University of California
Santa Barbara

Space, Place, and Public Health Surveillance

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Geography

by

Kevin J. Konty

Committee in charge:

Professor Stuart Sweeney, Chair
Professor Susan Cassels
Professor Richard Church

September 2019

The Dissertation of Kevin J. Konty is approved.

Professor Susan Cassels

Professor Richard Church

Professor Stuart Sweeney, Committee Chair

September 2019

Space, Place, and Public Health Surveillance

To Stella and Clyde and Kymberly. I have more time now.

Acknowledgements

This dissertation would not be possible without the generosity, patience, and support of many people as well as their confidence that I would finish despite all the evidence to the contrary. This is particularly true of my wife Kymberly who I can never repay for her patience. Our second date was a trip to visit the UCSB Geography department 20 years ago and she never doubted my ability to finish. My kids Clyde and Stella have had a dad who was "working" on his dissertation all their lives and they have been so supportive and kind and insightful.

I would also like to thank Stuart Sweeney. Stuart is the reason I have actually finished this dissertation. We have always worked so well together and through our conversations about technical matters and policy and content I have realized that I am fundamentally a geographer, whatever that means. He has stuck with me for 20 years and his presence has been both a pleasure and a major contributor to my work in public health. Eight years ago he casually suggested changing my dissertation to something from public health. He has also kindly provided bread and bed and breakfast and, perhaps most importantly, a model family. Through him I have gained an appreciation for (and an example of) the value of pursuing a purposeful life. And thank you Laura!

I would also like to thank the UCSB Geography department particularly Rick Church and Susan Cassels for their advice and support. The faculty, staff, and fellow students have always been so generous and optimistic about my prospects including Dan Montello and Waldo Tobler who would simply ask how I was progressing after not seeing them for several years. Likewise my colleagues Nick Nagle, Kathryn Grace, Greg Husak, Adrienne Domas, Meri Marsh, and Martin Herald. I couldn't have been luckier.

I was also very lucky to stumble into Geography and then into Public Health. Public health is incredibly interesting, meaningful, and I have directly experienced the potential for geographers to contribute to the field. In circling back to finishing my dissertation I am most indebted to my fellow traveler and dear friend Donald Olson and my longtime colleague and counsel Sophia Day, as well as colleagues Tsu-Yu Tsao, Alejandro Cajigal, Marc Paladini, Thomas Cannell, and Nikos Soulakis. Farzad Mostashari brought me into public health and it changed my life for the better.

Almost lastly, I have many friends and family who have never wavered in their support while rolling their eyes at how ridiculous I can be including my brother Mark who is so fun to talk to about social science and sports, my sister Desiree who is the most loving and generous person I know, my doubting dad Larry who I am pretty sure was just slyly prodding me along, my insightful Sociologist sister-in-law Melissa Fry, and big hearted brother-in-law Brandon Harris. Special thanks to my mother-in-law Ellie, whose pride when I finished felt like my own mother's pride, and all my other Southern California in-laws. Oh, and my friends Dax Zimmerman, Gentry Kirby, Kenneth Shapiro, and Jeff Krupman.

Finally, I would like to thank the organizers of Chaire Quetelet for their seminar on local demography in Louvain-La-Neuve, Belgium, 2015; UCSB's "un-conference" Spatial Information for Human Health in 2015, and the Robert Wood Johnson Foundation's support to attend the All In: Data for Community Health series of conferences. These meetings advanced the projects herein.

Curriculum Vitæ

Kevin J. Konty

Education

2019	Ph.D. in Geography (Expected), University of California, Santa Barbara.
1996	M.A. in Applied Statistics, The Wharton School of the University of Pennsylvania
1994	M.S. in Mathematics, California Institute of Technology
1993	B.S. in Mathematics, Tulane University

Professional Experience

2015-present	**Director, Research and Analytics** *New York City Department of Health and Mental Hygiene* and *New York City Department of Education* Office of School Health
2005-2015	**Director, Methodology Unit** and **Senior Scientific and Statistical Advisor** *New York City Department of Health and Mental Hygiene* Bureau of Epidemiology Services, Division of Epidemiology
2004-2005	**Statistician/City Research Scientist IV** *New York City Department of Health and Mental Hygiene* Bureau of Epidemiology Services, Division of Epidemiology
2003-2004	**Demographer** *United States Census Bureau*, Suitland, MD Administrative Records and Methodological Research Branch Population Division
1999-2003	**Research Assistant** *Department of Geography* University of California at Santa Barbara
1996-1998	**Senior Statistician / Litigation Consultant** *William E. Wecker Associates*, Novato, CA
1995-1996	**Statistician / Litigation Consultant** *Analysis and Inference, Inc.*, Swarthmore, PA

Publications

Journal Articles (49)

"Quantifying the impact of influenza among persons with Type 2 Diabetes Mellitus: A new approach to determine medical and physical activity impact" *Journal of Diabetes Science and Technology*, 2019 (In Press)
S. Samson, **K. Konty**, W.N. Lee, T. Quisel, L. Foschini, D. Kerr, J. Liska, H. Mills, R. Hollingsworth, M. Greenberg, and A. Beal

"Zip Code-level estimates from a local health survey: Added value and limitations." *Urban Health*, 2019 (In Press)
Q. Bi, A. Levanon-Seligson, F. He, **K. Konty** and H. Gould.

"The effects of student, school and neighborhood poverty on the association between fitness and absenteeism in New York City middle school youth. *Preventive Medicine*, 2019:127
E D'Agostino, S. Day SE, **K. Konty**, M. Larkin, and K. Wyka.

"Influenza Surveillance Using Wearable Mobile Health Devices." *Online Journal of Public Health Informatics*, 2019 30;11(1)
K. Konty, B. Bradshaw B, W.N. Lee, A. Signorini, and L Foschini .

"Using Indirect Measures to Identify Geographic Hot Spots of Poor Glycemic Control - Cross-Sectional Comparisons with an A1C Registry." *Diabetes Care*, 2018: https://doi.org/10.2337/dc18-0181
D. Lee, Q. Jiang, B. Tabaei, B. Elbel, C. Koziatek, **K. Konty** and W. Wu.

"The Association of Health-Related Fitness and Chronic Absenteeism Status in New York City Middle School Youth." *Journal of Physical Activity and Health*, 2018:20(3), pp.1-9
E D'Agostino, S. Day SE, **K. Konty**, M. Larkin, S. Saha, and K. Wyka.

"A Bayesian Evidence Synthesis Approach to Estimate Disease Prevalence in Hard-To-Reach Populations: Hepatitis C in New York City." *Epidemics*, 2018: https://doi.org/10.1016/j.epid
H.F. Tan, S. Makela, D. Heller, **K. Konty**, S. Balter, T. Zheng, and J. Stark.

"The association of fitness and school absenteeism across gender and poverty: a prospective multilevel analysis in New York City middle schools." *Annals of Epidmiology*, 2018: 28(3):189-196
E. D'agostino, S. Day, **K. Konty**, M. Larkin, S. Saha, and K. Wyka.

"Individual-Level Fitness and Absenteeism in New York City Middle School Youths, 2006-2013." *Preventing Chronic Disease*, January 2018:15(5)
E. D'agostino, S. Day, **K. Konty**, M. Larkin, S. Saha, and K. Wyka.

"Evaluating and implementing temporal, spatial, and spatio-temporal methods for outbreak detection in a local syndromic surveillance system." PloS one. 2017 Sep 8;12(9):e0184419

R. Mathes, R. Lall, A Levin-Rector, J. Sell, M. Paladini, **K. Konty**, D. Olson, and D. Weiss.

"Relationship between recreational resources in the school neighborhood and changes in fitness in New YorK City public school students." *Journal of Urban Health*, 2017:94(1):20-29

C. Bezold, J. Stark, A. Rundle, **K. Konty**, S. Day, J. Quinn, K. Neckerman, and A. Diez Roux.

"Ten-year cancer incidence in survivors of the World Trade Center disaster." *American Journal of Industrial Medicine*, 2016:59(9):709-21

J. Li, R. Brackbill, T. Liao, B. Qiao, J. Cone, M. Farfel, J. Hadler, A. Kahn, **K. Konty**, L. Stayner, and S Stellman.

"Potential reductions in premature mortality in NYC from raising the minimum wage to $15." *American Journal of Public Health*, 2016:106(6)

T. Tsao, **K. Konty**, G. Van Wye, O. Barbot, J. Hadler, N. Linos, and M. Bassett.

"A school-level proxy measure for individual poverty using school-level enrollment in free and reduced price meals." *Journal of School Health*, 2016, 86(3): 204-214.

S. Day, K. Hinterland, C. Myers, L. Gupta, T. Harris, and **K. Konty**.

"Evaluation of the New York City Green Carts program." *AIMS Public Health*, 2016, 2(4): 906-918.

S. Farley, R. Sacks, R. Dannefer, M. Johns, M. Leggat, S. Lim, **K. Konty**, and C. Nonas.

"Enhancing disease surveillance with novel data streams." *EPJ Data Science*, 2015, 4:17

B. Althouse, S. Scarpino, L. Meyers, J. Ayers, M. Bargsten, J. Baumbach, L. Castro,H. Clapham, D. Cummings, S. Del Valle, S. Epperson, S.Eubank, G. Fairchild, L. Finelli, N. Generous, D. George, D. Harper, L. Herbert-Dufresne, M. Johansson, **K. Konty**, T. Merlin, G. Millinovich, J. Miller, E. Nsoesie, D. Olson, M. Paul, P. Polgreen, R. Priedorsky, J. Read, I. Rodriguez-Barraquer, D. Smith, D.Swerdlow, D. Thompson, A. Vespignani, and A. Wesolowski.

"Invasive Pneumococcal Disease after the introduction of 13-valent conjugate vaccine in children, New York City, 2007-2012." *JAMA Pediatrics*, 2015, 169(7):646-652

J. Rosen, A. Farnham, C. Zimmerman, V. Papadouka, **K. Konty**, and J. Zucker. *Recipient the 2017 NYC DOHMH award for most impactful research article*

"The impact of neighborhood park access and quality on body mass index among adults in New York City." *Preventive Medicine*, 2014, 64: 63-68.

J. Stark, K. Neckerman, G. Lovasi, C. Richards, J. Quinn, C. Weiss, M. Bader, **K. Konty**, T. Harris, and A. Rundle.

"The effect of changes in physical fitness on academic performance in New York City youth." *Journal of Adolescent Health*, 2014, July 31.

C. Bezold, **K. Konty**, S. Day, M. Larkin, L. Harr, S. Saha, M. Napier, M. Berger, T.Harris, and J. Stark.

"Severe obesity among NYC public elementary and middle school children from 2006-2010." *Preventing Chronic Diseases*, 2014;11:130439.
S. Day, **K. Konty**, M. leventer-Roberts, C. Nonas, and T. Harris.

"Estimating the prevalence of Hepatitis C infection in New York City using surveillance data." *Epidemiology and Infection*, 2014:142.
S. Balter, J. Stark, K. Bornschlagel, J. Kennedy and **K. Konty**.

"Description of a school nurse visit syndromic surveillance system and comparison to emergency department visits, New York City." *American Journal of Public Health*, 2014:104(1).
E. Wilson, J. Egger, **K. Konty**, M. Paladini, D. Weiss, and T. Nguyen.

"Reaching children never previously vaccinated for influenza through a school-located vaccination program." *American Journal of Public Health*, 2014:104(1).
S. Kansagra, V. Papadouka, A. Geevarughese, M. Hansen, **K. Konty**, and J. Zucker.

"Reassessing Google Flu Trends data for detection of seasonal and pandemic influenza: A comparative epidemiological study at three geographic scales." *PLoS Computational Biology*, 2013:9(10). (**R**ecipient the 2014 ISDS award for outstanding research article)
D. Olson, **K. Konty**, M. Paladini, C. Viboud and L. Simonsen.

"Reportable disease incidence following Hurricane Sandy, New York City, 2012" *Disaster Medicine and Public Health Preparedness*, 2013:7(5):513-521.
S. Greene, E. Wilson, **K. Konty**, and A. Fine.

"Neighbourhood food environments and body mass index among New York City adults." *Journal of Epidemiology and Community Health*, 2013, 2013 Sep;67(9):736-42.
J. Stark, K. Neckerman, G. Lovasi, **K. Konty**, J. Quinn, P. Arno, D. Viola, T. Harris, C. Weiss, M. Bader, and A. Rundle.

"Socio-economic status, neighborhood food environments and consumption of fruits and vegetables in New York City." *Public Health Nutrition*, 2013, July:16(7):1197-1205.
D. Jack, K. Neckerman, O. Schwar-Soicher, G. Lovasi, J. Quinn, C. Richards, M. Bader, C. Weiss, **K. Konty**,P. Arno, D. Viola, B. Kerker, and A. Rundle.

"Aesthetic amenities and safety hazards associated with walking and bicycling in New York City." *Annals of Behavioral Medicine*, 2013, Feb;45:S76-85.
G. Lovasi, O. Schwar-Soicher, K. Neckerman, **K. Konty**, B. Kerker, J. Quinn and A. Rundle.

"Seroprevalence of pandemic influenza viruses among a population based sample of New York City adults." *Emerging Infectious Diseases*, 2012, Nov,18(11).
I. Weisfuse, T. Tsibane, **K. Konty**, J. Egger, E. Waddell, S Rahmat, E. Harris, D. Olson, and C. Basler.

"Neighborhood walkability and active travel(walking and cycling) in New York City." *Journal of Urban Health*, 2012, September.

L. Freeman, K. Neckerman, O. Schwar-Soicher, J. Quinn, C. Richards, M. Bader, G. Lovasi, C. Weiss, **K. Konty**, P. Arno, D. Viola, B. Kerker, and A. Rundle.

"Change in trans fatty acid content of fast food purchases associated with New York City's restaurant regulation." *Annals of Internal Medicine*, 2012, 157(2): 81-86. (**Recipient the 2014 NYC DOHMH award for most impactful research article.**)
S. Angell, L. Cobb, K. Curtis, **K. Konty**, and L. Silver.

"A case study of the impact of inaccurate cause-of-death reporting on health disparity tracking: New York City premature cardiovascular mortality." *American Journal of Public Health*, 2012, 102(11).
L. Johns, A. Madsen, G. Maduro, R. Zimmerman, **K. Konty**, and E. Begier.

"Usefulness of school absenteeism data for predicting influenza outbreaks, United States." *Emerging Infectious Diseases*, 2012, 18(8):August.
J. Egger, A. Hoen, J. Brownstein, D. Olson, and **K. Konty**.

"The effect of school dismissal on rates of influenza-like illness in New York City schools during the Spring 2009 novel H1N1 outbreak." *Journal of School Health*, 2012, 82:March.
J. Egger, **K. Konty**, E. Wilson, O. Barbot, D. Weiss, A. Karpati, and R. Platt.

"Obesity in K-8 students – New York City, 2006-07 to 2010-11 school years." *MMWR Morb Mortal Wkly Rep.*, December 16, 2011;60(49):1673-1678.
M. Berger, **K. Konty**, S. Day, L. Silver, C. Nonas, B. Kerker, C. Greene, T. Farley, and L. Harr.

"Public health implications of rapid Hepatitis C screening with an oral swab for community-based organizations serving high-risk populations." *American Journal of Public Health*, 2011, 101(11):2151-2155.
A. Drobnik, C. Judd, D. Banach, J. Egger, **K. Konty**, and E. Rude.

"Dual-Frame weighting of RDD And cell phone interviews at the local level." *Proceedings of the American Association of Public Opinion Research*. 2010.
M. Battaglia, D. Eisenhower, S. Immerwahr, and **K. Konty**.

"Monitoring temporal changes in the specificity of an oral HIV test: A novel application for use in postmarketing surveillance." *PLoS One*, 2010, 5(8): e12231.
J. Egger, **K. Konty**, J. Borrelli, J. Commiskey, and S. Blank.

"Case fatality rates based on population estimates of influenza-like illness due to novel H1N1 Influenza: New York City, May-June 2009." *PLoS One*, 2010, 5(7): e11677.
J. Hadler, **K. Konty**, K. McVveigh, A. Fine, D, Eisenhower, B. Kerker, and L. Thorpe.

"Overreporting of deaths from coronary heart disease in New York City hospitals, 2003." *Preventing Chronic Disease*, 2010, 7(3).
R. Agarwal, J. Norton, **K. Konty**, R. Zimmerman, M. Glover, A. Lekiachvili, H. McGruder, A. Malarcher, M. Casper, G. Mensah, and L. Thorpe.

"Outbreak of 2009 pandemic influenza A (H1N1) at a New York City School." *New England Journal of Medicine*, 2009, 361(27):2628-2636.
J. Lessler, N. Reich, and D. Cummings, and **New York City Swine Flu Investigation Team.**

"The severity of pandemic H1N1 influenza in the United States , April-July 2009." *PLoS Medicine*, 2009, Dec;6(12):e1000207.
A. Presanis, D. De Angelis, **New York City Swine Flu Investigation Team**, A. Hagy, C. Reed, S. Riley, B. Cooper, L. Finelli, P. Biedrzycki, and M. Lipsitch.

"A scan statistic for continuous data based on the normal probability model." *International Journal of Health Geographics*, 2009, 8:58.
M. Kulldorff, L. Huang, and **K. Konty**.

"Swine-origin influenza A (H1N1) virus infections in a school — New York City, April 2009." *MMWR Morb Mortal Wkly Rep.*, May 8, 2009;58(17):470-472.
H. Jordan, M. Mosquera, **New York City Swine Flu Investigation Team**, H. Nair, A. France.

"Searching for better flu surveillance? A brief communication arising from Ginsberg et al. Nature 457, 1012-10" *Nature Precedings*, 2009, 2009.3493.1.
D Olson, A. Baer, M. Coletta, L. Deyneka, R Gentry, A Ising, E. Murray, M. Paladini, J. Pendarvis, K. Soetebier, **K. Konty**, J. Schulmann, J. Engel, J. Gunn, R. Rolfs, and F. Mostashari.

"Letter regarding: Piegorsch, W.W., Cutter, S.L., and Hardisty, F. (2007). Benchmark analysis for quantifying urban vulnerability to terrorist incidents. *Risk Analysis*, 27(6), 1411-1425." *Risk Analysis*, 2009, 29(9):1119-1121
K. Konty, J. Egger, B. Kerker, B. Maldin, M. Raphael.

"Population prevalence of diagnosed and undiagnosed HIV and associated high-risk behaviors among New York City adults: The New York City Health and Nutrition Examination Survey, 2004." *AIDS*, 2008, 22(2):281-287.
T. Nguyen, C. Gwynn, S. Kellerman, E. Begier, M. Pfeiffer, **K. Konty**, T. Frieden, L. Torian, and L. Thorpe.

"Effective conversion of initial item nonresponse to questions on age and income in a telephone survey." *Proceedings of the American Association of Public Opinion Research.* 2008.
S. Immerwahr, D. Eisenhower, and **K. Konty**.

"Monitoring the impact of influenza by age: emergency department fever and respiratory complaint surveillance in New York City." *Plos Medicine*, 2007, 4(8):1349-1360.
D. Olson, R. Heffernan, M. Paladini, **K. Konty**, D. Weiss, and F. Mostashari.

"Robust point-pattern inference from spatially censored data." *Environment and Planning A*, 2005, 37(1):141-159.
S. Sweeney and **K. Konty**.

“Population forecasting with nonstationary multiregional growth matrices.” *Geographical Analysis*, 2002, 24(3):289-313.
S. Sweeney and **K. Konty**.

Official Reports

“Disparities in Asthma-related Morbidity among Children in New York City.”
New York City Department of Health and Mental Hygiene: Epi Data Brief, Dec. 2018.

“Methodology updates to the New York City Community Health Survey.”
New York City Department of Health and Mental Hygiene: Epi Research Report, Sept. 2012: 1-12.

“The New York City Community Health Survey Atlas, 2010.”
New York City Department of Health and Mental Hygiene, 2012.

“Child obesity risk: Nutrition and physical activity.”
New York City Department of Health and Mental Hygiene: NYC Vital Signs 10(4), 2011

“Childhood overweight, physical activity and ”screen-time” in New York City”
New York City Department of Health and Mental Hygiene: Epi Data Brief (1), October 2010

“Including New Yorkers who can only be reached by cell phones in the Community Health Survey: Results from the 2008 cell phone pilot survey”
Epi Research Report, May, 2010

“Childhood obesity is a serious helth concern in New York City: Higher level of fitness associated with better academic performance”
NYC Vital Signs 8(1), 2009

“Health among Native Americans in New York City: Summary of the Community Health Survey 2002-2006”
New York City Department of Health and Mental Hygiene, 2008

Invited Conference Talks and Panels

“Developing Use Cases to Tackle Legal Barriers and Make the Value Case for Data Sharing”, *All In For Health Annual Meeting,* Denver, CO September 2018

“Big Cities, Big Data, Big Lessons! Leveraging Multi-Sector Data in Public Health to Address Social Determinants of Health”, *Annual Meeting of the American Public Health Association,* Atlanta, GA November 2017.

“All In: Data for Community Health”, *Robert Wood Johnson Foundation and Academy Health,* Denver, CO April 2017.

"Spatial: the un-conference 2015, Spatial information for human health", *University of California,* Santa Barbara, CA December 2015.

"High-resolution maps of childhood obesity (that are safe fro public release?)" *New York City Department of Health and Mental Hygiene Epi Grand Rounds,* New York, NY October 2015.

"Geo-spatial statistics: Some comments." *NSF Census Research Network,* New York, September 2014.

"Social Observatories: Promise and Perils for 21st Century Data Collection." *Annual Meeting of the American Sociological Association,* San Francisco, August 2014.

"Bridging Statisticians with Public Health Practice Professionals in Disease Surveillance." *2014 Joint Statistics Meetings,* Boston, August 2014.

Next Generation Surveillance for the Next Pandemic. *Santa Fe Institute Workshop,* Santa Fe, NM May 2014.

"Childhood obesity estimates: The effects of new scales." *New York City Department of Health and Mental Hygiene Epi Grand Rounds,* New York, NY April 2014.

"Syndromic surveillance following Hurricane Sandy." *New York City Department of Health and Mental Hygiene Epi Grand Rounds,* New York, NY October 2013.

"Large scale surveys for local purposes: some practical considerations." *NSF Measuring People in Place,* Boulder, CO, October 2012.

"The American Community Survey in public health practice: a local perspective." *National Academy of Sciences,* Washington DC, June 2012.

"Zip code-level estimates for the New York City Community Health Survey." *New York City Department of Health and Mental Hygiene Epi Grand Rounds,* New York, NY March 2011.

"Surveys for public health practice at the neighborhood scale." *UCLA Center for Health Policy Research,* Los Angeles CA, May 2010.

"An efficient method for targeting small areas in health surveys." *National Cancer Institute/California Health Interview Survey Methods Workshop: Maintaining and Enhancing Representativeness of State Health Surveys,* Washington DC, November 2009.

Expert Panel on Geospatial Research, Surveillance and Capacity Building In support of CDC's Healthy Communities Goal, *Centers for Disease Control,* Atlanta, May 2008.

"Health monitoring in New York City: emergency preparedness through better use of public health data systems." *25th Behavioral Risk Factor Surveillance System Conference,* Atlanta, March 2007.

Abstract

Space, Place, and Public Health Surveillance

by

Kevin J. Konty

The last two decades have seen a broadening of the scope of public health that has resulted in a “spatial turn” in a field that is inherently spatial. The increased role for Geography is particularly clear when considering local public health *practice*. Several factors have contributed to this including increased interest in the Social Determinants of Health which are sometimes characterized as attributes of place. This has resulted in a focus on neighborhood or “context” effects on health including attributes of the built environment, demographic characteristics, and environmental measures. Simultaneously, there has been rapid acceleration in the availability and timeliness of data at fine geographic, demographic, and temporal resolution, and an emphasis on open data and inter-sectoral collaboration. This has enabled the timely, systematic characterization of communities or neighborhoods.

This dissertation presents three projects demonstrating the value of geography in public health surveillance in New York City. Surveillance is one of the core functions of public health and has been called the “essential feature of epidemiologic practice.” Traditionally, surveillance has been used for communicable disease outbreak detection and emergency response. However, in the current data context, the systematic, timely, and detailed characterization of populations and communities including health behaviors, outcomes including chronic disease, and social determinants of health is possible and facilitates all aspects of public health practice and policy making. The projects presented here further that goal by developing methods for the detailed characterization of spatial

and demographic patterns.

First, the construction of area-based poverty measures (ABPMs) for surveillance is presented. ABPMs can be used to measure and monitor disparities and track progress towards published goals. Second, a system for the timely characterization of child mental health outcomes is developed. The system can be used to target mental health services and evaluate the impact of ThriveNYC, an extensive preventive mental health program. Finally, a quantile regression framework for child BMI is presented. This has the potential to greatly increase the amount of information used in characterizing childhood obesity including spatial and demographic patterns. Together these projects provide actual use cases for the central role of geography in local government.

Contents

Curriculum Vitae **vii**

Abstract **xv**

List of Tables **xix**

List of Figures **xx**

1 Space, Place, and Public Health Surveillance **1**
1.1 Public Health Surveillance . . . 4
1.2 Setting . . . 6
1.3 Projects and their policy context . . . 6
1.4 Tables and Figures . . . 9

2 Area-based Poverty Measures **11**
2.1 Introduction . . . 13
2.2 Background . . . 14
2.3 Data and Methods . . . 30
2.4 Results . . . 40
2.5 Discussion . . . 45
2.6 Tables and Figures . . . 48

3 Spatial patterns of child mental health burden **59**
3.1 Introduction . . . 61
3.2 Background and Data . . . 64
3.3 Methods . . . 78
3.4 Results . . . 80
3.5 Discussion . . . 85
3.6 Tables and Figures . . . 88

4 Childhood obesity **106**
4.1 Introduction . . . 108
4.2 Background and Data . . . 108
4.3 Methods . . . 112
4.4 Results . . . 115
4.5 Discussion . . . 117
4.6 Tables and Figures . . . 120

5 Conclusion **127**

List of Tables

2.1 Poverty groups by geographic scale, 2013-17 48
2.2 One-year changes in Census Tract poverty designation 49
2.3 One-year changes in poverty designation, Census Tract vs NTA 50
2.4 Five-year changes in poverty designation, Census Tract vs NTA 51

3.1 ICD 9 codes for mental health diagnosis. Note: Codes not included – 315 (Specific delays in development) and 316 (Psychic factors associated with diseases classified elsewhere). 89
3.2 ICD10 codes for mental health 92
3.3 Regular Expressions for syndromic surveillance coding of mental health . 95
3.4 NYC Emergency Department Visits 96

4.1 AIC for alternative model specifications. 120
4.2 Parameter estimates and p-values for AR(2) model with poverty covariate. 120
4.3 AR(2) effects given $t_{i,j} - t_{i,j-1} = 1$ and $t_{i,j} - t_{i,j-2} = 2$. 121
4.4 Parameter estimates and p-values for AR(1) model. 121
4.5 Parameter estimates and p-values for AR(2) model. 121
4.6 Parameter estimates and p-values for AR(1) model with poverty covariate. 122

List of Figures

1.1 Syndromic Surveillance Impact Pyramid 9
1.2 Health Impact Pyramid 10

2.1 Neighborhood tabulation areas 52
2.2 Poverty maps under four different geographies 53
2.3 Coefficient of variation due to sampling error, Census Tracts vs NTAs . 54
2.4 Areas remaining, entering, or leaving poverty under 1-year transitions, Census Tracts vs NTAs 55
2.5 Areas remaining, entering, or leaving poverty under 5-year transitions, Census Tracts vs NTAs 56
2.6 tbd 56
2.7 ABPM-based disparity measure, Tracts vs NTAs 57
2.8 ABPM-based disparity for tracts using 4 different assignment rules . . . 57
2.9 Persistently poor and wealthy census tracts with transitioning tracts . . 58
2.10 Asthma by (NTA household poverty or change in poverty) correlations . 58

3.1 General psychiatric syndrome ED visits , all ages 97
3.2 General psychiatric syndrome ED visits by age group 97
3.3 General psychiatric syndrome ED visits, aged 5-12 98
3.4 ED visit rates school days and non-school days, 2008 and 2016. 99
3.5 Population rates of ED visits for mental health 100
3.6 Population rates of mental health-related ED visits, SPARCS vs SYNDROMIC 101
3.7 Population rate quintiles of mental health-related ED visits, SPARCS vs SYNDROMIC 102
3.8 Proportion of mental health-related visit quintiles, SPARCS vs SYNDROMIC 103
3.9 Area-specific temporal correlations, SPARCS vs SYNDROMIC 104
3.10 Quarter-specific areal rank correlations, SPARCS vs SYNDROMIC . . . 105
3.11 Measured area-based mental health ED visit disparities using population rates, SPARCS vs SYNDROMIC 105

4.1 Visual association between obesity and poverty. 123
4.1.1 Childhood obesity prevalence 123
4.1.2 Extreme poverty prevalence 123
4.2 Fit and validation by age group. 124
4.2.1 Unconditional model. 124
4.2.2 Conditional model, AR(1). 124
4.2.3 Conditional model, AR(2). 124
4.3 Fit by school district. 125
4.3.1 Unconditional model. 125
4.3.2 Conditional model, AR(1). 125
4.3.3 Conditional model, AR(2). 125
4.4 Fit diagnostics for AR(2) with poverty covariate. 126
4.4.1 Fit and validation by age group 126
4.4.2 Extreme poverty prevalence 126

Chapter 1

Space, Place, and Public Health Surveillance

Despite extensive national (CDC) and international (WHO) organizational structures, public health practice essentially occurs at the local level. Core functions of local public health practice including disease surveillance, emergency preparedness and planning, population and community health assessment, data collection efforts such as vital records and disease registries, and the implementation and evaluation of programs and interventions are inherently spatial and encounter geographical considerations often at a very fine scale. As such, geography has always played a role in public health practice.

Three trends have increased the need for geographic perspectives in public health and provided an opportunity for Geography to play a more central role in public health practice. First, the last two decades has seen a broadening of the scope of public health. This may represent a return to an earlier, more empowered, version of public health. Fairchild (2010) quotes a 1916 book entitled *The New Public Health* that presented a vision for public health: "The old public health was concerned with the environment; the new is concerned with the individual. The old sought the sources of infectious disease in the surroundings of man; the new finds them in man himself. The old public health... failed because it sought them... in every place and in everything *where they were not.* (Hill, 1916)." The "New Public Health" greatly narrowed the scope of public health and the role of government in healthcare. However, a (re)new(ed) appreciation of the role that contextual factors (Diez-Roux, 1998) and neighborhood effects (Diez-Roux, 2001) play in health has led to resurgence of Geography in local public health. Central to this is the acknowledgement of the importance of the social determinants of health (SDH). Methodological advances in causation allowing for causal modeling in non-experimental contexts with observational data have allowed for an assessment of SDH's role in disease etiology (Braveman and Gottlieb, 2014). Related to this is an increased interest in health disparities– increasingly viewed as a social justice issue (Marmot et al., 2008).

Second, the last two decades have seen an explosion in the quantity of data available to public health including increased spatial, demographic, and temporal resolution and increased timeliness. This surge of data is particularly pronounced within public health agencies reflecting the recognized authority of government to collect data such as disease registries and vital statistics and the HIPAA regulatory framework which provides an extensive public health exception for the use of such data for public health purposes. The potential for this data to transform public health is recognized but technological developments are constrained by the technical capacity within agencies (Khoury and Ioannidis, 2014). A key feature of efforts to process large administrative datasets within local government is the ability to link records across data systems (Tseliou et al., 2018). Integrated data systems and near real-time data sources will transform public health surveillance providing timely characterization with greater spatial and demographic detail.

Third, health care reform has incentivized health care organization such as hospital and provider systems and payers including private and public insurers to characterize and address the health outcomes of those under care. For example, the Affordable Care

Act requires all non-profit hospitals to conduct Community Heath Needs Assessments (see, for example, (Cottage Health, 2015)) to meet Internal Revenue Service requirements (Pennel et al., 2015). These changes have led to the development of Population Health as a distinct subfield within health and medicine that is confusing to public health practitioners (Diez-Roux, 2016; Gourevitch et al., 2019). Likewise, payment reform has increased the value of addressing health conditions preventively. Incentive structures for reducing repeat emergency department use has resulted in health care organizations proactively acknowledging and addressing the SDH of their customers and linking them to social services.

A by-product of these trends has been an increased need for interagency and cross-sectoral collaboration. Many social determinants of health are under the purview of other health and human services and education agencies, and police departments. Meaningful action to address SDH often falls under planning agencies or directly within mayor's office and city councils. Likewise, healthcare organization, payers, and public health agencies cooperate through Regional Health Information Organizations or Health Information Exchanges. The breadth of the new (old) public health is extensive, with a notion of health that subsumes other agencies' authority. However, it should be noted that other fields have experienced a broadening of scope or softening of boundaries. In Education, the Community School movement (Dryfoos et al., 2005) and the Whole School, Whole Community, Whole Child model (Lewallen et al., 2015) views health, housing, and child welfare as necessary components of education. The increased interdisciplinarity of government presents another opportunity for Geography to use its inherent interdisciplinarity to provide a useful frame to bring together information from disparate sources to characterize place and the people that inhabit it.

One result of this is the increased demand for Spatial Demography (Wachter, 2005) to characterize populations with high levels of demographic detail at fine spatial and temporal scales, to account for the change of these populations through time, and to do so routinely. Administrative records systems at government agencies capture encounters, events, or use of services; they rarely contain information describing the entire population from which the observations are drawn. While this demand is seen in other fields, it is particularly acute in public health practice. For example, public health reporting relies on denominators for the construction of rates; interventions are often targeted at specific geographic areas or demographic subgroups, communicable disease processes are inherently spatial; and emergency preparedness requires highly detailed information.

The importance of spatial concepts and techniques to public health is thoroughly recognized within Geography (McLafferty and Murray, 2017; Spielman and Yoo, 2009) and there has been increased recognition or repackaging from within academic public health (Ostfeld et al., 2005). A goal of this dissertation, is to demonstrate the value of a geographic perspective within local public health practice. To that end, I present three papers examining geographic considerations in public health surveillance in the current data context and recognizing the broadened scope of public health. The three papers serves as proofs of concept in the further development of public health surveillance systems.

1.1 Public Health Surveillance

Last's 2001 Dictionary of Epidemiology defines surveillance as the "systematic ongoing collection, collation, and analysis of data and the timely dissemination of information to those who need to know so that action can be taken" and calls surveillance "the essential feature of epidemiologic practice". While surveillance systems were originally developed for outbreak or aberration detection, their utility has expanded in the current data context. Surveillance information is now used to target programs and interventions, serve as an input in policy formation, provide a framework for evaluation, enable the development of health outcome and health disparity measures, allow for the tracking of public health goals such as Healthy People 2020 or Take Care New York 2024, and characterization of communities such as neighborhoods or demographic subgroups.

Data developments have allowed for more sources of data to be used for surveillance. For example, syndromic surveillance systems based on emergency room patient logs were developed to respond to bioterrorism concerns after the World Trade Center attacks (Heffernan et al., 2004). The system has expanded to include pharmacy and over-the-counter drug sales, ambulance dispatch data, school absenteeism and nurse visit data, and health clinic and urgent care encounter data. More novel data sources have been incorporated into disease surveillance systems including internet search query data and social media such as twitter feeds and yelp reviews for outbreak detection (Althouse et al., 2015). Simultaneously, these same sources have been repurposed to characterize chronic conditions, behavioral health, and other non-communicable diseases (Ayers et al., 2014; Lall et al., 2017). While surveillance has long been a core feature of public health, the current data context has allowed for a broader scope and increased coverage and completeness allowing systems to inform a wider array of public health actions (Thorpe, 2017).

Individual public health surveillance systems generally target a single level within the "surveillance pyramid" (Presanis et al., 2009) (see Figure 1.1). For example, in the case of influenza, a person could be uninfected, infected but asymptomatic, symptomatic but self treated, medically attended, in need of urgent care or an emergency department at a hospital, hospitalized, or dead. As you move the up the pyramid there are generally fewer cases observed. In two of the examples presented here we surveil the emergency department level for mental health and asthma-related visits. Since neither is communicable, our goal is to characterize the disease in a timely manner, hopefully at a level of detail that informs public health decision-making. The last example concerns surveillance of childhood obesity. This is essentially at the bottom two levels of the pyramid where we are estimating the prevalence of obesity and severe obesity conditions.

1.2 Setting

The setting for this dissertation is the New York City Department of Health and Mental Hygiene (NYC DOHMH) and the New York City Office of School Health (OSH). NYC DOHMH is the oldest and one of the largest public health entities in the U.S. NYC DOHMH has long been considered a progressive public health department that is empowered by New York State law allowing NYC DOHMH to create laws and regulations through the board of health. NYC DOHMH developed the first syndromic surveillance systems and established the International Society for Disease Surveillance. The Office of School Health is a joint program of the NYC DOHMH and the New York City Department of Education. OSH has over 2000 employees including all school nurses, school doctors, mental health consultants and programs addressing asthma, diabetes, vision, oral health, mental health, reproductive health, health services and accommodations, and vaccine exemptions. OSH is also responsible for health and physical education policy and the collection of physical fitness and biometric data.

