Developing a neighborhood deprivation index for the TBI Model Systems National Database

Raj Kumar, PhD MPH

Assistant Professor

Brain Injury Research Center

Icahn School of Medicine at Mount Sinai

TBIMS Analytic SIG

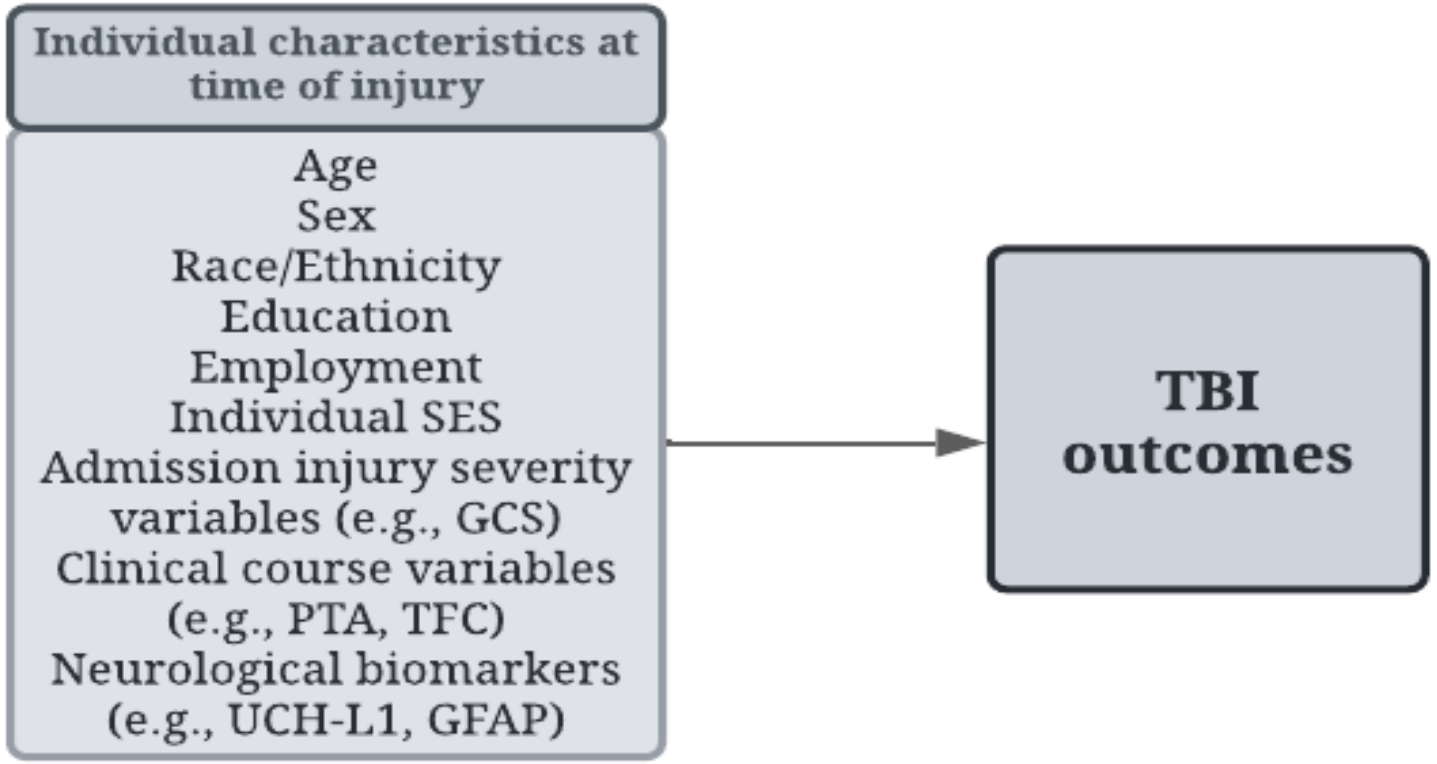
November 22nd, 2024

Mount Sinai *Brain Injury Research Center*

Background

- An axiom in the TBI field is: "***No two TBIs are the same***"
- This has historically referred to the heterogeneity of TBI clinical presentationon multiple personal levels:
  - Demographics (e.g., age, sex, personal socioeconomic status)
  - Mechanisms of injury (e.g., falls, motor vehicle, assaults)
  - Neuroanatomic insult (e.g., focal lesions to regions of the brain, diffuse axonal injuries)
  - Pathophysiology (e.g., over- or under-active inflammation, neurotrophic factors)
  - Acute and chronic symptoms (e.g., presence or absence of seizures, migraine, depression)

Current TBI IMPACT Prognostic Model has room for improvement



The cumulative variance explained across all variables considered (demographics, injury severity, secondary insults, CT characteristics, lab values, etc.) is only ~35%.

Motivation for study

- We endeavored to expand beyond personal factors alone, and develop a neighborhood socioeconomic index that could be employed in the TBI Model Systems National Database
- Other neighborhood socioeconomic indices indeed exist in the literature (i.e., Area Deprivation Index, Social Vulnerability Index); however, they were designed for other disease populations and may not have all the components relevant to TBI populations.
  - There is no consensus in components across indices, which motivated our TBI Model Systems network of researchers to create our own
- We also recognize that there is variability in socioeconomics within a given neighborhood
  - The intersection between neighborhood and individual socioeconomics is important to illustrate a more complete picture of socioeconomic circumstances

Study aims

- Our methods-focused study aimed to develop two separate, but related, objective measures of neighborhood socioeconomic status
  1. We first created a census-based composite neighborhood socioeconomic deprivation index (NSDI) based on geocoded residential addresses of TBIMS enrollees
