elbow, 311 (48%) had confirmed fractures and 84 had elbow joint effusions.

Children

Of the 778 children included in the analysis, 289 (37%) could fully extend their elbow, and of these patients all but two were followed up. We found 12 fractures (all identified at first visit) and six effusions in those with full elbow extension none of which required operative intervention.

Of the 491 children who could not fully extend their injured elbow, 210 (43%) had confirmed fractures and 59 had elbow joint effusions.

Test characteristics

A reference standard was determined in 1736 of the 1740 patients. Overall, test sensitivity and specificity

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a2428

The elbow extension test

The seated patient, with exposed and supinated arms, is asked to flex their shoulders to 90 degrees and then fully extend and lock both elbows. Injured and uninjured sides are compared visually and those with equal extension recorded as "full extension."

for detecting elbow fracture are shown in the table. A "worst case" sensitivity analysis, assuming that fractures were present in the four patients who were lost to follow-up and in all patients with effusions, gave an overall sensitivity of 95.3% for the detection of fracture.

The negative predictive values and negative likelihood ratios are also shown in the table.

In practice, adults who could fully extend their elbow after acute injury had a 1.6% (95% confidence interval 0.5 to 3.7) chance of fracture. In children the risk was 4.2% (2.2 to 7.4), despite the greater prevalence of fracture in adults (316/958, 33%) than in children (222/778, 29%: χ^2=3.98, P=0.046, df=1). The proportion of patients with a fracture who were not able to fully extend their elbow (sensitivity) was significantly greater in adults (311/316, 98.4%) than in children (210/222, 94.6%: χ^2=6.23, P=0.013, df=1). The specificity of the test did not differ between adults (306/642, 47.7%) and children (275/556, 49.5%: χ^2=0.39, P=0.53, df=1).

DISCUSSION

In this study we found that the elbow extension test, used in routine clinical practice, has a high sensitivity and negative predictive value for elbow fracture. The test was able to rule out a fracture and the need for radiography in about a quarter of patients presenting with acute elbow injury. This finding is useful, as over a third of patients with elbow injury[3-5] are able to fully extend their elbow at presentation. Patients who could not fully extend their elbow had a nearly 50% chance of radiologically confirmed fracture.

The low negative likelihood ratio confirms that this is a powerful test to rule out fracture in adults,[6] but the test does not exceed the sensitivity of 99% that we had previously judged as being clinically desirable. Ninety nine per cent sensitivity is a challenging standard, and our test has similar properties, in terms of sensitivity and specificity, to established clinical decision rules for other joints.[7] Ultimately, application of this test will rely on physicians' judgment, informed by the risk and consequences of false negatives, and by the availability of a gold standard diagnostic test (radiography) and follow-up. Most false negative results are likely to be minor or occult fractures that require no change in treatment.[8] However, we advise caution in the use of the elbow extension test as a single clinical decision rule for universal use, in view of the two olecranon fractures in adults, and the risk of occult supracondylar fractures in children.[9] The false negative rate is also higher in children than adults.

STRENGTHS AND LIMITATIONS

This study was carried out by usual practitioners in emergency departments during routine assessment of patients, reflecting the probable application of this test in real practise. The sample size was sufficient to meet our objectives, with suitably narrow confidence intervals. A high follow-up rate was essential to the study design, and ensured that a sensitivity analysis made no significant difference to the results.

It is possible that, our follow-up protocol might not have identified all patients with a fracture undetected by the test, and the recall criteria used are not validated. However, clinically significant injuries are unlikely to have been missed using this low threshold for patient recall, and a review of the database found no evidence of subsequent reattendance in patients who were discharged.

We did not assess interobserver agreement, and there was no mechanism to record or analyse equivocal results. While this may have contributed to the worse performance of the test in children than in adults, an under appreciation of the normal hyperextension in some children's elbows, or inadequate comparisons to the uninjured limb, are other possible explanations.

COMPARISON WITH PREVIOUS STUDIES

The incidences of full elbow extension and fracture in our study were similar to those reported in previous smaller studies.[3 4 10] The sensitivity of the test was also consistent with these studies, but with much narrower confidence intervals.