1.3 Projects and their policy context

The three projects presented here are directly applicable to current Office of School Health initiatives, and, more broadly to NYC DOHMH public health practice. They were chosen for two reasons: to serve as proofs of concept for methods that enhance surveillance efforts and to directly inform local policy once they have been implemented. The second chapter, **Area-based poverty measures in public health practice** reviews the development of area based poverty measures in New York City including choice of spatial scales, selection of appropriate input data, and updating rules. Current work on ABPMs at NYC DOHMH has only considered the retrospective one-time use of ABPMs to characterize trends. The chapter assesses current data policies as it applies to the prospective or repeated use of ABPMs with existing surveillance systems. Systematic updating of ABPMs allows for their use as context-level variables in hierarchical models of health outcomes. Further, it is possible to use a similar framework for the development of other SDHs with the goal of establishing a surveillance system of SDH themselves. This could facilitate the use of SDH in public health, allow for the assessment of SDH measures, allowing for the simultaneous tracking of change in SDH measures with change in health outcomes. The chapter characterizes trends in asthma disparities by ABPM, a current focus of NYC DOHMH programs.

The third chapter **Spatial patterns of child mental health burden: Measurement and monitoring of child mental health burden** uses two data sources to characterize spatial and temporal patterns of mental health-related emergency department (ED) visits. The first source, SPARCs, contains the diagnostic codes typically used for measuring trends or disparities in ED usage. SPARCs data has a lengthy time lag making it not viable for prospective use. The second data source, syndromic surveillance, is near real-time but generally does not have diagnostic codes, Instead chief complaints are processed and syndromes are assigned. If the syndromic information accurately reflects spatial patterns of mental health then it could be used prospectively or to characterize their recent past. The spatial patterns can be used to target interventions or efforts to increase access to mental health services. They can also be used to inform allocation of resources for the school mental health consultant program, a key initiative of the Mayor's ThriveNYC program. Additionally, since prevention of acute mental health events is a goal of ThriveNYC, accurate and timely characterization of emergency department visits can be used as an outcome measure to assess impact relative to resource allocation. We also characterize patterns of mental health ED visits during school-hours on school-days and construct disparities. ThriveNYC has a need for timely characterization of potential impacts, particularly when deciding on the renewal of extremely expensive initiatives.

The fourth chapter **Child obesity: Longitudinal information, local government, and public policy** describes the development of a system to monitor childhood obesity at fine spatial and demographics scales using quantile regression and to use the system to evaluate public policy. The NYC FITNESSGRAM includes approximately 900,000 unique measurements of height and weight per year in kindergarten through 12th grade. The longitudinal information describes individual growth trajectories. Detailed home and school information allow for analysis at fine spatial resolution allowing for the characterization of New York City neighborhoods and the investigation of community-level effects. This approach increases the level of information available to evaluate programs and policies, and to assess the relationship of BMI to the built environment. Childhood obesity rates are closely watched with respect to policy but current methods based on repeated cross-sections are inadequate for evaluation purposes. Using longitudinal growth trajectories as an input, quantile regression allows one to empirically reproduce growth charts (by including only age effects), and by incorporating lags and school, place, or program specific variables it provides a powerful framework for both policy evaluation and data quality assessment. The chapter describes several distinct advantages over current approaches to policy evaluation based on cross-sectional framing using CDC growth charts. Adoption of the methods with their use of growth trajectories would return surveillance of growth to its original form in the 19th century when Quetelet developed BMI while monitoring individual growth (Tanner and Tanner, 1981).

1.4 Tables and Figures

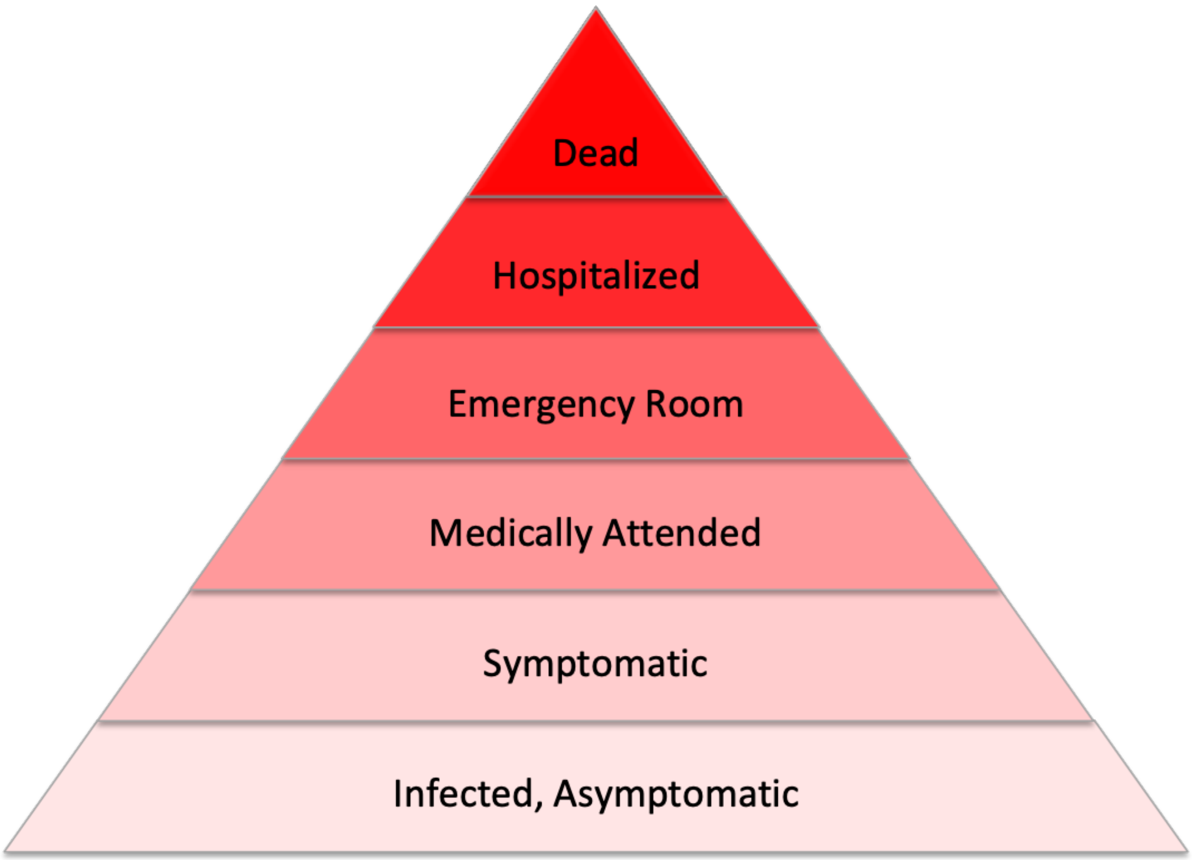


Figure 1.1: Syndromic Surveillance Impact Pyramid

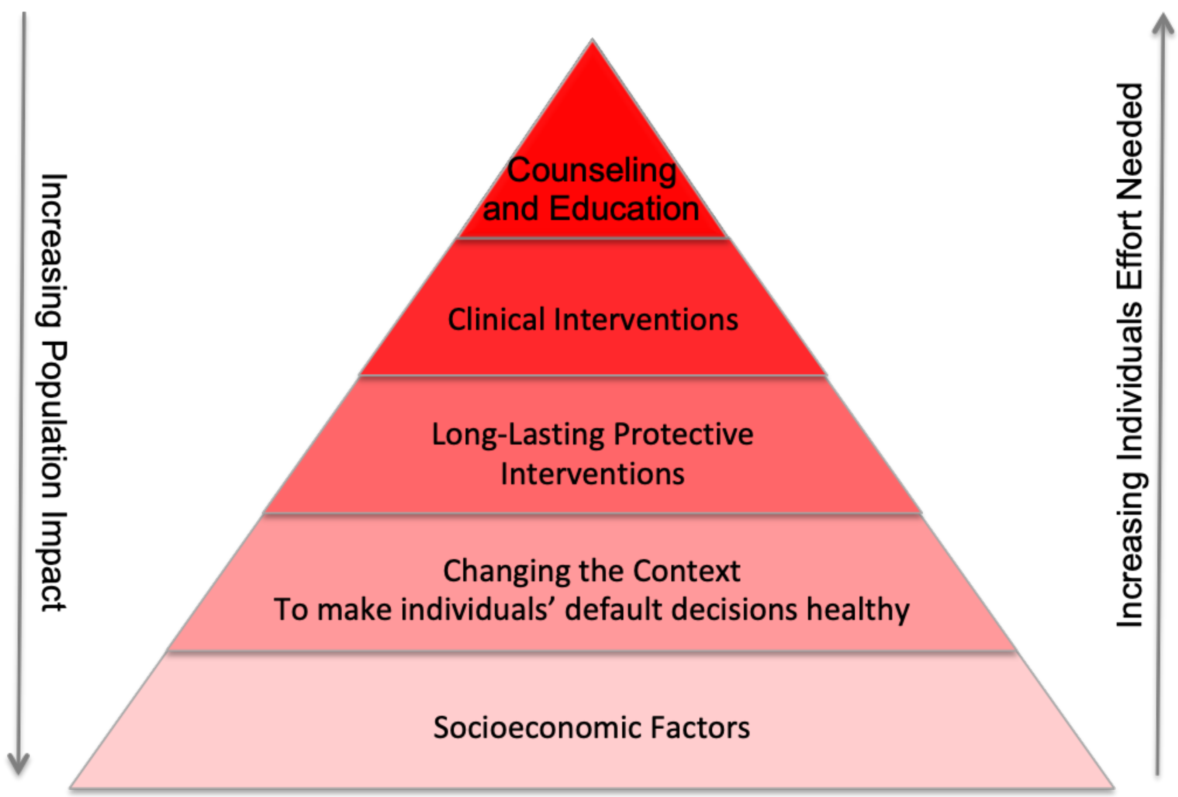


Figure 1.2: Health Impact Pyramid

Chapter 2

Area-based poverty measures in public health practice

Abstract

Area-based poverty measures (ABPMs) and other ecological measures capturing social determinants of health are increasingly used in public health practice. Area-based measures are particularly attractive for use with administrative records that lack high quality individual-level socio-economic or demographic information. In public health practice, this includes legally mandated disease reporting, disease and immunization registries, hospital claims data, and near real-time event data provided by hospitals, emergency medical services (ambulance), and pharmacies. Such information is now processed in near real-time by public health departments and serves as a key input to public health surveillance systems (Thorpe, 2017). In these contexts, ABPMs can be thought of as an imputation of underlying missing individual-level socio-economic data. However, another important use of ABPMs is to characterize the context in which individuals live. This provides information beyond individual-level variables and, importantly, allows for estimation of hierarchical models linking social determinants of health and outcomes. In

both cases, ABPMs provide a framework for the measuring and monitoring of health disparities.

In this paper, we review efforts by the New York City Department of Health and Mental Hygiene (NYC DOHMHM) to establish area based measures for use in public health surveillance and community health assessment. Specifically, we review the initial data policy established by NYC DOHMH for the use of ABPMs in public health practice and the broader academic and CDC literature upon which they are based. The approach we employ is entirely practical; focused on a concept/measure approach to system development. We attempt to establish a measure that reflects the underlying poverty concept. We argue that any good measure must balance the competing objectives of the two distinct viewpoints of the "area-based poverty" concept. Lastly, it is clear that any data policy has to be feasible (implementable) with extant data and capable of systematic updating in a transparent manner.

keywords: Area Based Poverty measures, syndromic surveillance, childhood asthma, health data policy, public health surveillance, population health

2.1 Introduction

Surveillance is a central activity of public health practice. Surveillance systems provide input for policy formation, enable the targeting of programs and interventions, facilitate the evaluation of interventions, allow for the measurement and monitoring of disparities, increase situational awareness during emergencies, and provide transparency allowing for public reporting of the impact of publicly funded programs (Thorpe, 2017). Data available for surveillance systems include vital statistics, disease and other health registries, administrative records such as claims data and electronic health records, novel data sources such as social network and search data, and health surveys (Althouse et al., 2015; Birkhead et al., 2015; Heffernan et al., 2004).

Many of these sources do not have access to socio-economic or other data addressing the social determinants of health (Krieger, 1992). Further, increased interest in social determinants and health disparities has increased interest in characterizing the context in which individuals live even when individual-level information is available. Area-based poverty measures (ABPMs) and other ecological measures capturing social determinants of health are increasingly used in public health surveillance and have been proposed to address these issues. This paper addresses a practical problem in the construction of public health surveillance systems. Namely, we examine construction of ABPMs including the assessment of potential data sources, the selection of specific measures and geographic scales, and the development of systems to routinize reporting. Recognizing that the resulting ABPMs will be used to publicly report on change in health outcomes, feasibility, validity, and transparency are key criteria.

Specifically, we review the initial data policy established by the New York City Department of Health and Mental Hygiene (NYC DOHMHM) for the use of ABPMs in public health practice. The review serves as an assessment of the policy's capacity to enhance NYC DOHMH's current surveillance systems (Toprani and Hadler, 2016). We argue that establishing a single measure must acknowledge the competing objectives of the two distinct viewpoints of the "area-based poverty" concept– 1) the use of ABPMs in the absence of socio-economic data and 2) the use of ABPMs to describe the context in which individuals live. Although, we focus on details of the currently proposed measure our discussion is relevant to other attempts to establish area-based measures for use in public health practice or, more broadly, efforts by local governments to implement policies and programs across agencies.

The next section discusses the relevant policy context including various concepts touched upon in the development of ABPMs including social determinants, health disparities, poverty and other area-based measures, geography and scale, and the notion of neighborhoods. We then describe the current data policy, input data sources, and surveillance systems that will use ABPMs. We then assess possible alternative specifications focusing on childhood asthma in New York City and propose a specific use in public health surveillance.

2.2 Background

The last two decades has seen a broadening of the scope of public health. Some have argued that this represents a return to an earlier vision of public health (Fairchild et al., 2010) that existed prior to the rise of an individual-focused, clinical view of health and disease that existed for most of the past century and that resulted in public health playing a reduced role in service to clinical medicine often acting as a referral service (Colgrove et al., 2008, pages 5-9). Diez-Roux links this to the advent of germ theory and a "unicausal" theory of disease that perpetuated the view that risk is individually determined (Diez-Roux, 1998) and, as such, to be addressed clinically. The rise of chronic diseases over the last several decades led to multimodal views of causation that include health behaviors but a prevailing ”methodological individualism” characterized such behaviors as individual choice ”disassociated from the social constructs that shape and constrain them” (Diez-Roux, 1998). Nonetheless, further development of multimodal causal models supported a more central role for social context in disease etiology (Braveman and Gottlieb, 2014).

2.2.1 Social determinants of Health

The health-related characteristics of social context are termed the social determinants of health (SDH) and are defined as “the conditions in which people are born, grow, live, work and age” and “the fundamental drivers of these conditions.” (Braveman and Gottlieb, 2014; WHO Comm of Soc Det of Health, 2008). SDH include broad socioeconomic factors such as wealth, income, and education status and evidence has accumulated that these are causes of a wide variety of health outcomes (Braveman and Gottlieb, 2014; WHO Comm of Soc Det of Health, 2008). Although SDH can be thought of as the individual characteristics that describe these “conditions,” they are often thought of collectively. For example, the Centers for Disease Control’s guidance on incorporating SDH’s into public health practice defines SDH as the “conditions in the places where people live, learn, work, and play” (Centers for Disease Control and Prevention, 2019).

Although it is common for research papers on SDH to refer to their importance as a recent development in public health occurring over last two decades (Graham, 2004; Braveman et al., 2011), SDH have been recognized far longer. In 1920, for example, Winslow discussed SDH as one of the “untilled fields of public health,” arguing that addressing social determinants is a key function of public health, and even recognizing the difficulties of establishing causation. “No one can perhaps tell just how far poverty in such cases is the real and effective cause of the failure to achieve and maintain a normal standard of physical health. It is clear, however, that there is a certain standard of income below which the maintenance of health is impossible" (Winslow, 1920). Winslow argues for public health that directly addresses this. John Graunt's 1665 "Observations on the Bills of Mortality" gives an earlier example. Chapter 10, entitled "of the inequalities of parishes," concerns variation in sizes of parishes as opposed to disparities in death rates but then observes that "whereas now in the greater Out-Parishes many of the poorer Parishioners through neglect do perish." In both cases, it is implied that action can be taken to alleviate this.

Acknowledging the importance of SDH in disease etiology is particularly important for public health departments embedded in local governments that implement policies directly influencing local environment across multiple domains including zoning, transportation planning, education, housing, law enforcement, and social services. Recognizing the importance of SDH in disease etiology increases the importance of public health practice in health and medicine while emphasizing the importance of inter-agency cooperation (Braveman et al., 2011). Likewise, a role for local government addressing fundamental causes of disease through SDH raises important issues of equity (Braveman and Gottlieb, 2014).

2.2.2 Health Disparities

It has long been established that for a variety of reasons, health conditions covary with SES and other SDH (Barr, 2014). Further, the past two decades have seen a shift in the focus of public health to include addressing these health disparities. For example, the United States Department of Health and Human Services Healthy People 2010, the organizing framework of national public health efforts, had "eliminate health disparities" as one of two overarching goals (US Department of Health and Human Services, 2000). Healthy People 2020 amended this to “Achieve health equity, eliminate health disparities, and improve health for all” (US Department of Health and Human Services, 2000). A key aspect of the Healthy People program is measurement of progress and, as such, monitoring of disparities has become a core function of public health.

The centrality of health disparities in public health practice is recognized from the local to global scale. *Take Care New York*, the “roadmap” for public health in New York City, sets goals for leading health indicators and establishes an overall target and an equity target.(Mettey et al., 2015). The World Health Organization’s Commission on Social Determinants of Health called for achieving health equity within a generation (WHO Comm of Soc Det of Health, 2008). The accompanying editorial argues for viewing health equity as a social justice or humans rights issue (Marmot et al., 2008). This view is key to understanding their current importance in practice. An early definition described health disparities as differences in health that “are not only unnecessary and avoidable but, in addition, are considered unfair and unjust” (Whitehead, 1990). This directly suggests a role for government which is responsible for delivering justice. Braveman’s 2006 study of the health disparities makes this implicit by suggesting a definition focused on policy– a “‘health disparity/inequality’ is a particular type of potentially avoidable difference in health or important influences on health that can be shaped by policies”. Choosing specific disparities has, “...important policy implications with practical consequences. It can determine not only which measurements are monitored by national, state/provincial, and local governments and international agencies, but also which activities will receive support” (Braveman, 2006).

Like the social determinants of health, health disparities can be constructed using individual characteristics without reference to place. Nonetheless, disparities are often constructed using geographic units. For example, the 2008 WHO report on SDH highlighted a 28 year difference in life expectancy between two neighborhoods in Glasgow, Galton (54 years) and Lenzie (82 years), 12 kilometers apart (WHO Comm of Soc Det of Health, 2008). The same WHO table reports a 17 year life expectancy between Washington D.C. (63) and Montgomery County, Maryland (80) (WHO Comm of Soc Det of Health, 2008). Likewise the 11-year range in neighborhood life expectancy in New York City is annually updated (Li et al., 2018) and widely reported (Tavernise and Sun, 2015).

Geographically-based health disparities can be approached in three distinct ways. Like the life-expectancy example above, health outcomes can be measured for the geographical units within a specific area and comparisons can be made between the top and bottom areas. This actually makes no explicit reference to SDH, but rather is interpreted after the fact by reference to SDH in these areas. For any measure, there will always be areas with the highest and lowest rates so determining how much of the difference is unjust or can be addressed by policy is difficult. It has been noted that many impoverished neighborhoods in other United Kingdom cities besides Glasgow have higher life expectancies and smaller discrepancies leading some to refer to the disparity in the WHO report as the "Glasgow Effect" (Reid, 2011). Here, the disparity in Glasgow meets the definition of a disparity by comparison to other cities but, conversely, smaller discrepancies in other cities could use Glasgow to argue that their unequal rates are not true disparities.

A second way to approach disparities through areal differences is ecological analysis in which health outcomes and SDH measures are both estimated over specific geographies and the relationship between them is examined at that scale. This approach attempts to establish how much of an areal-scale difference may be due to the SDH measure. This has raised concerns such as the ecological fallacy (Piantadosi et al., 1988), the modifiable areal unit problem (Fotheringham and Wong, 1991), and other issues of applying non-spatial approaches to spatial data (Spielman and Yoo, 2009). Other research has suggested that the ecological fallacy (and MAUP) concerns are overstated and based on the false premise that all ecological analyses are attempts to establish individual relationships (Macintyre and Ellaway, 2000). The use of areal measures to establish context is certainly an instance when this is not true, especially considering that these measures often appear in hierarchical models alongside individual measures (Diez-Roux, 1998). Nonetheless, monitoring the relationship between SDH measures and health outcomes using areal measures may be subject to these issues. Further, it is not clear how to monitor the full relationship between the spatial distributions of health outcomes and SDH to characterize trends in health disparities for use in surveillance. This approach though can yield valuable insights about the SDH, health outcome relationship and can provide a framework for estimating the potential benefits of policy. For example, in New York City the ecological approach has been used to roughly estimate potential impact of raising the minimum wage on premature mortality (Tsao et al., 2016) and to establish the strength of the relationship between health outcomes and rates of incarceration by neighborhood (Reilly et al., 2019).

The third approach, taken here, is to establish area-based measures *a priori* and identify the geographic areas representing advantaged and disadvantaged communities. Trends in health outcomes for the two groups are then monitored as the disparity. This has many practical advantages:

- individuals are assigned to categories analogous to health disparities constructed at individual level such as racial disparities
- context information is provided beyond individual-level factors
- assignment of areas to categories can be monitored
- it is transparent in that areas of interest for a specific SDH are determined prior

to estimation of health outcomes– grouping areas into categories allow for public release of data that is often not possible when reporting individual areas,

- estimates are more stable because they are based on larger samples.

The approach is also consistent with the underlying notion of health disparities being based on the "...clear (albeit usually implicit) assumption that the relevant differences are those between better- and worse-off social groups selected *a priori* based on who historically has been more and less advantaged in a society." (Braveman, 2006).

This approach was developed by Krieger as the "Public Health Disparities Geocoding Project" and motivated by the lack of individual information on administrative records (Krieger, 1992; Krieger et al., 2002, 2003). The initial output of that project was variables for attachment to individual records for processing into disparities in the absence of more direct approaches. The fact that this entailed an assignment of spatial units into categories was only a consideration in that different choices of spatial units yielded different categories (Krieger et al., 2005). Although, Krieger addressed several questions in the development of the ABPMs including identifying candidate variables and spatial scales for measurement, the exercise was severely constrained. Potential measures had to exist and be reported at the scales. Relying on Census long form data meant that monitoring the resulting assignments was not addressed other than the suggestion of updating measures every 10 years. Krieger identified the optimal measure and scale as the one yielding the greatest disparities over example health outcomes, suggesting a four or six category poverty-based measure assigned at the census tract scale (Krieger et al., 2003). Over several years the resulting ABPMs were applied to various health and social outcomes (Krieger et al., 2005, 2003).

The motivation, data context, and approach to the ABPMs developed in the Health Disparities Geocoding project differ from this paper in that Krieger was essentially imputing individual-level SES values for use with medical records, claims, and other administrative data that lacked individual measures (Krieger et al., 1997). Here our approach is explicitly spatial, we are attempting to characterize geographic areas and communities by measuring and monitoring SDH, health outcomes, disparities, and the relationship among these. That is, we are recognizing the role of context and attempting to establish its role in health (Diez-Roux, 2001; Diez Roux, 2004). Nonetheless, despite this difference in motivation, the issues are the same as in Krieger including identifying data sources, choosing specific SDH measures, and identifying relevant spatial scales. Additionally, we need to ensure that both health outcomes and SDH measures can be routinely updated so that the changes in each can be compared through time (Diez Roux, 2004). For SDH measures that will generally be taken from external demographic sources such as the Census Bureau, this relies on ongoing efforts by external agencies to produce estimates at appropriate scales.

2.2.3 Poverty and Other Area-based Measures

Poverty measures have played an important role in public policy for decades (Fisher, 1992). Although it is widely recognized that poverty measures often fail to capture what they intend to capture and that they may not achieve comparability between places, a large swath of federal policy and funding is based on such measures (Krampner et al., 2017). In New York City, measuring poverty with respect to underlying individual conditions has a long history. The current federal poverty level was developed in New York by a former New York City Department of Health employee (Orshansky, 1965; Fisher, 1992), while the market-basket approached used in its development was developed in New York in the early 1900s (More, 1907).

Although the approach, to define a poverty-level based on the cost of specific basket of goods given a specific household configuration, is widely used it suffers from scale (or spatial focus) issues. In the United States, underlying costs were taken as national averages and annual adjustments are equivalent regardless of location. This results in a poverty threshold that actually reflects different levels of deprivation depending on location within the United States. Further, measurement of income is problematic, ignoring accumulated wealth, tax credits, and the value of social programs that address poverty, housing and other social determinants. Recognizing these issues, New York City developed new poverty measures that account for these issues (Krampner et al., 2017). The new measures, however, require more detailed individual-level data and, as such, are more difficult to regularly update than current poverty measures and require a lengthy survey module or extensive administrative record linkage. Even aggregated across spatial units, the new measures would require a dedicated effort to maintain estimates and establish trends.

There is an extensive apparatus for producing federal poverty measures reflecting their importance in allocating federal funds and establishing eligibility for federal, state and local programs. For example, federal funding for the National School Lunch Program has been partially based on income thresholds since the program's inception in 1946 eventually making use of the federal poverty line (Gunderson, 2013). Current federal poverty measures are based on the Current Population Survey's Annual Social and Economic Supplement (U.S. Census Bureau, 2007) and the Small Area Income and Poverty Estimates which are derived from the American Community Survey (U.S. Census Bureau, 2014, 2017). Between them, consistently updated aggregate poverty estimates are produced with high temporal and spatial granularity. Further, the definitions used to generate the estimates can be implemented at the individual-level with just a few questions allowing for their use in health surveys.

2.2.4 Area-Based Measures

Because of their timeliness and high spatial and temporal resolution, ABPMs are an attractive option for the use in public health surveillance systems. Because of the limitations of poverty measures, and with the recognition of a broad array of social determinants of health, many other area-based measures have been proposed. These include measures based on other approaches to poverty (Subramanian and Kranes, 2015), segregation (Reardon and O'Sullivan, 2004), income or wealth inequality (Krieger et al., 2016), gentrification (Austensen et al., 2016), racial and demographic composition (Krieger et al., 2002; Delmelle, 2019), community loss (Albrecht and Abramovitz, 2014), and other directly measured social determinants such as reliance on public housing (Yim et al., 2018) or community rates of incarceration (Reilly et al., 2019).

The goal of this project is to develop a framework to use area-based measures to measure and monitor health disparities. For any candidate measure, geographic areas representing least- and most-disadvantaged groups are identified. Typically, a continuous measure is estimated at some geographic scale, the areas are binned into a small number of groups, and the two end groups are used to construct the disparity. For nominal measures such as neighborhood typologies (Delmelle, 2019), types are selected to represent these groups. This is similar to measuring, for example, the black/white health disparity. Intermediate steps include establishing the scale, deciding upon a continuous measure, and placing areas in a discrete number of groups. In addition to monitoring disparities, these measures can also be used for both area-based imputation and to provide additional context-level variables in hierarchical models. Surveillance systems that include area-based social determinant measures would store this information for these types of uses

2.2.5 Geographical Considerations

Several geographic considerations have been raised about area-based measures. These issues have been discussed from both the epidemiology/public health (Diez-Roux, 1998, 2001; Diez Roux, 2004) and geography perspectives spielman2009, spielman2012, kwan2012. The epidemiological viewpoint is focused on the validity of ABPMs for individual-level imputation or establishing the causal effects of neighborhood characteristics on health beyond individual characteristics. These efforts are fundamental to establishing the role social determinants of health, especially when they are considered as characteristics of place. Establishing causal effects generally requires longitudinal data; this emphasizes the need to be able monitor ABPMs to characterize change through time.

Geographic concerns focus more on issues of aggregation, scale, and ecological inference. Key geographic concerns include the modifiable area unit problem (MAUP) (Fotheringham and Wong, 1991) and the uncertain geographic context problem (UGCP) (Kwan, 2012). MAUP is concerned with the fact that different aggregations to a given scale can yield different correlations; or even under a fixed scale may yield different correlations under rotation of the lattice. The key point is that the areal boundaries function as an aggregator and the resulting correlations are dependent on where they are placed. UGCP states the important idea that if a phenomenon exists in space, such as a "neighborhood effect" or a place-based view of a "social determinant of health," it is not clear at what scale the phenomenon should be measured. A third related concern in using ABPMs is the ecological fallacy (Piantadosi et al., 1988) which warns that ecological relationships do not necessarily reflect underlying individual measures. Fourth, it has been shown that several area-based measures may display high-levels of scale dependence. Reardon has demonstrated that several widely used spatial segregation measures can vary widely as scales change (Reardon and O'Sullivan, 2004). Many other proposed measures are similarly scale dependent. For example, the Index of Concentration at the Extremes (ICE) (Massey, 2001) which has recently been proposed for public health monitoring (Krieger et al., 2016), can vary from one extreme to the other within nested geographies.

2.2.6 Neighborhoods

Acknowledging these geographic considerations, one possible way forward is to identify a relevant spatial scale and specific spatial configuration at that scale. It seems clear that the mechanisms by which context or place impact health outcomes, must come from an individual's interaction with and experience of that context. The concept of neighborhood captures in general the notion of a locally relevant spatial unit. Because of this, research on SDHs and spatial effects often conflate "neighborhood" with the entire effort. For instance, Duncan begins the 2nd edition of the book *Neighborhoods and Health* with, "The field of neighborhoods (sometimes referred to spatial epidemiology)..." (Duncan and Kawachi, 2018), while Diez-Roux states "...only recently have health researchers focused on investigating how spatial contexts, or more specifically neighborhood and community-level factors affect the health of residents." (Diez-Roux and Mair, 2010). Diez-Roux's reference to "community-level factors" and later to "and residential areas more broadly" serves more to acknowledge that neighborhoods are difficult to identify and that data is not always available at a neighborhood-scale than it is to suggest that there is a separate spatial concept (geographic context) that might at times be more appropriate.

It is widely recognized that the definition of the term neighborhood can be vague and imprecise (Spielman and Yoo, 2009) and so neighborhood definitions themselves can be seen as uncertain and malleable. Some recent work that directly focuses on the concept of neighborhood (Spielman et al., 2013; Patricios, 2002) makes reference to sociological work by Park and Burgess in the 1920s (Park and Burgess, 1925) or the urban planning notions of Perry (Perry, 1929). In the former, neighborhoods are "natural areas," "cultural areas with local sentiments and traditions," with separate "ecological, cultural, and political" dimensions. Perry, acknowledging this, identified the urban form that produces such areas including "elementary schools, small parks and playgrounds, and local shops." Other recent work does not bother to discuss historical notions at all (Bernard et al., 2007; Diez-Roux and Mair, 2010). Here, neighborhoods are simply the local areas that impact a person's health. With goal of identifying such areas, Bernard 2007 describes neighborhoods as "environments for accessing resources" and identifies five 'domains' that define neighborhoods: physical, economic, institutional, community organization, and local sociability;" essentially arriving at Park and Burgess.

One reason researchers might not formally define or recognize historical concepts of neighborhood, even when they are themselves presenting a conceptualization, is that the term neighborhood carries a common-sense definition that has existed since the 15th century that already plainly includes the dimensions identified by the Chicago school. In looking at neighborhood effects of health, neighborhood could be considered the area that impacts health through an individual's interaction and experience of it. This is why it could be incorrectly considered all of 'spatial epidemiology,' as it is in Duncan 2018. Defined this way, neighborhoods are clearly dynamic and neighborhood definitions/designations and neighborhood effects should be expected to vary by individual and change through time. This greatly complicates measuring neighborhood effects on health and is a key consideration when attempting to develop public surveillance systems.

From an epidemiological and research viewpoint, there are three possible ways forward: 1) use pre-defined neighborhood definitions deferring to historic, official, or administrative definitions or to previous research, 2) create health-specific regionalizations and use these to identify neighborhoods, or 3) define individual-level neighborhoods based on actual human-environment interaction. Most public health research and reporting use established neighborhood definitions reflecting its position within or need to interact with local government. It is important to recognize that local government administration routinely organizes itself, forms policy, and publicly reports using area definitions often based on historic notions of neighborhood. Local public policy reinforces these definitions through zoning, school assignment, identification of administrative areas such as policy precincts, and provision of resources. Further, in lieu of alternatives, federal official statistics at the local scale defers to such definitions as well. For example, the Census Bureau's Public Use Microdata Areas (PUMAS) are defined to match New York City Community Districts to the extent possible given constraints (Bradley et al., 2016).

Nonetheless, recognizing neighborhood and population dynamics, data-driven regionalizations have been proposed that identify areas by optimizing over pairwise correlations within a set of target SDH and health outcome measures (Spielman and Logan, 2013). Such algorithms allow for control of scale and variation of size and the resulting designations create functional neighborhoods. This also addresses known issues with using the American Community Survey at high spatial resolution (Spielman et al., 2014; Spielman and Folch, 2015; Sperling, 2012), discussed further below. However, these approaches ignore administratively-defined neighborhood designations potentially limiting the availability of auxiliary or outcome data, are dependent on the target measures, and may result in changing areas across the input sources limiting their utility for surveillance of SDH.

A third way forward is to use an ego-centric approach to neighborhood definitions based on an activity space concept (Hägerstraand, 1970). Directly measuring an individuals interaction with their environment perhaps best addresses the mechanism by which neighborhoods impact health. This approach has become increasingly popular in epidemiological work on neighborhood health effects (Duncan et al., 2014; Chaix, 2018) and has featured a repackaging of the activity space as it applies to health effects and time geography more generally as the "healthscape" (Rainham et al., 2010). Some of the research using the activity space concept is concerned with non-neighborhood spatial effects as opposed to a new individual-based neighborhood concept (Hurvitz and Moudon, 2012). Further, use of activity space data is often processed in reference to area-based SDH measures, such as the proportion of time spent in a high-poverty or low-walkability "neighborhood" which themselves need to be defined (Duncan et al., 2016). As such, publicly available public health surveillance system of social determinants of health will contribute to the development of these approaches.

2.2.7 Local public policy

The overarching goal of this study is to review the actual implementation of area-based measures in New York City to inform the formation of local public policy. To that end, the study presents efforts to establish a public health surveillance system for social determinants of health; to use the surveillance system to monitor disparities, evaluate interventions to address SDH and health outcomes by addressing SDH; to further establish the relationship of SDH to health outcomes; to transparently report this to the public; and to facilitate use of the system by researchers and community-based organizations. Because the focus is on construction of a public health surveillance system, several criteria must be met. These criteria include: ensuring the existence of input data, establishing the ability to regularly update the measures and produce health outcome and other data at the identified scale, identifying algorithms to process inputs and update estimates, and establishing rules for use.

This study considers two separate initiatives – or local policies – of the New York City Department of Health and Mental Hygiene (NYC DOHMH) both motivated by growing interest in the Social Determinants of Health. The first was the creation and implementation of an area-based poverty measure within New York City (Toprani and Hadler, 2016) following the Health Disparities Geocoding Project (Krieger et al., 2002). Like Krieger, initial guidelines for area-based measures in New York City focused on research use. The resulting data policy led to several issues when attempting to use the measure with existing surveillance systems or using it to establish an SDH-specific surveillance system. The second initiative was a Robert Wood Johnson-funded Data Across Sectors for Health (DASH) project focused on identifying a geography to integrate data from multiple city agencies and local health care organizations (RWJF Data for Health Advisory Committee, 2015). The second effort directly addresses a key issue in working with SDHs: that SDHs include measures not traditionally under the purview of health departments such as social support, incarceration, housing, and children's services.

The next section presents input data sources for both SDH and outcomes, identifies candidate geographies including the Neighborhood Tabulation Area used in the DASH project, introduces potential processing rules to establish change through time of ABPMs and to measure and monitor disparities, and describes some approaches to comparing candidate geographies and processing rules. These comparisons are presented in following section, as well as a running example using childhood asthma data. Ultimately, the recommended approach arises from practical considerations rather than from a formalism such as an optimization model. A primary consideration is feasibility– that the recommended approach can be implemented prospectively. The last section discusses potential use of the surveillance systems and possible modifications that may come from future research.

2.3 Data and Methods

2.3.1 Current Data Policy

Initial work on area-based poverty measures in New York City followed Krieger 2002. Using census long-form data, Krieger considered several candidate SES measures and geographic scales ultimately choosing the scale and measure that resulted in the largest measured disparity. Krieger found the clearest disparities using an area-based poverty measure based on classifying geographic areas by the percentage of households into a small number of groups and comparing the least to most impoverished. Disparities were most pronounced at the finest geographic scale possible, the census tract, and Krieger recommended using this scale when possible.

Adapting Krieger's method resulted in slightly modified ABPMs with different thresholds for classification reflecting a different distribution in New York City as compared to Massachusetts. A description of the effort to establish ABPMs for New York City was released as an official report (Toprani and Hadler, 2016), and initial results using the measure were reported (Toprani et al., 2016). Two sets of thresholds were established to create two variables, *Poverty4* and *Poverty6*. *Poverty4* assigns geographic areas based on the percentage of households living in poverty using the cut-points (10%, 20%, 40%). *Poverty6* divides the lowest and highest poverty areas into two additional groups. In both cases disparities are measured by comparing the lowest to highest poverty areas.

