

# Social Network Analysis of Peer Influence on Adolescent Smoking

Gregory J. Lambert, Patrick D. Finley, Thomas W. Moore

*Sandia National Laboratories*

Joint Statistical Meetings 2013

Montréal, Québec, Canada

August 3 – 8, 2013

This work was funded by the U.S. Food and Drug Administration through a contract with the U.S. Department of Energy/Sandia National Laboratories (funding document 224109011).

The information in this presentation is not a formal dissemination of information by the FDA and does not represent agency position or policy.

The contents are the responsibility of the authors alone.

- Adolescent cigarette smoking is strongly correlated with peer relationships
- How peers influence the behavior of other adolescents is not well understood
- Assortativity on smoking in adolescent friendship networks could indicate:
  - Peer influence (contagion)
  - Homophilous association (birds-of-a-feather)

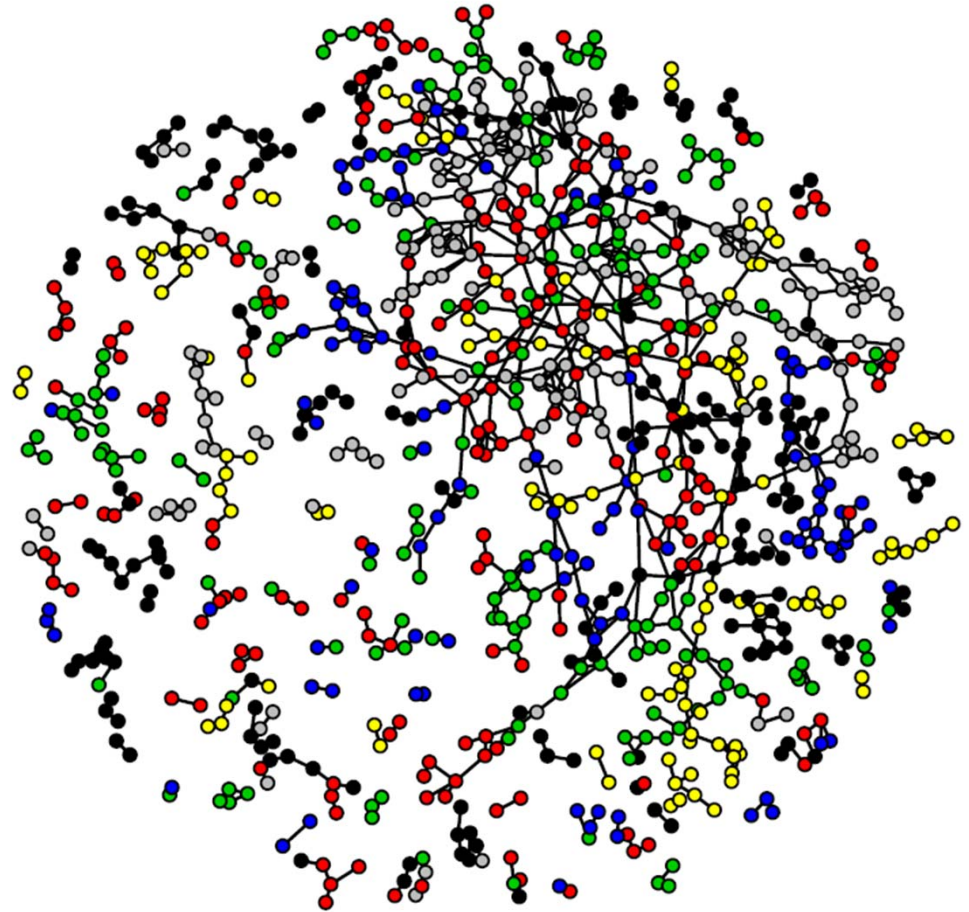
# Exponential Random Graph Model (ERGM)

