6-22-2018

Elucidating the Role of Neighborhood Deprivation in Hypertensive Disorders of Pregnancy

Kelly M. Winter
Florida International University, kwint003@fiu.edu

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

ELUCIDATING THE ROLE OF NEIGHBORHOOD DEPRIVATION

IN HYPERTENSIVE DISORDERS OF PREGNANCY

A dissertation submitted in partial fulfillment of the

requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PUBLIC HEALTH

by

Kelly Michelle Winter

2018

To: Dean Tomás R. Guilarte
    Robert Stempel College of Public Health and Social Work

This dissertation, written by Kelly Michelle Winter, and entitled Elucidating the Role of Neighborhood Deprivation in Hypertensive Disorders of Pregnancy, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

_____
Jeffrey Onsted

_____
Tan Li

_____
Wasim Maziak

_____
Mary Jo Trepka, Major Professor

Date of Defense: June 22, 2018

The dissertation of Kelly Michelle Winter is approved.

_____
Dean Tomás R. Guilarte
Robert Stempel College of Public Health and Social Work

_____
Andrés G. Gil
Vice President for Research and Economic Development
Dean of the University Graduate School

Florida International University, 2018

DEDICATION

I am living proof that it takes a village. I dedicate this dissertation to my husband,

Anthony, the most brilliant, hilarious, loving, inspiring partner I could ask for. Thanks for

encouraging me to be braver and more proud of our mutant status and for building the

coolest life with me. I'm so grateful we don't ever have to go through anything alone.

Kristin, thanks for being my best gal and the first person to love me unconditionally.

Your genius and strength impress me every day, and I'm honored to be "Aunt Kelly" to

your kind, well-rounded, beautiful children. Let's start our book! Ginny Lee, you are the

mentor and role model I've always wanted — Southern in all the good ways (humor,

creativity, resilience) yet decidedly un-Southern in your mind-set. Thanks for guiding me

in the ways of The Force. May I be as wise as Papa, as cool as Uncle Mikie, and as caring

and witty as Angela, my oldest friend. Trenyse, you are my twin and my hero. Shoutouts

to my CDC friends who encouraged me to keep going — particularly Blanche, Michelle,

Stefanie, Ida, Mae, Johanzynn, Carolina, Jenique, Hillary, Keun, Calvin, Erin, Onalee,

Heather, Abbey, and Alicia. To the grade school teachers who reinforced my love of

learning — and spent their own money to buy me supplies and send me to camps — I

will pay it forward. Steve Smith and Ron Laughridge, thank you for serving as father

figures and guiding me at a crucial time. I owe my very survival to my case worker, Judy

Peery, and my guardian ad litem, whose name I wish I knew. You two protected the

5-year-old me from a world of hurt and made my later success possible. And finally, to

children everywhere who are abused or bullied, know that it truly does get better. Don't

sink to their level. In time, you will find people who love and value you. You are stronger

than you realize, and what you have endured will enhance your empathy if you let it.

ACKNOWLEDGMENTS

ABSTRACT OF THE DISSERTATION

ELUCIDATING THE ROLE OF NEIGHBORHOOD DEPRIVATION

IN HYPERTENSIVE DISORDERS OF PREGNANCY

by

Kelly Michelle Winter

Florida International University, 2018

Miami, Florida

Professor Mary Jo Trepka, Major Professor

This dissertation examined risk factors for hypertensive disorders of pregnancy (HDP) — specifically whether neighborhood socioeconomic deprivation exacerbates individual socioeconomic disadvantage (deprivation amplification) to increase the likelihood of developing HDP. To select the optimal areal unit at which to investigate HDP, geographic proxies for neighborhoods were explored.

A thematic review qualitatively examined nontraditional neighborhood boundaries identified through internet sources. Data from 2008–2012 Miami-Dade County, Florida birth records (n=121,421) and the U.S. Census Bureau were used for the remaining analyses. Ordinary least squares (OLS) and geographically weighted regression (GWR) analysis empirically compared the proportion of HDP prevalence explained by six areal units: census block groups, census tracts, ZIP code tabulation areas (ZCTAs), and three types of natural neighborhood — census units clustered based on an eight-item Neighborhood Deprivation Index. Multilevel logistic regression examined

relationships between HDP, neighborhood deprivation, and individual-level factors. Odds ratios (OR) and adjusted odds ratios (aOR) were calculated.

The thematic review found 22 potential alternatives to census boundaries developed through techniques such as crowd-sourcing and qualitative research. In the sensitivity analysis, census tracts aggregated at the scale of ZCTAs performed twice as well as any other model ($GWR^2 = 0.27$) and were used as the Aim 3 unit of analysis. In the multilevel logistic regression, HDP was associated with moderate (aOR=1.13; CI: 1.05, 1.21) and high neighborhood deprivation (aOR=1.16; CI: 1.07, 1.26).