CONCLUSIONS

Patients with recent elbow injury who cannot fully extend their elbow should be referred for radiography. Those who are able to fully extend do not need radiography, provided the practitioner is confident that olecranon fracture is not present, that caution is used in children, and that the patient can return for reassessment if their symptoms have not resolved in 7-10 days.

We thank Beth Newstead, Charlotte Pagram, Julie Small and the reception and clinical staff of participating hospitals for their assistance and support of this study.

Contributors: See bmj.com

Funding: Research grant from the College of Emergency Medicine. The research was independent of the funders.

Competing interests: None declared.

Ethical approval: Each site obtained approval from local ethics committee. All eligible patients were recruited after written informed consent had been obtained.

Provenance and peer review: Not commissioned; externally peer reviewed.

WHAT IS ALREADY KNOWN ON THIS TOPIC

No clinical decision rule exists for deciding which patients with acute elbow injury require radiography

The elbow extension test has been proposed as a simple test to rule out the need for radiography, but it has not been validated in routine practice

WHAT THIS STUDY ADDS

The elbow extension test can be used in routine practice.

The test effectively rules out the need for radiography in patients with a recent elbow injury and full joint extension; caution should be used in children and in patients with suspected olecranon fracture

1 Stiell IG, Greenberg GH, McKnight RD, Nair RC, McDowell I, Reardon M, et al. Decision rules for the use of radiography in acute ankle injuries. Refinement and prospective validation. *JAMA* 1993;269:1127-32.
2 Stiell IG, Wells GA, McDowell I, Greenberg GH, McKnight RD, Cwinn AA, et al. Use of radiography in acute knee injuries: need for clinical decision rules. *Acad Emerg Med* 1995;2:966-73.
3 Docherty MA, Schwab RA, Ma OJ. Can elbow extension be used as a test of clinically significant injury? *South Med J* 2002;95:539-41.
4 Hawksworth CR, Freeland P. Inability to fully extend the injured elbow: an indicator of significant injury. *Arch Emerg Med* 1991;8:253-6.
5 Dildar S. Inability to fully extend and supinate the injured elbow: an indicator of significant injury. *Todays Emergency* 2007;12:6-7.

6 Worster A, Innes G, Abu-Laban RB. Diagnostic testing: an emergency medicine perspective. *CJEM* 2002;4:348-54.
7 Bachmann LM, Kolb E, Koller MT, Steurer J, Riet G. Accuracy of Ottawa ankle rules to exclude fractures of the ankle and mid-foot: systematic review. *BMJ* 2003;326:417-23.
8 Gorzack A, Mackway-Jones K. Repeat radiography is not needed for traumatic elbow effusions with no fracture on initial x-ray. Best Bets Topics, 1999 (modified 2003). www.bestbets.org.uk.
9 Griffith JF, Roebuck DJ, Cheng JC, Chan YL, Rainer TH, Ng BK, et al. Acute elbow trauma in children: spectrum of injury revealed by MR imaging not apparent on radiographs. *Am J Roentgenol* 2001;176:53-60.
10 Lennon RI, Riyat MS, Hilliam R, Anathkrishnan G, Alderson G. Can a normal range of elbow movement predict a normal elbow x ray? *Emerg Med J* 2007;24:86-8.

Accepted: 15 September 2008

Risk of Parkinson's disease after hospital contact for head injury: population based case-control study

Kathrine Rugbjerg,[1] Beate Ritz,[2] Lise Korbo,[3] Nick Martinussen,[1] Jørgen H Olsen[1]

[1]Institute of Cancer Epidemiology, Danish Cancer Society, 49 Strandboulevarden, DK-2100 Copenhagen, Denmark
[2]Department of Epidemiology and Environmental Health Sciences, School of Public Health, University of California at Los Angeles (UCLA), USA
[3]Department of Neurology, Bispebjerg Hospital, Copenhagen
Correspondence to: K Rugbjerg rugbjerg@cancer.dk

Cite this as: *BMJ* 2008;337:a2494
doi:10.1136/bmj.a2494

ABSTRACT

Objective To determine whether a hospital contact for a head injury increases the risk of subsequently developing Parkinson's disease.

Design Population based case-control study.

Setting Denmark.