Given a directed random graph  $Y$  consisting of a set of  $n$  nodes and  $m$  dyads

where

$Y_{ij} = 1$  if the nodes  $(i, j)$  are connected

$Y_{ij} = 0$  if not connected



# Exponential Random Graph Model (ERGM)

ERGM<sup>1</sup> generates a probabilistic distribution of the ties in a network

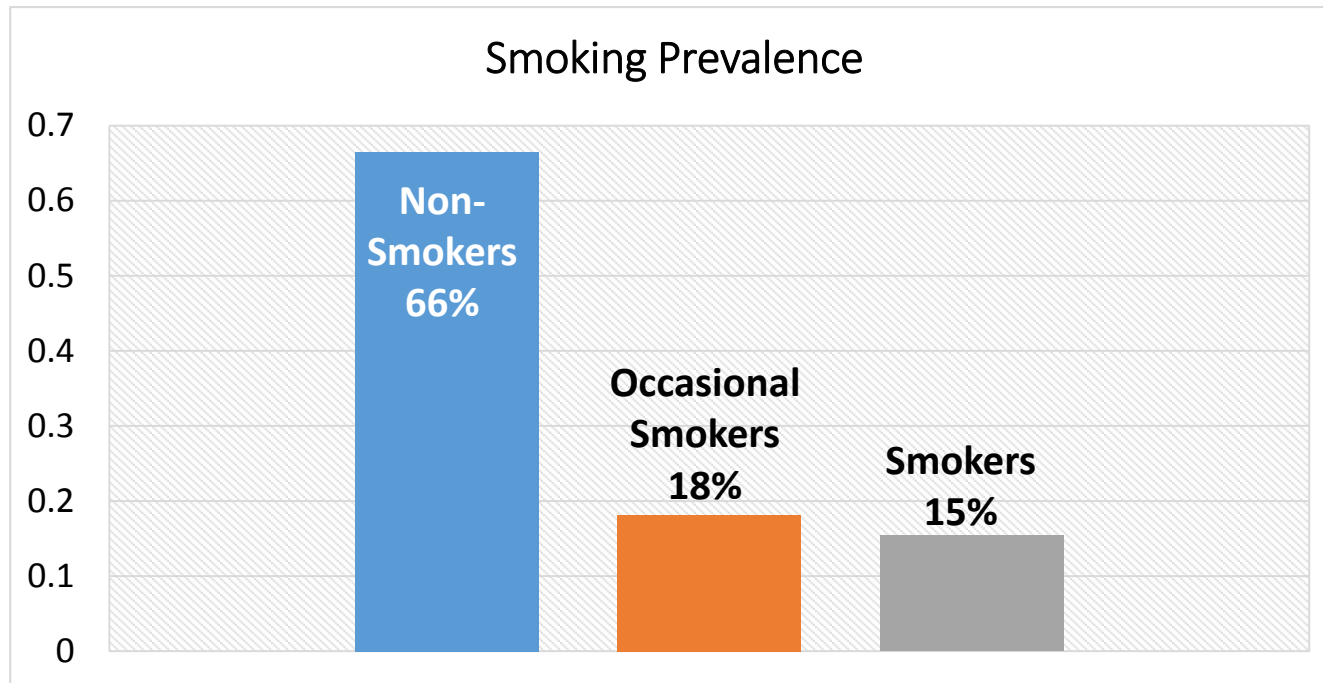
$$P(Y = y|\theta) = \frac{1}{c(\theta)} \exp\{\theta^T s(y)\}$$

- $Y$  is a random adjacency matrix for the network and  $y$  is a particular realization of  $Y$
- $s(y)$  represents any set of statistics depending on the observed network and nodal attributes
- $\theta$  is a vector of coefficients
- $c(\theta) = \sum_{all} \exp\{\theta^T s(y)\}$  is a normalizing constant

<sup>1</sup> Frank & Strauss (1986) and Wasserman & Patterson (1996)