Compared with mothers with private insurance, uninsured women (aOR=1.69; CI: 1.56, 1.84) and Medicaid recipients (aOR=1.12; CI: 1.05, 1.18) had higher HDP odds. Non-Hispanic Black women's HDP odds were 1.58 times those of non-Hispanic White women. Cross-level interactions — between neighborhood deprivation and educational attainment and neighborhood deprivation and insurance status — did not reach statistical significance.

Private sector neighborhood boundaries hold promise for developing new public health tools. Because they are relatively easy to generate from census data, natural neighborhoods may balance tradition and innovation. While no evidence of deprivation amplification was found, results suggested that individual-level and neighborhood deprivation are HDP risk factors. Interventions that target expectant mothers in deprived neighborhoods — particularly non-Hispanic Black and Hispanic women who lack health insurance — may help reduce HDP prevalence and disparities.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

aOR	Adjusted odds ratio
API	Application program interface
BG	Block group
BMI	Body mass index
CI	Confidence interval
CT	Census tract
DOH	Department of Health
GED	General equivalency diploma
GWR	Geographically weighted regression
HDP	Hypertensive disorders of pregnancy
MAUP	Modifiable areal unit problem
MLR	Multilevel logistic regression
MSA	Metropolitan statistical area
NDI	Natural deprivation index
NN	Natural neighborhood
OLS	Ordinary least squares
OR	Odds ratio
PCA	Principal components analysis
SES	Socioeconomic status
US	United States
ZCTA	ZIP code tabulation area

CHAPTER I.

INTRODUCTION

i. **Background and significance**

Hypertensive disorders of pregnancy (HDP) pose serious risks to maternal and fetal

health, both during and after pregnancy. Encompassing gestational hypertension, pre-

eclampsia, eclampsia, chronic hypertension, and chronic hypertension with superimposed

pre-eclampsia, HDP is the second-leading cause of maternal deaths worldwide and

seventh in the United States (Centers for Disease Control and Prevention, n.d.; Hutcheon,

Lisonkova, & Joseph, 2011). Associated fetal health risks include intrauterine growth

restriction, preterm birth, low birth weight, and stillbirth (Dadelszen, Stones, & Mathai,

2016; Duley, 2009; Flenady et al., 2011; Say et al., 2014; Task Force on Hypertension in

Pregnancy, 2013; Wolf et al., 2004). Estimates of U.S. annual expenditures related to this

group of illnesses range from $1.6 billion to $2.2 billion (Agency for Healthcare

Research and Quality, n.d.; Stevens et al., 2017).

Individual-level HDP risk factors include extremes of maternal age (<18 years and

>35 years), nulliparity, multifetal gestation, pre-existing conditions such as diabetes and

kidney disease, obesity and low socio-economic status (SES) (Dadelszen et al., 2016;

Hutcheon et al., 2011; Lo, Mission, & Caughey, 2013). Recently, however, research has

begun to consider how upstream factors, or social determinants, might influence

reproductive health outcomes, including HDP (Culhane & Elo, 2005; Vinikoor-Imler,

Gray, Edwards, & Miranda, 2011a; Vinikoor-Imler, Messer, Evenson, & Laraia, 2011b).

One key upstream factor is neighborhood deprivation, whose detrimental effects are

1

theorized to extend beyond — and possibly also exacerbate — those of individual-level poverty (Cohen et al., 2000; Macintyre, Maciver, & Sooman, 1993; Murray et al., 2006; Pickett, 2001). To date, however, there has been a dearth of research on the extent to which area-level characteristics such as neighborhood deprivation — also known as "neighborhood disadvantage" in some studies — might interact with individual-level factors to influence risk of maternal outcomes such as HDP.

Despite the ever-expanding body of research linking neighborhood deprivation to negative health behaviors and outcomes (Buka, 2003; Masi, Hawkley, Harry Piotrowski, & Pickett, 2007; Morenoff et al., 2007; Nkansah-Amankra, 2010), methodological inconsistencies have prevented this field from generating recommendations for policies and interventions (Cutchin, Eschbach, Mair, Ju, & Goodwin, 2011). Neighborhoods are a societal construct rather than a natural phenomenon; thus, there is no inherent unit of analysis (Stock & Ellaway, 2013). Most studies conducted by U.S. academic or government researchers have relied on census boundaries as proxies for neighborhoods, largely because they are linked to a free, robust dataset that is updated on a consistent timetable (Messer & Kaufman, 2006; Spielman, Folch, & Nagle, 2014; Yen, Michael, & Perdue, 2009). Yet some studies have questioned these administrative units' real-world validity — which geographers refer to as "ground truth" (Boyle & Willms, 1999; Hart & Waller, n.d.; Kim, Ali, Sur, Khatib, & Wierzba, 2012; Messer & Kaufman, 2006). Recently, private companies have begun to develop alternative neighborhood boundaries. By combining geospatial and statistical analysis with qualitative methods, they have attempted to more accurately reflect how consumers view their communities, with the ultimate goal of boosting profits (Forbes, 2015; M. Graham, 2008; Hayden, 2014; Wahl,