Public health use of ABPMs to monitor disparities generally require three inputs, the assignment of geographic area to poverty groups, numerator data at that scale over a specified time period, and population data at that scale over the same time period. Initial results focused on measured disparities in 1990, 2000, and 2010 using death records as numerators and taking denominators directly from the corresponding decennial census. As in Krieger 2002, the 1990 and 2000 ABPM designations were derived from Census long form data at various geographic scales. 2010 categories were established using the American Community Survey 5-year sample for 2008-2012. Recommendations for prospective use of ABPMs based on these efforts specify using the ACS 5-year estimates centered on the year of interest. In prospective surveillance, this isn't always possible due to the release schedule of the ACS 5-year estimates. For example the 2013-2017 ACS 5-year estimates became available in June 2018 and, as such, 2016, 2017,and 2018 estimates could not be based on the ACS designation centered on those years. 2016 estimates would obtain a centering ACS designation in June 2019. Long-term trends faced the additional issue that for 2001-2007, no ACS is available. For these years, use of the designation based on 2000 long form data was recommended.

Lastly, use of Census Tracts was recommended, whenever possible, based on the sharpness of measured disparities as compared to coarser geographies. It isn't always feasible to produce census tract estimates due to the additional inputs required. For example, numerators that are based on near real-time surveillance systems, such as the NYC syndromic surveillance system (Heffernan et al., 2004), or annual telephone surveys, such as the New York City Community Health Survey have zip code as the finest geographic scale. For syndromic surveillance, this reflects the need to abstract geographic information from emergency room intake forms in a standardized manner across hospitals in near real-time. For telephone surveys, this reflects recall and item non-response concerns; people generally do not know the census tract they live in and a high proportion of respondents are not willing to share their full address in telephone survey situations.

Through the use of ABPMs at NYC DOHMH, two important issues have been identified. First, yearly changes to categories may affect the estimation of trends by poverty status in a misleading way. That is, change in an estimator for a specific poverty level might reflect changes in the assignment of areal units to that level. For example, under an assumption of no change observed at all, changes to assignment will change the estimator.

Second, uncertainty in the American Community Survey (ACS), may lead to misclassification of poverty status and associated mismeasurement of disparity. We examine the role that uncertainty from the ACS plays in the identification of high and low poverty areas at three-scales (tract, ZCTA, PUMA). We use uncertainty estimates provided by ACS to examine misclassification and propagation of the uncertainty into disparities measures. We then examine how measurement of disparity trends might be misleading due to this uncertainty. Next, three possible solutions to trend mismeasurement will be examined: first, we characterize current recommendations from NYC; second, we look at trends over fixed geographies defined within the time period of interest; third, we classify areas based on a persistence metric, so that persistently poor neighborhoods are compared to persistent rich ones.

2.3.2 Candidate Geographies and the American Community Survey

For public health surveillance, geographies must either be fixed or based on input data that are regularly updated in a timely manner. Prospective surveillance systems that incorporate changing geographies rely on future updates as well, requiring investment of resources to insure their production. For a local public health department, adapting national efforts can alleviate the need for such planning while also providing a framework for other jurisdictions to establish similar systems.

In the present study, geographies and poverty estimates are taken from annual releases of American Community Survey five year estimates (U.S. Census Bureau, 2014, 2009a). Due to reliance on the ACS, candidate geographies must be constructed from available ACS releases and so must be constructed from census tracts– the smallest geographic unit– or Zip Code Tabulation Areas (ZCTAs) a census-block based geography that currently has separate annual reporting for the ACS. We consider four geographies: census tracts (CTs), zip code tabulation areas (ZCTAs), neighborhood tabulation areas defined by the New York City Department of City Planning (NTAs), and public use microdata areas (PUMAs) which have been assigned to approximate Community Districts. Descriptions of these geographies from the 2013-2017 5 year ACS are given in Table 1. CTs nest within NTAs which nest within PUMAs. ZCTAs are roughly the same scale as NTAs but with higher variation in size and less coherence in their relationship to historic New York City neighborhoods.

The Neighborhoods Tabulation Area (NTA)

The establishment of NTAs for reporting outcomes at a local scale represents a recent development in New York City government. Interest in reporting at a neighborhood scale increased with increased interest in social determinants and hierarchical models. The same trend increased interagency dependency. Simultaneously, open data laws, other attempts to increase transparency, and increased interest in data driven decision making in government created demand for data release at fine spatial scales. Issues with the use of census tracts for these purposes were raised including the low precision of ACS estimates, a lack of connection between census tract and traditional notions of neighborhoods or activity space, and the increased difficulty of producing tables for release that ensured adequate anonymity protections as required by law. In response, researchers (Spielman and Folch, 2015) and local governments (County of San Diego Health and Human Services Agency, 2015; Baltimore City Health Department, 2019) established new geographies or methods to produce new geographies at a scale between census tracts and PUMAs.

In New York City, these trends led directly to a project focused on identifying a geography to integrate data from multiple city agencies and local health care organizations funded by the Robert Wood Johnson Foundation “Data Across Sectors for Health”

(DASH) initiative (RWJF Data for Health Advisory Committee, 2015). The effort integrated SDH measures not traditionally under the purview of health departments such as social support, incarceration, housing, and children's services. The chosen geography, the Neighborhood Tabulation Area (see Figure 2.1), had been established in 2007 by the New York City Department of City Planning (DCP) as part of PlaNYC, the sustainability plan for New York City (City of New York, 2007), to produce highly detailed population projections. The primary goal of the effort was to characterize areas that users could relate to in the context of the city's neighborhoods.

The U.S. Census Bureau developed and defined PUMAs for the entire nation in the 1980s. In New York City PUMAs were constructed in close cooperation with DCP to approximate the city's 59 community districts– a key geography in New York City government responsible for review of zoning and planning decisions, coordination of local service delivery, and provision of local input to the city's budget. The nationwide PUMA program required jurisdictions to use whole census tracts with a minimum population of 100,000 persons (to preserve the anonymity of respondents in the public use microdata file). These constraints resulted in 55 PUMAs in New York City each aligning with individual community districts or the merger of two contiguous Community Districts to meet population requirements. Given the importance of PUMAs for the provision of census data and their approximate coterminality with Community Districts, NTAs were created as subdivisions of PUMAs, using whole census tracts as building blocks, thereby preserving the geographic hierarchy used for data provision.

NTAs needed to have a minimal population threshold, initially set at 15,000, to provide an adequate base for doing population projections. In addition, when subdividing PUMAs, the selection of census tracts was informed by neighborhood designations as they appeared in the 1969 Plan for the City of New York, which was the basis for community district boundaries that went into effect in 1977 (New York City Planning Commission, 1969). Additional sources containing historical neighborhood definitions within New York City such as the Encyclopedia of New York (Jackson et al., 2010) were also consulted and Community Board planners at the New York City Department of City Planning also provided input in drawing-up NTA boundaries.

Ultimately, 188 NTAs (figure 1) were defined with an average population size of 45,000. Due to constraints described above, it is important to recognize that NTAs were not intended to be definitive in their representation of neighborhood boundaries. Indeed, in order to meet criteria, adjacent NTAs sometimes include different sections of a large neighborhood, or an NTA may subsume multiple neighborhoods. Despite the obvious risks inherent in the subjective selection and assignment of neighborhood names, this exercise proved useful in that data users can readily relate to neighborhood names.

Four geographies were considered as candidates for defining poverty groups for ABPM-based surveillance of disparities: PUMAs (n=55 areas), ZCTAs (n=182), NTAs (n=188), and census tracts (n=2,123). Two additional areas, actual Community Districts (CDs, n=59) and the United Hospital Fund (UHF) neighborhood (n=42) are routinely used in public health practice including surveillance and reporting but are not considered here. They are both coarser than those under consideration and are not based on census tracts.

2.3.3 Candidate Processing Rules and the American Community Survey

The primary goal of public health surveillance and monitoring is to discern changes or measure trends in health outcomes. Because of this, approaches to ABPM-based surveillance need to consider change across three input sources: the poverty estimates used to designate poverty categories, the outcome data used as numerators, and the population estimates used for denominators. It is clear that assessing change in the poverty status of specific geographic areas would inform our understanding as disparity measures based on ABPMs. For example, it is possible that a narrowing disparity using a fixed geography over a long time period could reflect gentrification in the areas considered poor as opposed to improvements in the actual poor/rich disparity. For this reason, initial guidance recommended that for each year studied, ABPMs be redefined using the preferred source for that year as discussed above. For the years prior to 2005, the 2000 long form was recommended and for years 2005 to present it was recommended to use the available ACS 5-year ABPM designation that most aligned with that year. For example, 2009 would be based on the designation established using the 2007-2011 5-year sample, while 2016 would currently be based on the most recent release (2013-2017) but would be based on 2014-2018 after its release. This leads to disparity trends with each year's poor/rich difference potentially based on different sets of areas.

However, there is considerably more uncertainty in the ACS 5-year samples than in the 2000 Census long form (Spielman et al., 2014). ABPM designations may have high-levels of misclassification specifically due to sampling error and design effects. Changes in designation between release years may reflect this rather than underlying change such as gentrification. Further, whereas designations based on the census long form could only be calculated every ten years, the ACS annual release schedule can result in dramatic annual changes in designations that might not reflect true change. This is complicated by the fact that with an annual release of 5-year data, consecutive years share 4 years of data over which change cannot be assessed. This attenuates changes in designation, making it difficult to establish how much of a 1-year change could be resulting from uncertainty alone. Nonetheless, we investigate 1-year changes as well as 5-year changes based on non-overlapping samples.

Two alternative processing strategies are included as candidates. First, we construct ABPM-based disparity estimates using fixed geographies from either the initial or center year in question. It is unclear how best to employ this in a surveillance framework. One possibility would be to establish ABPM designations once every 10 years as with the Census long form; suggesting the use of 2008-2012 data. Second, ABPM-based disparity estimates use areas that are identified as persistently poor or rich based on changes over the study period. ABPM designations across several years are used to determine areas consistently estimated to be in rich/poor categories. As with the first alternative, it isn't clear how best to identify persistence prospectively for use in surveillance. One possibility is to establish designations using two non-overlapping ACS 5-year releases and then update them when a new overlapping sample becomes available. One feature of the second alternative is that the processing results in an understanding of those areas that have both left and entered high and low poverty areas, allowing for some assessment of change. Further, disparity trends based on persistent measures can be augmented with the same trends for the changing areas.

2.3.4 Outcome and Population Data

Outcome data used in our case study come from New York State's Statewide Planning and Research Cooperative System (NYS Department of Health Bureau of Health Informatics, 2014), New York's all-payer claims database (Miller et al., 2015). SPARCS contains records for all emergency department (ED) visits and hospitalizations in all New York State hospitals since 2003. The data is highly detailed with all diagnostic and procedure codes, as well as spatial and demographic information. SPARCS contains low quality race data like most administrative records systems that do not explicitly collect race information. A pseudo-identifier allows users to assess multiple visits by individuals across years. The data has been geocoded using patient addresses and can be mapped to all candidate geographies. Multiple day hospitalizations are analyzed by admit date as opposed to discharge date.

Asthma is of particular interest in developing ABPMs for public health practice for several reasons. Adverse asthma events result when a child who is 1) prevalent with asthma and 2) not adequately 'controlled' is 3) 'triggered' to have an asthma attack. Each of these requirements are highly correlated with poverty and spatial features. For reasons that are not understood, underlying asthma prevalence varies by poverty and space. Asthma control requires medication adherence, which in turn requires adequate health care access. This results in a distinct spatial pattern reflecting the decreased health care of access for those in poverty. Asthma attacks result from asthma triggers exacerbating prevalent and uncontrolled asthma. Triggers include spatial phenomenon such as air quality, pest and rodent infestation and other housing quality issues, and social stressors such as homelessness, crime, and deprivation. Because in almost all cases asthma attacks are preventable with appropriate medications, nearly all adverse asthma outcomes including emergency department visits are be considered "avoidable" and the costs associated with such visits reflect potential savings. These savings could be realized by addressing Social Determinants of Health such as housing quality and improving access to care. Altogether, we would expect asthma to have clear relationship with ABPMs.

We tabulate all asthma-related ED visits and hospitalizations to New York City school-aged children from 2006-2016 for all candidate geographies. Asthma-related visits are defined using the principal diagnosis field as defined by the International Classification of Diseases (ICD) codes, version 9 and 10. Although asthma is an extremely complicated disease it is among the most straightforward diseases to classify using ICD9 and ICD10 codes with a limited set of codes all within a specific diagnostic branch in the hierarchical system. Nonetheless, the switch from ICD9 to ICD10 in October 2015 may have resulted in changes that impact trend estimates. These would be adjusted by a multiplier called the "comparability ratios" (Anderson et al., 2001) but these are not yet published for diagnostic codes. To address this, we construct trends quarterly from 2006 to 2016 allowing for an assessment of the impact but do not attempt correction.

Post-censal population estimates are not produced by the census at the sub-county-level in New York City. As such, yearly population estimates are available by age, race, and gender for each of the five New York City boroughs from the census county estimates. Annual sub-county population estimates are produced by the Epidemiology division of NYC's Department of Health and Mental Hygiene; they are released approximately two months subsequent to the release of borough postcensal population estimates. The sub-county estimates are produced for ZCTAs, Community Districts, and PUMAs. The approach to the estimates is to first adjust population sizes within PUMAs using housing unit change, similar to the housing unit method (Smith and Lewis, 1980) and then to tightly control these changes to the borough age-sex-race-ethnicity total using the Iterative Proportional Fitting Procedure (IPFP, or "raking") (Wong, 1992). The resulting sub-county estimates are hierarchically consistent with each other and borough totals with population growth being allocated to areas of growth and by previous year's age-race-sex distributions. Although no individual census tract population estimates are produced, census tract-based ABPM by county estimates are created for each new release of ACS data for census tract, ZCTA, and NTA.

2.3.5 Methods

Poverty categories were identified from area-based poverty estimates using published ACS estimates for census tracts and PUMAs. Categories for neighborhood tabulation areas were created from the census tract tables. Uncertainty estimates were taken directly from ACS documentation published tables which include 90% confidence intervals based on resampling weights. When necessary for simulation, standard errors were back-calculated from these confidence intervals using the standard normal distribution. Uncertainty estimates for NTAs follow census guidance on margins of errors for derived estimates (chapter 8, (U.S. Census Bureau, 2018))

Reliability of estimates for areas at different scales were assessed using the coefficient of variation (CV), the standard errors as a proportion of the estimate, or relative standard error (CV presented as a percentage) and with reference to Census guidance on the use of estimates (U.S. Census Bureau, 2009a, 2018). To estimate misclassification of poverty categories, simulations were performed using uncertainty estimates. These were then compared to actual observed change. Spatial patterns of geographic areas leaving and entering high poverty categories relative to persistently high poverty areas were assessed visually.[1]

Population-level asthma ED-visit and hospitalization rates are calculated directly using standard statistical methods used in public health including comparisons and tests for trend (Fleiss et al., 2013). Trends are constructed using all candidate scales and processing rules. Finally, estimated poverty and asthma rates were characterized using scatterplots with reference to mapped asthma rates. Data processing and statistical analysis were conducted in SAS and R.

2.4 Results

Figure 2.2 presents area-based poverty designations over four geographies. The two shades of red correspond to high poverty, with dark red being the highest poverty category in the *poverty6* classification and the lighter shade being the second highest poverty. Together these correspond to the highest poverty category for *poverty4*. Similarly, the lowest poverty area is shaded dark blue with light blue the second lowest poverty. Like-

[1]While the clustering could be quantified and statistically tested using a local Moran's I, all the patterns displayed in the results are strong enough that a formal test seems unnecessary.

wise, the two form the low poverty category for *poverty4*. The middle two groups for both *poverty4* and *poverty6* are in grey with the darker shade corresponding to higher poverty. Given fixed cutoffs regardless of geographic scale finer geographies identify a higher proportion of the city's population being labeled high or low poverty. This is a straightforward result of aggregation, with mean poverty for combined areas tending to the overall mean.

Table 2.1 displays the number of areal units and the percentage of the population in each poverty group for each of the geographies. Census tract geographies classify a higher proportion of the population in the tails of the distribution. Disparities measured using ABPMs tend to display the widest disparities at the census tract scale (Toprani and Hadler, 2016; Krieger et al., 2002) . While this may seem to be a feature of using smaller geographies, two important issues are apparent. First, greater uncertainty in the ACS for census tracts may lead to misclassification impacting disparity measures. Second, yearly changes to categories may effect the estimation of trends by poverty status in a misleading way.

We examine the role that uncertainty from the ACS plays in the identification of high and low poverty areas. Figure 2.3 depicts level of uncertainty as measured by the relative standard errors (RSE) for census tracts and neighborhood tabulation areas for percentage of children living in poverty. The map's shading identifies areas that do not meet an RSE of 20%. This cutoff is actually looser than the census itself recommends: "while there is no hard and fast rule, for the purposes of this handbook, estimates with CVs of more than .15 are considered cause for caution when interpreting patterns in the data" (U.S. Census Bureau, 2009b). Only 20.1% of New York City census tracts meet a relative standard error cutoff of 20%, whereas 94.7% of NTAs meet it.

The impact of this uncertainty would be high levels of misclassification due to sampling error and change in poverty classifications due to sampling error as opposed to genuine change. This can be seen by constructing year-to-year transition tables for the candidate geographies using a simulation with iterations sampled from the distribution implied by the point estimate and its uncertainty. Table 3.2 presents two 1-year transition tables at the census tract scale. The top panel simulates the second year under the assumption of no change but accounting for the uncertainty in the first years sample. The lower panel present the observed changes from the 2008-2012 to 2019-2013 ACS releases for Census Tracts. In the simulation, only 65% of the New York City population is classified into the same *poverty6* categories in year 2, suggesting enormous change despite the fact that we assumed no change. From the lower pane we see that only 73.3% of the New York City population changed poverty assignment in year two. This suggests that the apparent change observed at the Census Tract may be illusory. Table 3.3 compares the observed 1-year change transitions for census tracts (top) and NTAs (bottom). NTA estimates are considerably more stable with less than 5% of the population changing poverty assignment over the two surveys. .

The rationale for reclassifying areas each year was to capture underlying change in those areas, such as gentrification in poor areas or increasing homogenization of wealthy areas. This can be examined by mapping the areas that have changed and assessing their spatial patterns. Figure 2.4 presents year-to-year change in the *poverty4* areas classified as poor from the 2008-12 to the 2009-13 survey release, with light blue being areas classified as poor in both years, red being areas newly categorized as poor, and dark blue being areas leaving the high poverty category potentially due to gentrification. Areas of NYC that are gentrifying are already well-known and can serve as a prior expectation on that patterns visible in the maps. One would expect that the blue areas – those leaving poverty designation – would be spatial clustered and would occur in known pockets of gentrification. But this is not the case. Another practical issue that becomes apparent when considering this map is the potential difficulty of reviewing changes in classification;

there are many census tracts to review with characterization difficult due to their abstract construction and limited available auxiliary information. The right panel presents the same year-to-year change for NTA estimates confirming the static nature of high poverty areas. Only two NTAs enter high poverty. Their correspondence to actual neighborhoods allows these transitions to be more easily reviewed.

One difficulty in examining year-to-year changes is that observed changes in categories resulted from ACS surveys pooled across five years that share four of the five years. Changes in categories result from differences in the first and last year of the six-year period covering the two surveys. Two sets of non-overlapping 5-year ACS surveys are available to assess this potential problem. Transition tables for 2008-2012 to 2013-2017 for census tracts and NTA are presented in Table 2.4. For census tracts, 51.6% of the population changes poverty designation over the 5-year period and 38.9% change *poverty4* designation between the two survey years. Ideally this would reflect actual change in the underlying poverty of the census tract. However, from Table 2.4,we would expect 34.4% and 25.7% of the tracts to change based on survey error alone under the assumption of no change, over half of the observed change. This suggests high levels of misclassification in the census tract assignments to poverty categories. NTA classifications are less volatile with change similar to the change observed in 1-year census tract estimates. No NTA changed by more than one category over the 5-year period. Figure 2.5 displays the geographic pattern of high-poverty areas for the two 5-year periods for census tracts (left panel) and NTAs (right panel). Although untested statistically, it is difficult to discern a pattern from the census tract changes whereas the NTA changes are compact, coherent (largely comport with known areas of gentrification), and clearly characteristic of an underlying smooth process of change (change areas border the no change areas).

Trends in asthma emergency department visits for high and low-poverty Neighborhood Tabulation Areas measured quarterly are presented as Figure 2.6 for fixed geographies based on 2008-2012 ACS. NTAs and census tract asthma rates by poverty group mirror each other in magnitude, seasonal and temporal trend and relationship between poverty groups. The proposed disparity measure is the relative risk of an ED visit for children in poor areas in reference to wealthy areas. This is the ratio of the top and bottom lines in Figure 2.6 and is presented as Figure 2.7 for 2 key age groups 5-12 and 13-19 year-olds. The trend in this disparity can be used to monitor public health efforts, as well as an input into a cost/savings calculation giving the potential value of reducing ED rates in the poorest areas. NTA based disparity estimates display more variance probably reflecting reduced sample size using fixed category cutoffs. A widening of the disparity for poor children age 5 to 12 is evident in both series but more pronounced by NTA.

To assess the candidate assignment rules, persistently poor areas were identified using 2008-12 and 2013-17 ACS releases and a dataset constructed with quarterly estimates for: 1) persistence-based assignment, 2) the varying-by-year assignment rule following current city guidelines, 3) fixed assignment based on 2008-12, and 4) fixed assignment based on 2013-17. Trend estimates for the measured disparity are given as Figure 2.8 for 5-12 year olds. The blue and red lines correspond to the two fixed geographies, the green line corresponds to the persistently poor and rich areas, and the purple line to the varying assignment. Although, greater variance was expected for the varying assignment rule, the measured disparity tracked both fixed assignment rules. The varying assignment rule is equivalent to a fixed assignment in 2010, 2015, and 2016. The persistent poverty-based estimates identify greater disparities by identifying coherently defined sets of high and low poverty areas.

To capture some of the dynamics in the varying assignment rule, we again plot the asthma ED visit rates in the persistently rich/poor areas (red lines) as Figure 2.9 together with those areas that were in the 2008-12 time varying assignment but were not deemed persistent (green lines) and those entering the calculation in 2013-17 period (blue lines). Clearly, the persistently high and low poverty areas experience the highest and lowest rates of asthma ED visits resulting in an increased disparity measure. Interestingly, the areas leaving and entering poverty are not distinguishable again suggesting change that does not reflect underlying change.

Lastly, we present scatterplots of the estimated household poverty rate in 2013-17 by asthma ED rate (figure 10, left panel) and the change in poverty rate by change in asthma EDs over the 2008-12 to 2013-17 time periods (right panel). Although there is a clear poverty asthma relationship, variance in asthma greatly increases as poverty increases. This should suggest that additional variance should be introduced by varying the assignment rule, but this wasnt seen in results. Finally there is no discernible pattern in the change scatterplot suggesting that uncertainty in the poverty change estimate may be masking actual change.

2.5 Discussion

We have described NYC's effort to develop area-based poverty measures for use in surveillance systems. Examining the initial policy established for the use of ABPMs in New York, we found that disparity estimates were susceptible to uncertainty in the assignment of poverty groups due to low reliability and that updating rules introduced uncertainty in trends that may obscure rather than highlight trends. This is especially true for smaller geographies such as census tracts for which SDH estimates rely on small sample size.

Ultimately, the choice of approach depends on the utility of the measures to support the development of local public policy and the delivery of services, programs, and interventions. A wide variety of local initiatives address social determinants of health including inter-agency community-level interventions; many targeted based on perceived or established relationships between SDH and health outcomes. Inter-agency community-based interventions include the identification of Community Schools for expansion of mental health services, housing interventions that provide integrated pest management to residences with children with asthma, housing placement for homeless families that takes continuity of schooling into account, changes to bail requirements for incarcerated adolescents, and monitoring of health outcomes for those placed in affordable housing.

Practical considerations were key to discerning between candidate geographies and processing rules. It is known that census tract estimates fail to meet Census recommendations for reliability for most indicators including household poverty in New York City. Further, while census tracts might provide the most accurate imputations for datasets lacking poverty information, they do not comport with theoretical notions of context like neighborhood or activity space. They are also more difficult to maintain and process and, in many situations, may not allow for public release of health data due to low sample sizes. NTAs improve on all of these criteria while slightly moderating the disparity measure for childhood asthma. This is somewhat addressed by the use of persistently poor geographic areas. NTAs correspondence with historical notions of neighborhood make them excellent candidates for characterizing context or place. Monitoring change at this scale is also more manageable and privacy concerns with data release are alleviated.

ABPMs are a single low-information example of a social determinant of health measure. A variety of other measures have been proposed as important causes and correlates of health conditions. Interest in incorporating these measures into public health practice is increasing rapidly. One possible approach is to bring together SDH measures as a surveillance system itself to be used alongside traditional public health data systems. Such an effort is analogous to the current practice of maintaining the demographic information used for denominators. In this approach, the goal is to observe changes in the determinants themselves, essentially maintaining ecological data about neighborhoods as panel data. NTAs provide a geographic frame to construct this.

Spatial information has always played a central role in public health practice, especially at the local scale. Recent developments, including a spatial turn in academic public health, have resulted in renewed interest in the effects that neighborhoods have on population health (Diez-Roux, 2001; Krieger et al., 2003). Geographers have given considerable attention to important spatial issues with the construction of geographic units for these purposes. Continued interest in the social determinants of health, especially conceived as attributes of place and integrating data from sources outside of health, present a major opportunity for Geography.

2.6 Tables and Figures

Table 1: Poverty Groups by Geographic Scale, 2013-2017										
% households below poverty		<5%	5-10%	10-20%	20 to 30%	30 to 40%	>40%			
	Poverty6	1	2	3	4	5	6			
	Poverty4	1	1	2	3	4	4			
# of Geographic Areas with pop>0								Total	Mean(pop)	CV(pop)
	Census Tracts	203	419	706	412	218	166	2124	3,965	0.56
	ZCTA	10	44	72	35	16	6	183	46,616	0.58
	NTA	2	40	77	40	19	10	188	44,782	0.49
	PUMA	0	7	24	15	6	3	55	153,073	0.22
% of population										
	Census Tracts	7.8	19.6	31.5	20.8	11.9	8.5			
	ZCTA	1.7	18.8	39.3	23.8	13.2	3.1			
	NTA	0.9	18.0	41.1	24.8	10.7	4.5			
	PUMA	0.0	13.3	45.0	26.5	9.8	5.5			

Table 2.1: Poverty groups by geographic scale, 2013-17

Apparent transition due to sampling error assuming no change, Census Tract Poverty Designations									
1-year change				**Year 2 assuming no changes**					
2008-2012		**1**	**2**	**3**	**4**	**5**	**6**	**Total**	
	1	**521,888**	138,077	132				660,097	
	2	284,073	**906,677**	266,443	1,412			1,458,605	
	3	65,246	413,145	**1,643,125**	320,780	8,191		2,450,487	
	4		20,393	311,099	**948,697**	331,556	16,485	1,628,230	
	5			6,168	316,448	**642,701**	174,544	1,139,861	
	6		18	91	16,181	108,807	**673,088**	798,185	
Total		871,207	1,478,310	2,227,058	1,603,518	1,091,255	864,117	**8,135,465**	
						unchanged poverty 6		**5,336,176**	**65.6%**
						unchanged poverty 4		**6,041,677**	**74.3%**

1-year change of Census Tract Poverty Designations									
1-year change				**2009-2013**					
2008-2012		**1**	**2**	**3**	**4**	**5**	**6**	**Total**	
	1	**538,316**	163,840	12,554	1,645			716,355	
	2	120,583	**1,085,413**	317,965				1,523,961	
	3	10,190	225,758	**1,830,095**	303,534	2,036		2,371,613	
	4	28	6,656	224,945	**1,176,823**	353,584	442	1,762,478	
	5			2,937	132,487	**706,224**	164,621	1,006,269	
	6				5,555	57,699	**622,378**	685,632	
Total		669,117	1,481,667	2,388,496	1,620,044	1,119,543	787,441	**8,066,308**	
						unchanged poverty 6		**5,959,249**	**73.9%**
						unchanged poverty 4		**6,465,992**	**80.2%**

Table 2.2: One-year changes in Census Tract poverty designation

1-year change of Census Tract Poverty Designations								
1-year change 2008-2012	**2009-2013**							
	1	**2**	**3**	**4**	**5**	**6**	**Total**	
1	**538,316**	163,840	12,554	1,645			716,355	
2	120,583	**1,085,413**	317,965				1,523,961	
3	10,190	225,758	**1,830,095**	303,534	2,036		2,371,613	
4	28	6,656	224,945	**1,176,823**	353,584	442	1,762,478	
5			2,937	132,487	**706,224**	164,621	1,006,269	
6				5,555	57,699	**622,378**	685,632	
Total	669,117	1,481,667	2,388,496	1,620,044	1,119,543	787,441	**8,066,308**	
					unchanged poverty 6		**5,959,249**	**73.9%**
					unchanged poverty 4		**6,465,992**	**80.2%**

1-year change of Neighborhood Tabulation Area Poverty Designations								
1-year change 2008-2012	**2009-2013**							
	1	**2**	**3**	**4**	**5**	**6**	**Total**	
1	**45,545**	24,753					70,298	
2		**1,298,216**	122,451				1,420,667	
3		22,830	**2,748,852**	61,512			2,833,194	
4			63,859	**2,223,930**	44,931		2,332,720	
5					**1,064,759**	53,577	1,118,336	
6						**291,220**	291,220	
Total	45,545	1,345,799	2,935,162	2,285,442	1,109,690	344,797	**8,066,435**	
					unchanged poverty 6		**7,672,522**	**95.1%**
					unchanged poverty 4		**7,750,852**	**96.1%**

Table 2.3: One-year changes in poverty designation, Census Tract vs NTA

5-year change of Census Tract Poverty Designations

5-year change 2008-2012	2009-2013 1	2	3	4	5	6	Total	
1	**296,181**	334,708	81,392	3,887	30		716,198	
2	252,777	**721,708**	507,275	36,651	3,571	1,979	1,523,961	
3	83,507	464,221	**1,323,891**	429,562	66,968	3,464	2,371,613	
4	8,257	60,844	530,021	**805,623**	271,533	86,200	1,762,478	
5		1,274	83,398	357,460	**366,342**	197,795	1,006,269	
6	133		10,836	46,704	237,165	**390,741**	685,579	
Total	640,855	1,582,755	2,536,813	1,679,887	945,609	680,179	**8,066,098**	
					unchanged poverty 6		**3,904,486**	**48.4%**
					unchanged poverty 4		**4,926,931**	**61.1%**

5-year change of Neighborhood Tabulation Area Poverty Designations

5-year change 2008-2012	2013-2017 1	2	3	4	5	6	Total	
1	**0**	69,932					69,932	
2	78,927	1,054,209	287,531				1,420,667	
3		341,614	2,409,894	81,686			2,833,194	
4			623,038	1,667,807	40,590		2,331,435	
5				227,289	782,712	107,919	1,117,920	
6					41,073	248,797	289,870	
Total	78,927	1,465,755	3,320,463	1,976,782	864,375	356,716	**8,063,018**	
					unchanged poverty 6		**6,163,419**	**76.4%**
					unchanged poverty 4		**6,461,270**	**80.1%**

Table 2.4: Five-year changes in poverty designation, Census Tract vs NTA

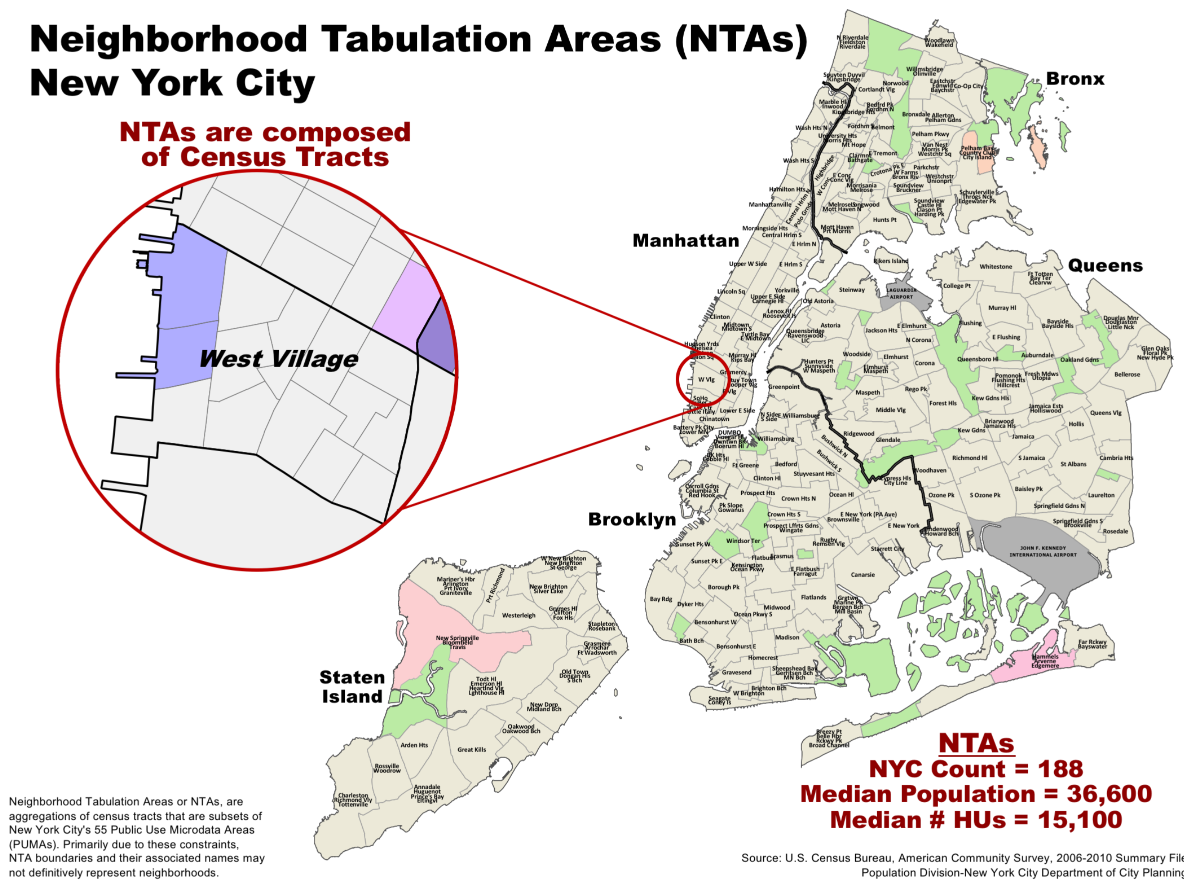


Figure 2.1: Neighborhood tabulation areas

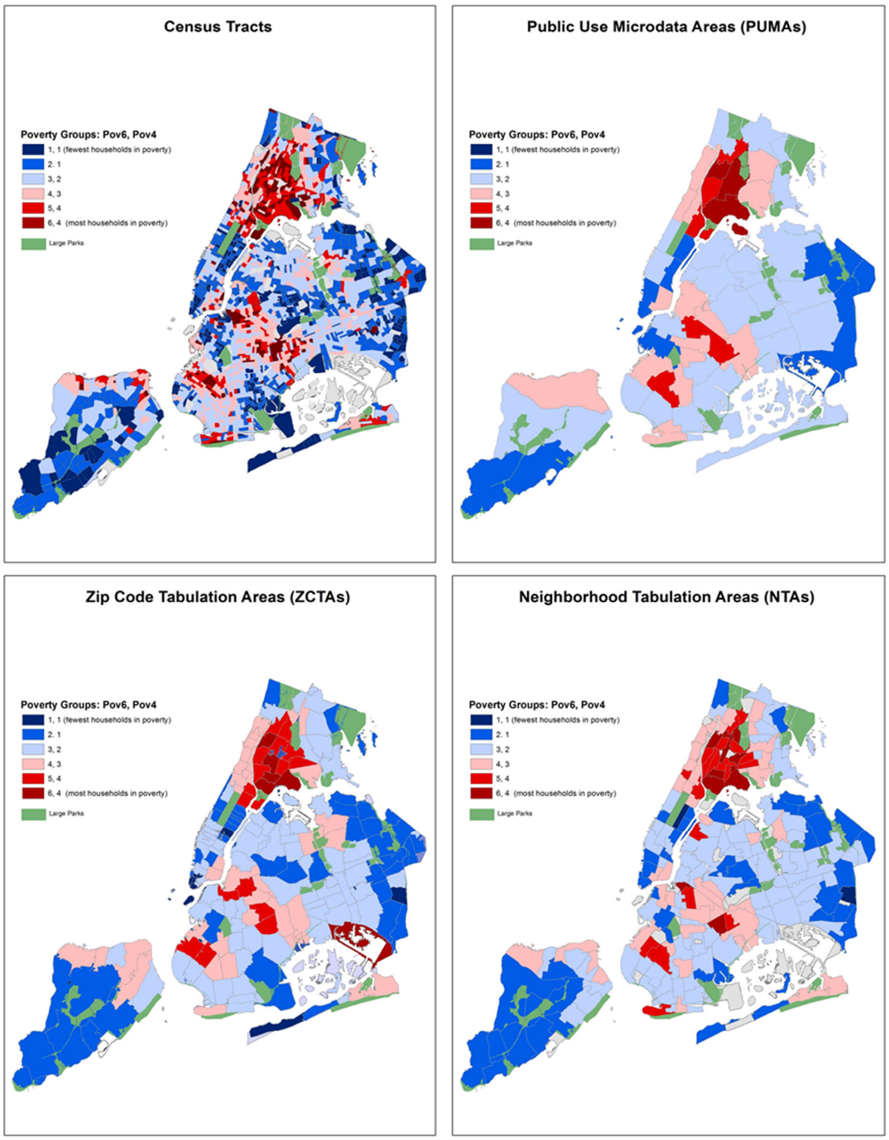


Figure 2.2: Poverty maps under four different geographies

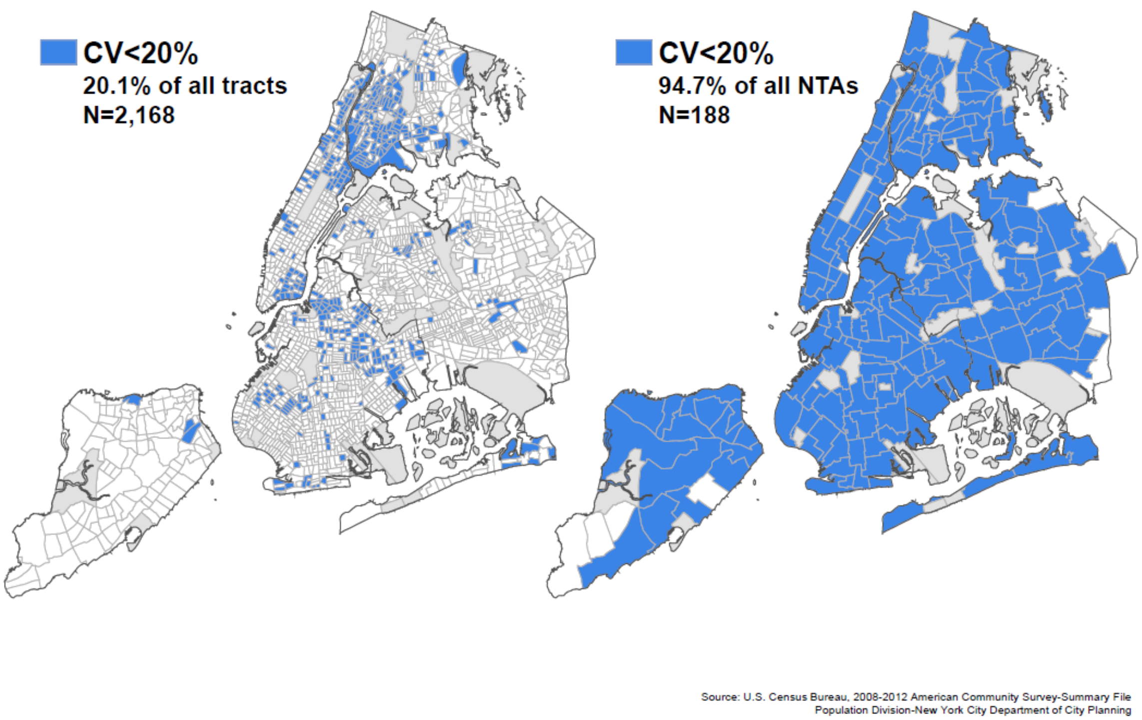


Figure 2.3: Coefficient of variation due to sampling error, Census Tracts vs NTAs