# Smoking Prevalence in Add Health

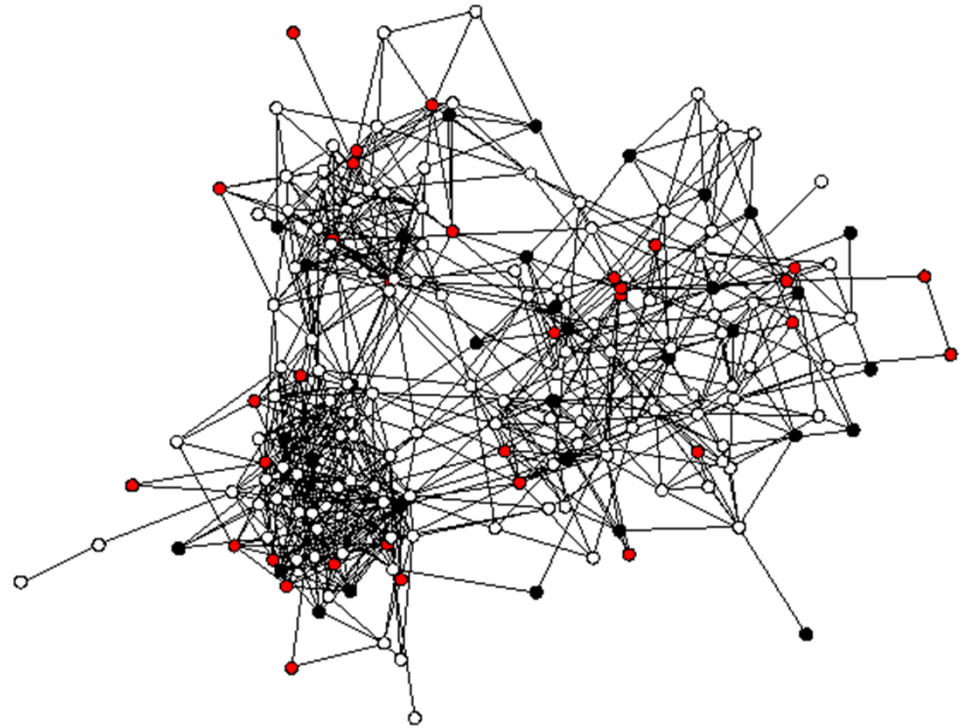
- Add Health is a longitudinal study of a nationally representative sample of adolescents in grades 7-12 in the United States.
- Example school district paired junior and senior high school: 682 adolescents (nodes) and 3718 friendship nominations (edges)