Participants 13 695 patients with a primary diagnosis of Parkinson's disease in the Danish national hospital register during 1986-2006, each matched on age and sex to five population controls selected at random from inhabitants in Denmark alive at the date of the patient's diagnosis (n=68 445).

Main outcome measures Hospital contacts for head injuries ascertained from hospital register; frequency of history of head injury.

Results An overall 50% increase in prevalence of hospital contacts for head injury was seen before the first registration of Parkinson's disease in this population (odds ratio 1.5, 95% confidence interval 1.4 to 1.7). The observed association was, however, due almost entirely to injuries that occurred during the three months before the first record of Parkinson's disease (odds ratio 8.0, 5.6 to 11.6), and no association was found between the two events when they occurred 10 or more years apart (1.1, 0.9 to 1.3).

Conclusions The steeply increased frequency of hospital contacts for a head injury during the months preceding the date at which Parkinson's disease was first recorded is a consequence of the evolving movement disorder rather than its cause.

INTRODUCTION

Parkinson's disease is a movement disorder characterised mainly by rigidity, bradykinesia, postural instability, and tremor.[1 2] The symptoms are related to a relative deficiency of the neurotransmitter dopamine, causing imbalances in the related neural circuitry following the accelerated death of dopaminergic neurones in the substantia nigra of the brain.[3] Apart from a few patients with genetically caused parkinsonism, the reason for the degeneration is unknown, although several non-genetic risk factors have been examined. One such risk factor is previous injury to the head, a hypothesis first put forward by James Parkinson in 1817 (see bmj.com).[4] Here, we report the results of the largest population based case-control study of Parkinson's disease subsequent to hospital contact for a head injury.

METHODS

Danish hospital register

The Danish national hospital register was instituted on 1 January 1977 and contains individual information on all admissions for medical conditions other than psychiatric diseases to hospitals in Denmark.[5] Any contact of a Danish resident with the hospital system generates a record in the hospital register, which includes the personal identification number of the patient, permitting accurate linkage between registers.

Study populations

We identified 13 739 patients with a first time diagnosis of Parkinson's disease in the files of the hospital register during the period 1986-2006. After exclusion of patients who were citizens of Greenland (n=1) and patients who were younger than 35 years at the time of first hospital admission for Parkinson's disease (n=43), we were left with a case group of 13 695 patients: 7423 men and 6272 women. To validate the diagnosis of Parkinson's disease listed in the hospital register, we used a continuously updated national prescription database started on 1 January 1995, which covers all prescribed drugs dispensed at any pharmacy in the country.[6]

For each patient, we chose five control subjects at random from the Danish central population register from among all inhabitants of the same sex and year of birth who were alive at the index date of their respective case (incidence density sampling). Although we aimed to recruit five control subjects for each case, 26 cases were matched with either two, three, or four controls, yielding a total of 68 445 controls.

Register information on head injuries

We re-linked cases and linked controls to the files of the Danish hospital register to ascertain hospital contacts for head injury that occurred before the index dates and after 1 January 1977. As a measure of severity, we ranked head injuries in the following order: concussion < fractured skull < intracranial haemorrhage/cerebral contusion. We included both primary and supplementary diagnoses of head injuries. However, using inpatient and outpatient

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a2494

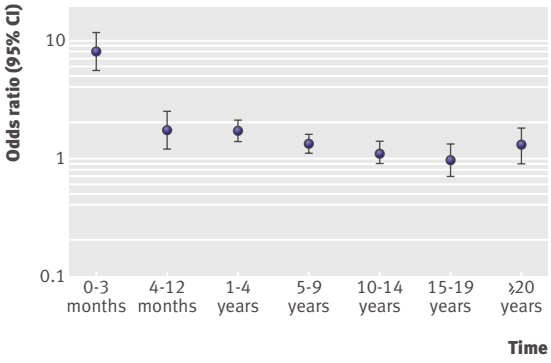
information to verify cases of head injury also implies that the study is unable to evaluate the consequences of mild and perhaps repeated trauma to the head as these do not usually lead to a hospital contact.

