Statistical Inference of Social Media Data on Neighborhood Violence Prevention and Its Impact on the Field of Biostatistics

White, A.M., Lu, N., He, H., Silenzio, V & Tu, X.M. University of Rochester Medical Center

This picture shows how homeowners take pride in their home and street. This street is revitalized with new curbs, new sidewalks, and new paved streets. It brings beauty to homes and community. It shows that neighbors communicate with each other. It shows that this is a safe block.
-Block Club
Health and Social Connectivity

Examples:
Spread of disease
Social isolation & health status
Social determinants of health (e.g., where you live)
Systems coordination and population health

**Societal**
- Poverty
- High crime levels
- High residential mobility
- High unemployment
- Local illicit drug trade
- Weak institutional policies
- Inadequate victim care services
- Inadequate community cohesion

**Community**
- Psychological/personality disturbance (d/o)
- Alcohol/substance abuse
- Victim of child maltreatment or current abuse
- Violent behavior—past or current
- Suicidal behavior—past or current
- Access to lethal means

**Relationship**
- Exposure to poor parenting or violent parental conflict
- Fractured family structures
- Family history of suicide
- Current relationship/marital turmoil—participant in intimate violence
- Financial, work stress; under- or unemployed
- Friends & family that engage in violence

**Individual**
- Unstable social infrastructure
- Economic insecurity
- Discrimination: gender; race; other
- Policies that increase inequalities
- Poverty
- Weak economic safety nets
- Cultural norms that support violence
- Access to lethal methods (firearms)
Community-based Participatory Research: Community capacity-building in health

CBPR is research conducted by, for, and with communities on issues that are relevant to the communities and with the goal of bringing positive change in the community.

CBPR is a research approach that is geared at enabling community members to participate not as ‘research subjects’ but as research collaborators, in full participation at every phase of the research process.

In CBPR, research is not an end to itself but rather a means to empower communities through the participatory research process and to bring positive social change through mobilizing community-led evidence-based action.

(adapted from LOKA 2002 and Israel et al, 1996 by Shakya & Murtaza, July 15, 2009)
Promise and Challenge of Social Network Information on Health Behaviors and Health

In the context of ‘big data’ analytics of social network data (natural helping, violence and wellness) via social media, describe:

• Statistical issues arising for causal inference models

• Leverage theories of U-statistics and Functional Response Models to effectively address fundamental analytic issues that will lead to misinterpretations of study findings if unacknowledged and unaddressed.
“Sometimes he gets on my nerves but that man—there’s nothing he wouldn’t do for you. You could come in and ask him something…it may take him a couple of days to get there, but he’s a good man”

The Rock Garden highlights how passion and creativity can transform. It also has a memorial plaque & apple tree in honor of one of our Block Club Leaders who passed away. So much is about vision and perspective. Possibility thinking vs. staying in a disgruntled frame of mind. The Rock Garden signifies hope.

“My thing is to show them love and concern. You can’t look at everybody with a troubled life—you can’t look at them and downsize them.”

Raspberry patch separated by a wooden fence in the middle. The plants and trees honor those who have died. The fence was a suggestion of one of our youth so we’d be able to “eat off of both sides of the raspberries”. There are memorials to friends & family who have died to help youth & adults deal with the loss in a healthy way.
Are There Connections Among Formal and Informal Adult Helping Networks?: Identify, Foster & Map Information Flow

Formal Prevention/Promotion Program – Organization - Provider

Natural Helper

Individuals/ Family Members
Ordinary adult residents are intentionally initiating key functions in neighborhoods to maintain mental wellness. Their approaches are transferable, can be replicated, and may play a bridging role in young residents’ having positive “formal helping system” engagement.

Residents crafting of their own “learning collaborative” on NHing generated positive changes. NHs efforts to reduce youth and community violence, build mental wellness, and reduce neighbor and intergenerational mistrust or disrespect was expanded or sustained (where fatigued).

Intergenerational transmission of traditional “natural helping” approaches is fractured. NHs wish to understand where ties to youth, young adults and young families are strongest and how to keep these networks strong.

Formal helping systems can help and hurt NH attempts
‘Big Data’ Hypotheses: Relationship between NH and Community Violence and Health

Lu, Hua & White et al, NSF grant [in review]
Challenges in Modeling Social Media Data

- Traditional statistical models such as regression (R) and structural equation models (SEM) are defined for modeling relationships between variables, measures of attributes of a subject such as age and alcohol abuse.

- An **independent** sample of subjects is required to provide inference about model parameters, e.g., standard errors, p-values and confidence intervals.

- Network connections are defined by pairs of subjects, and the relationships are **dynamic** in the sense that one subject can form different relationships with others, violating independence across subjects, a **foundation for inference** for all popular statistical models such as R and SEM.
Challenges in Modeling Social Media Data

Two different data structure, left is the social network data, right is the regular data structure.
Leverage theories of U-statistics and Functional Response Models

The U-statistics (U) and functional response models (FRM) are uniquely positioned to address the fundamental flaws of traditional models.

U and FRM model between-subject dynamics such as network connection:

\[ R \text{ (SEM)}: \quad Y_i = h(X_i); \]
\[ U \text{ (FRM)}: \quad f(Y_i, Y_j) = h(X_i, X_j). \]

Thus, unit of analysis is an individual for R (SEM), but a pair of individuals for U (FRM).
Leverage theories of U-statistics and Functional Response Models

Current approaches to SN analysis utilizes re-sampling methods such as Bootstrap.

Such methods are again developed for modeling individual-level variables, yielding biased estimates, when applied to between-subject relationships as in SNA.
Pilot study

- We conducted a pilot study to examine the performance of traditional and FRM models in SNA.
- We considered the network density, as this relatively simple measure of network connectivity embodies the prototypical issue with traditional models.
- The network density measures the average connectivity between two subjects, or nodes, in a SN.
Pilot Study Findings
The between-subject dynamic relationship in SNA invalidates the foundation for inference in current statistical paradigm.

A new set of statistics and regression models need to be developed to provide valid inference for SNA.

Premised on the U-statistics framework, FRM is a class of regression models for between-subject dynamic relationships such as SN density and other parameters of interest.
The end