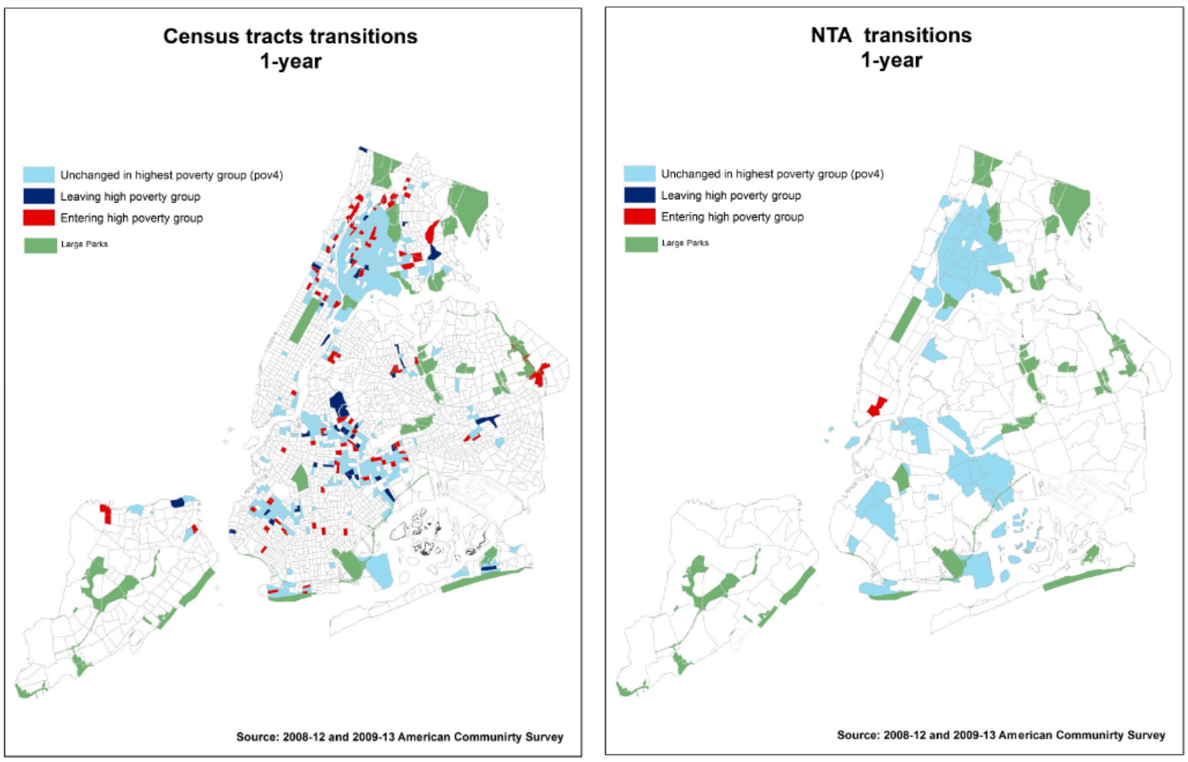


Figure 2.4: Areas remaining, entering, or leaving poverty under 1-year transitions, Census Tracts vs NTAs

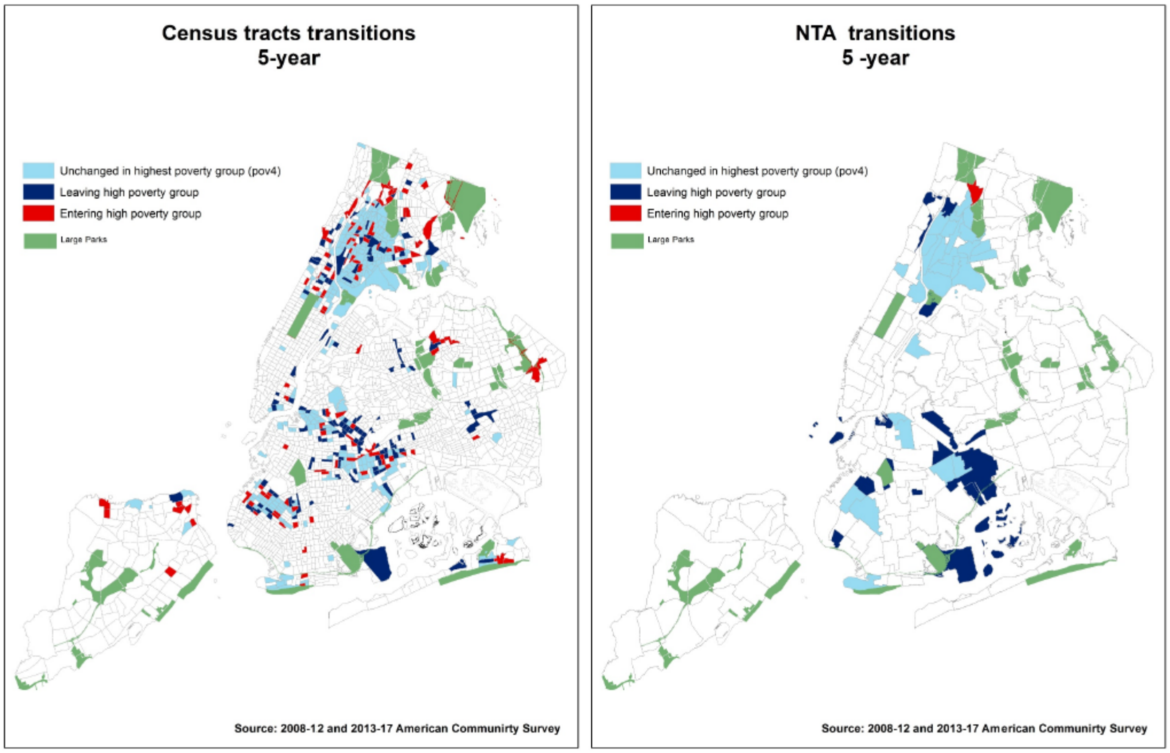


Figure 2.5: Areas remaining, entering, or leaving poverty under 5-year transitions, Census Tracts vs NTAs

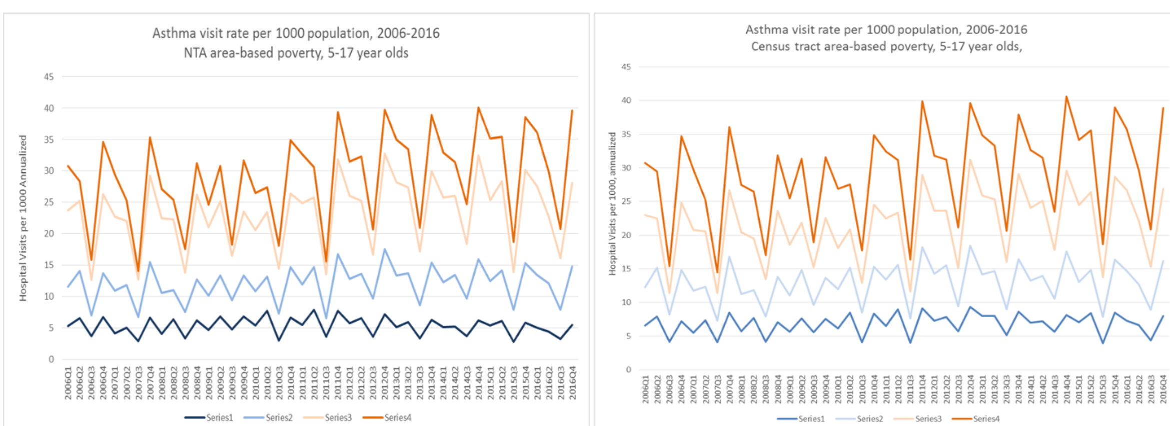


Figure 2.6: tbd

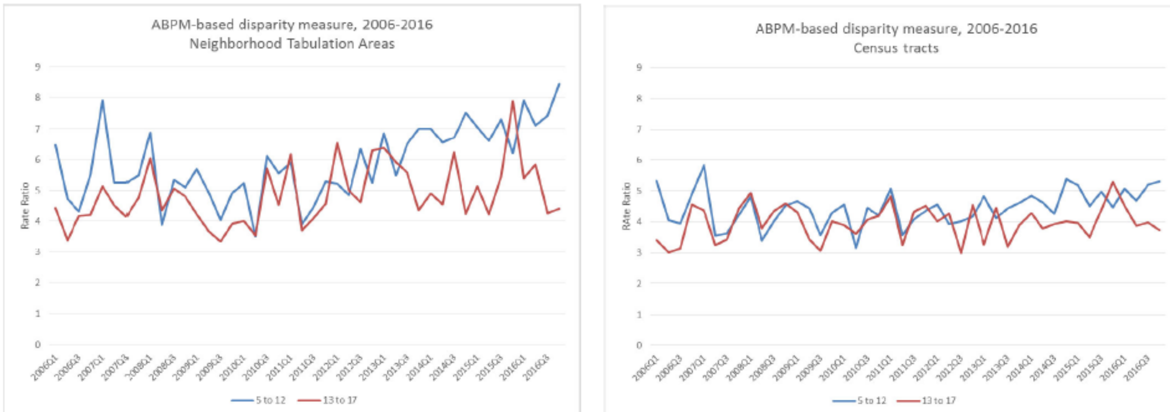


Figure 2.7: ABPM-based disparity measure, Tracts vs NTAs

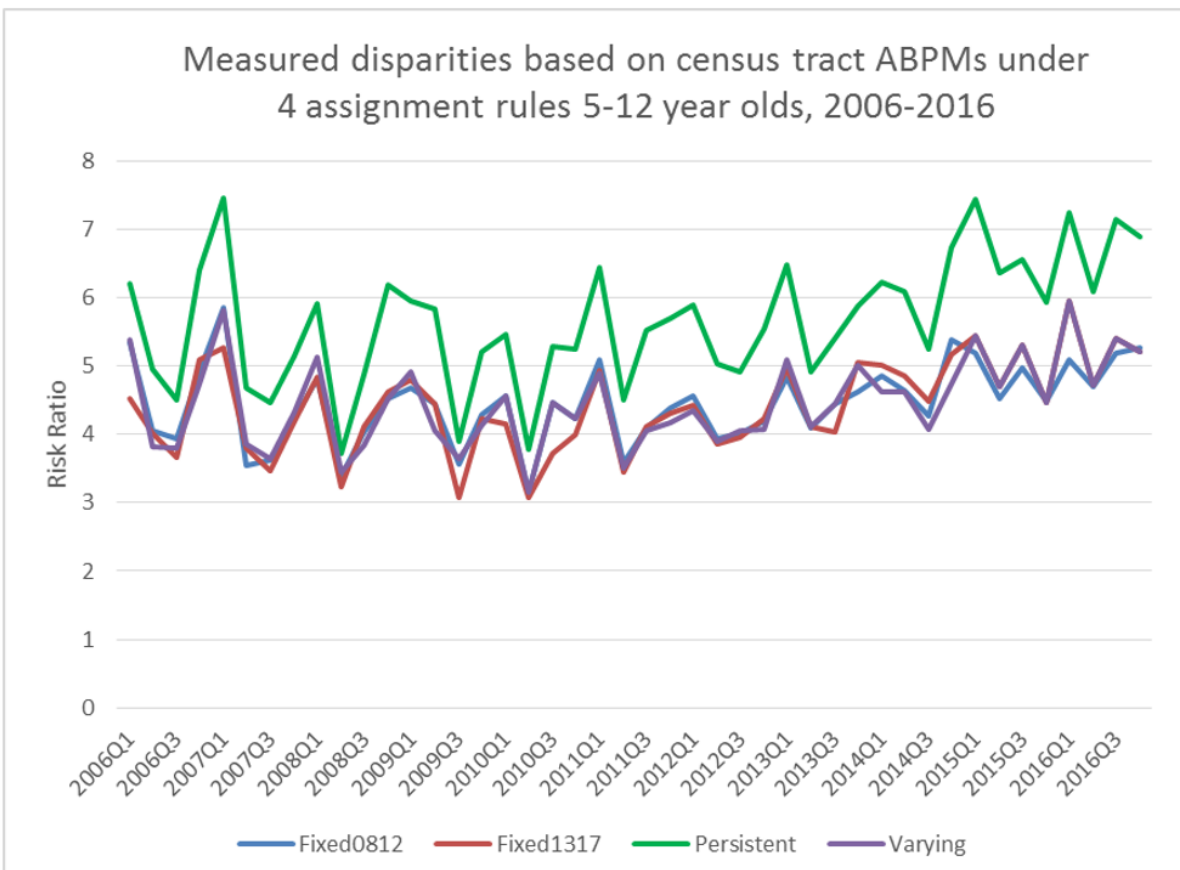


Figure 2.8: ABPM-based disparity for tracts using 4 different assignment rules

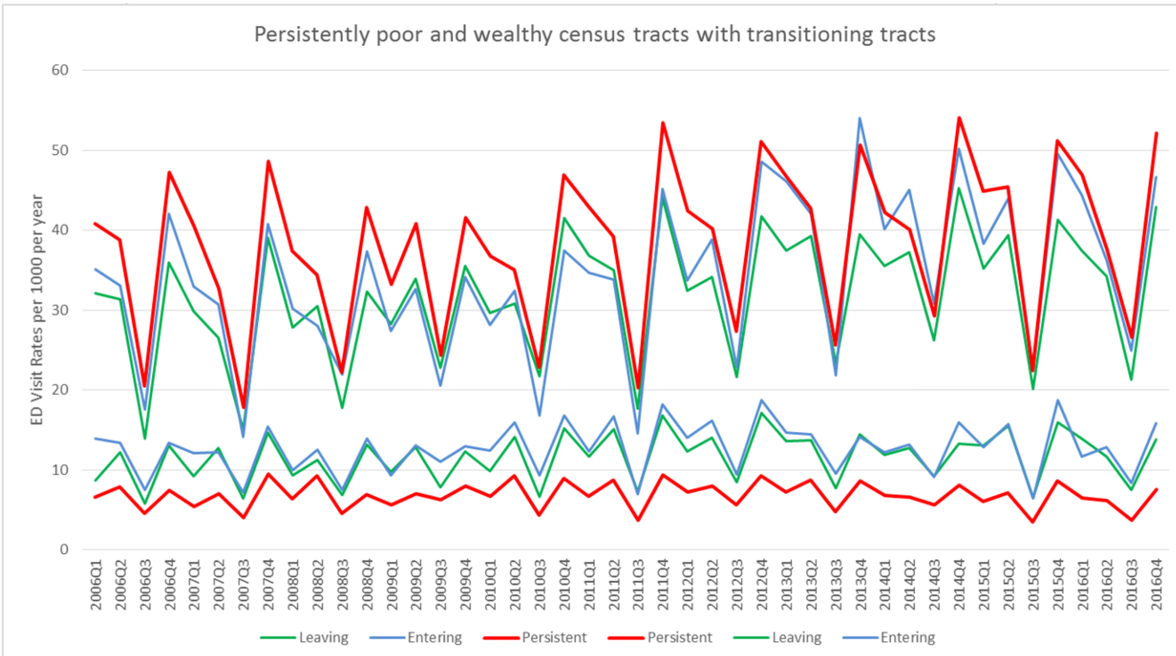


Figure 2.9: Persistently poor and wealthy census tracts with transitioning tracts

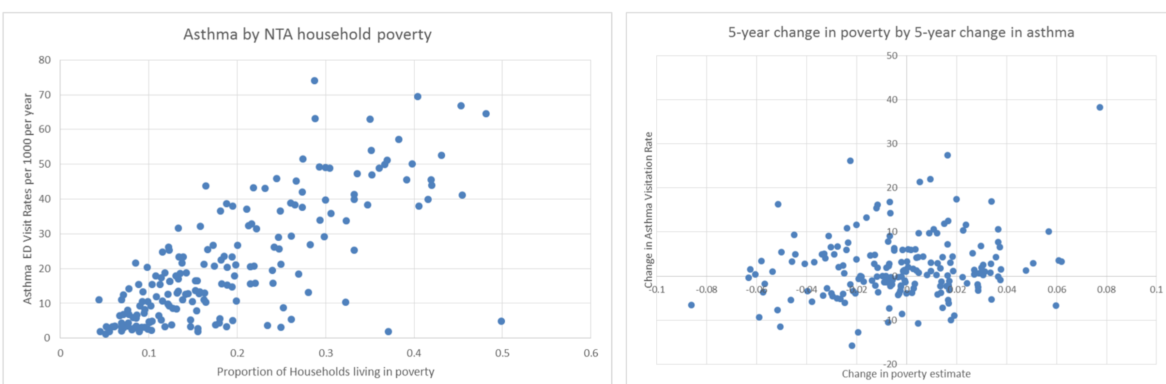


Figure 2.10: Asthma by (NTA household poverty or change in poverty) correlations

Chapter 3

Spatial patterns of child mental health burden: Measurement and monitoring of child mental health disparities in New York City

Abstract

From 2001-2011, mental health-related hospitalizations and emergency department (ED) visits increased among United States children nationwide (Simon and Schoendorf, 2014). During this period, mental health-related hospitalizations among NYC children increased nearly 23% (Mills and Davila, 2016). Much of this burden is "avoidable" (Hsia and Niedzwiecki, 2017) in the sense that the condition could have been addressed preventatively. Further, because hospital emergency departments are ill-equipped to deliver pediatric mental health care, many of these visits are not helpful to the patient (Grupp-Phelan et al., 2007); potentially resulting in additional hospital and ED burden for the

same untreated condition. Surveillance systems that measure and monitor patterns and trends of pediatric and adolescent mental-health ED visits and hospitalizations have the potential to greatly enhance efforts to address these issues and to inform mental health-related public health policy. Such systems enhance efforts to target interventions, support evaluation of programs and policy, enable the estimation and monitoring of health disparities, and provide detailed characterization of mental health patterns that can be used to inform decision makers and their resulting policy formation.

In this study, we describe the spatial pattern and trends in Mental Health (MH) burden in emergency departments in New York City from 2006-2017. We do this by developing a surveillance system that can be employed prospectively. The system relies on two data sources: the New York Statewide Planning and Research Cooperative System (SPARCS) and the New York City emergency department syndromic surveillance system. Using this system, area-based child mental health disparity metrics are constructed at three geographic scales, census tract, Neighborhood Tabulation Area and Zip Code Tabulation Area. These metrics together with additional geographic information characterizing communities, service availability, and the built environment can be used to evaluate citywide and targeted public health policy.

We find that the resulting surveillance system can be used to identify patterns of mental health service utilization. Importantly, surveillance based on natural language processing of chief complaints found in syndromic surveillance data reflects patterns similar to those found in the SPARCS system that use standardized diagnostic coding. This suggest the potential for prospective monitoring of child mental health in New York City.

keywords: childhood mental health, NYC SPARCs, syndromic surveillance systems.

3.1 Introduction

The global burden of disease due to mental health has been increasing over the last two decades, with 10-20% of children and adolescents worldwide currently affected by a mental health condition (Kieling et al., 2011). The early onset of mental health disorders, their impact at the individual, family and community levels, and their long-lasting effects throughout the life course contribute to their role as a leading cause of morbidity and disability both in the United States (US) and abroad(Kieling et al., 2011; Perou et al., 2013). Surveillance data from 1994 through 2011 show that the prevalence of mental health conditions in the US have increased, with an estimated 13-20% of children experiencing a mental disorder each year (Perou et al., 2013). The annual rate of mental health-related emergency department (ED) visits in the US has increased over 80% among children and adolescents (Grupp-Phelan et al., 2007; Pittsenbarger and Mannix, 2014; Simon and Schoendorf, 2014).

In New York City (NYC), mental health related inpatient hospitalization rates among children are greater among adolescents age 13-17 years than younger children, and increased roughly 23% between 2000 and 2013 (Mills and Davila, 2016). Appropriate preventive treatment of mental health conditions in children has impacts well beyond the condition treated, and can significantly impact physical health, educational outcomes, adult mental health status and economic trajectories (Kieling et al., 2011; Kessler et al., 1995). ThriveNYC, an ambitious and comprehensive governmental initiative to better understand and address the problems of mental health among New Yorkers, was begun in 2015 (McCray et al., 2015). A key objective of the initiative has been to support innovation and evidence-based mental health care practices and create more equitable and responsive systems for better collecting, sharing, and using information and data related to mental health (McCray et al., 2015).

An important input into these efforts is data that can be used to characterize mental health status and outcomes. Surveillance systems that measure and monitor patterns and trends of pediatric and adolescent mental-health ED visits and hospitalizations have the potential to greatly enhance efforts, including ThriveNYC, to understand underlying mental health patterns and inform mental health-related public health policy. The systems can be used to enhance efforts to target interventions, support evaluation of programs and policy, enable the estimation and monitoring of health disparities, and provide detailed characterization of mental health patterns for use by decision makers in the formation of new policies.

Although, an extensive literature on MH burden and disparities already exists, there are issues with the existing evidence base used in these studies. The first limitation is the lack of comparability in their results because they are based on different data sources, methods and designs. For example, in the 8 studies cited here that construct trends or disparity measures[1]. , a variety of time frames, age cutoffs and case definitions were used. Additionally, some studies used retrospective cross-sectional survey designs based on diagnosis code and reason for visit data, while others used routinely collected information intended for billing claims. In almost all cases the source of data used came from survey designs, surveillance systems, or administrative records that were not originally designed to measure mental health burden or disparities. All of these sources of non-comparability are consequently in the resulting estimates and are particularly problematic from the perspective of monitoring of trends. A key design element in any mental health disparity monitoring system is that it can be routinely updated and provide comparable estimates through time. None of the reviewed studies reviewed, or the sources of data they used, would satisfy that need.

[1](Simon and Schoendorf, 2014; Pittsenbarger and Mannix, 2014; Perou et al., 2013; Grupp-Phelan et al., 2007; Hakenewerth et al., 2013; Alegría et al., 2015; Carubia et al., 2016; Christodulu et al., 2002)

The first aim of this study is to report on the design and use of a surveillance system that addresses the shortcomings of prior systems for monitoring MH burden and disparities. The surveillance system measures the mental health burden based on ED visits and hospitalizations using two data sources: the New York State Planning Research Cooperative System (SPARCS) and the New York City syndromic surveillance system. SPARCS contains data from all ED visits and hospitalizations in the state. It is coded using standard ICD9 and ICD10 codes that most national mental health trend estimates are based on, but it is released with a considerable lag. Syndromic surveillance data registers the chief complaint for almost all ED visits in New York City in the form of an open text string. Syndromic surveillance data is received at the end of each day and is processed in near real-time. Although initially developed for bioterrorism and outbreak detection, it has proved very useful for chronic conditions. For example, public health surveillance of mental health conditions using hospital data has been useful within the North Carolina syndromic system (Hakenewerth et al., 2013).

While the system developed will inform public health practice in New York City as a whole, the motivation for this project is to demonstrate how the system can be used to support local public policy formation and analysis at a variety of geographic scales. First, surveillance system information is used to characterize the spatial pattern of child mental health ED visits and hospitalizations. Those results are then combined with area-based poverty measures (ABPMs) (Krieger et al., 2003) to estimate child mental health disparities at the neighborhood-level as defined by the Neighborhood Tabulation Areas (NTAs) and Zip Code Tabulation areas. The resulting disparity estimates provide a basis for monitoring the target population during the implementation of the ThriveNYC project. Second, the system is used to explore an important local policy issue, the possible overuse of 911 calls or Emergency Medical Services (EMS) as a response to behavioral issues in schools. This issue has been suggested as a key outcome measure for several of the ThriveNYC initiatives. We will attempt to identify and characterize areas of the city with elevated burdens of Mental Health during school hours.

The plan of the paper is as follows. The next section provides an overview of child mental health disparities, a discussion of the origins and policy context that resulted in the overuse of 911 by schools, and a review of datasets available for surveillance. The methods section describes the approaches used for data processing, the identification of spatial patterns, and the construction of prevalence trend estimates and disparity measures at various spatial scales. The final two sections present the results and a discussion.

3.2 Background and Data

3.2.1 Child mental health disparities

There are various risk factors that can contribute to MH disorders. These include genetic background, deficiencies in psychosocial or educational environment, exposure to harmful substances, and childhood exposure to adverse events - including domestic violence, neglect, abuse, family financial strain or divorce - as well as certain community conditions, such as unsafe neighborhoods (McCray et al., 2015; Alegría et al., 2015; Aneshensel, 2009; Kieling et al., 2011). Each additional exposure to one of these adverse events further increases the risk of developing MH problems in childhood and all are associated with chronic diseases and threats to MH in adulthood (Alegría et al., 2015; Kieling et al., 2011). Many risk factors are more common in low-income communities, and children from low-income families are disproportionately likely to experience MH problems. Poverty has multiple indirect effects on children's emotional and behavioral development. This results in the need for disproportionate MH services among low-income children. Further, the longer a child lives in poverty the greater the likelihood the child will develop MH disorders (Alegría et al., 2015; Bringewatt and Gershoff, 2010).

MH disparity is defined as the disproportionate amount of psychopathology among people of a disadvantageous social standing. Variation in MH disparities among subgroups in a population can be understood by examining social inequities. Low socioeconomic status (SES) has been shown to be the strongest predictor in early childhood of the development of emotional problems by the age of 18 (Bringewatt and Gershoff, 2010). In NYC, most of the children with MH disorders live in poverty; for example, roughly 70% of children ages 2-12 whose parents reported their child being diagnosed with at least five common MH disorders live in poverty (McCray et al., 2015). Among poor children, MH has been associated with impaired cognitive development, low self-esteem, discrimination, and poor mental and physical health in adulthood (Wickham et al., 2017). Children manifesting these problems are frequently given psychiatric labels that connote internal pathological conditions. In reality, many of these symptoms would not develop but for the environmental circumstances.

Race and ethnicity are also correlated with adverse MH outcomes and disparities. In NYC, the distribution of MH prevalence, diagnosis, and treatment varies strongly by income as well as by racial and ethnic groups. Evidence suggests that the receipt of treatment for MH problems is lower for blacks and Latinos. However, while blacks are half as likely as whites to receive community-based MH care, they are twice as likely to be hospitalized for MH illness (McCray et al., 2015). Economic opportunity, urban design, neighborhood effects, and public safety must also be accounted for when examining how mental illness varies throughout NYC neighborhoods. Neighborhood effects in the form of safety, walkability, aesthetics, noise, housing quality, and social cohesion can cause or alleviate stress that modifies the risk of MH conditions (Alegría et al., 2015; Aneshensel, 2009). These stressors distributed unequally and tend to be more prevalent in neighborhoods that have suffered from structural inequality and racial discrimination (McCray et al., 2015), The chance of psychiatric hospitalization in NYC varies by income and neighborhood. Children from the lowest income neighborhoods are twice as likely to be hospitalized for MH disorders compared to their high-income neighborhood counterparts (McCray et al., 2015). Variations among neighborhoods may also reflect a lack of options for residents to address their mental health needs including access to care (Calman et al., 2006).

The provision of MH services is influenced by patterns of social organization, status, age, poverty, race, and other factors. Low-income communities are not only at a greater risk for developing MH problems, but the communities are also less equipped to treat them (McCray et al., 2015). Lack of access to MH services is problematic in low-income, urban communities. In NYC, according to a study of three of the five boroughs MH treatment slot capacity, there are available slots for only 12% of children ages 5-17 who have treatment needs in the Bronx, Brooklyn, and Staten Island (Citizens' Committee for Children of New York, 2012). Even though people of color and those in poverty bear the greatest MH burden, they are the least likely to get help (McCray et al., 2015; Acri et al., 2016; Bringewatt and Gershoff, 2010).

Socio-economic inequalities contribute to discrepancies across all kinds of health outcomes. However, MH outcomes are particularly neglected since traditionally they have not been treated with as much urgency as physical health problems. Some estimates indicate that nearly 30% of children with emotional, mental, or behavioral conditions experienced problems with access to medical care (e.g., delays, unmet needs) compared to 17% of all children (Child and Initiative, 2012). Furthermore, there is a shortage of qualified professionals to meet patient demand, particularly for low-income patients (Acri et al., 2016; Carubia et al., 2016; Citizens' Committee for Children of New York, 2012). Estimates from 2012 indicate that there are only 8,300 qualified professionals in the US, whereas the projected need is 30,000 (Carubia et al., 2016). Additionally, even when MH specialists are available, there is an overall inconstancy of care. Consequently, people living in communities that are underserved by qualified professionals are turning elsewhere for MH treatment. At the national level, nearly 75% of children exhibiting psychiatric symptoms are seen by a pediatrician rather than a MH specialist (Acri et al., 2016).

Additional barriers of access to MH services that parents face prevent the provision of appropriate care for their children. Many of these barriers are particularly common in low-income communities and include financial, logistical, and legal obstacles, lack of information and a misunderstanding about different agency responsibilities in serving children, a general distrust of the system, as well as fear and stigma associated with MH problems (McCray et al., 2015; Bringewatt and Gershoff, 2010). In NYC, MH services are not located based on need, but are located in areas of concentrated wealth, meaning there are even fewer options for low-income communities (McCray et al., 2015; Acri et al., 2016; Citizens' Committee for Children of New York, 2012).

MH disparities are also exacerbated by disparities in the healthcare system. Health insurance is a major determinant in a person's ability to access medical care, and race/ethnicity are closely linked to insurance status. For example, blacks and Hispanics are more likely to have public insurance or be uninsured (Calman et al., 2006), both of which limit their access to medical care. However, insurance coverage does not necessarily ensure access to healthcare. Even conditioning on differential insurance coverage, separate and unequal systems of care exist within health care institutions (Calman et al., 2006). Although many low-income children are eligible for Medicaid, they often do not receive the proper MH screening to which they are entitled (Bringewatt and Gershoff, 2010). Medicaid also does not reimburse many prevention-related services that are necessary for addressing MH needs. In the end, neither private nor public coverage guarantees access to MH services. Worse still, uninsured patients are generally charged a hospitals' highest fees, while insurance companies routinely negotiate discounted hospital rates on behalf of those covered by their plans. However, the government offers no regulation of the fees hospitals charge to the uninsured (Bringewatt and Gershoff, 2010; Calman et al., 2006). In NYC, because people of color are more likely to be uninsured, they are more likely to incur medical debt or delay the medical treatment that they are unable to afford (Calman et al., 2006).

Public hospitals are far more likely to care for poor, underinsured, and uninsured patients (Calman et al., 2006). These hospitals are already more likely to be overburdened since they serve poorer communities that lack many options for medical treatment. Such burdens create additional healthcare challenges in these communities, including waitlists that delay service initiation, and poorly coordinated care that provide ineffective services (Carubia et al., 2016). The average time it takes to get an appointment ranges from 4-6 weeks, but is sometimes as long as 12 weeks (Citizens' Committee for Children of New York, 2012). In NYC, most community-based outpatient MH care providers are reported to operate overcapacity and without adequate resources (Carubia et al., 2016).

Evidence suggests that patients are presenting to ERs in increasing numbers as a way to have their non-urgent MH needs taken care of because they are unable to get care in other settings (Carubia et al., 2016; Simon and Schoendorf, 2014). In NYC, the percentage of hospitalizations among children and adolescents for psychiatric conditions was higher than the national average (11% compared to 10% nationally) (Mills and Davila, 2016). As a consequence, overburdened ERs in poor communities experience a decrease in quality of care. Longer wait times are also a problem in these communities and often result in patients leaving prior to receiving assessment or intervention (Kessler et al., 1995). Patients also postpone screening for and treatment of illness until their symptoms become serious enough, thus requiring more intensive and costly curative services.

National policy has repeatedly attempted to address this. The Mental Health and Substance Abuse Parity Act of 1998 tried to put mental health and substance abuse treatment on equal footing with respect to insurance policy. This resulted in a number of unintended consequences including insurance plans dropping coverage of psychiatric conditions, increased premiums, cuts to other benefits, and increased control over reimbursement. The updated laws – the Mental Health Parity and Addiction Equity Act of 2008 and the Affordable Care Act – attempted to correct these issues (Mechanic and Olfson, 2016). For New York City, all children living in poverty are eligible to receive public insurance complying with parity laws. As a result most New York City children have insurance that covers mental health services. These services though are generally not available.

3.2.2 Local Policy Background

In 2004, the Daily News published a story highlighting NYC schools' increased use of 911 calls – Emergency Medical Service (EMS) – to handle misbehaving students students (Morgan and Gendar, 2004). Doctors were noticing increasing numbers of students being sent to the emergency departments (EDs) by schools. NYC Department of Education (DOE) responded by stating that there are procedures in place which are to: 1) attempt to calm the student, 2) call school safety, 3) call the child's parent, and only as a last resort, 4) call EMS. Further DOE noted that the instances mentioned in the news story were extreme situations where principals were making the best decisions they could. DOE refused to release any information on how many students are sent to the ER every year. In fact, instances like those described in the news story were not uncommon in NYC public schools and the chain of events that take place when a disruptive student is sent to the ER is not only increasingly common but can be very traumatizing for the student. The underlying problem is that the guidance provided by DOE is vague and there are no clear restrictions on when not to use 911.

The only guidance given by DOE as to when 911 calls should be made in the school-setting is provided in the following three Chancellor's Regulations: (1) Security in Schools – regulation A-412, which states that 911 should be called if an individual requires medical attention; (2) Suicide Prevention/Intervention – regulation A-755, which states that 911 should be called when staff has knowledge of a suicide attempt or where appropriate if a staff member becomes aware of suicidal behavior or ideation; and (3) School Health Services – regulation A-701, which states that if the student's condition warrants more emergency care than can be given in the school, 911 must be called.