Analyses

We compared the frequency of a history of a previous head injury in Parkinson's disease patients with that of their population controls. We expressed the association as an odds ratio derived from a conditional logistic regression analysis for matched sets. We estimated risks for various intervals between the hospital contact for a head injury and the first registration of a diagnosis of Parkinson's disease (0-3 months, 4-12 months, 1-4 years, 5-9 years, 10-14 years, 15-19 years, ≥20 years) (see bmj.com).

RESULTS

The average age of patients at their first hospital contact for Parkinson's disease was 73.0 years (72.5 for men and 73.6 for women). Of the patients, 566 (4.1%) were reported as having had at least one hospital contact for a head injury before the index date. For the 68 445 population controls (37 101 men and 31 344 women), the corresponding figure was 1904 (2.8%). Thus, a head injury of any type was significantly more prevalent among people in whom Parkinson's disease was subsequently diagnosed than among population controls (odds ratio 1.5, 95% confidence interval 1.4 to 1.7).



Risk of Parkinson's disease after hospital admission for head injury, by time between head injury and first hospital contact for Parkinson's disease

As indicated in the figure, the increased overall risk of Parkinson's disease seemed to be due almost entirely to head injuries that occurred during the three months before a first hospital contact for the disease (odds ratio 8.0, 5.6 to 11.6). Head injuries that had occurred within four months to nine years before were associated with only a modestly increased risk of Parkinson's disease (1.5, 1.3 to 1.7), and those occurring even more distantly in time showed no association (≥10 years: 1.1, 0.9 to 1.3). Similar analyses done separately for each of the three subentities of head injuries showed a risk pattern comparable to the one for all types of head injuries combined (table). The estimated risk for concussion, representing the mildest forms of head injury, showed some increase in all time windows from zero to nine years before diagnosis of Parkinson's disease, but we found a clear negative trend with increasing time between the two events ($P<0.0001$) (table).

For the subset of people for whom the head injury was the main diagnosis and the primary reason for the hospital contact, we combined type of injury and length of stay in hospital as a proxy measure for severity. In this analysis, we excluded head injuries registered within one year of a first record of Parkinson's disease, as we considered that such injuries might have been a result of the evolving movement disorder rather than a risk factor. We found no indication of an increased risk of Parkinson's disease after an accident that resulted in a fractured skull or intracranial haemorrhage/cerebral contusion, irrespective of length of stay in hospital for the injury. (See bmj.com for all results.)

Risk of Parkinson's disease after hospital contact for head injury by type of head injury and time between medical events

Type of head injury	No of cases/controls (n=13 695/68 445)	Odds ratio (95% CI)
Concussion		
Latency:		
0-3 months	54/41	6.6 (4.4 to 9.9)
4-12 months	38/99	1.9 (1.3 to 2.8)
1-4 years	106/298	1.8 (1.4 to 2.2)
5-9 years	133/480	1.4 (1.1 to 1.7)
10-14 years	76/320	1.2 (0.98 to 1.5)
15-19 years	53/254	1.0 (0.8 to 1.4)
≥20 years	41/161	1.3 (0.9 to 1.8)
Total	501/1653	1.5 (1.4 to 1.7)
Fractured skull		
Latency:		
0-3 months	6/3	10.0 (2.5 to 40)
4-12 months	2/3	3.4 (0.6 to 20)
1-4 years	6/27	1.1 (0.5 to 2.7)
5-9 years	11/53	1.0 (0.5 to 2.0)
10-14 years	10/47	1.1 (0.5 to 2.1)
15-19 years	4/22	0.9 (0.3 to 2.6)
≥20 years	4/15	1.3 (0.4 to 4.0)
Total	43/170	1.3 (0.9 to 1.8)
Traumatic intracranial haemorrhage/cerebral contusion		
Latency:		
0-3 months	19/7	13.6 (5.7 to 32)
4-12 months	6/27	1.1 (0.4 to 2.4)
1-4 years	12/64	0.9 (0.5 to 1.7)
5-9 years	14/86	0.8 (0.5 to 1.4)
10-14 years	10/45	1.1 (0.6 to 2.2)
15-19 years	3/28	0.5 (0.2 to 1.8)
≥20 years	4/10	2.0 (0.6 to 6.4)
Total	68/267	1.3 (0.98 to 1.7)