2008; Walker, 2015). Many of these online boundary tools have open-source elements, allowing private citizens a chance to refine and combine them in new, promising ways (Adams, 2016; Gelernter, Ganesh, Krishnakumar, & Zhang, n.d.). However, academic researchers might not be aware that these resources are available — often at no cost.

There is also a potential "middle ground" between using traditional census boundaries and the uncharted waters of the aforementioned private sector tools: Several researchers have re-aggregated census data into more socioeconomically homogeneous units or "natural neighborhoods" (NN) with the goal of more closely approximating actual communities (Bissonnette, Wilson, Bell, & Shah, 2012; Parenteau & Sawada, 2011). In this instance, the word "natural" is not used in the typical, vernacular sense, and the authors are in no way suggesting that neighborhood deprivation is natural. Instead, "natural" in this instance denotes "meaningful" or "locally relevant" (Pickett, 2001; Ross, Tremblay, & Graham, 2004). It also refers to the process by which these neighborhoods are formed: ArcGIS software searches for natural statistical patterns or clusters based on values of one or more variables, (e.g., neighborhood deprivation) (Esri Demographics, n.d.). The software attempts to maximize within-group similarity and between-group variability. To the best of the author's knowledge, NNs have not been used in previous HDP research.

The importance of choosing the appropriate geographic unit for a given health outcome cannot be overstated. Even among studies using census units, the choice of a particular unit has been inconsistent, and validity can be jeopardized by the modifiable areal unit problem (MAUP) — a type of spatial misclassification bias that arises from

imposing artificial boundaries onto an area (Cutchin et al., 2011; Openshaw, 1984; Schlossberg, 2003). Past research has underscored the need for sensitivity analyses to determine the appropriate geographic unit to use when measuring relationships between neighborhood factors and health outcomes (Diez Roux, 2007; Hayward & Parent, 2009; Jelinski & Wu, 1996; White & Borrell, 2011). The author is not aware of any past studies that have empirically compared NNs constructed at different scales from the same data.

Many neighborhood deprivation studies share another potential methodological pitfall — a tendency to control or adjust for individual-level factors that might be mediators rather than confounders (Blakely & Subramanian, 2006; Blumenshine, Egerter, Barclay, Cubbin, & Braveman, 2010). Even fewer studies have explored the possibility of cross-level interactions, in which area- and individual-level variables synergistically influence health outcomes (e.g., deprivation amplification) (Nogueira, Gama, Mourão, Marques, & Padez, 2014; Ross, Oliver, & Villeneuve, 2013). Moreover, while some studies have examined associations between neighborhood deprivation and HDP, few have employed a multilevel framework and, to the best of our knowledge, none have examined HDP and deprivation amplification (Vinikoor-Imler, Gray, Edwards, & Miranda, 2011a; Vinikoor-Imler, Messer, Evenson, & Laraia, 2011b). Investigating risk factors simultaneously at the individual and neighborhood levels — as well as interactions between those factors — is a necessary step toward understanding root causes of HDP disparities in a large, multicultural urban area such as Miami-Dade County, Florida.

This dissertation seeks to advance the discipline of neighborhood health research by offering two paths toward a more valid unit of analysis: alternative geographic units that

might more closely approximate real-world neighborhoods and a straightforward, empirical strategy for selecting the appropriate unit to use when studying a particular health outcome. Reaching consensus on this matter is vital in order to facilitate the knowledge synthesis necessary to shape health policy and interventions. This dissertation also seeks to enhance our understanding of the pathways through which neighborhood deprivation influences HDP risk and the degree to which it differentially affects those living in poverty. Such information can be used to guide HDP intervention strategies and local policies concerning resource allocation.

ii. Specific aims / hypotheses

1.) To conduct a thematic review of neighborhood boundaries developed by private sector organizations to explore their potential utility in public health research.

2.) To conduct a sensitivity analysis to compare the strength of association between HDP prevalence and neighborhood deprivation at six different units of analysis — block group, census tract, ZIP Code Tabulation Area (ZCTAs), and three types of natural neighborhood formed by aggregating block groups and census tracts at three different scales — using structural equation modeling and geographically weighted regression. By comparing the R^2 values for the six geographic units, I will determine the unit of analysis to use for Aim 3.