  2. We created a second measure characterizing the degree of consistency between the individual's socioeconomic status and that of their neighborhood (*Neighborhood:Individual NSDI residual*)

We analyzed data from the year 2, 5, and 10 year post-injury follow-up interviews

Neighborhood data sources

- We characterized neighborhood socioeconomic circumstances using data from the US Census' American Community Survey (ACS)
  - We used five-year average ACS estimates between 2015-2019
- We focused on eight ACS variables spanning domains of: education, income/poverty, (un)employment, and single parenthood

Variable Name	Definition	Formula
Percent Unemployed	The percentage of civilian unemployed (people 16 and over)	= (# civilian unemployed / # in labor force)
Percent Single Head of Household	The percentage of single parent headed households with children <18	= ((# male household + # female household) / # in family households).
Percent with No High School Diploma or GED	The percentage of people >=25 years old without a high school diploma or GED	= (No schooling completed + Nursery school + Kindergarten + 1st through 11th grade + 12th grade, no diploma)/Total in CensusTract
Percent with Bachelor's or Higher	The percentage of people >=25 years old with a bachelor's degree or higher	= (Bachelor's degree + Master's degree + Professional school degree + Doctorate degree) / Total in CensusTract
Percent Below Poverty	The percentage of households with incomes in the past 12 months below poverty level	= (Income in the past 12 months below poverty level / Total for income versus poverty level)
Percent SNAP	The percentage of households that received Food Stamps/SNAP in the past 12 months	= (Household received Food Stamps-SNAP in the past 12 months / Total for receipt of SNAP)
Median Household Income	Median household income in the past 12 months (in inflation-adjusted dollars)	This variable was contained in the data set with no need for calculations
Median Family Income	Median family income in the past 12 months (in inflation-adjusted dollars)	This variable was contained in the data set with no need for calculations