Despite media coverage these practices continued to increase, and in 2012, the New York Times published a story highlighting the schools' use of EMS to handle misbehaving students (Winerip, 2012). The article tells the story of a 2nd grade student, Gabriel, who had been sent to the ER by his public elementary school multiple times over the school year. Gabriel had been diagnosed with attention deficit disorder, opposition defiant disorder, and was supposed to be receiving several special education services that the school did not have enough money to provide. In December of 2013, a lawsuit was filed on behalf of six children and their parents. The families represented in the lawsuit were black or Latino, lived in low-income neighborhoods, and the children all had disabilities and had been repeatedly removed or threatened with removal by EMS. The lawsuit alleged that school personnel resorted to calling EMS in response to tantrums and behavioral problems, that schools lacked procedures and systems to support student MH, and that parents were not involved in the decision-making process with regard to whether to call EMS.

In December 2014, a settlement was agreed to with NYC that, in addition to the monetary relief awarded to the plaintiffs, provided for systematic changes in the DOE. These changes included increased trainings and additional resources for schools, and were designed to help schools better handle students in serious MH crisis and prevent unwanted outcomes. New DOE policies and protocols were created for the removal of students engaged in serious disruptive behaviors by EMS, which were adopted by the NYC DOE as a Chancellor's regulation, A-411, issued in May 2015 titled "Behavioral crisis de-escalation/intervention and contacting 911". A-411 was implemented in August at the start of the 2015-16 school year and introduced narrowed guidance for NYC public school staff about when EMS should be called for the first time.

Nonetheless, there is still concern about the use of EMS by school staff and school safety agents to address behavioral issues. This reflects the broader policy context that has emerged over the last two decades in New York City across multiple sectors including Education, Public Safety, Medicine, and Public Health. In education sector, increased emphasis on "high stakes" testing has been accompanied by expanded autonomy by principals. This has led many schools to de-emphasize aspects of education not related to content covered by the testing. Areas de-emphasized include physical and health education, the arts, and programs designed to promote social and emotional development. Principals in New York City have the ability to ignore state-regulated physical and health education requirements without penalty, in pursuit of improved math and English scores. It is not surprising that the increased use of 911 and suspension to address behavioral issues flourished in this context. Removal of students was motivated not from the perspective of the disruptive child's mental well-being, but as an attempt to decrease disruptions to other students thereby improving instruction and ultimately test scores.

In 1998, Mayor Giuliani led a campaign to transfer responsibility for school safety from the Board of Education to the New York City Police Department. Placing school safety outside of the education sector gives care of disruptive children to a sector that does not view its actions or responsibilities from the perspective a child's health or educational needs. For instance, school safety has a limited role which is to minimize incidents, thus ignoring the impact that specific actions might have on education and MH. This policy by definition increases the level to which public students are policed and allows discipline issues to be viewed as legal infractions. Further from 1998 to 2017, the number of School Safety officers increased from 2,000 to 5,000 at which point they outnumbered guidance counselors by 1,000. Placing school safety outside of the education sector also allows principals to use 911 without bearing the responsibility or cost of doing so. When students or parents have issues with treatment received by School Safety offices, they have to be filed with Bureau of Internal Affairs at the NYPD.

During this same period, the public health sector has been increasingly focused on the social determinants of health, including poverty, adequate housing, legal services, and education. By viewing these factors as inextricably linked to health, public health interventions have broadened into areas traditionally served by other agencies. In NYC this has caused tension between Mayoral agencies and the DOE. In the medical sector, a movement towards prevention and away from fee-for-service care has also increased tension. Previously, city agencies including public health and education generated referrals to clinical care. Addressing health issues preventatively within schools conflicts with a medical model that treats illness and responds to incidents on an individual fee-based basis.

In January 2015, the NYC Mayor's Office launched ThriveNYC: A Mental Health Roadmap for All (ThriveNYC) (McCray et al., 2015). ThriveNYC is based on principles developed through comprehensive research, the experience of other cities and countries, input from hundreds of local organizations working to promote MH, and individual New Yorkers with mental illnesses. The principles were advanced in part through 54 initiatives, which represented an investment of $850 million over four years. Together they comprise an entirely new and more holistic approach to MH in NYC, and set a foundation for responding to this public health challenge in the years ahead.

Notably, one of the guiding principles of ThriveNYC is "Act Early", which emphasizes prevention and early intervention; staples of the public health approach. This guiding principle informs the design and implementation of ThriveNYC MH initiatives and represents an innovative approach to MH service delivery. In addition to improving the culture, services, and access for MH, several of the initiatives being introduced directly support schools in complying with regulation A-411 and aim to reduce the number of these ER visits for disruptive students. Of the 54 initiatives, three are housed directly within the New York City Office of School Health, a joint program of New York City Departments of Health and Education. Those three initiatives focus on delivery mental health services and linking students to community resources. The proposed outcome metric that will be used to monitor progress on the initiatives is reduction in the use of 911 by schools and the associated mental health burden from ED visits and hospitalizations.

3.2.3 Child mental health data

Public Health monitoring and surveillance requires information that is regularly updated and isn't subject to changes in coding practice or health seeking behavior. For mental health, information on conditions and outcomes can be taken from surveys, administrative records, or surveys of administrative records. For children, surveys often involve interviews with the "most knowledgeable care provider" or a "sufficiently knowledgeable care provider" and generally rely on questions that establish a past or current diagnosis. These questions can only establish prevalence of diagnosis rather than the underlying prevalence of the condition. Adolescents (and adults) are often queried directly allowing for direct screening. However, most surveillance systems based on surveys such as the Youth Risk Behavior Surveillance System (YRBS) are multi-purpose health surveys with limited space for medical health modules and question development and validation for mental heath-related survey items is difficult (Kessler et al., 2002). For example, the Kessler 6 score for non-specific psychological distress involves six questions and may vary by language and race (Kessler et al., 2002; Kim et al., 2016).

In New York City, the Kessler 6 has been implemented in the NYC Community Health Survey (CHS), an annual survey of adults used for surveillance and monitoring. The YRBS is a biannual survey of high school students that has been in the field since 1997 and generally contains a mental health module. Since the YRBS is a national survey, implemented by states and localities, the responses can be compared among location and to nation trends. Other child surveys include the NYC Child Health Survey and the NYC Child Health, Emotional Wellness, and Development Surveyhave collected mental health information for younger children relying on reports/interviews from sufficiently knowledgeable care providers. These surveys are not repeated regularly enough to provide surveillance information. Although the YRBS provides useful information, its design does not allow for estimates beyond a YRBS-specific geographical scale that breaks the city into 8 areas, borough by a smaller geography currently termed "neighborhood health action centers. "

Administrative records record medical events including diagnoses from emergency department admission files, electronic health records, hospital billing files, and screening forms. Administrative records of diagnoses suffer from their reliance on the availability of mental health services and differential health seeking behavior. There are also no current population-representative datasets that integrate data from electronic health records from primary care and mental health providers. New York State Medicaid records contain this information for children on Medicaid, which is over half of New York City youth, but confidentiality restrictions preclude its use as an input to a surveillance system. Surveys of administrative records, include national sentinel systems that construct estimates from, for instance, samples of emergency department or hospital records and then perform chart review to assess the accuracy of classifications are expensive and difficult to employ at the local scale. .

The mental health surveillance systems developed here, relies on two data sources that reflect the use of hospital emergency departments to treat acute mental health events. As such, surveillance takes place at a higher level than primary care in the "surveillance pyramid" discussed in chapter 1. The pyramid with death at the tip, and outcomes arranged by severity level forming layers beneath to represent, moving downwards, increased prevalence and less severity. Generally, hospitalizations and emergency department visits comprise the next two levels, followed by medically attended cases and prevalent cases not medically attended. Focusing surveillance on more severe events may be particularly useful for monitoring mental health disparities. For mental health, hospitalizations and ED visits are often labeled "potentially avoidable" reflecting a lack of appropriate preventive mental health services (Hsia and Niedzwiecki, 2017). By focusing, on this level we are able to measure the mental health burden on hospitals and emergency rooms due to a lack of mental health services, directly reflecting disparities. The two data systems are the New York State Planning Research Cooperative System (SPARCS) and the New York City syndromic surveillance system. SPARCS contains data for all ED visits and hospitalizations in the state and is coded using standard ICD9 and ICD10 codes that are the basis for most national mental health trend estimates. The drawback of SPARCS is that it is released with considerable lag. Syndromic surveillance data registers almost all ED visits in New York City based on chief complaint which is recorded as an open text string. Syndromic surveillance data is received at the end of each day and is processed in near real-time.

SPARCS data was accessed from 2006 through the end of 2016. The data is recorded as a line listing with one record per visit for all ED visits and hospitalizations in New York City. The data contains date, time of day, age, detailed address information, diagnostics and procedural codes, and other details. SPARCS is produced by the New York State Department of Health and based on billing claims. SPARCS represents an initial step in the development of an all payers claims database that exist in many states and underpin studies in the Health Services Research field. SPARCS is released at a specific time and so contains a lag ranging from 3 to 15 months. Additional latency results from processing requirements such as geocoding addresses.

The New York City syndromic surveillance systems was originally developed to be used for outbreak detection and bioterrorism response (Pavlin et al., 2003). National security concerns and significant advances in information systems led to the widespread use of syndromic surveillance across multiple public health objectives (Heffernan et al., 2004; Mostashari and Hartman, 2003), and syndromic surveillance data have been validated against traditional laboratory and mortality surveillance for influenza (Olson et al., 2007; Buehler et al., 2008), and representative population survey data (Metzger et al., 2004; Hakenewerth et al., 2009). The New York City syndromic surveillance system is updated each day with the previous day's visits. In existence since 2002, we use data from 2006 through 2017 for this study, although the proposed surveillance framework could, in fact, be implemented prospectively in near real-time. For mental health outcomes, the timeliness is less essential but it is notable that a syndromic-based mental health surveillance can be brought up-to-date at any time. Since syndromic surveillance is based on the chief complaint text string it is to be expected that sensitivity and specificity of classification of mental health will be low. A key requirement in the development of a syndromic surveillance system is that the patterns obtain using the system matches the underlying true pattern that can only be assessed with less timely data (or not assessed at all).

Auxiliary information used in this study includes the locations of schools in which

ThriveNYC programs have been targeted – Community Schools and the Prevention and Intervention (PIP) Program. This includes the New York City Community School Mental Health program which is targeted within the DOE's Community School program. Community Schools represent an approach to educational reform that emphasizes the link between school and community and focuses efforts on increasing ties to the community (Oakes et al., 2017) and delivering services such as additional vision screening, asthma case management, and providing mental health managers. Mental health managers work within individual schools to coordinate the delivery of in-school mental health services and develop relationships with community-based mental health organizations. The PIP program represents an expansion of the Community School programs chosen because they represented a disproportionate share of suspensions and mental health issues. A third school-based ThriveNYC program exists covering the balance of schools and employing mental health consultants who cover portfolios of schools and seek to link schools with community-based services.

Additional auxiliary information includes shapefiles identifying New York City's Mental Health Provider Shortage Areas (MHPSAs). These areas are national designations that enable clinics to receive Federally Qualified Health Center and additional funds A national formula drives the identification of these areas that may not effectively reflect shortage areas in specific geographies (Oakes et al., 2017). Additionally, the designation of MHPSAs incentivizes clinics working in these areas to expand their services. It will not be clear how MHPSA designation impacts mental health without additional timing information which is not available. Lastly, official area-based poverty measures (ABPMs), neighborhood tabulation areas (NTAs), population estimates by ABPM and NTAs, and DOE school calendars were also used in this study.

3.3 Methods

New York City syndromic surveillance and SPARCs data were processed into diagnostic and syndrome categories following groupings developed elsewhere. New York City mental health syndrome codings initially followed North Carolina (Hakenewerth et al., 2013) but were subsequently modified in an iterative procedure by the Bureau of Children Youth and Families in the NYC Department of Health of Mental Hygiene. The resulting syndrome coding constitutes official city definitions and are given in Table 3.3. Classification of SPARCS data using ICD9 codes followed Simon 2014. In October 2015, ICD10 codes for classifying morbidity went into effect. National recommendations directly map ICD9 to ICD10 codes for mental health conditions and do not recommend further action when constructing trends. ICD9 and ICD10 codes and definitions are given in Tables 3.1 and 3.2. Additionally, data was classified by age group into four age groups: 5 to 9,10 to 12, 13 to 17, and 18 to 19. These age groupings represent the smallest groups for which annual population estimates are available. Both datasets are ostensibly censuses of NYC hospital emergency department visits. The NYC syndromic system has varied from 90 to 98% coverage over its existence and contains all check-in records at the ED's intake. SPARCs data by law covers all visits seen at EDs and hospitals in New York State. Nonetheless, SPARCs does not include patients who eloped before being seen. Provision of pediatric mental health services in emergency care settings is incredibly inefficient and may result in relatively high proportions of patients not being seen (Tucci et al., 2015).

Spatial information exists in both syndromic surveillance and SPARCS data. Since its inception, the New York City syndromic surveillance has collected zip code information for use in spatial cluster detection. However, this system is voluntarily provided by hospitals with the requirement that it is processed daily. Abstraction of zip code is relatively easy and does not require detailed geocoding procedures. The system collects no address information. There are 181 populated zip code tabulation areas (ZCTAs) in New York City. A number of other zip codes have been linked with specific ZCTAs. Because the syndromic surveillance system covers all visits to NYC emergency departments, child mental health visits are observed from children throughout the world. We subset the records to include only those providing a zip code that can be linked to a ZCTA after initial processing. SPARCs data includes full address information that is processed upon receipt from New York State. The data set includes geographic coordinates, in addition to all standard geographies used in public health practice. We focus on NTAs, a scale roughly equivalent to ZCTAs with lower variation in size, and census tracts that have ABPM designations and are highly specific geographically. Since one of our aims is to detect areas of the city with a high reliance on EDs for mental health services during school hours, census block and actual coordinates may provide additional details.

Both SPARCs and syndromic data are processed at the individual level and classified as mental health or non-mental health visits. ABPM categories are attached by specific geographies: NTA, ZCTA, and census tract. Two types of measures can be constructed for surveillance. The ideal measure for surveillance is the population-based rate of ED mental health visit. This requires trust in both numerators and denominators. For SPARCs, we calculate the actual count of visits encoded as mental health visits following national best practices and population rates can be directly computed. Uncertainty in this setting with no sampling error would rely on a super-population view of the data generating mechanism; New York City as a realization of possible NYC realizations. In the case of rates, these are straightforward analytic calculations. More complex statistics can be assessed using resampling. For syndromic estimates, low sensitivity and specificity result in counts that are not thought to be actual counts and are not generally publicly reported. Instead syndromic data is usually viewed as the percent of visits that were classified as mental health visits. When making trends this only relies on the assumption of stable misclassification of visits based on chief complaints. Calculations based on syndromic surveillance do not yield estimates that can directly represent medical or economic burden. Linking retrospective SPARCs burden estimates to prospective syndromic estimates provides a way forward but is not pursued here.

Overall time series are presented, broken out by age group, ABPM designation. We then establish the correspondence between syndromic and SPARCs data using correlation across temporal and spatial patterns. Trends in ED visits rates, overall, by poverty and ABPM status are estimated and described. We then present a disparity measure based on the relative increase in school hour mental health burden between rich and poor neighborhoods and suggest one approach to evaluation of the population-level impact of ThriveNYC mental health program. Finally, we construct detailed maps of all quantities discussed with the aim of identifying areas or clusters with elevated rates of mental health burden during school hours. Although these quantities involve more detailed definitions than over all totals, uncertainty estimates for rates and relative risks can be calculated directly or by resampling.

3.4 Results

Table 3.4 gives yearly counts for mental health-related and total emergency department visits for children aged 5-19 for both SPARCS and syndromic surveillance data. Although both data systems receive nearly all visits and therefore should be very close in total visits, differences do exist. For example, total visit counts are lower in the syndromic surveillance system for 5-12 year olds but greater for 13-17 year-olds. This could reflect a number of issues including lower completeness of syndromic surveillance by hospital, higher syndromic surveillance counts per hospital because visits are captured at intake rather than discharge, and differences in the ability to specify New York City in both system. SPARCs records a higher proportion of total visits as mental-health related ranging from 3.4 to 4.5 percent for 5-12 year-olds and 7 and 11 percent for 13-17 year olds versus ranges of 1.8 to 3.1 percent for 5-12 year-olds and 4.4 and 8.7 percent for syndromic surveillance. This reflects reduced sensitivity in syndromic due to its reliance on a short text string instead of diagnostic codes. Increases in the proportions of visits that are mental health-related visits are clearly evident in both data systems.

Figure 3.1 gives daily counts and 28-day moving averages for all ages for mental health visits from 2006 to 2016 using syndromic surveillance coding. Daily mental health-related visits range from 130 to 205 in New York City over the period covered. There is a clear day of week pattern, a mildly increasing trend over the observed period, and a slight but discernible seasonal pattern with fewer visits in summer and during the winter holiday periods.

Figure 3.2 presents the same information for three pediatric age groups, 0 to 4, 5 to 12, and 13 to 17. There are almost no mental health-related visits in children younger than 5. The older age groups display the same pattern as the all ages chart but with more pronounced day of week, and seasonal patterns, Figure 3.3 presents the data for 5 to 12 year olds for the period from July 1st, 2014 to June 30th, 2015 covering the 2014/2015 school year. The summer recess and three school breaks are shaded grey. There are very few mental health-related visits during these periods. Weekly peaks during school periods are about seven times higher than weekly troughs, as opposed to the doubling seen in time series for all ages given in figure 3.1. Altogether, school days have greatly increased counts of mental health-related ED visits.

Spatial patterns during school days and school hours for 2008 and 2016 are given in figure 3.4. The maps present population-based rates per 1000 children per year by Neighborhood Tabulation Areas (NTA) for 5 to 17 year olds using SPARCs and classified by ICD9 or ICD10 codes. The top row corresponds to 2008 and the bottom row to 2016 while left panel presents ED visit rates for school hours during school days and the right panel presents visits during school hours on non-school days. All maps are shaded on the same scale. A clear spatial pattern is evident on school days with a number of generally poor NTAs having visits in excess of 18 per 1,000 children as compared with a rates below 6 in a variety of wealthier areas. During non-school days, most NTAs in the city have rates below 6 per 1000 with a few areas exceeding 6 per 1000. The same pattern has persisted over the study period with 2008 and 2016 clearly displaying similar patterns.

The disparity, clear from figure 3.4, can be summarized using disparity measures based on area-based poverty measures developed in the previous chapter. Figure 3.5 presents the numerators and denominators for the rate ratio disparity measure that expresses the rate for persistently poor areas relative to the rate for persistently rich areas. The poor areas are given in solid red for school days and dashed red for non-school days with wealthy areas presented in blue. The rates for poor areas are consistently 3 to 4 times elevated over wealthy areas during school days. Both groups have increased from 2008 with slight declines from 2014 to 2016. The non-school day rates are also elevated to a similar magnitude suggesting an underlying poverty pattern that is only slightly modified by school days and school hours.

Advancing the goal of prospective surveillance of child mental health outcomes requires establishing the utility of near real-rime data to characterize the patterns presented. We compare patterns establish using SPARCs to syndromic surveillance based patterns. For practical reasons, syndromic surveillance is typically analyzed as proportions of total visits. We examine syndromic surveillance proportions and population-based rates as compared to SPARCs rates. The only geographical information available in syndromic surveillance data is patient zip code as given at intake and the location of the hospital. Characterizing spatial patterns relies on zip code information assigned to zip code tabulation areas (ZCTAs).

Figure 3.6 presents spatial patterns of population-based rates at the ZCTA scale for SPARCs (left) and syndromic surveillance (right) on the same scale for 5-17 year-olds for 2016 all days and all hours. The same pattern is evident in both data streams with lower rates clear in syndromic reflecting the lower sensitivity of text processing as compared to diagnostic coding.

Figure 3.7 presents the same data using quintiles to show correspondence in the rank patterns of the two data streams. There is high correspondence between SPARCs and syndromic spatial patterns with discordance in Staten Island (shape at lower left) reflecting known problems with chief complaint coding practice at Staten Islands largest hospital. Figure 3.8 presents quintile patterns for proportion of visits that are mental health-related. Notable here is the lack of correspondence between proportion-based metrics and population-based metrics and the lack of discernible poverty pattern. This suggests that syndromic surveillance using proportions of mental health-related ED visits will not adequately capture underlying spatial or temporal patterns of mental health visits and, as such, syndromic surveillance of mental health cannot be directly incorporated into NYCs syndromic surveillance which is based on proportions.

To quantify the correspondence between SPARCs and syndromic surveillance over the 2006 to 2016 period we construct correlation measures in two ways. First, for individual ZCTAs we calculate the correlations, for each spatial unit, between SPARCs and syndromic rates and proportions across time. Secondly, for each time unit we calculate the Spearman rank correlation between maps constructed using syndromic and SPARCs rates and proportions. Figure 3.9 presents maps with correlations between SPARCs and syndromic surveillance measures at the ZCTA geographic and quarter temporal scale. The upper left panel displays correlations by ZCTA for the SPARCs and syndromic population-based rates between 2006 and 2016. Correlations are generally above .7 with some lower correlation in some areas including Staten Island reflecting coding issues. The upper right panel gives the same correlation between 2011 and 2016 to determine if the correspondence might be improving over time. The average correlation in the map increases slightly with same pattern and areas of low correlation. The bottom panel presents within ZCTA correlation between SPARCs population-based rates and syndromic proportions. Correlations are lower than for maps based on rates but still generally high.

Figure 3.10 presents time series of correlation between maps based on SPARCs and syndromic rates and proportions. This measures agreement between maps by quarter for three comparisons. The series for a comparison of population-based rates (blue lines) has high correlations, in the .7 to .8 range, for 5 to 12 year olds and .6 to .7 range for 13 to 17 year olds. Correlation based on proportions (orange lines) are lower in both cases and the correlation between SPARCs rates and syndromic proportions (grey line) is lower still, about .15 lower for 5 to 12 year-olds and .3 lower for 13 to 17 year olds – giving correlation below .5 and usually in the .2 to .4 range. This suggests that syndromic proportions are not adequate for capturing the patterns observed in SPARCs data.

The use of syndromic data to characterize spatial and temporal patterns in adolescent mental health ED visits relies on its direct correspondence to SPARCs population-based rate data which would be the preferred measure were it available. An important use of the system being developed here is the measure and monitoring of disparities. Figure 3.11 presents the disparity measures by quarter for SPARCs and syndromic based on population-based rates. Syndromic (green) tracks the disparity as measured in SPARCs, given in (orange). Interestingly, the measures drift apart beginning in quarter 4 of 2015 for both 5 to12 year olds (dotted) and 13 to 17 year olds (solid). This may reflect the switch from ICD9 to ICD10 codes in October 2015. Syndromic data can be processed in near real-time and an additional year is presented here suggesting improvements in the disparity measure.

3.5 Discussion

It has long been recognized that substantial mental health disparities exist between socio-economic and race groups across multiple dimensions including exposure to risk factors (Barr, 2014). Mental health disparities are complicated, reflecting inequity not just in the prevalence of child mental health conditions but also an inverse disparity in the provision of services. For example, white children are more likely to (have their caregivers) report being diagnosed with a mental health condition and to be receiving mental health services (Lu, 2017; McGuire et al., 2006), but black, Hispanic, and poor children are more likely to have mental health disorders and to be seen in emergency settings (Christodulu et al., 2002; Alegría et al., 2015). Emergency treatment for child mental health conditions is extremely expensive and mostly ineffective (Tucci et al., 2015) . Further, untreated child mental health conditions are associated with adverse mental and physical health outcomes into adulthood which incur enormous costs for the public sector (in the form of Medicaid expenditures and sequelae of untreated mental health conditions) and for the untreated individual and their family Breslau et al. (2006).

Within this context, public health action including surveillance, policy development, and evaluation can play a critical role. Our point here is to demonstrate the utility of a mental health surveillance system combining two data sources, one in near real-time based on a low-specificity, low-sensitivity classification at a suboptimal geographic scale and a retrospective one with official diagnostic coding using full address information with geocoding. We have demonstrated a strong correspondence between the datasets both spatially and temporally. The system developed represents an additional use case for syndromic surveillance systems in public health practice Lall et al. (2017). We have also used the system to construct disparity measures based on ABPMs and presented trends through time. Because pediatric mental health emergency departments are generally considered avoidable, our findings here can play an important targeting role for mental health services. This is particularly relevant in New York City where mental health service expansion is one of the key initiatives of the current mayor suggesting an evaluation use for the system.

The surveillance framework proposed here is easily automated and can run prospectively. Receipt of new SPARCs data requires some processing but allows for updating of disparities measured in two separate systems. Characterization of current child mental health patterns is of direct value; providing information to leadership and policy-makers, enabling targeting of interventions, and allowing for the evaluations of programs. Sub-syndromes within the set of all mental heath conditions, such as suicide, routinely receive heightened concerned from public health and city leadership. This reflects discerned patterns often based on anecdote or driven by the press. The surveillance system presented provides a valuable tool in response. We presented this use-case for the system by identifying areas of the city where schools may be overusing emergency medical services to address behavior problems.

A number of ongoing initiatives are bringing improvements to this data and will further enhance surveillance of mental health and could be incorporated into the surveillance system presented. This presents opportunities for several possible future refinements and uses of the system. First, the recent establishment of a retrospective dataset linking SPARCs to syndromic surveillance will allow for direct assessment of sensitivity and specificity of syndromic surveillance classification to diagnostic measures. Second, a record linkage between the SPARCs system and school records can be employed to directly estimate MH burdens by school, and individual student-level linkages allow for the assessment of the impact of mental health conditions on absenteeism and other educational outcomes. Third, ongoing improvements to the SPARCs system include bringing it in line with the all-payers claims databases from other states (Love et al., 2010) will likely result in reduced delay in reporting and increased standardization. Further development of New York Citys syndromic surveillance system will result in increased completeness of Mode of Arrival and Disposition variables that may allow for an assessment of the experience of children delivered to hospitals as well as increased completeness for ICD10 coding and new fields such as nurse triage notes that could prove valuable in assessing mental health visits. Fourth, although privacy restrictions make it difficult to openly conduct research using Medicaid information, such data is available to public health practitioners with as short as a two week lag and may enable the monitoring at the individual level of the consumption of mental health services both preventative care and in emergency settings. Finally, data collected by New York City's emergency medical service and the Department of Education may provide important additional information.

3.6 Tables and Figures

Table 3.1: ICD 9 codes for mental health diagnosis

Psychosis (290-299)

Organic psychotic conditions

290. Senile and presenile organic psychotic conditions
291. Alcoholic psychoses
292. Drug psychoses
293. Transient organic psychotic conditions
294. Other organic psychotic conditions (chronic)

Other Psychoses

295. Schizophrenic psychoses
296. Affective psychoses
297. Paranoid states
298. Other nonorganic psychoses
299. Psychoses with origin specific to childhood

Neurotic disorders, personality disorders, and other nonpsychotic mental disorders (300-314)

Neurotic disorders

300. Neurotic Disorders

Personality disorders

301. Personality disorders

Sexual deviations and disorders

302. Sexual deviations and disorders

Psychoactive substance

303. Alcohol dependence syndrome
304. Drug dependence
305. Nondependent abuse of drugs

Other (primarily adult onset)

306.Physiological malfunction arising from mental factors
307. Special symptoms or syndromes, not elsewhere classified
308. Acute reaction to stress
309. Adjustment reaction
310. Specific nonpsychotic mental disorders following organic brain damage
311. Depressive disorder

Mental disorders diagnosed in childhood

312. Disturbance of conduct, not elsewhere classified
313. Disturbance of emotions specific to childhood and adolescence
314. Hyperkinetic syndrome of childhood

Other psychological or physical stress not elsewhere classified

V62.84 Suicidal ideation
V62.85 Homicidal ideation

Table 3.1: ICD 9 codes for mental health diagnosis. Note: Codes not included – 315 (Specific delays in development) and 316 (Psychic factors associated with diseases classified elsewhere).

Table 3.2: ICD 10 codes for mental health diagnosis

Mental and behavioral disorders (F00-F99)

Organic, including symptomatic, mental disorders

F00. Dementia in Alzheimer's disease
F01. Vascular dementia
F02. Dementia in other diseases classified elsewhere
F03. Unspecified dementia
F04. Organic amnesic syndrome, not induced by alcohol and other psychoactive substances
F05. Delirium, not induced by alcohol and other psychoactive substances
F06. Other mental disorders due to brain damage and dysfunction and to physical disease
F07. Personality and behavioural disorders due to brain disease, damage and dysfunction
F09. Unspecified organic or symptomatic mental disorder

Mental and behavioural disorders due to psychoactive substance use

F10. Mental and behavioural disorders due to use of alcohol
F11. Mental and behavioural disorders due to use of opioids
F12. Mental and behavioural disorders due to use of cannabinoids
F13. Mental and behavioural disorders due to use of sedatives or hypnotics
F14. Mental and behavioural disorders due to use of cocaine
F15. Mental and behavioural disorders due to use of other stimulants, including caffeine
F16. Mental and behavioural disorders due to use of hallucinogens
F17. Mental and behavioural disorders due to use of tobacco
F18. Mental and behavioural disorders due to use of volatile solvents
F19. Mental and behavioural disorders due to multiple drug use and use of other psychoactive substances

Schizophrenia, schizotypal and delusional disorders

F20. Schizophrenia
F21. Schizotypal disorder
F22. Persistent delusional disorders
F23. Acute and transient psychotic disorders
F24. Induced delusional disorder
F25. Schizoaffective disorders
F28. Other nonorganic psychotic disorders
F29. Unspecified nonorganic psychosis

Mood [affective] disorders

F30. Manic episode
F31. Bipolar affective disorder
F32. Depressive episode
F33. Recurrent depressive disorder
F34. Persistent mood [affective] disorders

— continues on next page →

F38. Other mood [affective] disorders
F39. Unspecified mood [affective] disorder

Neurotic, stress-related and somatoform disorders

F40 .Phobic anxiety disorders
F41. Other anxiety disorders
F42. Obsessive-compulsive disorder
F43. Reaction to severe stress, and adjustment disorders
F44. Dissociative [conversion] disorders
F45. Somatoform disorders
F48. Other neurotic disorders

Behavioral syndromes associated with physiological disturbances and physical factors

F50. Eating disorders
F51. Nonorganic sleep disorders
F52. Sexual dysfunction, not caused by organic disorder or disease
F53. Mental and behavioural disorders associated with the puerperium, not elsewhere classified
F54. Psychological and behavioural factors associated with disorders or diseases classified elsewhere
F55. Abuse of non-dependence-producing substances
F59. Unspecified behavioural syndromes associated with physiological disturbances and physical factors

Disorders of adult personality and behavior

F60. Specific personality disorders
F61. Mixed and other personality disorders
F62. Enduring personality changes, not attributable to brain damage and disease
F63. Habit and impulse disorders
F64. Gender identity disorders
F65. Disorders of sexual preference
F66. Psychological and behavioural disorders associated with sexual development and orientation
F68. Other disorders of adult personality and behaviour
F69. Unspecified disorder of adult personality and behaviour

Disorders of psychological development

F82. Specific developmental disorder of motor function
F84. Pervasive developmental disorders
F88. Other disorders of psychological development
F89. Unspecified disorder of psychological development

Behavioral and emotional disorders with onset usually occurring in childhood and adolescence

F90. Hyperkinetic disorders
F91. Conduct disorders

— continues on next page →

F92. Mixed disorders of conduct and emotions
F93. Emotional disorders with onset specific to childhood
F94. Disorders of social functioning with onset specific to childhood and adolescence
F95. Tic disorders
F98. Other behavioural and emotional disorders with onset usually occurring in childhood and adolescence

Unspecified mental disorder

F99. Mental disorder, not otherwise specified

Other symptoms and signs involving emotional state

R45.850 Homicidal ideation
R45.851 Suicidal ideation (not constituting part of a mental disorder)

Table 3.2: ICD 10 codes for mental health diagnosis. Note: Codes not included – F70-F79 (Mental retardation), F80 (Specific developmental disorders of speech and language), F81 (Specific developmental disorders of scholastic skills), F83 (Mixed specific developmental disorders).

Table 3.3: Regular Expressions for syndromic surveillance coding of mental health.