3.) To quantify the relationship between neighborhood deprivation and individual-level HDP status in Miami-Dade County, Florida using multilevel logistic regression. I will also examine cross-level interactions between neighborhood deprivation and two measures of individual-level deprivation — low educational attainment (< high

5

school diploma/General Equivalency Diploma) and lack of access to health care (uninsured/Medicaid recipient).

- *Hypothesis 1a: Women living in neighborhoods with higher levels of deprivation will have higher odds of HDP compared with those in low-deprivation neighborhoods.*

- *Hypothesis 1b: Women with low educational attainment who live in deprived neighborhoods will have higher odds of HDP compared with women with low educational attainment in low-deprivation neighborhoods.*

- *Hypothesis 1c: Women who lack access to health care who live in deprived neighborhoods will have higher odds of HDP compared with women who lack access to health care in low-deprivation neighborhoods.*

iii.   Conceptual framework

The conceptual model for this dissertation (Figure 1) is based on the theory of "deprivation amplification," which posits that area-level resource scarcity and stressors serve to intensify the negative impacts of individual-level resource scarcity on health behaviors and outcomes (Kawachi & Berkman, 2003; Kneebone & Holomes, 2016; Macintyre, Maciver, & Sooman, 1993; Murray et al., 2006; Stafford & Marmot, 2003). In other words, living in deprived neighborhoods that lack sufficient resources (e.g., green spaces, sidewalks, high-quality medical facilities) is more detrimental to the health of residents who lack personal resources (e.g., money, education, a personal vehicle, health insurance) (Capps, n.d.; Kawachi & Berkman, 2003; Macintyre, 2000). Fewer doctors practice in disadvantaged areas, and those who do may severely limit the number of

Medicaid patients they serve (Dreier, Mollenkopf, & Swanstrom, 2014). Deprived neighborhoods also have fewer conventional banks and more predatory "fringe financial institutions" (e.g., check-cashing stores and title loan companies), which further exacerbate poor families' precarious financial situations (Dreier et al., 2014). Conversely, even in deprived areas, residents who are less disadvantaged on an individual level can venture beyond their own neighborhood to go to a park or a bank or to seek medical care (Dreier et al., 2014).

With more convenience stores than grocery markets and higher levels of stress-inducing factors such as environmental hazards, sub-standard housing and high rates of crime, deprived neighborhoods also promote unhealthy behaviors such as smoking, alcohol use, and overconsumption of unhealthy foods (Andrews et al., 2014; Cantrell et al., 2014; Dreier, Mollenkopf, & Swanstrom, 2014; Hogue, Hoffman, & Hatch, 2001; Karriker-Jaffe, 2013; Laraia, Messer, Evenson, & Kaufman, 2007; Pearce, Blakely, Witten, & Bartie, 2007; Sarkar, Gallacher, & Webster, 2013). These behaviors, in turn, increase the likelihood of overweight/obesity, chronic diseases, and poor pregnancy outcomes (Laraia et al., 2012; Nogueira, Gama, Mourão, Marques, & Padez, 2014; Wang, Kim, Gonzalez, MacLeod, & Winkleby, 2007).

This dissertation's conceptual model also draws inspiration from previous research linking reproductive health outcomes to neighborhood context through key intervening variables (Culhane & Elo, 2005). Unlike many models that consider neighborhood deprivation as a focal variable, the current framework emphasizes the interactions between neighborhood deprivation and individual-level variables that more directly influence the health outcome.

This dissertation examined relationships between HDP and neighborhood deprivation and individual-level deprivation — represented by insurance status and education status. Other individual-level medical and behavioral characteristics indicated to be HDP risk factors in the literature were included as lying on the causal pathway between race/ethnicity and HDP. It is important to emphasize that while race/ethnicity is not a viable construct from a biological standpoint, racism and discrimination as well as their associated negative psychological and physiological consequences are quite real (Freeman, 2003; Krieger, 2001; A. Smedley & Smedley, 2005; Yudell, Roberts, DeSalle, & Tishkoff, 2016).

