The Collective Dynamics of Smoking in a Large Social Network

Nicholas A. Christakis, M.D., Ph.D., M.P.H.
Department of Sociology
Harvard University

James H. Fowler, Ph.D.
Department of Political Science
University of California, San Diego

Presented by

Xiao Yuan
Contents

◆ Introduction

◆ Methodology
  — Source Data
  — Statistical Analysis

◆ Results
  — Network Analysis
  — Interpersonal Analysis

◆ Conclusion
Smoking Situations in U.S.

About 44.5 MILLION adults were smokers in USA in 2004

Prevalence of smoking has declined from 45% to 21% over the past four decades.
The Cost of Smoking

Accounts for an estimated 443,000 deaths, or nearly one of every five deaths, each year in the United States.

More deaths are caused by tobacco use than by all deaths from HIV, illegal drug use, alcohol use, motor vehicle injuries, suicides, and murders combined.
The Cost of Smoking

Causes an estimated 90% of all lung cancer deaths in men and 80% of all lung cancer deaths in women.

An estimated 90% of all deaths from chronic obstructive lung disease are caused by smoking.
About 443,000 U.S. Deaths Attributable Each Year to Cigarette Smoking*

- Lung Cancer: 128,900
- Ischemic Heart Disease: 126,000
- Chronic Obstructive Pulmonary Disease: 92,900
- Other Diagnoses: 44,000
- Stroke: 15,900
- Other Cancers: 35,300

*Average annual number of deaths, 2000–2004.
Source: MMWR 2008;57(45):1226–1228.
The Framingham Heart Study

12,067 people were evaluated in terms of smoking behavior and social-network ties over 32 years

- 1948 Original Cohort initiation 5209
- 1971 Offspring Cohort children and spouses 5124
- 1994 OMNI Cohort oversample 508
- 2002 Third-Generation Cohort children and spouses 4095


A cardiovascular study on residents of Massachusetts
Motivation

The extent to which smoking depends on how people are embedded in a social network and to which smoking behavior transcends direct dyadic ties are not known.

Whether smoking cessation also spreads more than one person to another.
Methods — Source Data

Used offspring cohort as the source of 5124 subjects → “egos”

Persons to whom subjects are linked as social contacts → “alters”

A total of 12,067 subjects and contacts were connected at some point

Connections were identified longitudinally

53,228 observed families and social ties, yielding 10.4 average ties per subject within the social network
Methods — *Source Data*

There are 3,542 unique friendships:
- subject-perceived friendship
- contact-perceived friendship
- mutual friendship

Coworkers and neighbors

Mean educational level was 1.6 years of college
Methods — *Statistical Analysis*

Dichotomous cut point: no cigarettes *vs.* ≥1 cigarettes/day

**Kamada-Kawai Algorithm**

A force-based (or force-directed) algorithm for drawing general connected, undirected graphs with good aesthetic layout.

Make nodes and the ties connecting them overlap as little as possible, thus producing interpretable images.
Kamada-Kawai Algorithm
Methods — *Statistical Analysis*

**Centrality**

Assesses whether subjects are near the center of the network.

a. The number of subjects’ ties
   — Subjects with more ties tend to be more central.

b. Eigenvector Centrality
   — A measure of the influence of a node in a network.
   — Assigns relatives scores to all nodes in the network based on the concept that connections to high-scoring nodes contribute more to the score of current node than equal connections to low-scoring nodes.
Methods — *Statistical Analysis*

**Clustering**

Compare whole observed network at each examination to simulated networks with same network topology and same prevalence, but with randomly distributed incidence of smoking.

The probability that a contact is a smoker should be higher in observed network than in the random networks, if given the given subject is a smoker.

**Reach of the cluster**

The point at which the probability that a contact is a smoker is no longer related to whether the subject is a smoker.
Clustering

The reason of clustering:

1. Subjects might choose to associate with contacts with similar smoking behavior — *homophily*.
2. Subjects and their contacts might have common unobserved contemporaneous exposures that cause their smoking behavior to covary — *confounding*.
3. Contacts might influence subjects — *induction*.
Methods — *Statistical Analysis*

**Longitudinal Logistic Regression**

Subject’s smoking status at time $t+1$ is a function of attributes such as age, sex, education level, smoking status at time $t$, and the smoking status of subject’s contacts at time $t$ and $t+1$.

$$Y(t+1) = b_0 + b_1 X_1(t) + b_2 X_2(t) + \ldots + b_nX_n(t+1)$$

**Time-lagged dependent variable**

Eliminates serial correlation in the errors and substantially controlled for subject’s genetic endowment and any intrinsic stable predilection to smoke.

**Time-lagged independent variable**

Helps to account for homophily. The key variable of interest is a contact’s smoking behavior at time $t+1$. 

Results — *Network Analysis*

Smokers in 2000 are less as compared with 1971, and are more likely to be at the periphery of the network.
The average risk of smoking in the observed network compare to in a random network:

Degree of separation = 1 61% higher
Degree of separation = 2 29% higher
Degree of separation = 3 11% higher
Degree of separation = 4 no excess

Thus the reach of the smoking cluster is three degrees.
Results — Network Analysis

Geographic distance does not appear to affect a contact’s smoking behavior
Smokers tend to tightly knit together, even when the incidence of smoking sharply declined.

Whole clusters stop smoking
Results — *Network Analysis*

Smokers’ centrality decreased over 32-year period.

Smokers become increasingly less central and more peripheral.
Results — *Interpersonal Model*

**Family**
- Spouses play an important role 67%
- Sibling 25%

**Friendship**
- subject-perceived friendship 36%
- contact-perceived friendship 15%
- mutual friendship 43%

**Educational Level**
- Persons with more education are more influential and more able to be influenced

**Coworkers**
- Effects are more stronger as the size of company decreases
Results — Interpersonal Model

![Graph showing the relationship between decrease in risk of smoking and different types of relationships.](image-url)
Results — *Interpersonal Model*

Effects in different smoking intensity:

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>siblings &amp; spouses</td>
<td>casual, moderate, heavy</td>
</tr>
<tr>
<td>friendship</td>
<td>heavy</td>
</tr>
<tr>
<td>coworkers</td>
<td>moderate, heavy</td>
</tr>
</tbody>
</table>
Conclusions

Network phenomena appear to be relevant to smoking cessation. Smoking behavior spreads through close and distant social ties, groups of interconnected people stop smoking in concert, and smokers are increasingly marginalized socially. These findings have implications for clinical and public health interventions to reduce and prevent smoking.
Questions?