General Psych [GENPSYCH]

```
(
   m/(\bEDP )|PSY|PHYCH|PY[CS]|CPEP|\bMENTAL  [HE]/
OR
   m/\b (29[0123456789]|30[0123456789]|31[0123456])(\b |\.\d )/
OR
   m/\b (V70\.?[12]|V71\.?X|V11|V40|V61\.?[012]|V62\.?64)/
)
AND NOT
   m/EPSY|IPSY|IOPSY|PSYCHOSIS|MENTAL RETARDATION/
```

Mood Disorders [MOOD]

```
   m/DEPRES|MOOD DIS|MOOD D\bO |MDD|MANI[AC]|BIPOLAR/
OR
   m/\b (296|311|F3[012349])(\b |\.?\d )/
```

Behavior Disorder [BEHAVE]

```
(
   m/AGG?ITA|AGG?RESS?|COMBAT|IRRATION|B[EA]HAV|EMOTIONAL|MELTDOWN|TEMPER\b
   |DEFIANT|CONDUCT|DISTRUPT|VIOLENT|ASSAULTIVE|HYPERA|ADHD|ATTENTION|KICKING|
   OUT OF CONTROL|IMPULSE CONTROL|PERSONALITY/
OR
   (
      m/\bACTING /
   AND
      agenum>1
   )
OR
   (
      m/THROW/
   AND
      m/CHAIR|THINGS/
   AND NOT
      m/THROW(ING)? UP/
   )
OR
   m/\b (312|F91)(\b |\.?\d )/
)
AND NOT
   m/ALLERG|INGESTION|ASTH[MN]A|PERIO|SKIN|EXTERNAL|ATTETION TO|SEXUAL BEHAVIOR
   |UNCERTAIN|SEIZURE|REACT|EYE|ECZEMA/
```

Psychosis [PSYCH]

```
(
   m/PSYCHOSIS|SCHIZ|PARANO|DISORG|CATATON|HALLUC|DELUS|VOICES/
OR
   (
      m/DISTURB/
   AND
      m/CONDUCT|EMOTION/
   )
OR
   (
      m/BORDERLINE/
   AND
      m/\bPD \b /
```

— continues on next page →

```
   )
OR
   (
       m/PERSONALITY/
   AND
        m/DISORDER\b /
   )
OR
   m/\b (29[578]|F2[029])(\b |\.?\d )/
)
```

Anxiety [ANX]

```
(
   m/ANXI|NERVOUS|NEUROTIC|NEUROSIS|HYSTERI|PANIC|ADJUSTMENT DIS|COMPULSIVE|
   CRISIS|DISSOC|PHOBIC|OBSESS/
OR
   m/\b (300|309|301|F43)(\b |\.?\d )/
)
AND NOT
   m/SICK?[EL]{2}|DEVICE|TYMPANIC|REACTION|JAUNDICE|DISEASE WITH CRISIS|

   ANXILLA|ANGIO/
```

Suicide/Suicidal Ideation [SI]

```
(
NOT
    m/(DEN(Y|IES|IED|YING)|\bDEN (Y(ING)?|[IE]{1,2}(S|D|R|\b |\d )[^S]|I\b
    |IAL)|DN(EI|IE)E?(D|S)[^S]|\bNOT ?)*(\.|-|:|\|)? *(ANY|ACTIVE|OBV(IOUS)?)?
    *(S(\/?I|U|ELF| *(AND|\/|\+|\.| ) *H)|(H\/?I? *(AND|\/|\+|\.| ) *(SI?)))/
AND
   (
        (
              m/\bS (H )?I(\b |DEA)|\bS (\.|\/)?(H )?I\.?(\b |DEA)|SU[CDI]{2,}|
              SUI?C(I{0,2}D|IC?)|SUSCID/
        AND NOT
              m/SUDD|SUCCESS/
        )
   OR
        (
              m/WA{1,2}NT(S|ED)*/
        AND
              m/\bDIE \b /
        )
   OR
        (
              m/ID[EAI]{1,3}T[IAU]{0,2}ON/
        AND
            (
                   m/SELF/
            OR
                    m/SU[CDI]+|\bS \//
            )
        )
   OR
        (
              m/\bKILL |([^NG]|\b )HU{1,2}R{1,2}T|SHOO?T|\bHA {1,2}RM|
              MUT(I|U)LAT|STAB[^L]|INFLICK?/
        AND
              m/SELF/
```

— continues on next page →

```
        AND NOT
              m/EPI PEN/
        )
   OR
        m/\b (V62\.?84|R45\.?851|E95[0-9]|T14\.?91)\b /
   )
```

Mental Health [MH]

```
(
    GENPSYCH
OR
    ANX
OR
    MOOD
OR
    BEHAVE
OR
    PSYCH
OR
    SI
)
```

Psychiatric Medications [PSYCHMEDS]

```
    m/MED(S|ICATION)|REFIL|R[AU]N OUT/
```

Table 3.3: Regular Expressions for syndromic surveillance coding of mental health

New York City Emergency Department Visits for School-aged Children 2006-2017

5 to 12 Years Old

	SPARCS			**Syndromic Surveillances**		
YEAR	Total Visits	Mental Health	Proportion of Vistis	Total Visits	Mental Health	Proportion of Vistis
2006	249,273	6,781	2.72%	230,861	5,107	2.21%
2007	254,932	7,591	2.98%	230,623	5,041	2.19%
2008	247,323	7,680	3.11%	241,593	5,881	2.43%
2009	291,420	7,009	2.41%	290,813	5,354	1.84%
2010	258,333	7,067	2.74%	250,423	6,136	2.45%
2011	270,520	8,343	3.08%	262,775	6,618	2.52%
2012	272,012	8,462	3.11%	260,112	6,880	2.65%
2013	271,267	9,352	3.45%	258,726	7,466	2.89%
2014	269,202	9,233	3.43%	259,053	7,187	2.77%
2015	262,894	7,949	3.02%	256,318	6,291	2.45%
2016	270,873	7,781	2.87%	265,478	6,172	2.32%
2017				249,360	7,567	3.03%

13 to 17 Years Old

	SPARCS			**Syndromic Surveillances**		
YEAR	Total Visits	Mental Health	Proportion of Vistis	Total Visits	Mental Health	Proportion of Vistis
2006	152,680	10,644	6.97%	162,065	7,176	4.43%
2007	158,159	12,478	7.89%	162,461	7,726	4.76%
2008	155,697	12,903	8.29%	173,564	9,513	5.48%
2009	171,266	12,845	7.50%	195,307	9,577	4.90%
2010	157,017	12,402	7.90%	175,124	10,386	5.93%
2011	157,741	14,552	9.23%	173,752	11,181	6.44%
2012	156,953	15,263	9.72%	170,322	12,228	7.18%
2013	150,857	15,539	10.30%	164,413	12,817	7.80%
2014	150,869	16,221	10.75%	166,314	13,163	7.91%
2015	147,652	14,753	9.99%	164,567	12,121	7.37%
2016	144,989	14,234	9.82%	162,258	12,010	7.40%
2017				158,618	13,720	8.65%

Table 3.4: NYC Emergency Department Visits

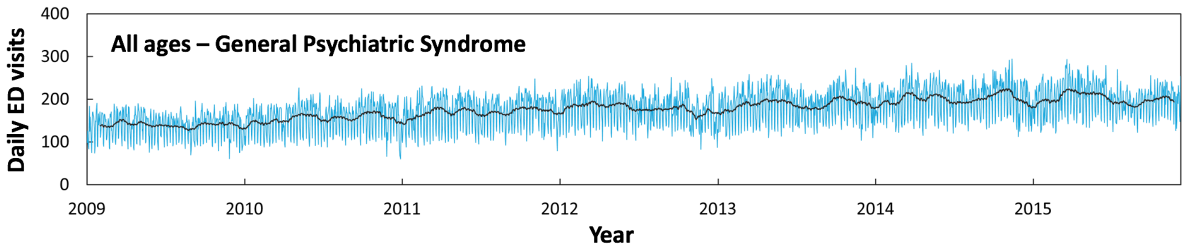


Figure 3.1: General psychiatric syndrome ED visits , all ages

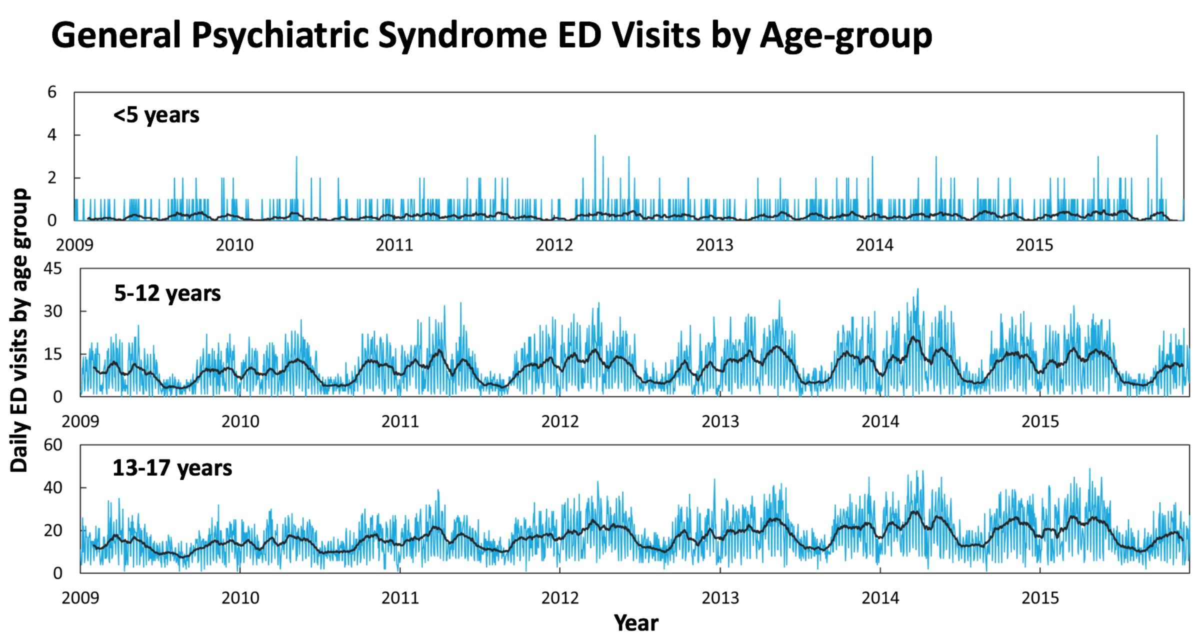


Figure 3.2: General psychiatric syndrome ED visits by age group

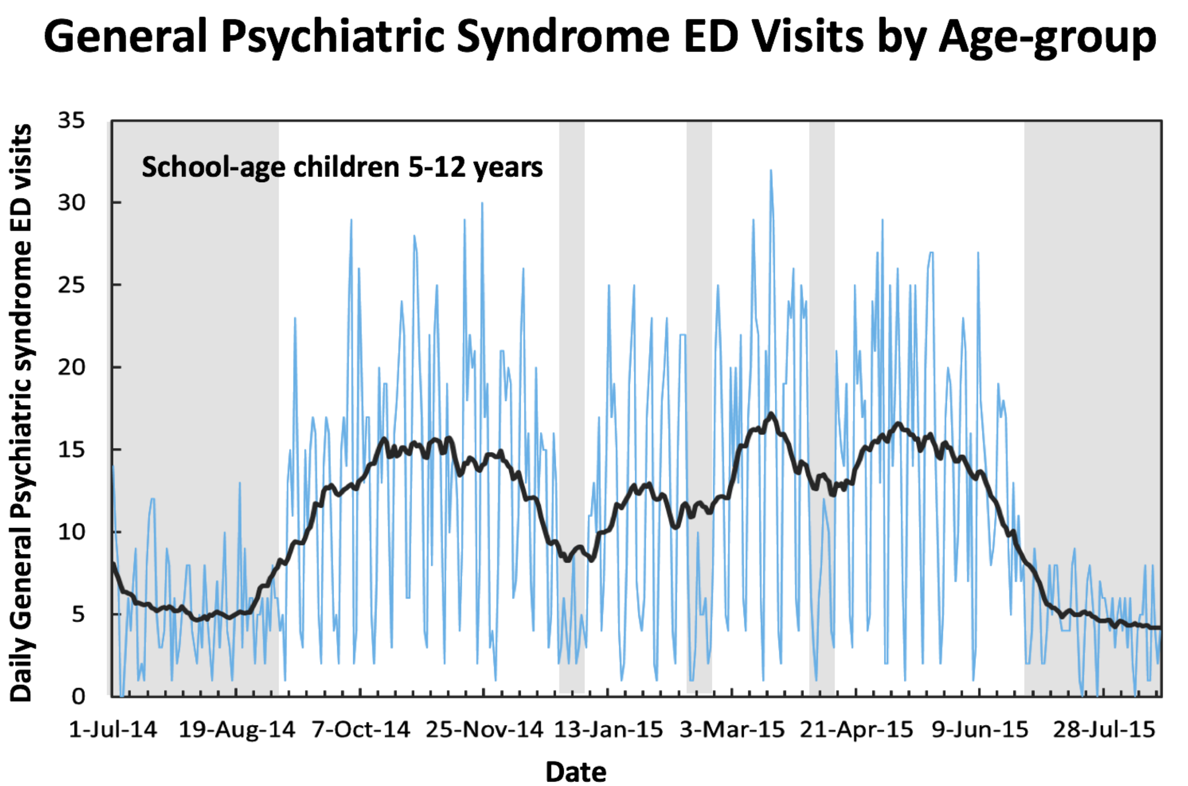


Figure 3.3: General psychiatric syndrome ED visits, aged 5-12

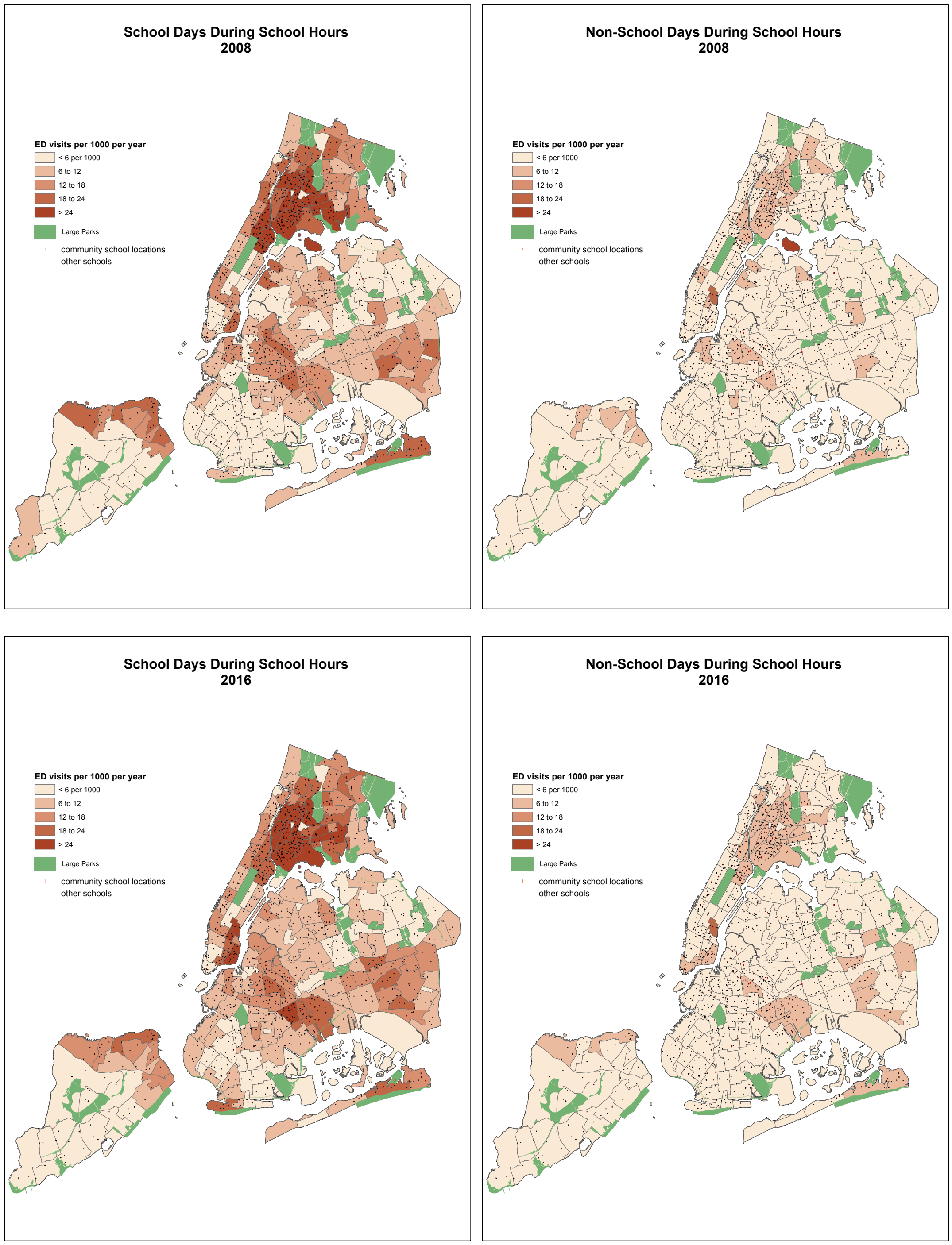


Figure 3.4: ED visit rates school days and non-school days, 2008 and 2016.

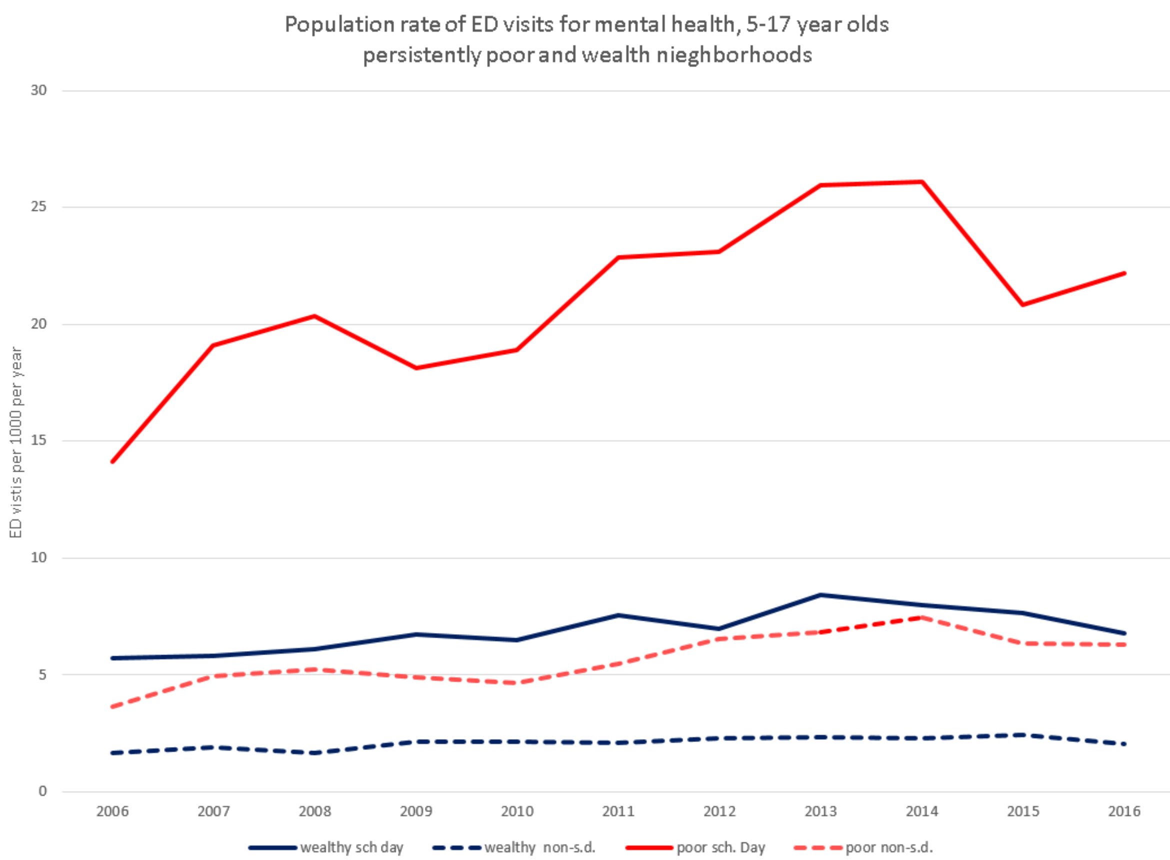


Figure 3.5: Population rates of ED visits for mental health, aged 5-17, poor vs wealth x school day vs non-school day

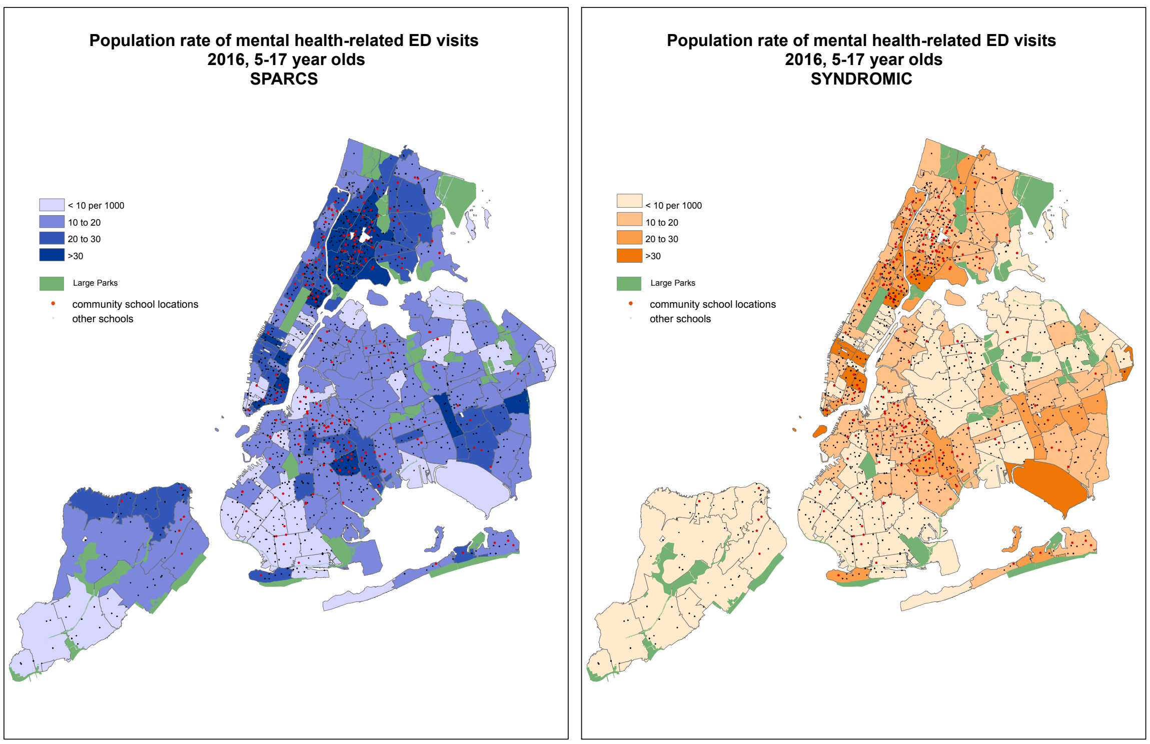


Figure 3.6: Population rates of mental health-related ED visits, SPARCS vs SYNDROMIC

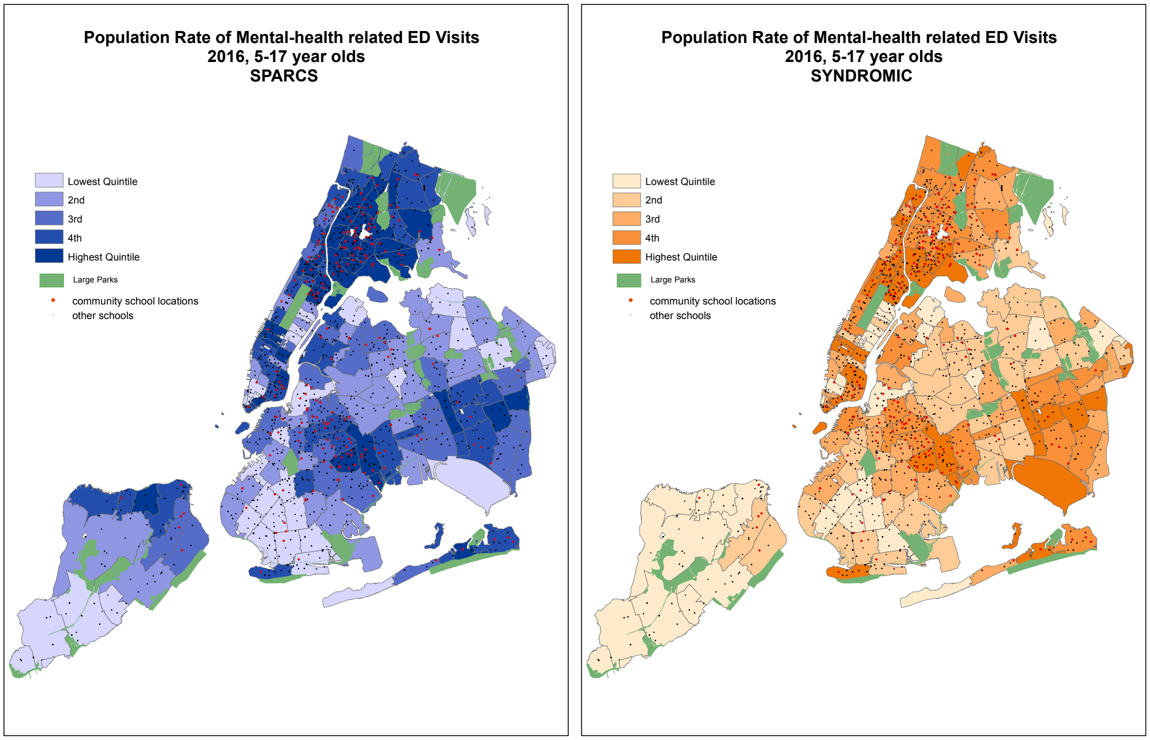


Figure 3.7: Population rate quintiles of mental health-related ED visits, SPARCS vs SYNDROMIC

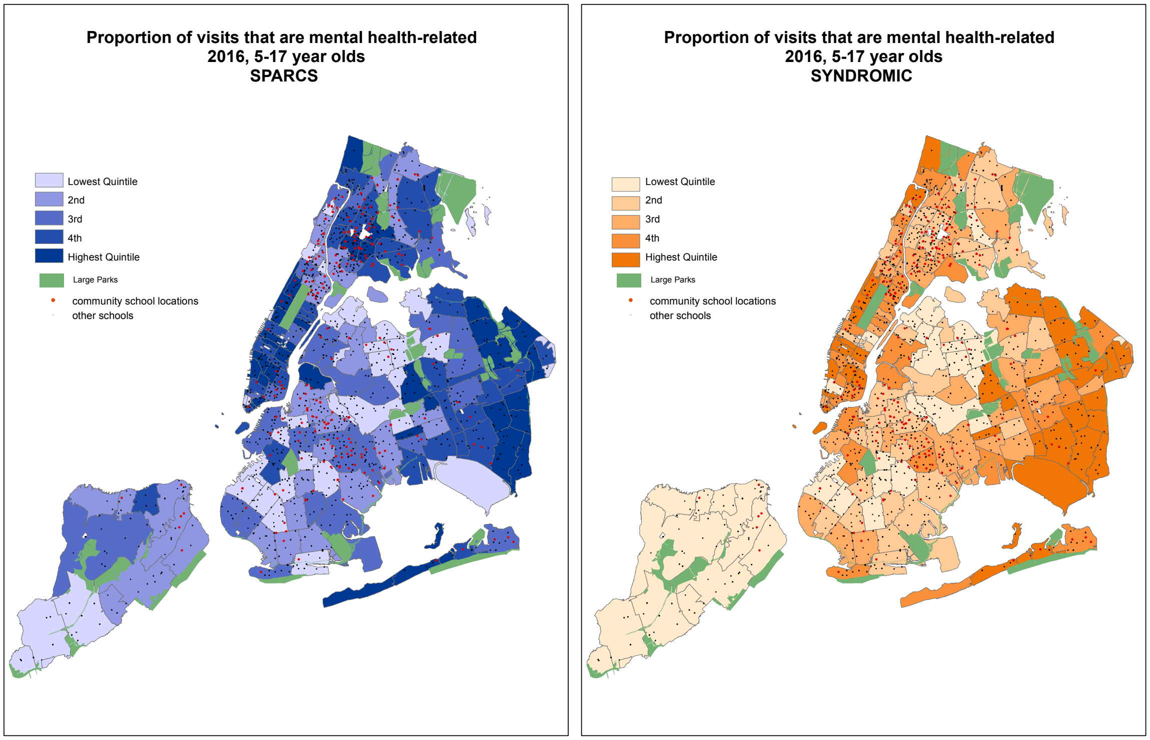


Figure 3.8: Proportion of mental health-related visit quintiles, SPARCS vs SYNDROMIC

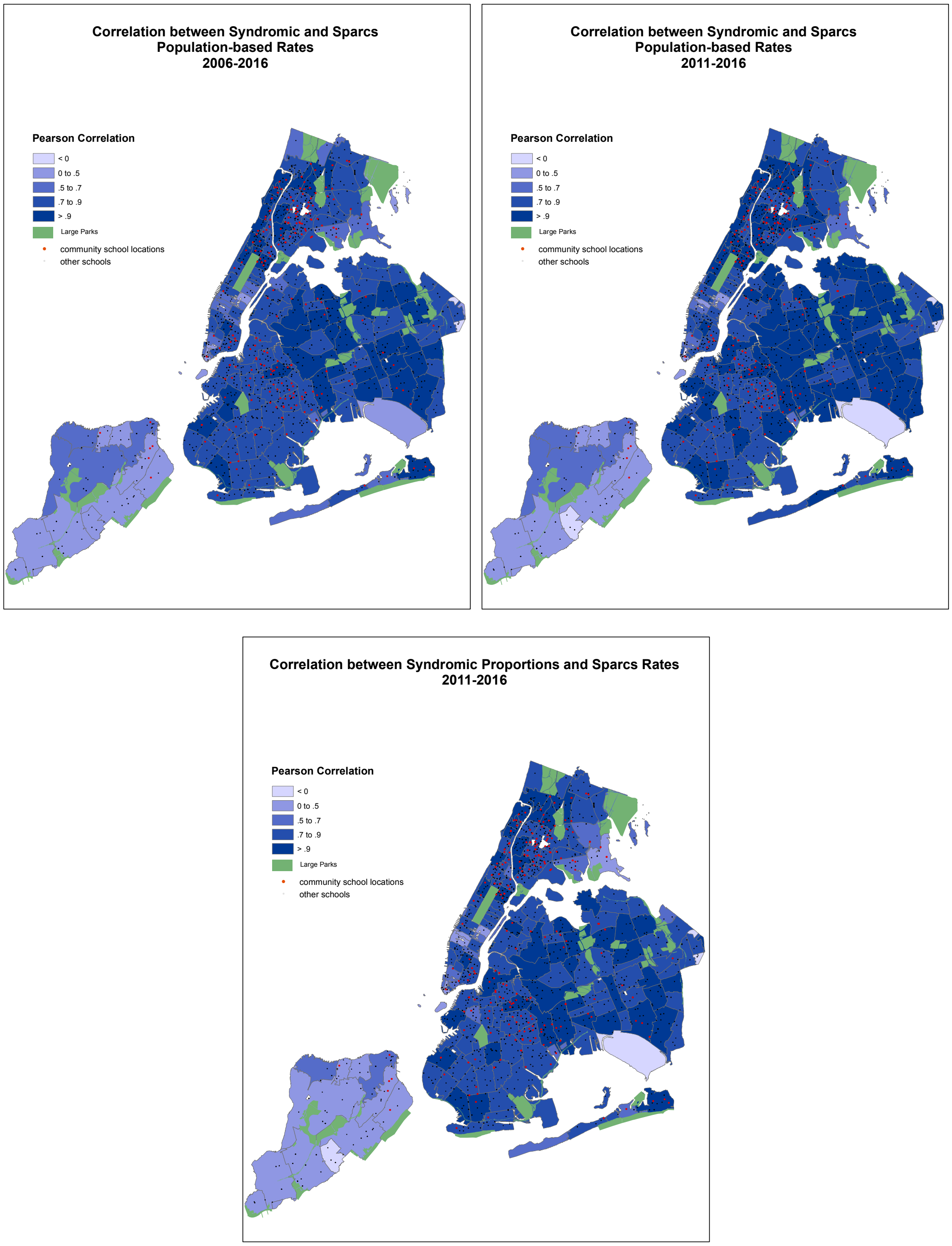


Figure 3.9: Area-specific temporal correlations, SPARCS vs SYNDROMIC

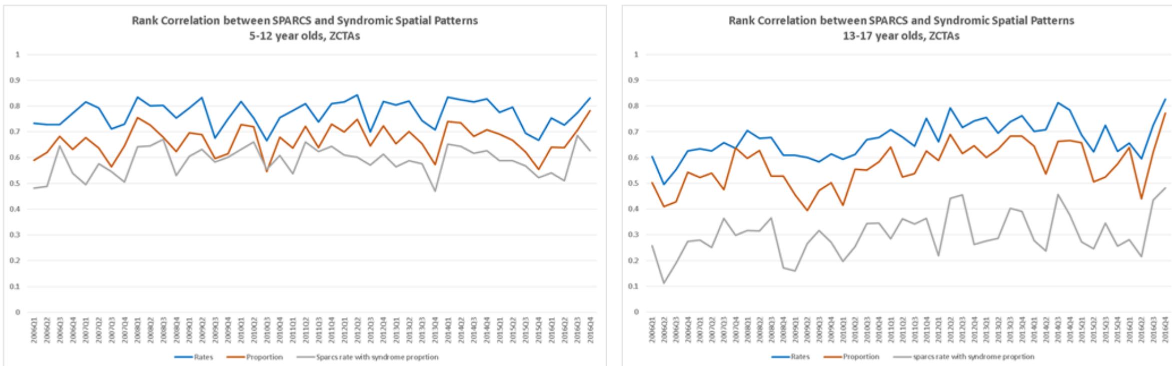


Figure 3.10: Quarter-specific areal rank correlations, SPARCS vs SYNDROMIC

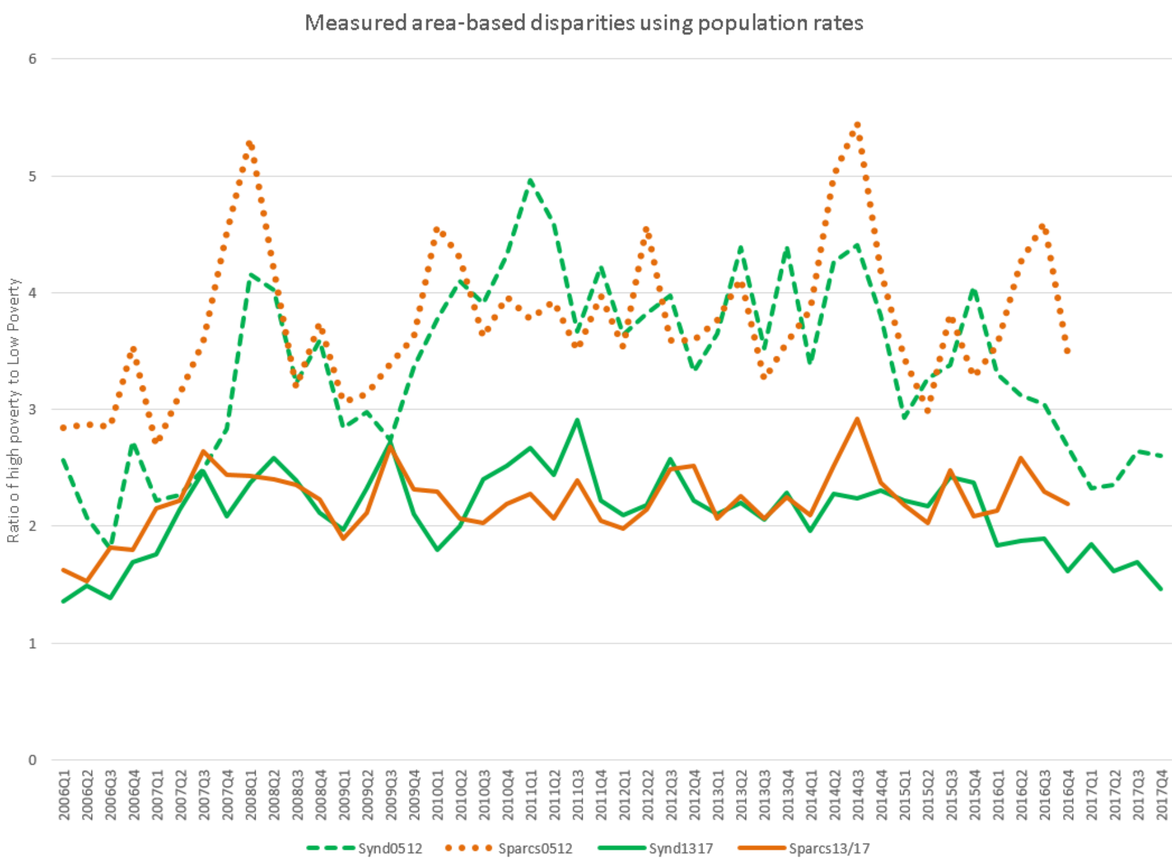


Figure 3.11: Measured area-based mental health ED visit disparities using population rates, SPARCS vs SYNDROMIC

Chapter 4

Childhood obesity: Longitudinal information, local government, and public policy

Abstract

School-based collection of height and weight data has become increasingly common. Three states currently mandate annual collection and several other jurisdictions including California and New York City (NYC) collect BMI as part of physical fitness assessments. This has resulted in the establishment of extremely large databases that share important characteristics including the ability to define longitudinal growth curves by student, low quality measurements as compared to clinical, and high coverage rates. Since 2006, height and weight measurement have been recorded for NYC public school students. Records are linked to registry, academic, and attendance data and across years resulting in a longitudinal dataset containing 9 cohorts with 2 million unique children. A high level of demographic and geographic detail allow for analysis of public policy at the local scale. We develop a quantile regression framework for BMI trajectories. Models consisting solely of age terms yield empirical curves similar to CDC growth charts; covariates modify these curves. Incorporating lag terms yields a distribution of possible growth trajectories and the effect of interventions can be explicitly quantified. We validate our approach using out-of-sample prediction. The framework enables the use of longitudinal growth trajectories to evaluate interventions. The use of BMI results in additional power over obesity status because of the additional information, and quantile regression can focus on the upper tail of BMI distributions. Finally, the model allows for longitudinal data quality assessment.

keywords: childhood obesity, NYC FITNESSGRAM, quantile regression.

4.1 Introduction

Spatial information has always played a central role in public health practice, especially at the local scale. Recent developments, including a spatial turn in academic public health, have resulted in renewed interest in the effects that neighborhoods have on population health and an increasingly important role for Spatial Demography (Diez-Roux, 2001; Krieger et al., 2003). This paper will describe current challenges involving the use of spatial information in public health practice in New York City.

Specifically, we describe the development of a system to monitor childhood obesity at fine spatial and demographics scales and to use the system to evaluate public policy. The data consists of administrative records describing longitudinal growth trajectories for New York City public school children. Detailed home and school information allow for analysis at fine spatial resolution allowing for the characterization of New York City neighborhoods and the investigation of community-level effects. The longitudinal trajectories can be matched with additional data sources describing school-based programmatic obesity reduction interventions and potentially obesogenic neighborhood-level built environment measures. We argue that quantile growth curve models can be used to evaluate whether and to what extent interventions alter the growth curve trajectories of children while controlling for demographic and neighborhood effects. We demonstrate the potential of the models and their interpretation using a neighborhood poverty indicator.

4.2 Background and Data

4.2.1 Childhood Obesity in New York City

In the United States, childhood obesity is a major public health concern and increasing trends since the 1980s have been documented(Ogden et al., 2014, 2002). Recently there has been much discussion about the current situation with some suggesting childhood obesity may be intractable(Skinner et al., 2016), others noting improvements(RWJF, 2012; Dietz, 2016), and still others re-focusing attention on severe obesity trends(Skinner and Skelton, 2014).

In New York City (NYC), child obesity has been a focus of public health for over a decade(Thorpe et al., 2004). Since 2006, basic anthropometric measures have been collected annually in public schools as part of the NYC FITNESSGRAM assessment. These measures are linked to enrollment, attendance, physical fitness and academic outcomes to form a longitudinal dataset that records individual growth trajectories while providing detailed geographic and demographic information. Using this data, obesity patterns and trends are monitored at various scales(Berger et al., 2011; Day et al., 2014) and links to academic performance have been established (Egger et al., 2009).

The NYC public school system is the largest and most diverse system in the US with approximately 1.1 million children each year. The student population is characterized by substantial socio-economic heterogeneity and residential segregation by income. Over 40% of NYC public school students speak a language other than English at home, many of whom live in various ethnic enclaves. Administrative school records that include student residential address can be used to describe the demography of households in NYC at arbitrarily fine geographic scale. Additionally, the data goes beyond simply locating residence. It captures home address, school address, identifies siblings, classmates and their addresses. It also records movement by families in NYC, both residential mobility and change in schools.

4.2.2 Child Obesity Data

The NYC FITNESSGRAM includes approximately 900,000 unique measurements of height and weight per year in kindergarten through 12th grade. For cross-sectional reporting, non-response is currently adjusted for through a post-stratification procedure that gives weights to those measured based on demographic and geographic characteristics. These weights are then used to produce tables at various demographic and geographic scales. Repeated cross sections are compared to monitor for population change. Comparison of these changes to policy is common (such as difference-in-difference methods). Nonetheless, these methods do not take advantage of the longitudinal nature of the data. The data provides a rich longitudinal resource with a large share of students followed having three or more measurements in consecutive years.

Data of this form are becoming increasingly common. Multiple jurisdictions including Philadelphia Robbins et al. (2015), California (Babey et al., 2011; Jin and Jones-Smith, 2015), and Texas (Welk et al., 2010) currently collect and report on child obesity using school-based collection of body composition data. Additionally, the Institute of Medicine has recommended school-based collection of physical fitness measures including body composition (Pate et al., 2012; Kohl III et al., 2013) and the Presidential Youth Fitness Program has endorsed FITNESSGRAM as a physical fitness assessment to be used in conjunction with an expanded physical education curriculum (Program, 2014; Welk and Meredith, 2010). Increasingly this data can be linked to academic and other student outcomes as local governments and school districts adopt integrated data systems.