Figure 1. Conceptual framework of relationships between hypertensive disorders of pregnancy, neighborhood deprivation, and individual-level factors.
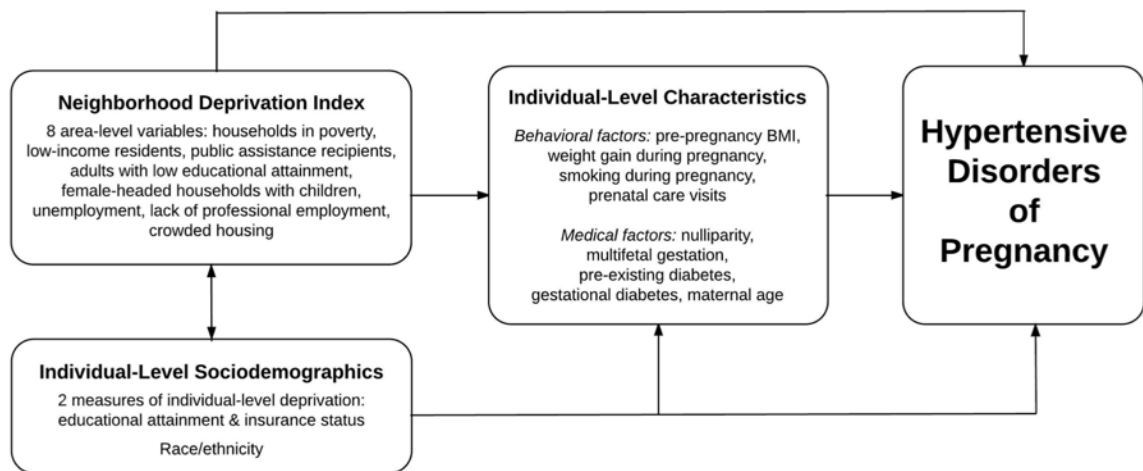
References

Adams, B. (2016). Wāhi, a discrete global grid gazetteer built using linked open data. *International Journal of Digital Earth*, 10, 490–503.

Andrews, J. O., Mueller, M., Newman, S. D., Magwood, G., Ahluwalia, J. S., White, K., & Tingen, M. S. (2014). The Association of individual and neighborhood social

cohesion, stressors, and crime on smoking status among African-American women in Southeastern US subsidized housing neighborhoods. *Journal of Urban Health*, *91*, 1158–1174.

Agency for Healthcare Research and Quality. (n.d.). 2014 National Diagnoses — Clinical Classification Software (CCS), Principal Diagnosis: #183 Hypertension complicating pregnancy, childbirth and the puerperium — Standard Errors (SE). Retrieved April 3, 2017, from hcupnet.ahrq.gov

Bissonnette, L., Wilson, K., Bell, S., & Shah, T. I. (2012). Neighbourhoods and potential access to health care: The role of spatial and aspatial factors. *Health & Place*, 18, 841–853.

Blakely, T., & Subramanian, S. V. (2006). Multilevel Studies. In J. M. Oakes & J. S. Kaufman, *Methods in Social Epidemiology* (pp. 316–340). New York: John Wiley & Sons Incorporated.

Blumenshine, P., Egerter, S., Barclay, C. J., Cubbin, C., & Braveman, P. A. (2010). Socioeconomic Disparities in Adverse Birth Outcomes. *American Journal of Preventive Medicine, 39*, 263–272.

Boyle, M. H., & Willms, J. D. (1999). Place effects for areas defined by administrative boundaries. *American Journal of Epidemiology, 149*, 577–585.

Buka, S. L. (2003). Neighborhood support and the birth weight of urban infants. *American Journal of Epidemiology, 157*, 1–8.

Cantrell, J., Anesetti-Rothermel, A., Pearson, J. L., Xiao, H., Vallone, D., & Kirchner, T. R. (2014). The impact of the tobacco retail outlet environment on adult cessation and differences by neighborhood poverty. *Addiction*, *110*, 152–161.

Capps, K. (n.d.). How Many Neighborhoods is Too Many for One Map? Retrieved September 9, 2016, from http://www.citylab.com/tech/2015/09/how-many-neighborhoods-is-too-many-for-one-map/403474/

Centers for Disease Control and Prevention. (n.d.). Pregnancy Mortality Surveillance System. Retrieved September 2, 2017, from https://www.cdc.gov/reproductivehealth/maternalinfanthealth/pmss.html

Cohen, D., Spear, S., Scribner, R., Kissinger, P., Mason, K., & Wildgen, J. (2000). "Broken windows" and the risk of gonorrhea. *American Journal of Public Health*, 90, 230–236.

Culhane, J. F., & Elo, I. T. (2005). Neighborhood context and reproductive health. *American Journal of Obstetrics and Gynecology, 192*, S22–S29.

Cutchin, M. P., Eschbach, K., Mair, C. A., Ju, H., & Goodwin, J. S. (2011). The socio-

spatial neighborhood estimation method: An approach to operationalizing the neighborhood concept. *Health & Place, 17*, 1113–1121.

Dadelszen, von, P., Stones, W., & Mathai, M. (2016). *The FIGO Textbook of Pregnancy Hypertension: An Evidence-Based Guide to Monitoring, Prevention and Management*. (L. A. Magee). London: The Foundation for the Global Library of Women's Medicine.