DISCUSSION

In this population based case-control study of more than 13 000 patients with a primary diagnosis of Parkinson's disease, a previous, medically confirmed diagnosis of a head injury treated at a hospital or clinic increased the risk of Parkinson's disease by 50% compared with age and sex matched population controls. This association was,

WHAT IS ALREADY KNOWN ON THIS TOPIC

Parkinson's disease is characterised by an insidious onset, with imbalance as an early sign

Parkinson's disease has been associated in several case-control studies with a self reported head injury event that occurred before the diagnosis of Parkinson's disease

WHAT THIS STUDY ADDS

A positive association was found between Parkinson's disease and a head injury before the diagnosis

However, the overall associations were due almost exclusively to injuries occurring months to a few years before hospital contact for Parkinson's disease

The statistical association between the two medical events reported in the literature probably has no causal basis and might be due to differential recall bias or inaccurate timing of events

however, due entirely to head injuries that had occurred less than 10 years before the first hospital contact for Parkinson's disease, particularly injuries occurring during the three months preceding the first hospital contact for Parkinson's disease.

If our finding of an overall association between a head injury and Parkinson's disease represented a true causal relation, we would have to hypothesise that the biological mechanism linking the injury to the disease must be rapid —that is, immediately inducing extensive cell death, preferentially in the substantia nigra of the brain. An often proposed mechanism for the purported link is that head injuries damage the blood-brain barrier, exposing the brain to inflammatory factors, toxins, and antigens.[7 8] According to McGeer and colleagues, this process, if uncontrolled, can result in chronic inflammation and activated microglia, leading to Parkinson's disease over a decade or two.[9] This hypothesised pathogenic process is clearly in conflict with the findings of our study.

We find it more reasonable to conclude that our findings of a positive association between head injuries and a hospital contact for Parkinson's disease can be explained by reverse causality. Recent findings suggesting that poor balance is an early sign of Parkinson's disease support this interpretation.[10]

Comparison with earlier studies

Our results contrast with the findings of several of the earlier interview based case-control studies,[7 8 11-15] as well as a previous medical record review study in the Mayo Clinic system.[16] These interview based case-control studies may, however, be influenced by recall bias. As patients with Parkinson's disease have an increasing number of accidental falls as the disease progresses, they might tend to recall and report head injuries more frequently when asked.

Limitations

The limitations of our study are lack of information on diagnostic details for patients with Parkinson's disease and lack of information on the date of first symptoms of Parkinson's disease or the date of first treatment with anti-parkinsonian drugs for the entire study group. Our linkage of a subgroup of patients to the files of a national prescription database revealed that 9% of cases never received treatment with anti-parkinsonian drugs, indicating some diagnostic misclassification. If this disease misclassification is non-differential with respect to exposure, it would tend to dilute—but not remove—a truly positive or negative association between a head injury and Parkinson's disease, leading to an underestimate of the risk.

Milder single and repeated head injuries were not included, as we assessed only those that resulted in at least an emergency room visit, if not a hospital visit. This limitation applies, however, to the same extent to our controls and thus should not affect our risk estimates. If Parkinson's disease is selectively caused by mild or repeated injuries to the head, this linkage study would not be able to detect such an association.

Contributors: See bmj.com.
Funding: This study was supported by grants from the National Institutes of Health, USA (grant No R01 ES013717) and the UCLA Udall Parkinson Disease Center of Excellence (grant No P50 NS038367). The funding source had no role in the design or analysis of the study or in the decision to submit the manuscript for publication.
Competing interests: None declared.
Ethical approval: The study protocol was approved by the Danish Data Protection Agency (No 2002-41-2112).
Provenance and peer review: Not commissioned; externally peer reviewed.