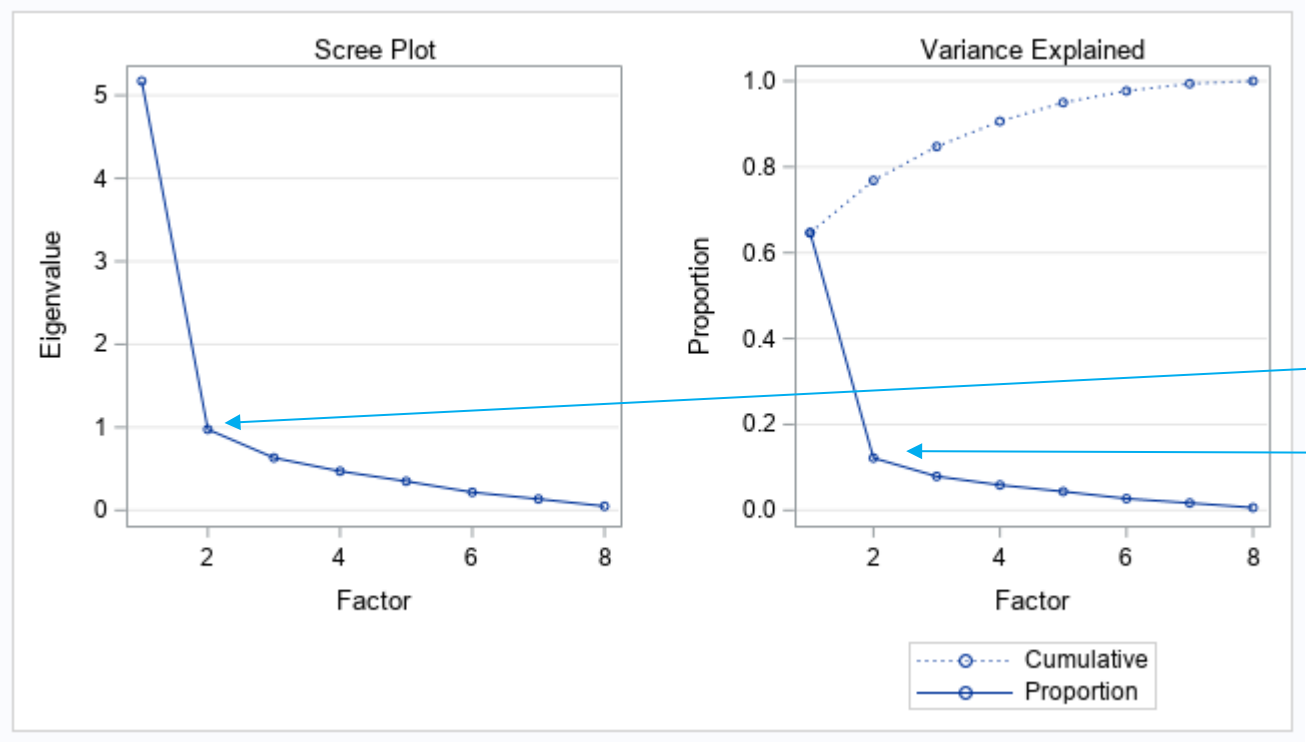
Individual socioeconomic data

- We evaluated three individual-level SES variables within the TBI Model Systems National Database corresponding to *similar constructs* to the neighborhood variables used:
  - Household income (less than $25,000; $25,000 - $50,000; $50,000 - $100,000; $100,000 - $150,000, $150,000 - $200,000; or $200,000 or more)
  - Years of education (continuous variable 1-20 years)
  - Unemployed (vs. competitively or specially employed, student, homemaker, retired)