Diez Roux, A. V. (2007). Neighborhoods and health: where are we and where do we go from here? *Revue D'épidémiologie Et De Santé Publique, 55*, 13–21.

Dreier, P., Mollenkopf, J. H., & Swanstrom, T. (2014). *Place Matters: Metropolitics for the Twenty-First Century* (3rd ed.). Lawrence, Kan.: University Press of Kansas.

Duley, L. (2009). The global impact of pre-eclampsia and eclampsia. *Seminars in Perinatology, 33*, 130–137.

Esri Demographics. (n.d.). Tool Reference: Spatial Statistics Toolbox — Grouping Analysis. Retrieved March 18, 2016, from http://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/grouping-analysis.htm

Flenady, V., Koopmans, L., Middleton, P., Frøen, J. F., Smith, G. C., Gibbons, K., et al. (2011). Major risk factors for stillbirth in high-income countries: a systematic review and meta-analysis. *The Lancet*, 377, 1331–1340.

Forbes. (2015, March 12). Location Intelligence Leads to Business Insights and Creates Competitive Advantage. Retrieved October 23, 2016, from http://www.forbes.com/sites/forbespr/2015/03/12/location-intelligence-leads-to-business-insights-and-creates-competitive-advantage/#7f0c6bc81f12

Freeman, H. P. (2003). Commentary on the meaning of race in science and society. *Cancer Epidemiology and Prevention Biomarkers*, *12*, 232s–236s.

Gelernter, J., Ganesh, G., Krishnakumar, H., & Zhang, W. (n.d.). Automatic gazetteer enrichment with user-geocoded data. Proceedings of the Second ACM SIGSPATIAL International Workshop on Crowdsourced and Volunteered Geographic Information, 87–94.

Graham, M. (2008). Lawsuit claims mapmaking firm owns your neighborhood. *Wired*. Retrieved from https://www.wired.com/2008/10/lawsuit-claims-mapmaking-firm-owns-your-neighborhood/

Hart, T. C., & Waller, J. (n.d.). Neighborhood boundaries and structural determinants of social disorganization: Examining the validity of commonly used measures. *Western Criminology Review, 14*, 16–33.

Hayden, T. (2014). *The Mobile Commerce Revolution: Business Success in a Wireless World*. Indianapolis, IN: Que Publishing.

Hayward, P., & Parent, J. (2009). Modeling the influence of the modifiable areal unit problem (MAUP) on poverty in Pennsylvania. *The Pennsylvania Geographer, 47*, 120–135.

Hogue, C. J. R., Hoffman, S., & Hatch, M. (2001). Stress and preterm delivery: a conceptual framework. *Pediatric and Perinatal Epidemiology*, *15* (Suppl. 21), 30–40.

Hutcheon, J. A., Lisonkova, S., & Joseph, K. S. (2011). Epidemiology of pre-eclampsia and the other hypertensive disorders of pregnancy. *Best Practice & Research Clinical Obstetrics & Gynaecology, 25*, 391–403.

Jelinski, D. E., & Wu, J. (1996). The modifiable areal unit problem and implications for landscape ecology. *Landscape Ecology, 11*, 129–140.

Karriker-Jaffe, K. J. (2013). Neighborhood socioeconomic status and substance use by U.S. adults. *Drug and Alcohol Dependence*, *133*, 212–221.

Kawachi, I., & Berkman, L. F. (Eds.). (2003). *Neighborhoods and Health*. New York, N.Y.: Oxford University Press.

Kim, D. R., Ali, M., Sur, D., Khatib, A., & Wierzba, T. F. (2012). Determining optimal neighborhood size for ecological studies using leave-one-out cross validation. *International Journal of Health Geographics*, 11, 10.

Kneebone E. & Holmes, N. (March 31, 2016). U.S. concentrated poverty in the wake of the Great Recession. *Brookings*. Retrieved from https://www.brookings.edu/research/u-s-concentrated-poverty-in-the-wake-of-the-great-recession/

Laraia, B. A., Karter, A. J., Warton, E. M., Schillinger, D., Moffet, H. H., & Adler, N. (2012). Place matters: Neighborhood deprivation and cardiometabolic risk factors in the Diabetes Study of Northern California (DISTANCE). *Social Science & Medicine*, *74*, 1082–1090.

Laraia, B., Messer, L., Evenson, K., & Kaufman, J. S. (2007). Neighborhood factors associated with physical activity and adequacy of weight gain during pregnancy. *Journal of Urban Health*, *84*, 793–806.