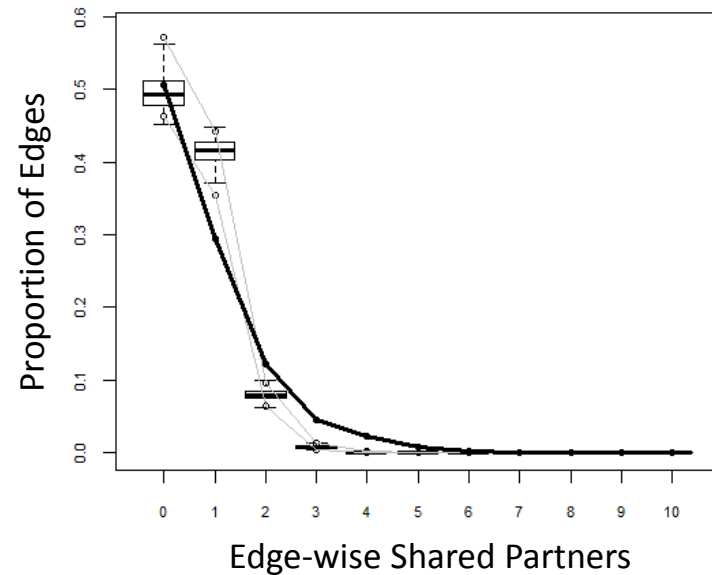
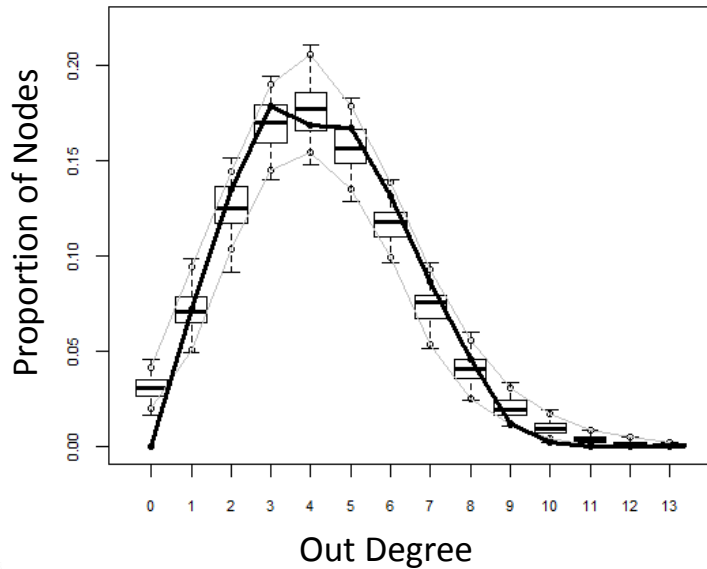
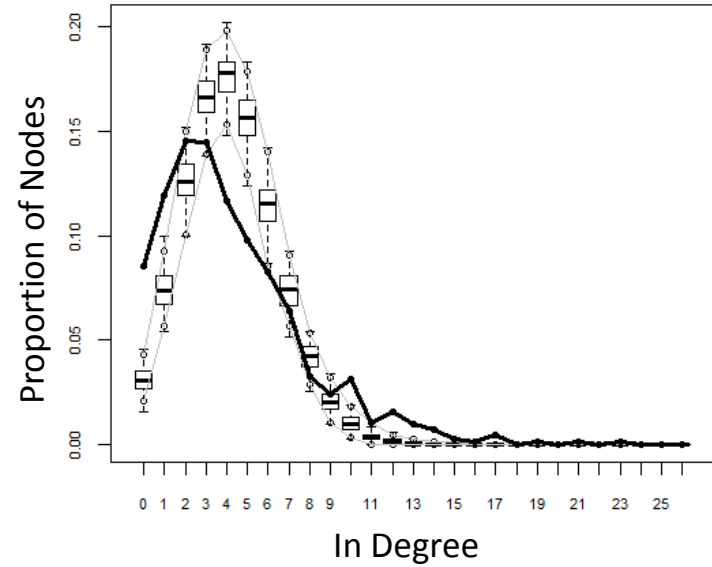
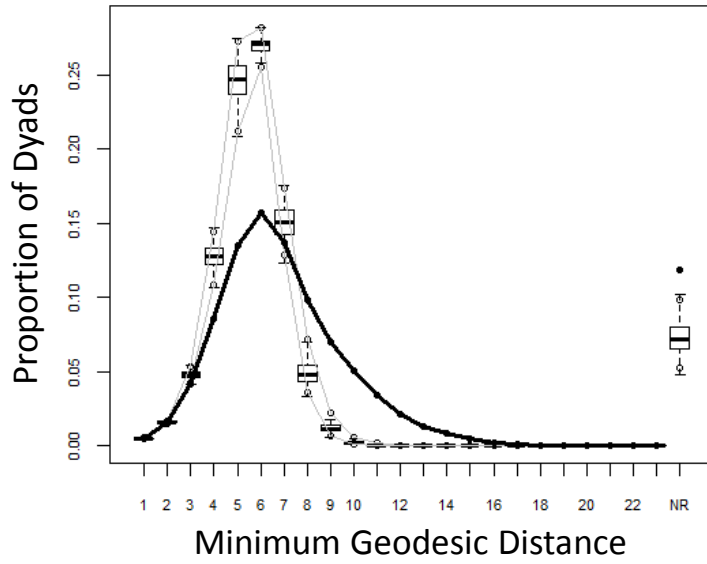
- Baseline model: Bernoulli or Erdos-Renyi
- Network and node terms are added to the model one at a time examining MCMC diagnostics and parameter estimates via AIC/BIC scores
  - Network terms: density, in-degree, out-degree
    - geometrically-weighted edgewise shared partner
    - geometrically-weighted dyad shared partner
  - Node terms: smoking assortativity mixing by grade level