1 Gelb DJ, Oliver E, Gilman S. Diagnostic criteria for Parkinson's disease. *Arch Neurol* 1999;56:33-9.
2 Hughes AJ, Ben-Shlomo Y, Daniel SE, Lees AJ. What features improve the accuracy of clinical diagnosis in Parkinson's disease: a clinicopathologic study. *Neurology* 1992;42:1142-6.
3 Braak H, Bohl JR, Muller CM, Rub U, de Vos RA, Del TK. Stanley Fahn Lecture 2005: the staging procedure for the inclusion body pathology associated with sporadic Parkinson's disease reconsidered. *Mov Disord* 2006;21:2042-51.
4 Parkinson J. *Essay on the shaking palsy*. London: Whittingham and Rowland, for Sherwood, Neely and Jones, 1817.
5 Andersen TF, Madsen M, Jorgensen J, Mellemkjoer L, Olsen JH. The Danish national hospital register: a valuable source of data for modern health sciences. *Dan Med Bull* 1999;46:263-8.
6 Gaist D, Sorensen HT, Hallas J. The Danish prescription registries. *Dan Med Bull* 1997;44:445-8.
7 Factor SA, Weiner WJ. Prior history of head trauma in Parkinson's disease. *Mov Disord* 1991;6:225-9.
8 Stern MB. Head trauma as a risk factor for Parkinson's disease. *Mov Disord* 1991;6:95-7.
9 McGeer PL, Itagaki S, Akiyama H, McGeer EG. Rate of cell death in parkinsonism indicates active neuropathological process. *Ann Neurol* 1988;24:574-6.
10 Gao X, Chen H, Schwarzschild MA, Logroscino G, Ascherio A. Perceived imbalance and risk of Parkinson's disease. *Mov Disord* 2008;23:613-6.
11 Dick FD, De PG, Ahmadi A, Scott NW, Prescott GJ, Bennett J, et al. Environmental risk factors for Parkinson's disease and parkinsonism: the Geoparkinson study. *Occup Environ Med* 2007;64:666-72.
12 Goldman SM, Tanner CM, Oakes D, Bhudhikanok GS, Gupta A, Langston JW. Head injury and Parkinson's disease risk in twins. *Ann Neurol* 2006;60:65-72.
13 Seidler A, Hellenbrand W, Robra BP, Vieregge P, Nischan P, Joerg J, et al. Possible environmental, occupational, and other etiologic factors for Parkinson's disease: a case-control study in Germany. *Neurology* 1996;46:1275-84.
14 Semchuk KM, Love EJ, Lee RG. Parkinson's disease: a test of the multifactorial etiologic hypothesis. *Neurology* 1993;43:1173-80.
15 Taylor CA, Saint-Hilaire MH, Cupples LA, Thomas CA, Burchard AE, Feldman RG, et al. Environmental, medical, and family history risk factors for Parkinson's disease: a New England-based case control study. *Am J Med Genet* 1999;88:742-9.
16 Bower JH, Maraganore DM, Peterson BJ, McDonnell SK, Ahlskog JE, Rocca WA. Head trauma preceding PD: a case-control study. *Neurology* 2003;60:1610-5.

Accepted: 19 September 2008

BMJ VIDEO

Find out more about the study and the Alexander technique at www.bmj.com/cgi/content/full/337/dec11_2/a2656

EDITORIAL
by Godlee and Groves

New format for *BMJ* research articles

It's nearly 10 years since we began abridging original research articles for readers of the print *BMJ*. Now we're going a step further, using the advantages of both web and print. The full, open access version of this original article is published online (doi:10.1136/bmj.a2656), along with our first *BMJ* research video and a podcast. Here are two abridged versions: an abstract prepared by the authors and a Short Cuts article written by the *BMJ*. Which version would you read and use? Which would you prefer if you were the author? Please tell us your views, as readers and researchers, by posting rapid responses to this article online (doi:10.1136/bmj.a2946).

Randomised controlled trial of Alexander technique for chronic and recurrent back pain: economic evaluation

Sandra Hollinghurst,[1] Debbie Sharp,[1] Kathleen Ballard,[3] Jane Barnett,[2] Angela Beattie,[1] Maggie Evans,[1] George Lewith,[2] Karen Middleton,[2] Frances Oxford,[3] Fran Webley,[2] Paul Little[2]

STUDY QUESTION **What is the difference in cost effectiveness between long and short courses in the Alexander technique, massage, and a general practitioner's prescription for exercise for patients with persistent back pain?**

ANSWER **An exercise prescription and six lessons in Alexander technique alone were each more than 85% likely to be cost effective at values above £20 000 per quality adjusted life year (QALY), but the Alexander technique performed better than exercise on the full range of outcomes. A combination of six lessons in Alexander technique and exercise was the most effective and cost effective option.**

Study design This cost effectiveness analysis compared the cost to the NHS of different interventions with patients' outcomes (Roland-Morris disability score, days free of pain, and quality adjusted life year (QALY)). Patients' and societal costs were analysed separately.