Statistical analyses

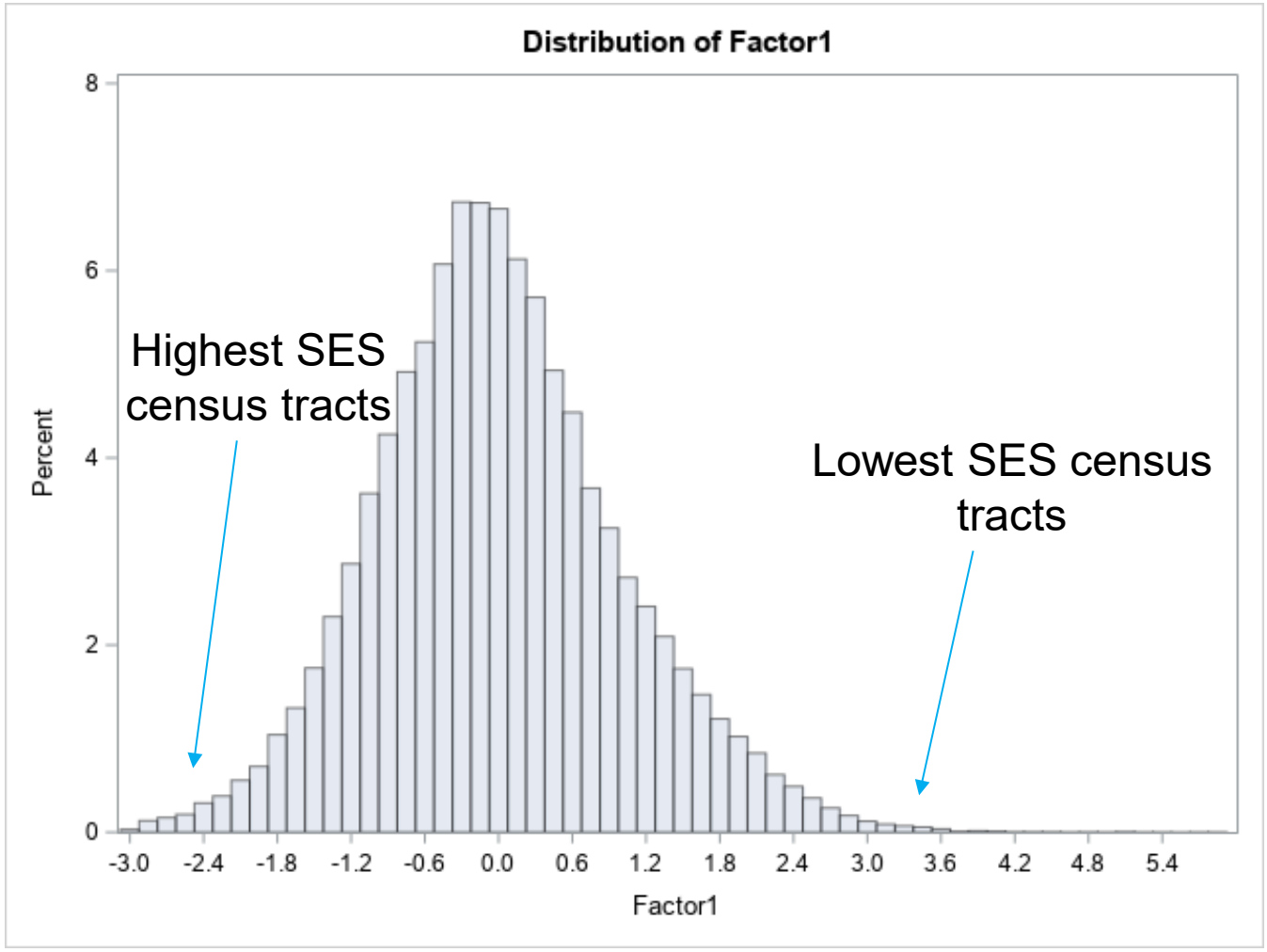
- For Aim 1, to create our neighborhood index, we used principal components analysis (PCA) to reduce the dimensionality of our eight variables into a concise neighborhood index, known as the TBIMS Neighborhood Socioeconomic Disadvantage Index (NSDI)
  - NSDI was run on all US census tracts (not just TBIMS sample) before it was linked to TBIMS
- For Aim 2, to quantify the intersection between neighborhood and individual socioeconomic status, we used residual analysis methods
  - We first ran a multivariable ordinary least squares (OLS) linear regression model with NSDI as the outcome with individual SES predictors of household income, education, and unemployment status
  - We calculated residuals defined as the difference between the observed and the predicted values of the NSDI at 2-, 5-, and 10-year follow-up time points post-injury