This process results in a fitted ERG model

- Comparison of a single outcome from the simulation to the original network is of limited value
- Our process:
  - Generated 200 networks from our fitted ERG model
  - Compared simulated network topology to the observed Add Health network



# Goodness-of-Fit Diagnostics





Statistically significant terms ( $\alpha = 0.1$ ) for the fitted ERGM

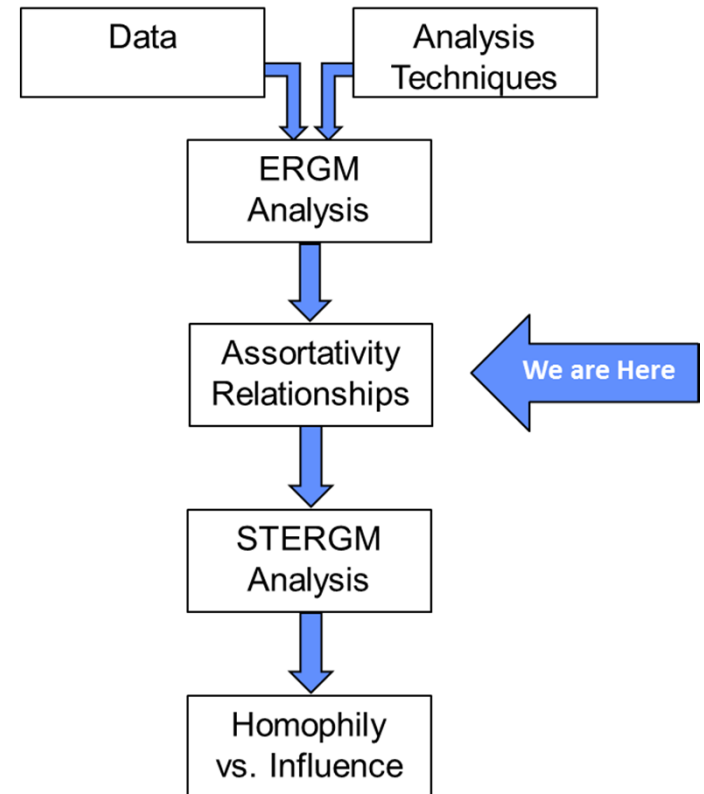
Term	Estimate	P-value	Log odds probability
edges	-4.93	<1e-04	0.007
gwesp	1.78	<1e-04	0.856
gwdsp	-0.16	<1e-04	0.460
reciprocity	3.86	<1e-04	0.979
10 <sup>th</sup> grade smoker → 9 <sup>th</sup> grade light smoker	-1.47	0.08	0.187
11 <sup>th</sup> grade light smoker → 10 <sup>th</sup> grade light smoker	-2.28	0.04	0.093
9 <sup>th</sup> grade light smoker → 10 <sup>th</sup> grade smoker	0.60	0.04	0.646
9 <sup>th</sup> grade light smoker → 11 <sup>th</sup> grade smoker	0.65	0.06	0.657
11 <sup>th</sup> grade light smoker → 12 <sup>th</sup> grade light smoker	0.84	0.03	0.698
11 <sup>th</sup> grade smoker → 12 <sup>th</sup> grade smoker	-0.88	0.10	0.293

## Conclusions

- Our ERGM analysis of Add Health data provided:
  - Reasonable goodness of fit
  - An objective way to look at metrics such as probabilities of connections among students of different grades and smoking behaviors
  - Awareness of constraints of working with limited data

## Future Directions

- Compare statistically significant model parameters for schools with different smoking prevalence
- Explore dynamic modeling to investigate peer influence : STERGM (Krivitsky 2012)
- Analysis of the impacts of contagion and homophily on adolescent smoking networks



- This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (<http://www.cpc.unc.edu/addhealth>). No direct support was received from grant P01-HD31921 for this analysis.