Source of effectiveness data Interventions were applied in a randomised controlled trial using a 4×2 factorial design (*BMJ* 2008;337:a884). A short course of six lessons in the Alexander technique, a longer course of 24 lessons, and six sessions of massage were compared with normal care. Half of each group also received a doctor's prescription for exercise and behavioural counselling from a practice nurse.

Data NHS costs comprised the cost of the intervention, primary care contacts, outpatient appointments, inpatient hospital stays, and medication. Personal costs were those for travel associated with back pain treatment, private treatment and over the counter drugs, prescription charges, loss of earnings, and expenditure on domestic help and care giving. Societal costs were for time taken off work or unpaid activities and use of informal care. Treatment outcomes were derived from routine records and participants' self completed questionnaires.

Main results Incremental cost to the NHS ranged from £100 (for normal care plus exercise) to £607 (for 24 lessons in the Alexander technique plus exercise) over 12 months. Benefits were additional pain-free days (8-20 per patient, by intervention group), improvements in the ability to perform daily activities (reduction in the disability score of 0.45-4.22 per patient, by group), and a gain in QALY of up to 0.065 per patient, by group. The best value single treatment was normal care plus exercise, at £61 per point reduction in the disability score, £9 per extra pain-free day, and £2847 per QALY gain. The best value dual treatment comprised six lessons of Alexander technique plus exercise, with an additional £64 per point reduction in disability score, £43 per pain-free day, and £5332 per QALY gain.

Limitations of study The missing quality of life data add to uncertainty around the QALY estimates and the data on lost productivity were incomplete. The factorial trial design complicated the economic evaluation, and individual group analysis was therefore used to present the main findings. The payments to teachers and therapists were relatively high and may not be generalisable, but this means that the results are conservative.

Competing interests None declared. All researchers are independent of the study funders, the Medical Research Council.

[1]Academic Unit of Primary Health Care, University of Bristol, Bristol BS8 2AA
[2]Primary Care Group, CCS Division, University of Southampton, Aldermoor Health Centre, Southampton SO16 5ST
[3]Society of Teachers of the Alexander Technique, Linton House, London NW5 1RS
Correspondence to: S Hollinghurst **s.p.hollinghurst@bristol.ac.uk**
Cite this as: *BMJ* 2008;337:a2656

Alexander technique and exercise are cost effective options for chronic back pain

Hollinghurst et al. Randomised controlled trial of Alexander technique lessons, exercise, and massage (ATEAM) for chronic and recurrent back pain: economic evaluation. *BMJ* 2008;337:a2656, doi:10.1136/bmj.a2656

Chronic low back pain is common, disabling, and expensive for both patients and third party payers. A prescription for home based exercise can help reduce symptoms and improve quality of life. So can Alexander technique lessons. But which treatment or combination of treatments represents the best value for money?

Using effectiveness data from a recent randomised controlled trial (*BMJ* 2008;337:a884), health economists from the UK have calculated that a simple exercise prescription backed up with brief counselling from a primary care nurse costs just £9 for each extra pain-free day per month, or an estimated £2847 for each additional quality adjusted life year. Six lessons in the Alexander technique also looked cost effective, at £13 for each extra pain-free day and £5899 for each extra quality adjusted life year.

More lessons (24 *v* 6), or the combination of lessons and exercise worked better but cost more in this analysis. Even so, the authors say that first choice for primary care patients should probably be six lessons in Alexander technique followed by an exercise prescription. This combination worked better than exercise alone, worked almost as well as 24 lessons of Alexander technique, and cost the NHS only £43 for each extra pain-free day or £5332 for each extra quality adjusted life year.

Alison Tonks associate editor, *BMJ* **atonks@bmj.com**
Cite this as: *BMJ* 2008;337:a2762

PROBABILITY THAT DIFFERENT INTERVENTIONS FOR BACK PAIN ARE COST EFFECTIVE