Aim 1: PCA output & deciding number of retained components



"Elbow" in these plots tells us very little incremental value of PC2, PC3, PC4 etc. PC1 is adequate to explain most of variance (~2/3)

PCA output: PC1 distribution of census tracts



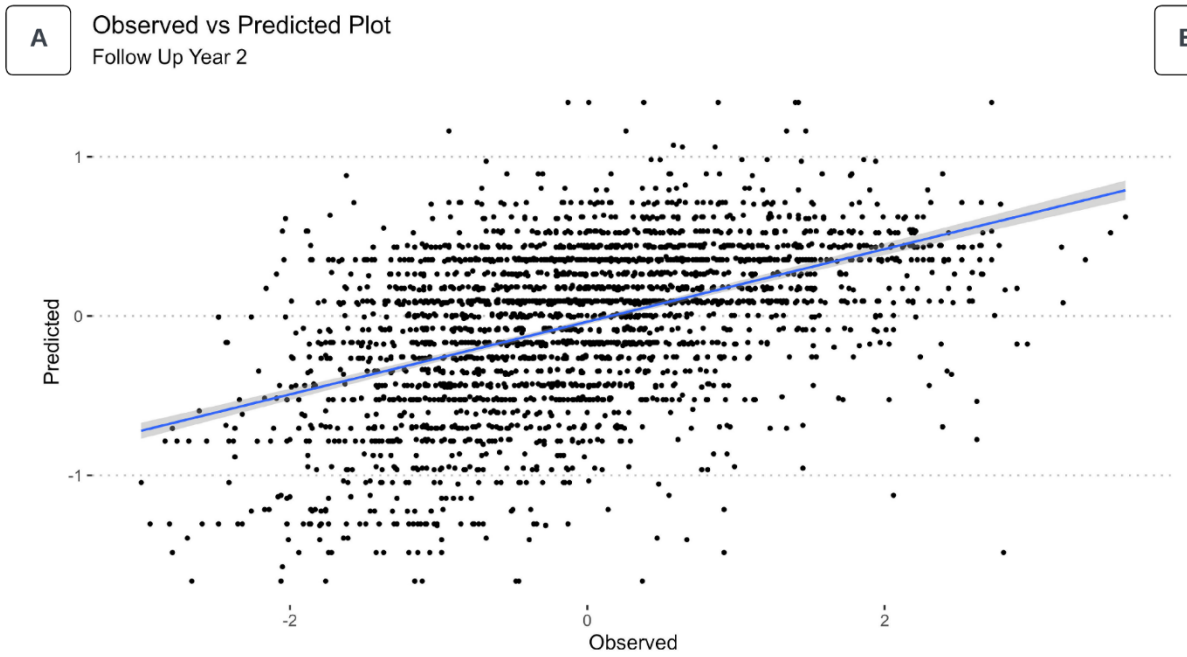
Moments			
N	72497	Sum Weights	72497
Mean	0	Sum Observations	0
Std Deviation	1	Variance	1
Skewness	0.30974648	Kurtosis	0.40731399
Uncorrected SS	72496	Corrected SS	72496
Coeff Variation	.	Std Error Mean	0.00371398