4.2.3 Local public policy

A number of local initiatives address childhood obesity including school- and community-level interventions; many targeted based on perceived or established patterns of obesity.

School-based interventions include changes to food served in schools, changes to the food environment, regulations on vending machines, and targeted programs and interventions addressing diet and exercise. Other policy efforts address food access in communities including establishing green carts and farmers markets, restrictive licensing, increasing access to fresh produce in existing stores and incentivizing the opening of new grocery stores in neglected areas. Finally, efforts to increase community walkability and safety, to enable active transportation (e.g. adding bicycle lanes), or that increase recreation opportunities by adding playgrounds, parks or other centers also serve to address obesity.

Collecting data describing targeted policies is a key challenge to linking demographic data, health outcomes and policy. Often multiple interventions are in effect simultaneously and at various geographic scales. Our preferred approach would record the timing of (all) school-based interventions while establishing a longitudinal dataset describing changes to the built environment resulting from community-based policy. There is currently no monitoring system of interventions or the built environment operating in NYC. Such systems would enable causal interpretations of policy evaluations by quantifying changes in the environment for comparison to changes in outcomes.

Within this data and policy context, extensive efforts are underway to evaluate school- and community-based obesity interventions. Examples from NYC include evaluations of the Breakfast-in-Classroom program(Corcoran et al., 2014), the installation of water jets in school cafeterias(Schwartz et al., 2016), and the format of the information provided to parents on their children's fitness assessments(Almond et al., 2016). In California, researchers have looked at Nativie-American obesity in response to the establishment or expansion of casinos on reservations (Jones-Smith et al., 2014). These evaluations each utilized the data form we address here and each took distinct analytic approaches. We propose a general approach that can be applied to a variety of evaluation questions.

4.3 Methods

The framing of current childhood obesity reporting and analysis relies on standard centile reference charts from the US Centers for Disease Control (CDC 2000; 2010) for height, weight, and BMI. The growth standards represent historic U.S. population age distributions for a given measure (height, weight, or BMI). Current CDC growth charts are estimated using the LMS model (Cole and Green, 1992); each age-group distribution is converted to normal using a Box-Cox transformation and the estimated location, scale, and skewness parameters are constrained to vary smoothly across age groups. For any raw measure, the age-specific LMS parameters are used to convert the measure to a z-score that relates to the standard reference centile chart. The notion of obesity or severe obesity among children by demographic subgroup or local district is made in reference to the share of the population that is above age-specific z-scores, or equivalently centiles, of the reference distribution.

Policy analysis to evaluate the effectiveness of city-wide, district, or school-based interventions to reduce obesity have relied on this same framing in relation to standard growth charts. A policy effect is conceived of as, $Pr(z > z^*|\text{no treatment}) - Pr(z > z^*|\text{treatment}) = \delta$, where z is the z-score in the study population and z^* is a threshold from the reference growth chart, δ is the policy effect size, and the probability model may include additional controls depending on the research design. It is recognized that the z-scores produced by current LMS methods perform poorly in the tails of the distribution (Flegal et al., 2009) and can display erratic behavior across age (Koenker, 2018). This limitation occurs within the subgroup most of interest– the obese and severely obese– and alternate thresholds have been suggested (Gulati et al., 2012) to address it. Further, a substantial number of children are observed at these extreme values; over 5 percent of NYC public school children are severely obese (Day et al., 2014).

Another key feature of current reporting and analysis that relies on standard growth charts is that all measurement is cross-sectional. As noted by Wei (2004) and Wei et al. (2005), it is possible to develop an alternative framing based on longitudinal analysis when repeated measures are available for each individual. Indeed, comparing individual growth paths to reference growth charts based on cross-sectional data could be inappropriate (Wei, 2004, see examples). This alternative framing relies on quantile regression longitudinal growth curve models. Defining BMI for individual i as Y_i and age as t_i, a baseline model of smooth quantile growth curves is:

$$Q_{Y_i|t_i,x_i}[\tau|t_i, x_i] = g_\tau(t_i) + x_i^T \gamma_\tau + e_i. \tag{4.1}$$

Estimation of the smooth term g_τ results in an empirical growth curve such that approximately τ, proportion of the observations lie below the fitted curve. Details of estimation are available in Koenker and Bassett Jr (1978); Koenker et al. (1994); Wei (2004). The model can also include covariates that act to shift the growth curves. In this case we define x_i as an individual-level covariate indicating policy treatment or no policy treatment. The policy effect in the quantile model is, $Q_{Y_i|t_i}[\tau|t_i, x_i = \text{treatment}] - Q_{Y_i|t_i}[\tau|t_i, x_i = \text{no treatment}] = \gamma_\tau$; that is, the difference between the τth growth curve percentile under treatment or no treatment. Note that the policy effects will vary depending on the growth curve percentile selected such that the impact of a policy could reflect a shift in the location and shape of the growth curve distributions under the policy.

Wei (2004) refers to model (4.1) as the unconditional model. The model can be extended to condition on one or more prior measures of BMI when longitudinal data is available. Following (Wei, 2004), a model for irregular follow-up measurement times and including one prior measure of BMI is,

$$Q_{Y_i(t_{i,j})}(\tau|t_{i,j},.,X_{i,j}) = g_\tau(t_{i,j}) + [\beta_{1,\tau} + \beta_{2,\tau}(t_{i,j} - t_{i,j-1})]Y_i(t_{i,j-1}) + X_{i,j}^T\gamma_\tau + e_{i,j}. \tag{4.2}$$

and with two prior measures is,

$$\begin{aligned} Q_{Y_i(t_{i,j})}(\tau|t_{i,j},.,X_{i,j}) = g_\tau(t_{i,j}) &+ [\beta_{1,\tau} + \beta_{2,\tau}(t_{i,j} - t_{i,j-1})]Y_i(t_{i,j-1}) \\ &+ [\beta_{3,\tau} + \beta_{4,\tau}(t_{i,j} - t_{i,j-2})]Y_i(t_{i,j-2}) + X_{i,j}^T\gamma_\tau + e_{i,j}. \end{aligned} \tag{4.3}$$

The autoregressive terms are approximated by a linear function of the elapsed time between current and past measures. The smooth term remains and predictions again yield empirical growth curves for the τ^{th} centile. Notice that the policy effect is now also indexed by j to indicate that the covariate references a particular period and could also be specified as a lagged term. As noted by Wei (2004, see pages 32 and 61), the interpretation of autoregressive estimate $\hat{\beta}(\tau)$ is complicated under heteroskedastic error distribution. If $(t_{i,j} - t_{i,j-1})$ and $(t_{i,j} - t_{i,j-2})$ are held fixed, the resulting estimates do reflect the scale of the autoregressive effects.

For each model it will be necessary to interpret the magnitude, sign, and significance of parameter estimates and overall model fit. Hypothesis testing of parameter estimates is based on standard errors and p-values produced using the R package **quantreg** (Koenker, 2015). To assess model fit and aid in model selection we compare AIC across models specification and for different values of τ.[1] We also use visual fit diagnostics suggested by Wei (2004), that are based on comparing the share of the population in a subgroup below a predicted quantile to the nominal quantile; that is, $\tau - \hat{\tau}$ where $\hat{\tau} = \sum_{i=1}^{n} n^{-1} I[Y_i(t) \leq \hat{Q}_{Y(t)}(\tau)]$. For example, for children aged 12.50 to 12.75 the share of children below the

[1] AIC is also used to select the smoothness penalty parameter λ. For each specification and value of τ we fit the model for a range of λ values and then choose the smoothness penalty that minimized AIC (Koenker, 2015; Koenker et al., 1994).

predicted 0.97 quantile for model (4.1) is 0.9717 and for model (4.3) is 0.9659, giving fit measures respectively of -0.0017 and 0.0041. Positive valued fit indicates the model predicted a threshold such that fewer individuals than expected were below the predicted quantile, and more than expected were above it. Another interpretation, and supposing $\tau = 0.97$ is a marker for severe obesity, is that a positive fit measure suggests that the subpopulation has a larger share of extremely obese children than predicted by the model. We produce graphical measure of fit by age groups from 7 to 18 in steps of quarter year and fit by school district. In addition to assessing fit with respect to the data used for model fitting we also assess it with respect to an equivalent sized out-of-sample data set for model validation. From the longitudinal NYC FITNESSGRAM data we randomly extract two sets of records of approximately 100,000 female students each. One data set is used for model fitting and the other is used for validation.

4.4 Results

As noted already, socio-economic status (SES) and childhood obesity are correlated and both have a strong spatial signal (see Figures 4.1.1 and 4.4.2). Several districts in east Brooklyn bordering Queens and in the south Bronx, are characterized by high prevalence of both child obesity and extreme poverty. The connection between the two is well documented in the literature (Singh et al., 2010, 2011; Jin and Jones-Smith, 2015). The populations living in those districts are likely targets for interventions and we propose the efficacy of those interventions could be evaluated using the methods proposed in the previous section.

Instead of evaluating a specific policy intervention, we will treat neighborhood poverty as if it is a designed intervention. Assume that an individual's SES status is the outcome of an overt policy that distributes income unequally across neighborhoods in such a way that poverty is clustered as depicted in Figure 4.1.1. Using the calibration data of 100,000 female student records, we fit the unconditional model (4.1) without a covariate and the AR(1) and AR(2) models with and without the covariate. The poverty covariate includes four levels – low, medium, high, or extreme poverty – with low being the reference category. Model selection is based on AIC (see Table 4.1) and the visual fit diagnostics in the domains of age and school district (see Figures 4.2, 4.3, and 4.4). The AR(2) model including the poverty covariate has the lowest AIC and thus the best overall fit.

The visual diagnostics concur with AIC and reinforce the selection of the AR(2) model. Figure 4.2 indicates the model fit across age groups for the unconditional, AR(1), and AR(2) models, each excluding the poverty covariate. All three fit almost perfectly in the age domain to both the training and validation data sets. The first is slightly worse in the tails for the validation data but there is no indication of gross overfitting to the training data.

The same model specifications are evaluated in the spatial domain in Figure 4.3 focusing on the 0.9 and 0.97 centiles since our interest is in obesity. For the unconditional model (Figure 4.3.1), the same areas of east Brooklyn and south Bronx have positive residuals, indicating a larger share of obese children in those districts than predicted by the model. After conditioning on prior BMI in the AR(1) and AR(2) models (see Figures 4.3.2 and 4.3.3), the spatial fit is improved but the direction persists with darker beige in the same districts. Thus, even after conditioning on having two prior BMI measures, children in those districts are becoming more obese than elsewhere in the city. That is, there are areas in the city where a child's BMI increases greater than a child from another area with same past BMI trajectory and equal age.

Visual fit diagnostics for the AR(2) model with the poverty covariate included are presented in Figure 4.4. The fit and validation are still excellent in the age domain, and now with the inclusion of a poverty covariate (our assumed multiple level spatial treatment), the spatial fit is noticeably improved with the direction of lack of fit reversing in some of the formerly high poverty areas. This means that the estimate for the poverty covariate has captured the spatial variation in childhood obesity. Since the AR(2) model with poverty covariate is of most interest for interpretation, we provide parameter estimates, p-values, and a summary of the autoregressive effects in Tables 4.2 and 4.3.[2]. The autoregressive effects are uniformly significant for the 0.25 centile and higher, and the effect sizes increase with centile. The relative magnitudes of the effects on the low tail and upper tail indicate that the lower tail has a relatively flatter trajectory than the upper tail. For the 0.9 and 0.97 centiles, the lagged BMIs contribute roughly 10% or more to BMI trajectory. The effect sizes for poverty also increase sharply with centile, with the values increasing almost exponentially moving from $\tau = 0.03$ to $\tau = 0.97$ for the effect of extreme poverty versus low poverty. This means the BMI distribution for children living in extreme poverty not only shifts rightward compared to low poverty children, but the distribution is more skewed with the 0.97 centile shifting 0.8 units (in BMI).

4.5 Discussion

It is already well established that for a variety of reasons, health conditions covary with SES and the link between childhood obesity and poverty is just a single instance of that broader pattern (Barr, 2014). Our point here is to demonstrate the utility of quantile growth curve modeling of longitudinal records to evaluate the size and significance of interventions. If our poverty experiment was an intervention, we would have clear evidence of its impact. We are only using a trivial example here and the model specifications could be more rich allowing AR(1) or AR(2) terms to vary with the policy

[2]Parameter estimates and p-value for the other models are provided in Tables 4.4, 4.5, and 4.6

covariate instead of a simple shift. It would also be possible to use interesting cross-over research designs that would use the information on a child's change in residence to treat neighborhood features as a kind of natural experiment. The key point in this paper is that we have demonstrated that it is possible to develop models of child BMI trajectories that incorporate longitudinal records and a hypothetical policy control without resorting to an external centile reference chart or the myriad assumption inherent in converting raw measures to the domain of reference charts. It is also important to note that because the quantile growth models can be used to characterize differences in the overall shape and location of distributions under policy treatment or no policy treatment, we can also translate our results back into the domain of CDC growth charts. That is, we can use the models to indicate how CDC defined severe obesity increases or decreases.

The efficacy of childhood obesity policy should be evaluated as the impact on individuals' BMI trajectories. Quantile regression of longitudinal growth curves can describe full BMI distributions while incorporating demographic and policy-specific information. Further, the impacts of obesity-specific policies are of varying importance across the distribution of BMIs. This is a key feature of quantile regression (Koenker, 2018) that has begun to receive attention at the research design or policy formation stage (Wang et al., 2018). As long as policy can be applied to individual children, either by their school or by the residence (community), the impact can be quantified. Likewise, changes in the built environment (community or school) can be incorporated into these models to assess their impact on individual growth trajectories.

The statistical model developed in this paper rely on that availability of highly detailed longitudinal administrative data. Such data are becoming increasingly common and the detail in these databases is also expanding with the establishment of integrated data systems. Our framework is particularly relevant when evaluating targeted policies that are typically evaluated in an ad hoc manner. Even rigorous evaluations of such policies routinely ignore the full policy context in which it was implemented. Our proposed approach takes advantage of the demographic, geographic, and temporal detail within school records. In countries without health registries, such as the United States, demographics taken from school records offer incredibly high coverage with high detail and the ability to see individuals change through time. Our analytic approach takes full advantage of longitudinal information potentially enabling causal interpretations of the link between policies and outcomes. By focusing on BMI the approach achieves greater power than using an obesity indicator, while the quantile regression allows us to quantify the impact on the right tail of the distribution where health risk is known to occur.

4.6 Tables and Figures

Specification	τ						
	0.03	0.1	0.25	0.5	0.75	0.9	0.97
Unconditional	583015	564794	559874	577525	618750	668551	720624
AR(1)	521279	461349	418484	403307	426148	478314	557445
AR(1)+poverty	521269	461345	418457	403208	425897	477696	555969
AR(2)	516225	457422	414855	399424	420345	468608	539327
AR(2)+poverty	516207	457407	414828	399316	420074	467997	538347

Table 4.1: AIC for alternative model specifications.

Parameter	τ						
	0.03	0.1	0.25	0.5	0.75	0.9	0.97
β_1	0.465	0.643	0.745	0.793	0.792	0.748	0.663
	0.000	0.000	0.000	0.000	0.000	0.000	0.000
β_2	-0.015	-0.003	0.006	0.018	0.030	0.041	0.051
	0.003	0.368	0.000	0.000	0.000	0.000	0.000
β_3	0.215	0.187	0.177	0.179	0.215	0.262	0.347
	0.000	0.000	0.000	0.000	0.000	0.000	0.000
β_4	-0.006	0.000	0.003	0.008	0.014	0.025	0.040
	0.059	0.806	0.008	0.000	0.000	0.000	0.000
$\gamma_{\text{medium poverty}}$	0.102	0.093	0.042	0.052	0.070	0.110	0.243
	0.027	0.000	0.014	0.001	0.000	0.000	0.000
$\gamma_{\text{high poverty}}$	0.050	0.069	0.028	0.058	0.096	0.184	0.365
	0.321	0.011	0.122	0.001	0.000	0.000	0.000
$\gamma_{\text{extreme poverty}}$	0.001	0.064	0.088	0.155	0.250	0.446	0.806
	0.986	0.025	0.000	0.000	0.000	0.000	0.000

Table 4.2: Parameter estimates and p-values for AR(2) model with poverty covariate.

Effect	τ						
	0.03	0.1	0.25	0.5	0.75	0.9	0.97
AR(1)	0.450	0.641	0.751	0.811	0.822	0.789	0.714
AR(2)	0.202	0.188	0.183	0.195	0.243	0.313	0.426

Table 4.3: AR(2) effects given $t_{i,j} - t_{i,j-1} = 1$ and $t_{i,j} - t_{i,j-2} = 2$.

Parameter	τ						
	0.03	0.1	0.25	0.5	0.75	0.9	0.97
β_1	0.616	0.792	0.897	0.954	0.988	0.991	0.954
	0.000	0.000	0.000	0.000	0.000	0.000	0.000
β_2	-0.012	0.001	0.009	0.024	0.041	0.061	0.082
	0.012	0.833	0.000	0.000	0.000	0.000	0.000

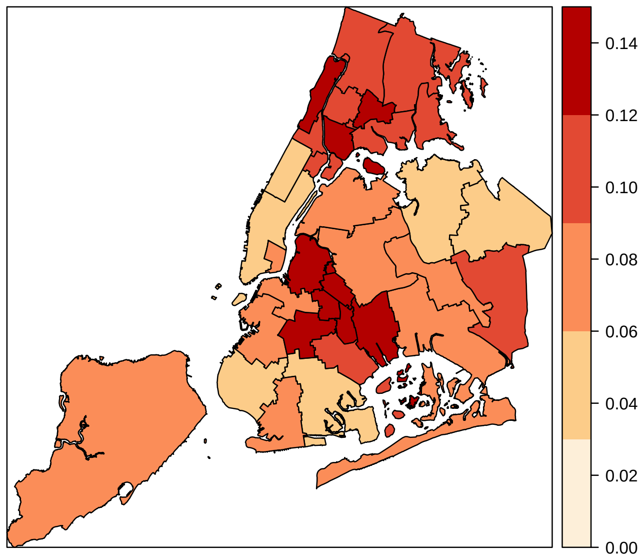
Table 4.4: Parameter estimates and p-values for AR(1) model.

Parameter	τ						
	0.03	0.1	0.25	0.5	0.75	0.9	0.97
β_1	0.461	0.643	0.745	0.794	0.794	0.757	0.665
	0.000	0.000	0.000	0.000	0.000	0.000	0.000
β_2	-0.013	-0.003	0.006	0.018	0.030	0.040	0.049
	0.006	0.330	0.000	0.000	0.000	0.000	0.000
β_3	0.216	0.188	0.177	0.179	0.214	0.258	0.351
	0.000	0.000	0.000	0.000	0.000	0.000	0.000
β_4	-0.006	0.000	0.003	0.008	0.015	0.026	0.041
	0.056	0.839	0.005	0.000	0.000	0.000	0.000

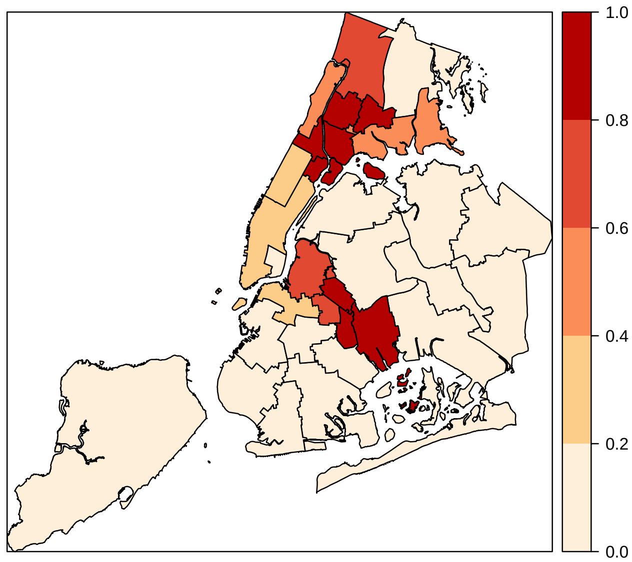
Table 4.5: Parameter estimates and p-values for AR(2) model.

Parameter	τ						
	0.03	0.1	0.25	0.5	0.75	0.9	0.97
β_1	0.619	0.793	0.896	0.953	0.986	0.987	0.950
	0.000	0.000	0.000	0.000	0.000	0.000	0.000
β_2	-0.012	0.000	0.009	0.024	0.041	0.060	0.080
	0.008	0.983	0.000	0.000	0.000	0.000	0.000
$\gamma_{\text{medium poverty}}$	0.101	0.075	0.050	0.055	0.068	0.111	0.384
	0.029	0.003	0.004	0.000	0.000	0.000	0.000
$\gamma_{\text{high poverty}}$	0.021	0.063	0.031	0.060	0.108	0.200	0.565
	0.677	0.021	0.109	0.000	0.000	0.000	0.000
$\gamma_{\text{extreme poverty}}$	0.018	0.059	0.105	0.149	0.261	0.473	1.061
	0.734	0.042	0.000	0.000	0.000	0.000	0.000

Table 4.6: Parameter estimates and p-values for AR(1) model with poverty covariate.

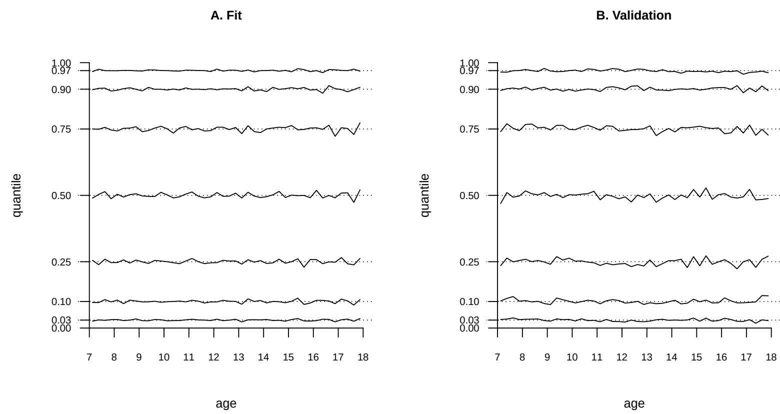


4.1.1: Childhood obesity prevalence

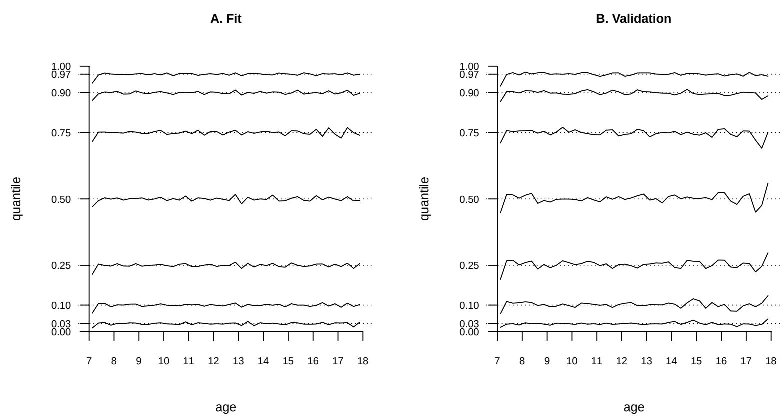


4.1.2: Extreme poverty prevalence

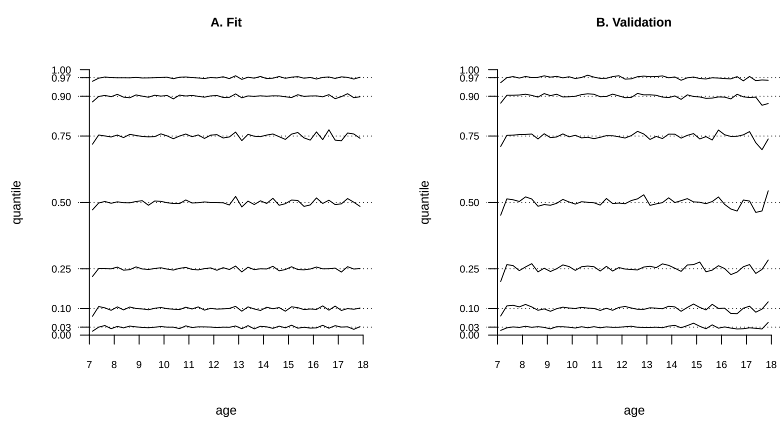
Figure 4.1: Visual association between obesity and poverty.



4.2.1: Unconditional model.

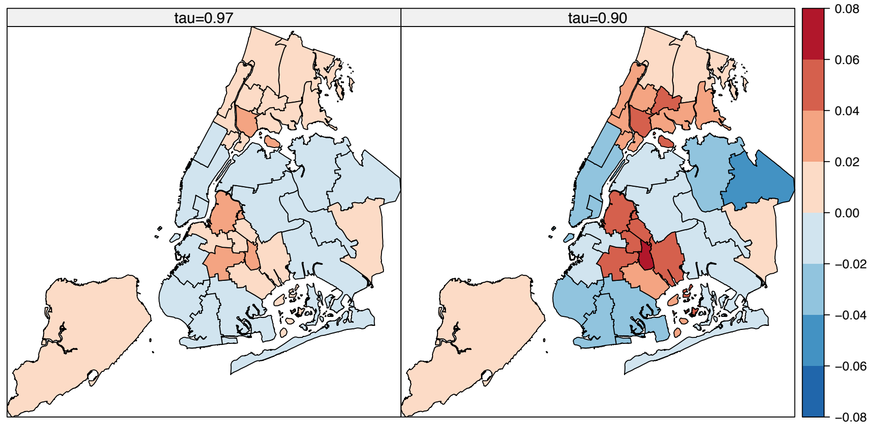


4.2.2: Conditional model, AR(1).

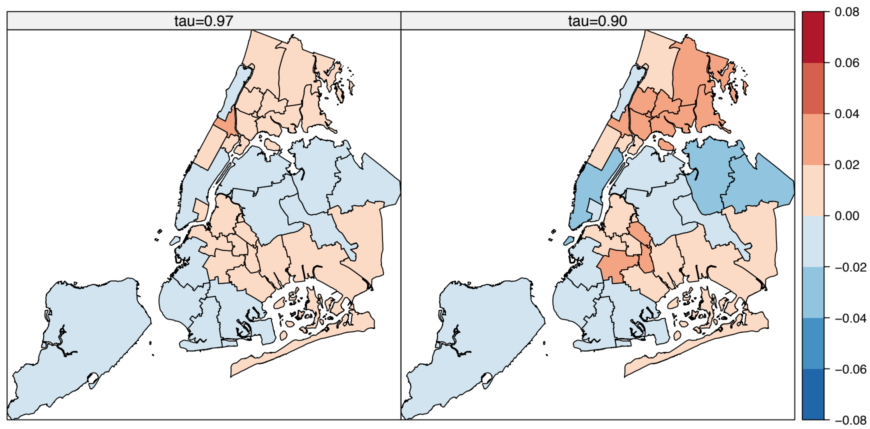


4.2.3: Conditional model, AR(2).

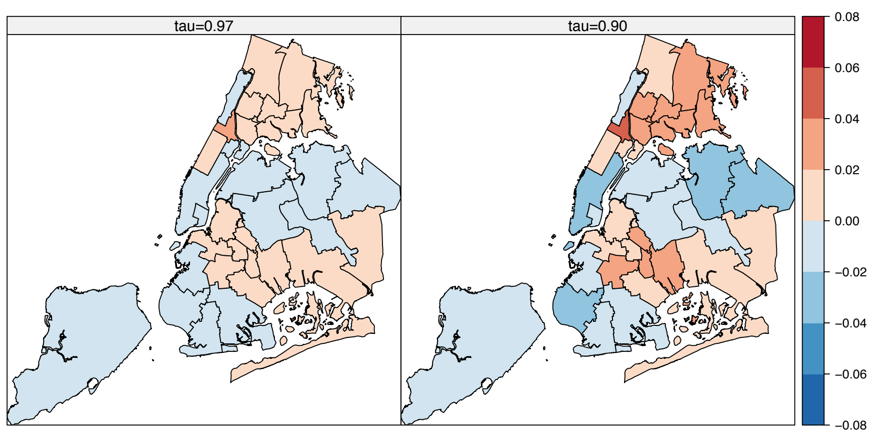
Figure 4.2: Fit and validation by age group.



4.3.1: Unconditional model.

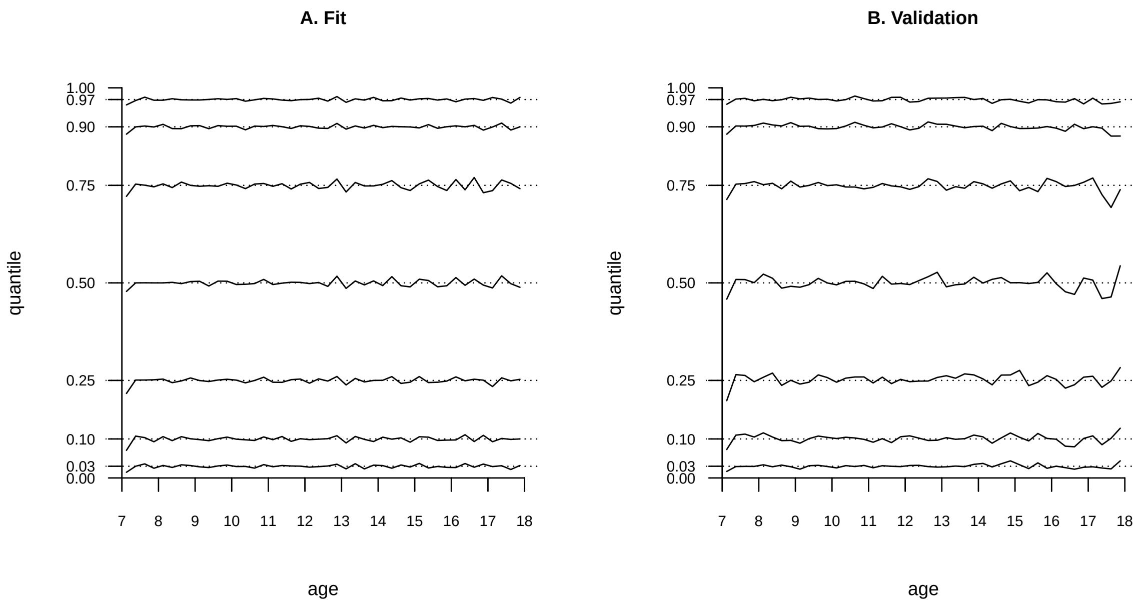


4.3.2: Conditional model, AR(1).

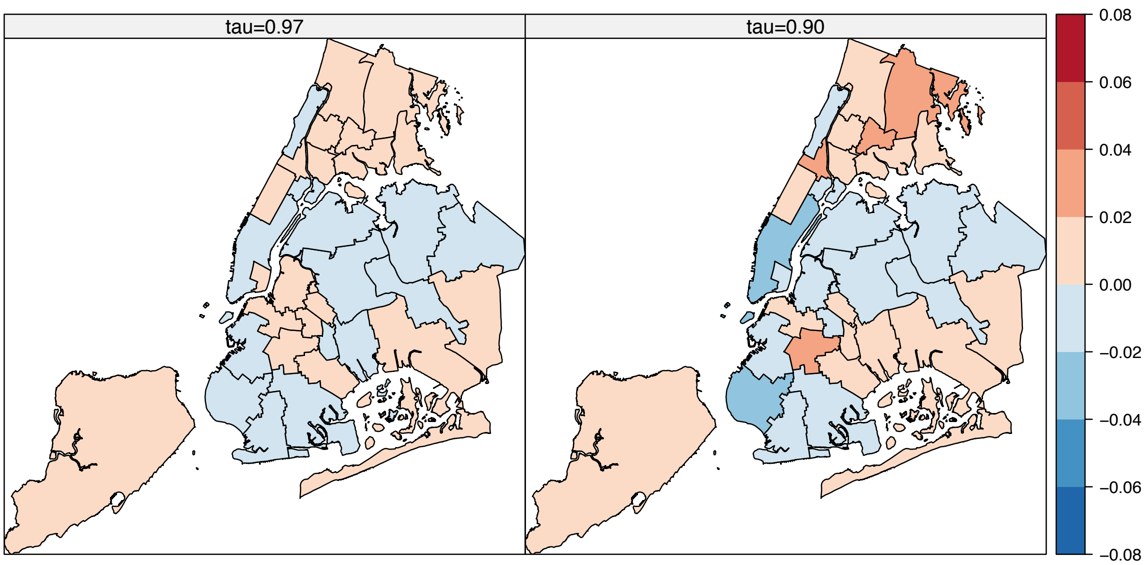


4.3.3: Conditional model, AR(2).

Figure 4.3: Fit by school district.



4.4.1: Fit and validation by age group



4.4.2: Extreme poverty prevalence

Figure 4.4: Fit diagnostics for AR(2) with poverty covariate.

Chapter 5

Conclusion

This dissertation has explored the increasing role of Geography in local public health *practice*. A number of factors have led to this "spatial turn" in public health, which has, in fact, always been spatial. This is particularly the case at the local scale in which policy is implemented, programs and services are delivered, disease control and outbreak response are conducted, and people and their communities are addressed. The factors include the broadening of the scope of public health including a (re)new(ed) appreciation of the role that the social determinants of health play in disease etiology, the related focus on context or neighborhood effects on health including the built environment and environmental measures, the availability and timeliness of data at finer geographic, demographic, and temporal resolution, and an emphasis on open data and intersectoral collaboration. All of these factors have contributed to an increased interest in public health surveillance including the surveillance of behavioral risk factors and social determinants themselves in addition to traditional disease surveillance.

The common thread in these trends as it applies to surveillance is the detailed characterization of local communities and the people that reside in them. Timely characterization across three domains– health outcomes, social determinants, and behavioral risk factors– using vital records, disease registries, survey-based surveillance systems, administrative records, and novel data streams has the potential to impact public and population health and enable a more preventive approach to clinical medicine. Such surveillance allows for the targeting of programs and policy, the measurement and monitoring of health disparities, the evaluation of programs and policies incorporating a longitudinal or panel approach to context-level variables, greater inter-agency cooperation, and increased transparency. Ultimately, the new surveillance seeks to provide more detailed "information to those who need to know" so "action can be taken" in a more timely and effective manner. These features are evident in the three projects presented here.

First, **Area-based poverty measures (ABPMs) in public health practice** reviewed the details of the construction of ABPMs in New York City and the utility of ABPMs for surveillance purposes. We reviewed the available scales using practical criteria necessary for surveillance systems including reliability and the ability to systematically update measures and acknowledging two separate use cases for ABPMs, imputation and use as context-level descriptors. On these grounds, the Neighborhood Tabulation Area provides a geographic scale with high-reliability that is continuously updated by the U.S. Census Bureau and exists within a geographical hierarchy at the conceptually relevant "neighborhood" scale. There are instances when census tracts and Zip Code Tabulation Areas may be preferable, as an imputation method (tracts) or when complete address information is not available (ZCTA). As an updating rule for the construction of disparity measures, using persistent poor and wealthy areas has more stable properties than alternatives. One possible way forward here is to use the consecutive non-overlapping American Community Survey five-year samples: 2008/12, 2013/17, 2018/22, etc. This would allow for updating areas every five years and allow for assessment of change. Disparities measured using persistence are clearer than other assignment rules.

The approach presented is applicable to other social determinants of health. In fact, New York City recently completed a Robert Wood Johnson Foundation-funded *Data Across Sectors for Health* project that brought together data from several city agencies and health organizations at the Neighborhood Tabulation Areas scale on a one-time basis. Systematic collection of the same data could help establish NYC's first surveillance system of the Social Determinants of Health, which in turn would allow for the evaluation of projects intervening on SDH as well as causal modeling of SDH change on health outcomes. Future work includes the construction and monitoring of disparities across a variety of health indicators using the ABPM approach presented here, further analysis of change in the NTA and census tract based assignments and an assessment of their ability to capture true change such as gentrification, and analysis of continuous relationship of poverty to heath outcomes.

The second project **Spatial patterns of child mental health burden: Measurement and monitoring of child mental health disparities in New York City** investigated the relationship between two data sources as used to characterize spatial and temporal patterns of mental health-related emergency department (ED) visits. The first, SPARCs, contains the diagnostic coding typically used to estimate trends in emergency department (ED) visits in the United States. However, SPARCs is not timely with current lags between 18 and 30 months. The second data system, syndromic surveillance, is based on ED entry logs and is coded using text processing of the patient's chief complaints. Although it is clear that such processing will likely have lower sensitivity and specificity for mental health related visits, it is not clear that diagnostic coding provides a gold standard. The project found high levels of correspondence between the two data streams with similar temporal and geographic patterns, and similar measured disparities. This suggests the potential to use syndromic for timely public health response and to provide feedback to ongoing programs. This is important for two reasons. First, the

ThriveNYC initiative is the most extensive and expensive mental health program ever undertaken in New York City and ED visits for mental health should provide a clear indicator of the programs success given its stated goal to provide preventive mental health services. Second, New York City schools have been repeatedly criticized for over-using hospital EDs to address behavioral issues. This has led to an explicit policy forbidding the activity; nonetheless the majority of hospital ED visits for child mental health occur on school days during school hours.