Lo, J. O., Mission, J. F., & Caughey, A. B. (2013). Hypertensive disease of pregnancy and maternal mortality. *Current Opinion in Obstetrics and Gynecology, 25*, 124–132.

Macintyre, S., Maciver, S., & Sooman, A. (1993). Area, class and health: Should we be focusing on places or people? *Journal of Social Policy, 22*, 213–234.

Macintyre, S. (2000). The social patterning of exercise behaviours: the role of personal and local resources. *British Journal of Sports Medicine*, *34*, 6–6.

Murray, C. J. L., Kulkarni, S. C., Michaud, C., Tomijima, N., Bulzacchelli, M. T.,

Iandiorio, T. J., & Ezzati, M. (2006). Eight Americas: Investigating mortality disparities across races, counties, and race-counties in the United States. *PLoS Medicine*, *3*, e260.

Nogueira, H., Gama, A., Mourão, I., Marques, V. R., & Padez, C. (2014). Pathways to childhood obesity: a deprivation amplification model and the overwhelming role of socioeconomic status (Vol. 1, pp. 1697–1708). Presented at the Sustainable City 2014, Southampton, UK: WIT Press.

Masi, C. M., Hawkley, L. C., Harry Piotrowski, Z., & Pickett, K. E. (2007). Neighborhood economic disadvantage, violent crime, group density, and pregnancy outcomes in a diverse, urban population. *Social Science & Medicine, 65*, 2440–2457.

Messer, L. C., & Kaufman, J. S. (2006). Using Census Data to Approximate Neighborhood Effects. In J. M. Oakes & J. S. Kaufman, *Methods in Social Epidemiology* (pp. 209–236). New York: John Wiley & Sons Incorporated.

Morenoff, J. D., House, J. S., Hansen, B. B., Williams, D. R., Kaplan, G. A., & Hunte, H. E. (2007). Understanding social disparities in hypertension prevalence, awareness, treatment, and control: The role of neighborhood context. *Social Science & Medicine*, 65, 1853–1866.

Murray, C. J. L., Kulkarni, S. C., Michaud, C., Tomijima, N., Bulzacchelli, M. T., Iandiorio, T. J., & Ezzati, M. (2006). Eight Americas: Investigating mortality disparities across races, counties, and race-counties in the United States. *PLoS Medicine*, 3, e260.

Nkansah-Amankra, S. (2010). Neighborhood Contextual Factors, Maternal Smoking, and Birth Outcomes: Multilevel Analysis of the South Carolina PRAMS Survey, 2000-2003. *Journal of Women's Health*, 19, 1543–1552.

Nogueira, H., Gama, A., Mourão, I., Marques, V. R., & Padez, C. (2014). Pathways to childhood obesity: a deprivation amplification model and the overwhelming role of socioeconomic status (Vol. 1, pp. 1697–1708). Presented at the Sustainable City 2014, Southampton, UK: WIT Press.

Openshaw, S. (1984). Ecological fallacies and the analysis of areal census data. *Environment and Planning A* 16, 17–31.

Parenteau, M.-P., & Sawada, M. C. (2011). The modifiable areal unit problem (MAUP) in the relationship between exposure to NO2 and respiratory health. *International Journal of Health Geographics*, 10, 58.

Pearce, J., Blakely, T., Witten, K., & Bartie, P. (2007). Neighborhood deprivation and access to fast-food retailing. *American Journal of Preventive Medicine*, *32*, 375–382.

Pickett, K. E. (2001). Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. *Journal of Epidemiology & Community Health*, 55, 111–122.

Ross, N. A., Tremblay, S., & Graham, K. (2004). Neighbourhood influences on health in Montréal, Canada. *Social Science & Medicine*, 59, 1485–1494.

Ross, N., Oliver, L., & Villeneuve, P. (2013). The contribution of neighbourhood material and social deprivation to survival: A 22-Year follow-up of more than 500,000 Canadians. *International Journal of Environmental Research and Public Health*, 10, 1378–1391.

Say, L., Chou, D., Gemmill, A., Tunçalp, Ö., Moller, A.-B., Daniels, J., et al. (2014). Global causes of maternal death: a WHO systematic analysis. *The Lancet Global Health*, 2, e323–e333.

Schlossberg, M. (2003). GIS, the US Census and Neighbourhood Scale Analysis. *Planning Practice and Research*, 18, 213.

Sarkar, C., Gallacher, J., & Webster, C. (2013). Built environment configuration and change in body mass index: The Caerphilly Prospective Study (CaPS). *Health & Place*, *19*, 33–44.

Smedley, A., & Smedley, B. D. (2005). Race as biology is fiction, Racism as a social problem is real: Anthropological and historical perspectives on the social contruction of race. *American Psychologist*, *60*, 16–26.