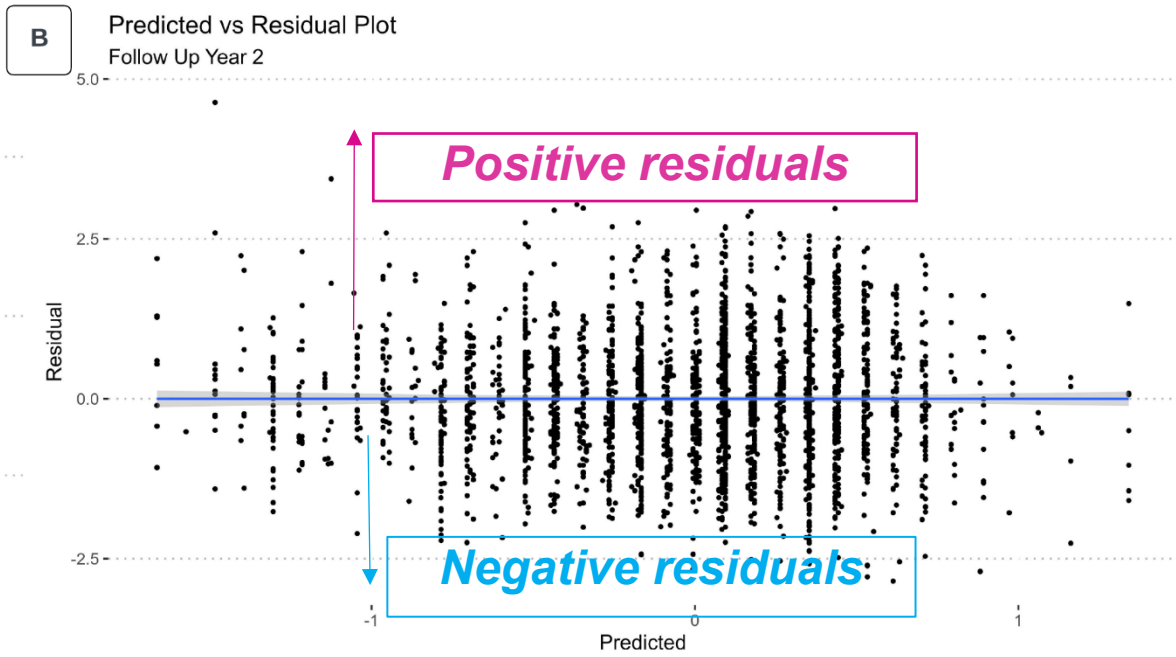
Basic Statistical Measures			
Location		Variability	
Mean	0.00000	Std Deviation	1.00000
Median	-0.05960	Variance	1.00000
Mode	.	Range	8.89072
		Interquartile Range	1.26184

Quantiles (Definition 5)	
Level	Quantile
100% Max	5.8455311
99%	2.5581767
95%	1.7733053
90%	1.3248413
75% Q3	0.5998058
50% Median	-0.0595968
25% Q1	-0.6620355
10%	-1.2098358
5%	-1.5565775
1%	-2.2428902
0% Min	-3.0451889

Note: Higher values=lower SES, why we are interpreting this as a neighborhood socioeconomic deprivation index

Aim 2: Observed vs. Predicted & Residual Plots





- The “observed” values of the NSDI are the dots, and the “predicted” values are represented by the line (panel A)

Residual = Observed value – Predicted value

Positive residuals: Obs > Predicted

Negative residuals: Obs < Predicted

Aim 2: Qualitative Characterization of Residual

TABLE 2 *Description of the sample by Neighborhood: Individual NSDI residual group*

***Neighborhood:Individual NSDI Residual Group* Categories**	**Studentized Residual Values[a] (Min, Max)**	**Qualitative Description**
Very low	(−2.86, −1)	*Much less neighborhood disadvantage than predicted based on individual SES*
Moderately low	(−1, −0.35)	*Somewhat less neighborhood disadvantage than predicted based on individual SES*
Near Zero	(−0.35, 0.35)	*Close neighborhood disadvantage to predicted based on individual SES*
Moderately high	(0.35, 1)	*Somewhat higher neighborhood disadvantage than predicted based on individual SES*
Very high	(1, 4.64)	*Much higher neighborhood disadvantage than predicted based on individual SES*

[a]By using a studentized residual, the residual value is interpreted in standard deviation units. For example, a person in the "very high" residual group would have an observed TBIMS-NSDI value over one standard deviation higher than their predicted value.