Future work here includes incorporation of a number of developments in the data systems. First, syndromic surveillance data has increasing completeness on ICD9 coding, nurse triage notes, mode of arrival, and disposition fields. This should allow for improved classification. Second, syndromic surveillance and SPARCs data have been linked to the extent that they can for 2016 and 2017 allowing for the construction of sensitivity and specificity estimates for syndromic mental health coding assuming SPARCs as a gold standard. Linking the two systems will also allow for the characterization of records that don't match. For example, what non-mental health ICD9 codes are most common when syndromic classifies a visit as mental health related. Third, New York State Medicaid information – which is available with an average lag of two weeks – is now available for linking to student records. This should allow for the direct identification of schools overusing EDs for behavioral issues.

Finally, **Child obesity: Longitudinal information, local government, and public policy** describes the implementation of quantile regression models for adolescent BMI. A quantile regression model produced solely with age by gender terms returns the age and gender specific growth curves for the population in the model. These growth curves can then be directly compared with current CDC growth curves which were established in mid 1980s and which are used to establish child obesity estimates. By constructing such models it is possible to identify geographic areas or demographic subgroups that are more obese than expected. By extending the model to include longitudinal information, individual growth trajectories can be estimated. Together, this leads to several improvements to current approaches. First, the quantity of information used is greatly increased. Currently, an obese indicator is constructed from a single age- and gender-specific threshold value. Monitoring prevalence then ignores changes in children unless they cross this threshold. Moving to a quantile regression framework increases the power to detect actual change across the distribution of children. This allows for improved targeting of programs and increased power to conduct evaluations, provides a longitudinal framework for looking at impact that includes a counterfactual distribution, and enables assessment of data quality.

Future work here should focus on realizing these benefits by applying the framework to real-world problems. For example, universal free lunch was implemented over the past several years in New York City, a longitudinal framework could incorporate the implementation to estimate the impact of the policy to various BMI quantiles. Likewise, data quality issues including "implausible" values can be assessed using the expected distribution of values. Geographic and demographic disparities can be constructed using a simple model and then mapping or characterizing deviations from the model. Lastly, the framework could be provided to other jurisdictions for their use. Interestingly, while this would greatly increase the amount of information being employed and improving the way it is employed, it would simply return such analyses to the original 19th century approach of demographers such as Adolphe Quetelet (1796-1874).

These examples demonstrate the value of considering space and place in public health surveillance. The increasing importance of Geography in public health will continue reflecting continued development of data systems and similar trends in other fields relevant to local government such as education, social services, and criminal justice. Ultimately, local agencies will rely on detailed characterizations of the communities they serve to inform and improve their public service provision. And community-based organizations and the public will rely on such characterizations as they advocate for efficient and just provision of services and more effective, responsive government. Geography is central to this effort.

Bibliography

Acri, M. C., L. A. Bornheimer, K. O'Brien, S. Sezer, V. Little, A. F. Cleek, and M. M. McKay (2016). A model of integrated health care in a poverty-impacted community in New York City: Importance of early detection and addressing potential barriers to intervention implementation. *Social Work in Health Care 55*(4), 314–327.

Albrecht, J. and M. Abramovitz (2014). The community loss index: A new social indicator. *URISA Journal 26*(1), 5–12.

Alegría, M., J. G. Green, K. McLaughlin, and S. Loder (2015). Disparities in child and adolescent mental health and mental health services in the US. Technical report, New York, NY: William T. Grant Foundation.

Almond, D., A. Lee, and A. E. Schwartz (2016). Impacts of classifying New York City students as overweight. *Proceedings of the National Academy of Sciences 113*(13), 3488–3491.

Althouse, B. M., S. V. Scarpino, L. A. Meyers, J. W. Ayers, M. Bargsten, J. Baumbach, J. S. Brownstein, L. Castro, H. Clapham, D. A. Cummings, et al. (2015). Enhancing disease surveillance with novel data streams: Challenges and opportunities. *EPJ Data Science 4*(1), 17.

Anderson, R. N., A. M. Miniño, D. L. Hoyert, and H. M. Rosenberg (2001). Comparability of cause of death between icd-9 and icd-10: preliminary estimates. *National Vital Statistics Report 49*(2), 1–32.

Aneshensel, C. S. (2009). Toward explaining mental health disparities. *Journal of Health and Social Behavior 50*(4), 377–394.

Austensen, M., I. Gould Ellen, L. Herrine, B. Karfunkel, G. Khun Jush, S. Moriarty, et al. (2016). State of new york city's housing and neighborhoods in 2015. `http://furmancenter.org/files/sotc/NYUFurmanCenter_SOCin2015_9JUNE2016.pdf`. Accessed: 2019-01-04.

Ayers, J. W., B. M. Althouse, and M. Dredze (2014). Could behavioral medicine lead the web data revolution? *JAMA 311*(14), 1399–1400.

Babey, S. H., J. Wolstein, A. L. Diamant, A. Bloom, and H. Goldstein (2011). A patchwork of progress: Changes in overweight and obesity among california 5th, 7th, and 9th graders, 2005-2010.

Baltimore City Health Department (2019). 2017 Neighborhood Health Profile Technical Notes. `https://health.baltimorecity.gov/sites/default/files/2019.01.22%20OpenBaltimore%20Baltimore%20City%202017%20Neighborhood%20Health%20Profiles%20Technical%20Notes.pdf`. Accessed: 2019-07-04.

Barr, D. A. (2014). *Health Disparities in the United States: Social Class, Race, Ethnicity, and Health*. JHU Press.

Berger, M., K. Konty, S. Day, L. Silver, C. Nonas, B. Kerker, C. Greene, T. Farley, and L. Harr (2011). Obesity in K-8 students-New York City, 2006-07 to 2010-11 school years. *MMWR. Morbidity and Mortality Weekly Report 60*(49), 1673.

Bernard, P., R. Charafeddine, K. L. Frohlich, M. Daniel, Y. Kestens, and L. Potvin (2007). Health inequalities and place: A theoretical conception of neighbourhood. *Social Science & Medicine 65*(9), 1839–1852.

Birkhead, G. S., M. Klompas, and N. R. Shah (2015). Uses of electronic health records for public health surveillance to advance public health. *Annual Review of Public Health 36*, 345–359.

Bradley, J. R., C. K. Wikle, and S. H. Holan (2016). Bayesian spatial change of support for count-valued survey data with application to the american community survey. *Journal of the American Statistical Association 111*(514), 472–487.

Braveman, P. (2006). Health disparities and health equity: Concepts and measurement. *Annual Review of Public Health 27*, 167–194.

Braveman, P., S. Egerter, and D. R. Williams (2011). The social determinants of health: Coming of age. *Annual Review of Public Health 32*, 381–398.

Braveman, P. and L. Gottlieb (2014). The social determinants of health: It's time to consider the causes of the causes. *Public Health Reports 129*(1_suppl2), 19–31.

Breslau, J., S. Aguilar-Gaxiola, K. S. Kendler, M. Su, D. Williams, and R. C. Kessler (2006). Specifying race-ethnic differences in risk for psychiatric disorder in a USA national sample. *Psychological Medicine 36*(1), 57–68.

Bringewatt, E. H. and E. T. Gershoff (2010). Falling through the cracks: Gaps and barriers in the mental health system for america's disadvantaged children. *Children and Youth Services Review 32*(10), 1291–1299.

Buehler, J. W., A. Sonricker, M. Paladini, P. Soper, and F. Mostashari (2008). Syndromic surveillance practice in the United States: Findings from a survey of state, territorial, and selected local health departments. *Advances in Disease Surveillance 6*(3), 1–20.

Calman, N. S., M. Golub, C. Ruddock, and L. Le (2006). Separate and unequal care in New York City. *Journal of Health Care Law & Policy 9*, 105.

Carubia, B., A. Becker, and B. H. Levine (2016). Child psychiatric emergencies: Updates on trends, clinical care, and practice challenges. *Current Psychiatry Reports 18*(4), 41.

Centers for Disease Control and Prevention (2000, May 30). CDC growth charts, United States. http://www.cdc.gov/growthcharts/.

Centers for Disease Control and Prevention (2010). Use of World Health Organization and CDC growth charts for children aged 0-59 months in the United States. *MMWR 59*(RR-9), 1–14.

Centers for Disease Control and Prevention (2019). Ten essential public health services and how they can include addressing social determinants of health inequities. `https://www.cdc.gov/publichealthgateway/publichealthservices/pdf/Ten_Essential_Services_and_SDOH.pdf`. Accessed: 2019-02-02.

Chaix, B. (2018). Mobile sensing in environmental health and neighborhood research. *Annual Review of Public Health 39*, 367–384.

Child and A. H. M. Initiative (September,2012). National Profile of Complementary and Alternative Medicine (CAM) Use for Children with Emotional, Mental, or Behavioral Conditions or Problems (2-17 years). Final report, Data Resource Center for Child & Adolescent Health.

Christodulu, K. V., R. Lichenstein, M. D. Weist, M. E. Shafer, and M. Simone (2002). Psychiatric emergencies in children. *Pediatric Emergency Care 18*(4), 268–270.

Citizens' Committee for Children of New York, I. (January,2012). New York City's Children and Mental Health: Prevalence and Gap Analysis of Treatment Slot Capacity. Final report, NYC Citywide Children's Committee and NYC Early Childhood Strategic Mental Health Workgroup.

City of New York (2007). Planyc: a greener, greater new york. `http://www.nyc.gov/html/planyc/downloads/pdf/publications/full_report_2007.pdf`. Accessed: 2019-01-04.

Cole, T. J. and P. J. Green (1992). Smoothing reference centile curves: The LMS method and penalized likelihood. *Statistics in Medicine 11*(10), 1305–1319.

Colgrove, J. K., G. E. Markowitz, and D. Rosner (2008). *The Contested Boundaries of American Public Health*. Rutgers University Press.

Corcoran, S. P., B. Elbel, and A. E. Schwartz (2014). The effect of breakfast in the classroom on obesity and academic performance: Evidence from New York City. Working Paper# 04-14.

Cottage Health (2015). Community health needs assessment report 2016. `https://www.cottagehealth.org/app/files/public/1991/Cottage_Population_Health_CHNA_Report_020217.pdf`. Accessed: 2019-08-20.

County of San Diego Health and Human Services Agency (2015). San Diego Community Health Profiles: Data Guide, Methodology, and Metadata. `https://www.sandiegocounty.gov/content/dam/sdc/hhsa/programs/phs/CHS/Community%20Profiles/Community%20Profiles%20Data%20Guide_2015.pdf`. Accessed: 2019-07-04.

Day, S. E., K. J. Konty, M. Leventer-Roberts, C. Nonas, and T. G. Harris (2014). Severe obesity among children in New York City public elementary and middle schools, school years 2006–07 through 2010–11. *Preventing Chronic Disease 11*, E118.

Delmelle, E. C. (2019). The increasing sociospatial fragmentation of urban america. *Urban Science 3*(1), 9.

Dietz, W. (2016). Are we making progress in the prevention and control of childhood obesity? It all depends on how you look at it. *Obesity 24*(5), 991–992.

Diez-Roux, A. (2016). On the distinction –or lack of distinction –between population health and public health. *American Journal of Public Health 106*(4), 619.

Diez-Roux, A. V. (1998). Bringing context back into epidemiology: Variables and fallacies in multilevel analysis. *American Journal of Public Health 88*(2), 216–222.

Diez-Roux, A. V. (2001). Investigating neighborhood and area effects on health. *American Journal of Public Health 91*(11), 1783–1789.

Diez Roux, A. V. (2004). Estimating the neighborhood health effects: The challenges of casual inference in a complex world. *Social Science & Medicine 58*(10), 1953–1960.

Diez-Roux, A. V. and C. Mair (2010). Neighborhoods and health. *Annals of the New York Academy of Sciences 1186*(1), 125–145.

Dryfoos, J. G., J. Quinn, and C. Barkin (2005). *Community Schools in Action: Lessons from a Decade of Practice*. Oxford University Press.

Duncan, D. T., F. Kapadia, S. D. Regan, W. C. Goedel, M. D. Levy, S. C. Barton, S. R. Friedman, and P. N. Halkitis (2016). Feasibility and acceptability of global positioning system (gps) methods to study the spatial contexts of substance use and sexual risk behaviors among young men who have sex with men in new york city: A p18 cohort sub-study. *PloS one 11*(2), e0147520.

Duncan, D. T. and I. Kawachi (2018). *Neighborhoods and Health.* Oxford University Press New York.

Duncan, D. T., S. D. Regan, D. Shelley, K. Day, R. R. Ruff, M. Al-Bayan, and B. Elbel (2014). Application of global positioning system methods for the study of obesity and hypertension risk among low-income housing residents in new york city: a spatial feasibility study. *Geospatial Health 9*(1), 57.

Egger, J., K. Konty, K. Bartley, L. Benson, D. Bellino, and B. Kerker (2009). Childhood obesity is a serious concern in New York City: Higher levels of fitness associated with better academic performance. *NYC Vital Signs 8*(1), 1–4.

Fairchild, A. L., D. Rosner, J. Colgrove, R. Bayer, and L. P. Fried (2010). The exodus of public health: What history can tell us about the future. *American Journal of Public Health 100*(1), 54–63.

Fisher, G. M. (1992). The development and history of the poverty thresholds. *Social Security Bulletin 55*, 3.

Flegal, K. M., R. Wei, C. L. Ogden, D. S. Freedman, C. L. Johnson, and L. R. Curtin (2009). Characterizing extreme values of body mass index–for-age by using the 2000 Centers for Disease Control and Prevention growth charts. *The American Journal of Clinical Nutrition 90*(5), 1314–1320.

Fleiss, J. L., B. Levin, and M. C. Paik (2013). *Statistical Methods for Rates and Proportions.* John Wiley & Sons.

Fotheringham, A. S. and D. W. Wong (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and planning A 23*(7), 1025–1044.

Gourevitch, M. N., L. H. Curtis, M. S. Durkin, A. Fagerlin, A. C. Gelijns, R. Platt, B. M. Reininger, J. Wylie-Rosett, K. Jones, and W. M. Tierney (2019). The emergence of population health in us academic medicine: A qualitative assessment. *JAMA network open 2*(4), e192200–e192200.

Graham, H. (2004). Social determinants and their unequal distribution: Clarifying policy understandings. *The Milbank Quarterly 82*(1), 101–124.

Graunt, J. (1665). *Natural and Political Observations Mentioned in a Following Index, and Made Upon the Bills of Mortality, 4th edition.* London: the Royal Society.

Grupp-Phelan, J., J. S. Harman, and K. J. Kelleher (2007). Trends in mental health and chronic condition visits by children presenting for care at US emergency departments. *Public Health Reports 122*(1), 55–61.

Gulati, A. K., D. W. Kaplan, and S. R. Daniels (2012). Clinical tracking of severely obese children: A new growth chart. *Pediatrics 130*(6), 1136–1140.

Gunderson, G. (2013). National school lunch act. `https://www.fns.usda.gov/nslp/history_5`.

Hägerstraand, T. (1970). What about people in regional science? *Papers in Regional Science 24*(1), 7–24.

Hakenewerth, A., J. Tintinalli, A. Waller, I. AI, and D. T (2013). Emergency department visits by patients with mental health disorders–North Carolina, 2008-2010. *MMWR. Morbidity and Mortality Weekly Report 62*(23), 469.

Hakenewerth, A. M., A. E. Waller, A. I. Ising, and J. E. Tintinalli (2009). North Carolina Disease Event Tracking and Epidemiologic Collection Tool (NC DETECT) and the National Hospital Ambulatory Medical Care Survey (NHAMCS): Comparison of emergency department data. *Academic Emergency Medicine 16*(3), 261–269.

Heffernan, R., F. Mostashari, D. Das, A. Karpati, M. Kulldorff, D. Weiss, et al. (2004). Syndromic surveillance in public health practice, New York City. *Emerging Infectious Diseases 10*(5), 858–864.

Hill, H. W. (1916). *The New Public Health*. Macmillan.

Hsia, R. Y. and M. Niedzwiecki (2017). Avoidable emergency department visits: A starting point. *International Journal for Quality in Health Care 29*(5), 642–645.

Hurvitz, P. M. and A. V. Moudon (2012). Home versus nonhome neighborhood: Quantifying differences in exposure to the built environment. *American Journal of Preventive Medicine 42*(4), 411–417.

Jackson, K. T., L. Keller, and N. Flood (2010). *The Encyclopedia of New York City*. Yale University Press.

Jin, Y. and J. C. Jones-Smith (2015). Associations between family income and children's physical fitness and obesity in California, 2010–2012. *Preventing Chronic Disease 12*, E17.

Jones-Smith, J. C., W. H. Dow, and K. Chichlowska (2014). Association between casino opening or expansion and risk of childhood overweight and obesity. *JAMA 311*(9), 929–936.

Kessler, R. C., G. Andrews, L. J. Colpe, E. Hiripi, D. K. Mroczek, S.-L. Normand, E. E. Walters, and A. M. Zaslavsky (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine 32*(6), 959–976.

Kessler, R. C., C. L. Foster, W. B. Saunders, and P. E. Stang (1995). Social consequences of psychiatric disorders, I: Educational attainment. *American Journal of Psychiatry 152*(7), 1026–1032.

Khoury, M. J. and J. P. Ioannidis (2014). Big data meets public health. *Science 346*(6213), 1054–1055.

Kieling, C., H. Baker-Henningham, M. Belfer, G. Conti, I. Ertem, O. Omigbodun, L. A. Rohde, S. Srinath, N. Ulkuer, and A. Rahman (2011). Child and adolescent mental health worldwide: Evidence for action. *The Lancet 378*(9801), 1515–1525.

Kim, G., J. DeCoster, A. N. Bryant, and K. L. Ford (2016). Measurement equivalence of the K6 scale: The effects of race/ethnicity and language. *Assessment 23*(6), 758–768.

Koenker, R. (2015). *Quantreg: Quantile Regression*. R package version 5.19.

Koenker, R. (2018). Quantile regression: 40 years on. *Annual Review of Economics 9*(1), 155–176.

Koenker, R. and G. Bassett Jr (1978). Regression quantiles. *Econometrica 46*(1), 33–50.

Koenker, R., P. Ng, and S. Portnoy (1994). Quantile smoothing splines. *Biometrika 81*(4), 673–680.

Kohl III, H. W., H. D. Cook, et al. (2013). *Educating the student body: Taking physical activity and physical education to school*. National Academies Press.

Krampner, J., J. Shin, V. Virgin, Q. Tran, E. Studer, and C. Li (2017). New york city government poverty measure 2005-2015. Technical report, New York City Center for Economic Opportunity.

Krieger, N. (1992). Overcoming the absence of socioeconomic data in medical records: Validation and application of a census-based methodology. *American Journal of Public Health 82*(5), 703–710.

Krieger, N., J. T. Chen, P. D. Waterman, D. H. Rehkopf, and S. Subramanian (2003). Race/ethnicity, gender, and monitoring socioeconomic gradients in health: A comparison of area-based socioeconomic measures-the public health disparities geocoding project. *American Journal of Public Health 93*(10), 1655–1671.

Krieger, N., J. T. Chen, P. D. Waterman, D. H. Rehkopf, and S. Subramanian (2005). Painting a truer picture of us socioeconomic and racial/ethnic health inequalities: The public health disparities geocoding project. *American Journal of Public Health 95*(2), 312–323.

Krieger, N., J. T. Chen, P. D. Waterman, M.-J. Soobader, S. Subramanian, and R. Carson (2002). Geocoding and monitoring of us socioeconomic inequalities in mortality and cancer incidence: Does the choice of area-based measure and geographic level matter? The Public Health Disparities Geocoding Project. *American Journal of Epidemiology 156*(5), 471–482.

Krieger, N., P. D. Waterman, J. T. Chen, M.-J. Soobader, and S. Subramanian (2003). Monitoring socioeconomic inequalities in sexually transmitted infections, tuberculosis, and violence: Geocoding and choice of area-based socioeconomic measures – The public health disparities geocoding project (US). *Public Health Reports 118*, 240–260.

Krieger, N., P. D. Waterman, J. Spasojevic, W. Li, G. Maduro, and G. Van Wye (2016). Public health monitoring of privilege and deprivation with the index of concentration at the extremes. *American Journal of Public Health 106*(2), 256–263.

Krieger, N., D. R. Williams, and N. E. Moss (1997). Measuring social class in us public health research: Concepts, methodologies, and guidelines. *Annual Review of Public Health 18*(1), 341–378.

Kwan, M.-P. (2012). The uncertain geographic context problem. *Annals of the Association of American Geographers 102*(5), 958–968.

Lall, R., J. Abdelnabi, S. Ngai, H. B. Parton, K. Saunders, J. Sell, A. Wahnich, D. Weiss, and R. W. Mathes (2017). Advancing the use of emergency department syndromic surveillance data, New York City, 2012-2016. *Public Health Reports 132*(1_suppl), 23S–30S.

Last, J. (2001). *A Dictionary of Epidemiology*. Oxford University Press.

Lewallen, T. C., H. Hunt, W. Potts-Datema, S. Zaza, and W. Giles (2015). The whole school, whole community, whole child model: A new approach for improving educational attainment and healthy development for students. *Journal of School Health 85*(11), 729–739.

Li, W., P. Zheng, M. Huynh, A. Castro, L. Falci, G. Maduro, E. Lee, Y. Sun, and G. Van Wye (2018). Summary of vital statistics 2016. Technical report, New York, NY: New York City Department of Health and Mental Hygiene.

Love, D., W. Custer, and P. Miller (2010). All-payer claims databases: state initiatives to improve health care transparency. Technical Report 1439, Commonwealth Fund.

Lu, W. (2017). Child and adolescent mental disorders and health care disparities: Results from the National Survey of Children's Health, 2011–2012. *Journal of Health Care for the Poor and Underserved 28*(3), 988–1011.

Macintyre, S. and A. Ellaway (2000). Ecological approaches: Rediscovering the role of the physical and social environment. *Social Epidemiology 9*(5), 332–348.

Marmot, M., S. Friel, R. Bell, T. A. Houweling, S. Taylor, C. on Social Determinants of Health, et al. (2008). Closing the gap in a generation: Health equity through action on the social determinants of health. *The Lancet 372*(9650), 1661–1669.

Massey, D. S. (2001). The prodigal paradigm returns: Ecology comes back to sociology. In A. Booth and A. Crouter (Eds.), *Does It Take a Village? Community Effects on Children, Adolescents, and Families.*, pp. 41–48. Mahwah NJ: Lawrence Erlbaum Associates.

McCray, C., R. Buery, and M. Bassett (2015). ThriveNYC: A mental health roadmap for all. Technical report, Mayor's Office: New York City.

McGuire, T. G., M. Alegria, B. L. Cook, K. B. Wells, and A. M. Zaslavsky (2006). Implementing the Institute of Medicine definition of disparities: An application to mental health care. *Health Services Research 41*(5), 1979–2005.

McLafferty, S. and A. T. Murray (2017). *Regional perspectives on public health*, pp. 161–173. Cham: Springer International Publishing.

Mechanic, D. and M. Olfson (2016). The relevance of the Affordable Care Act for improving mental health care. *Annual Review of Clinical Psychology 12*, 515–542.

Mettey, A., A. Garcia, L. Isaac, N. Linos, O. Barbot, and M. Bassett (2015). Take care new york 2020: Every neighborhood, every new yorker, everyone's health counts. Technical report, New York City Department of Health and Mental Hygiene.

Metzger, K. B., A. Hajat, M. Crawford, and F. Mostashari (2004). How many illnesses does one emergency department visit represent? Using a population-based telephone survey to estimate the syndromic multiplier. *Morbidity and Mortality Weekly Report 53*(Supplement), 106–111.

Miller, M., D. Love, R. Sullivan, J. Porter, and A. Costello (2015). All-payer claims databases: An overview for policymakers. `http://www.statecoverage.org/files/SCI_All_Payer_Claims_ReportREV.pdf`. Accessed: 2019-07-17.

Mills, C. and M. Davila (June,2016). Psychiatric hospitalizations among children and adolescents in New York City. Epi Data Brief 70, New York City Department of Health and Mental Hygiene.

More, L. B. (1907). *Wage-earners' Budgets: A Study of Standards and Cost of Living in New York City.* H. Holt and Company.

Morgan, R. and A. Gendar (2004). The new craze schools using ERs as dumping ground educators shuttling surly students to city's white coats. *New York Daily News.* Available at: http://www.nydailynews.com/archives/news/new-craze-schools-ers-dumpingground-educators-shuttling-surly-students-citywhite-coats-article-1.647686#ixzz2uMTckZNk (accessed – 2018-01-19).

Mostashari, F. and J. Hartman (2003). Syndromic surveillance: A local perspective. *Journal of Urban Health 80*(Suppl 1), i1–i7.

New York City Planning Commission (1969). *Plan for New York City 1969: A Proposal.* Department of City Planning, City of New York.

NYS Department of Health Bureau of Health Informatics (2014). Statewide planning and research cooperative system (sparcs) data governance policy and procedure manual. `https://www.health.ny.gov/statistics/sparcs/training/docs/sparcs_dgc_manual.pdf`. Accessed: 2019-05-20.

Oakes, J., A. Maier, and J. Daniel (2017). Community schools: An evidence-based strategy for equitable school improvement. Technical report, National Education Policy Center.

Ogden, C. L., M. D. Carroll, B. K. Kit, and K. M. Flegal (2014). Prevalence of childhood and adult obesity in the United States, 2011–2012. *JAMA 311*(8), 806–814.

Ogden, C. L., K. M. Flegal, M. D. Carroll, and C. L. Johnson (2002). Prevalence and trends in overweight among US children and adolescents, 1999–2000. *JAMA 288*(14), 1728–1732.

Olson, D. R., R. T. Heffernan, M. Paladini, K. Konty, D. Weiss, and F. Mostashari (2007). Monitoring the impact of influenza by age: Emergency department fever and respiratory complaint surveillance in New York City. *PLoS Medicine 4*(8), e247.

Orshansky, M. (1965). Counting the poor: Another look at the poverty profile. *Social Security Bulletin 28*, 3.

Ostfeld, R. S., G. E. Glass, and F. Keesing (2005). Spatial epidemiology: an emerging (or re-emerging) discipline. *Trends in Ecology & Evolution 20*(6), 328–336.

Park, R. E. and E. W. Burgess (1925). *The City.* University of Chicago Press.

Pate, R., M. Oria, L. Pillsbury, et al. (2012). *Fitness measures and health outcomes in youth.* National Academies Press.

Patricios, N. N. (2002). The neighborhood concept: A retrospective of physical design and social interaction. *Journal of Architectural and Planning Research 19*(1), 70–90.

Pavlin, J. A., F. Mostashari, M. G. Kortepeter, N. A. Hynes, R. A. Chotani, Y. B. Mikol, M. A. Ryan, J. S. Neville, D. T. Gantz, J. V. Writer, et al. (2003). Innovative surveillance methods for rapid detection of disease outbreaks and bioterrorism: Results of an interagency workshop on health indicator surveillance. *American Journal of Public Health 93*(8), 1230–1235.

Pennel, C. L., K. R. McLeroy, J. N. Burdine, and D. Matarrita-Cascante (2015). Nonprofit hospitals' approach to community health needs assessment. *American journal of public health 105*(3), e103–e113.

Perou, R., R. H. Bitsko, S. J. Blumberg, P. Pastor, R. M. Ghandour, J. C. Gfroerer, S. L. Hedden, A. E. Crosby, S. N. Visser, L. A. Schieve, et al. (2013). Mental health surveillance among children? United States, 2005–2011. *MMWR Surveillance Summary 62*(Suppl 2), 1–35.

Perry, C. A. (1929). City planning for neighborhood life. *Social Forces 8*, 98.

Piantadosi, S., D. P. Byar, and S. B. Green (1988). The ecological fallacy. *American Journal of Epidemiology 127*(5), 893–904.

Pittsenbarger, Z. E. and R. Mannix (2014). Trends in pediatric visits to the emergency department for psychiatric illnesses. *Academic Emergency Medicine 21*(1), 25–30.

Presanis, A. M., D. De Angelis, A. Hagy, C. Reed, S. Riley, B. S. Cooper, L. Finelli, P. Biedrzycki, M. Lipsitch, et al. (2009). The severity of pandemic h1n1 influenza in the united states, from april to july 2009: a bayesian analysis. *PLoS medicine 6*(12), e1000207.

Program, P. Y. F. (2014). Monitoring student fitness levels.

Rainham, D., I. McDowell, D. Krewski, and M. Sawada (2010). Conceptualizing the healthscape: Contributions of time geography, location technologies and spatial ecology to place and health research. *Social Science & Medicine 70*(5), 668–676.

Reardon, S. F. and D. O'Sullivan (2004). Measures of spatial segregation. *Sociological Methodology 34*(1), 121–162.

Reid, M. (2011). Behind the "glasgow effect". *Bulletin of the World Health Organization 89*(10), 701–776.

Reilly, K. H., E. Johns, N. Noyan, M. Schretzman, and T.-Y. Tsao (2019). Potentially avertable premature deaths associated with jail incarceration in New York City. *Journal of Community Health*, 1–7.

Robbins, J. M., G. Mallya, M. Polansky, and D. F. Schwarz (2015). Prevalence, disparities, and trends in obesity and severe obesity among students in the Philadelphia, Pennsylvania, school district, 2006–2010. In P. Vash (Ed.), *The Childhood Obesity Epidemic: Why Are Our Children ObeseAnd What Can We Do About It?*, pp. 29–42. CRC Press.

RWJF (2012). Declining childhood obesity rates – Where are we seeing the most progress? Health Policy Snapshot, Issue Brief.

RWJF Data for Health Advisory Committee (2015). Data for health: Learning what works. `https://www.rwjf.org/en/library/research/2015/04/data-for-health-initiative.html`. Accessed: 2019-05-20.

Schwartz, A. E., M. Leardo, S. Aneja, and B. Elbel (2016). Effect of a school-based water intervention on child body mass index and obesity. *JAMA Pediatrics 170*, 220–226.

Simon, A. E. and K. C. Schoendorf (2014). Emergency department visits for mental health conditions among US children, 2001-2011. *Clinical Pediatrics 53*(14), 1359–1366.

Singh, G. K., M. Siahpush, R. A. Hiatt, and L. R. Timsina (2011). Dramatic increases in obesity and overweight prevalence and body mass index among ethnic-immigrant and social class groups in the united states, 1976–2008. *Journal of Community Health 36*(1), 94–110.

Singh, G. K., M. Siahpush, and M. D. Kogan (2010). Neighborhood socioeconomic conditions, built environments, and childhood obesity. *Health affairs 29*(3), 503–512.

Skinner, A. C., E. M. Perrin, and J. A. Skelton (2016). Prevalence of obesity and severe obesity in US children, 1999-2014. *Obesity 24*(5), 1116–1123.

Skinner, A. C. and J. A. Skelton (2014). Prevalence and trends in obesity and severe obesity among children in the United States, 1999-2012. *JAMA Pediatrics 168*(6), 561–566.

Smith, S. K. and B. B. Lewis (1980). Some new techniques for applying the housing unit method of local population estimation. *Demography 17*(3), 323–339.

Sperling, J. (2012). The tyranny of census geography: Small-area data and neighborhood statistics. *Cityscape*, 219–223.

Spielman, S. E., D. Folch, and N. Nagle (2014). Patterns and causes of uncertainty in the american community survey. *Applied Geography 46*, 147–157.

Spielman, S. E. and D. C. Folch (2015). Reducing uncertainty in the american community survey through data-driven regionalization. *PloS one 10*(2), e0115626.

Spielman, S. E. and J. R. Logan (2013). Using high-resolution population data to identify neighborhoods and establish their boundaries. *Annals of the Association of American Geographers 103*(1), 67–84.

Spielman, S. E. and E.-h. Yoo (2009). The spatial dimensions of neighborhood effects. *Social Science & Medicine 68*(6), 1098–1105.

Spielman, S. E., E.-H. Yoo, and C. Linkletter (2013). Neighborhood contexts, health, and behavior: Understanding the role of scale and residential sorting. *Environment and Planning B: Planning and Design 40*(3), 489–506.

Subramanian, S. and S. Kranes (2015). Beyond meal status: A new measure for quantifying poverty levels in the city's schools. Technical report, New York City Independent Budget Office.

Tanner, J. M. and J. M. Tanner (1981). *A history of the study of human growth.* Cambridge university press.

Tavernise, S. and S. Sun (2015). Same city, but very different life spans. https://www.nytimes.com/interactive/2015/04/29/health/life-expectancy-nyc-chi-atl-richmond.html. Accessed: 2019-01-04.

Thorpe, L. (2017). Surveillance as our sextant. *American Journal of Public Health 107*(6), 847–848.

Thorpe, L. E., D. G. List, T. Marx, L. May, S. D. Helgerson, and T. R. Frieden (2004). Childhood obesity in New York City elementary school students. *American Journal of Public Health 94*(9), 1496–1500.

Toprani, A. and J. Hadler (March,2016). *Selecting and applying a standard area-based socioeconomic status measure for public health data: Analysis for New York City.* Number 68.

Toprani, A., W. Li, and J. L. Hadler (2016). Trends in mortality disparities by area-based poverty in new york city, 1990–2010. *Journal of Urban Health 93*(3), 538–550.

Tsao, T.-Y., K. J. Konty, G. Van Wye, O. Barbot, J. L. Hadler, N. Linos, and M. T. Bassett (2016). Estimating potential reductions in premature mortality in new york city from raising the minimum wage to $15. *American Journal of Public Health 106*(6), 1036–1041.

Tseliou, F., M. Rosato, and D. O'Reilly (2018). Mental health need: use of administrative data and record linkage to inform mental health policy and practice. *The Lancet 392*, S88.

Tucci, V., K. Siever, A. Matorin, and N. Moukaddam (2015). Down the rabbit hole: Emergency department medical clearance of patients with psychiatric or behavioral emergencies. *Emergency Medicine Clinics 33*(4), 721–737.

U.S. Census Bureau (2007). Current population survey, annual social and economic supplement.

U.S. Census Bureau (2009a). A compass for understanding and using american community survey data: What researchers need to know.

U.S. Census Bureau (2009b). A compass for understanding and using american community survey data: What users of data for american indians and alaska natives need to know.

U.S. Census Bureau (2014). American community survey design and methodology (version 2.0, january 2014). https://www.cdc.gov/publichealthgateway/publichealthservices/pdf/Ten_Essential_Services_and_SDOH.pdf. Accessed: 2019-01-04.

U.S. Census Bureau (2017). Small area income and poverty estimates (saipe) program: Saipe methodology 2017. https://www.census.gov/programs-surveys/saipe/technical-documentation/methodology.html. Accessed: 2019-01-04.

U.S. Census Bureau (2018). Understanding and using american community survey data: What all data users need to know. https://www.census.gov/content/dam/Census/library/publications/2018/acs/acs_general_handbook_2018.pdf. Accessed: 2019-07-14.

US Department of Health and Human Services (2000). *Healthy People 2010.* Office of Disease Prevention and Health Promotion, US Department of Health and Human Services.

Wachter, K. W. (2005). Spatial demography. *Proceedings of the National Academy of Sciences 102*(43), 15299–15300.

Wang, L., Y. Zhou, R. Song, and B. Sherwood (2018). Quantile-optimal treatment regimes. *Journal of the American Statistical Association 113*(523), 1243–1254.

Wei, Y. (2004). *Longitudinal Growth Charts Based on Semi-parametric Quantile Regression.* Ph. D. thesis, University of Illinois at Urbana-Champaign.

Wei, Y., A. Pere, R. Koenker, and X. He (2005). Quantile regression methods for reference growth charts. *Statistics in Medicine 25*(8), 1369–1382.

Welk, G. and M. D. Meredith (2010). *Fitnessgram and Activitygram Test Administration Manual-Updated 4th Edition.* Human Kinetics.

Welk, G. J., M. D. Meredith, M. Ihmels, and C. Seeger (2010). Distribution of health-related physical fitness in Texas youth: A demographic and geographic analysis. *Research Quarterly for Exercise and Sport 81*(sup3), S6–S15.

Whitehead, M. (1990). The concepts and principles of equity in health. Technical report, Copenhagen: WHO, Reg. Off. Eur.

WHO Comm of Soc Det of Health (2008). Closing the gap in a generation: Health equity through action on the social determinants of health: Final report of the commission on social determinants of health. Who report, Geneva: World Health Organization.

Wickham, S., M. Whitehead, D. Taylor-Robinson, and B. Barr (2017). The effect of a transition into poverty on child and maternal mental health: A longitudinal analysis of the UK Millennium Cohort Study. *The Lancet Public Health 2*(3), e141–e148.

Winerip, M. (2012). Keeping students' mental health care out of the E.R. *New York Times.* Available at: https://www.nytimes.com/2012/04/09/nyregion/trying-to-keep-students-mental-health-care-out-of-the-er.html (accessed – 2018-01-1).

Winslow, C. (1920). The untilled fields of public health. *Science*, 23–33.

Wong, D. W. (1992). The reliability of using the iterative proportional fitting procedure. *The Professional Geographer 44*(3), 340–348.

Yim, B., R. E. Howland, G. M. Culp, A. Zhilkova, O. Barbot, and T.-Y. Tsao (2018). Disparities in preventable hospitalizations among public housing developments. *American Journal of Preventive Medicine*.