Aim 2: Qualitative Characterization of Residual

TABLE 3 *Case exemplars illustrating the interpretation of the Neighborhood:Individual NSDI residual group at year 2 post-injury*

	Pseudo-ID†	Household annual income	Years of education	Unemployment status¥	Predicted TBIMS-NSDI ($\hat{y}$)	Observed TBIMS-NSDI (y)	Raw residual (y-$\hat{y}$)	Residual group𝔯
Example No. 1: Same individual SES characteristics, different neighborhood SES	101	$25,000–$49,000	12	Unemployed	0.17	−0.98	−1.15	Very low
	102	$25,000–49,000	12	Unemployed	0.17	−0.53	−0.70	Moderately low
	103	$25,000–49,000	12	Unemployed	0.17	0.11	−0.06	Near 0
	104	$25,000–49,000	12	Unemployed	0.17	1.04	0.87	Moderately high
	105	$25,000–49,000	12	Unemployed	0.17	2.12	1.95	Very high
Example No. 2: Different individual SES characteristics, same neighborhood SES	201	Less than $25,000	12	Unemployed	0.43	−0.34	−0.77	Moderately low
	202	Less than $25,000	11	Unemployed	0.52	−0.34	−0.86	Moderately low
	203	$25,000–49,000	11	Not unemployed	0.18	−0.34	−0.52	Moderately low
	204	$50,000–99,999	12	Not unemployed	−0.16	−0.34	−0.17	Near 0
	205	$50,000–99,999	19	Not unemployed	−0.79	−0.34	0.45	Moderately high

†: Pseudo-IDs were created to protect the anonymity of participants, but all other data elements were real data from actual TBI Model Systems participants.
𝔯: The residual group was calculated based on values of the studentized residual, which divides the raw residual by the standard deviation of the residuals.
¥: Participants were classified as either unemployed or not unemployed. The unemployed category is either "unemployed: looking" or "unemployed: not looking", and the latter category consists of all other codes for the employment variable.

Summary of descriptive results for residual variable

- We found age at injury was similar between residual groups, but proportion of males was higher in groups with *moderate or very high residuals* (i.e., higher neighborhood disadvantage than predicted)
- Other key variables associated with living in a higher disadvantaged neighborhood than predicted:
  - Black and Hispanic race/ethnicity
  - Medicaid insurance
  - Violent mechanism of injury
  - Residence in Urban areas
  - Geographically located in Northeast and South

Conclusions and next steps

- In this methods-focused study, we created two indices, the *TBIMS NSDI* and the corresponding Neighborhood: Individual SES residual variable
  - The *TBIMS NSDI* has been linked to the TBIMS NDB for future researchers to deploy for substantive research questions
  - Our residual variable provides a *supplemental* empirical measure that incorporates information about both the neighborhood and the individual
- Future directions:
  - We plan to conduct a follow-up study comparing the prognostication of TBI outcomes between: 1) *TBIMS NSDI*, 2) the Neighborhood: Individual SES residual, and 3) individual SES variables

Acknowledgements

I would like to thank my collaborators on this project:

- Andrew Delgado, PhD
- John Corrigan, PhD
- CB Eagye, MS
- Gale G. Whiteneck, PhD
- Shannon B. Juengst, PhD
- Librada Callender, PhD
- Jennifer Bogner, PhD
- Shanti Pinto, MD
- Amanda R. Rabinowitz, PhD
- Paul B. Perrin, PhD
- Umesh M. Venkatesan, PhD
- Amanda L. Botticello, PhD
- Anthony H. Lequerica, PhD
- Shameeke Taylor, MD
- Ross D. Zafonte, DO
- Kristen Dams-O'Connor, PhD