Spielman, S. E., Folch, D., & Nagle, N. (2014). Patterns and causes of uncertainty in the American Community Survey. *Applied Geography, 46*, 147–157.

Stafford, M., & Marmot, M. (2003). Neighbourhood deprivation and health: does it affect us all equally? *International Journal of Epidemiology*, *32*(3), 357–366.

Stevens, W., Shih, T., Incerti, D., Ton, T. G. N., Lee, H. C., Peneva, D., et al. (2017). Short-term costs of preeclampsia to the United States health care system. *American Journal of Obstetrics and Gynecology, 217*, 237–248.e16.

Stock, C., & Ellaway, A. (2013). *Neighbourhood Structure and Health Promotion*. (C. Stock & A. Ellaway). Boston, MA: Springer.

Task Force on Hypertension in Pregnancy. (2013). *Hypertension in Pregnancy*. Washington, D.C.: American College of Obstetricians and Gynecologists.

Vinikoor-Imler, L. C., Gray, S. C., Edwards, S. E., & Miranda, M. L. (2011a). The effects of exposure to particulate matter and neighbourhood deprivation on gestational hypertension. *Paediatric and Perinatal Epidemiology, 26*, 91–100.

Vinikoor-Imler, L. C., Messer, L. C., Evenson, K. R., & Laraia, B. A. (2011b). Neighborhood conditions are associated with maternal health behaviors and pregnancy outcomes. *Social Science & Medicine, 73*, 1302–1311.

Wahl, B. (2008, December 4). Mapping the World ... One Neighborhood at a Time - Part Two. Retrieved September 4, 2017, from https://www.directionsmag.com/article/2425#v

Walker, R. (2015). *Success with Big Data: From Data and Analytics to Profits*. New York: Oxford University Press.

Wang, M. C., Kim, S., Gonzalez, A. A., MacLeod, K. E., & Winkleby, M. A. (2007). Socioeconomic and food-related physical characteristics of the neighbourhood environment are associated with body mass index. *Journal of Epidemiology & Community Health*, *61*, 491–498.

Yudell, M., Roberts, D., DeSalle, R., & Tishkoff, S. (2016). Taking race out of human genetics. *Science*, *351*, 564–565.

White, K., & Borrell, L. N. (2011). Racial/ethnic residential segregation: Framing the context of health risk and health disparities. *Health & Place, 17*, 438–448.

Wolf, M., Shah, A., Jimenez-Kimble, R., Sauk, J., Ecker, J. L., & Thadhani, R. (2004). Differential risk of hypertensive disorders of pregnancy among Hispanic women. *Journal of the American Society of Nephrology, 15*, 1330–1338.

Yen, I. H., Michael, Y. L., & Perdue, L. (2009). Neighborhood environment in studies of health of older adults. *American Journal of Preventive Medicine, 37*, 455–463.

CHAPTER II

LITERATURE REVIEW

i.   **Hypertensive disorders of pregnancy**

Both the World Health Organization's Sustainable Development Goals and the U.S.

Department of Health and Human Services Healthy People 2020 have prioritized the

improvement of maternal and child health outcomes (Office of Disease Prevention and

Health Promotion, 2017; World Health Organization, 2017). Hypertensive disorders of

pregnancy (HDP) are a major source of morbidity and mortality for both mothers and

children, and — with the exception of eclampsia — rates of the illnesses that comprise

HDP have been increasing in recent decades (Ananth, Keyes, & Wapner, 2013; Fingar et

al., 2017; Hutcheon, Lisonkova, & Joseph, 2011). Of the types of HDP, pre-eclampsia —

by itself or superimposed on chronic hypertension — presents the greatest risk (Task

Force on Hypertension in Pregnancy, 2013). It is defined as systolic blood pressure (BP)

$\geq$ 160 mm Hg or diastolic BP of 110 mm Hg measured on two different occasions $\geq$ 4

hours apart while a woman is on bed rest (unless she has already started antihypertensive

therapy) with either proteinuria or, in absence of proteinuria, any of the following:

thyrombocytopenia (platelet count < 100,000/microliter); progressive renal insufficiency;

impaired liver function; pulmonary edema; new-onset visual or cerebral disturbance

(Task Force on Hypertension in Pregnancy, 2013). The American College of

Obstetricians and Gynecologists' Task Force on Hypertension in Pregnancy, proteinuria

is no longer a requirement for diagnosis of pre-eclampsia (Task Force on Hypertension in

Pregnancy, 2013). Both physicians and scientists disagree as to whether there is a

clinically significant distinction between mild and severe pre-eclampsia (Dadelszen, Stones, & Mathai, 2016; Task Force on Hypertension in Pregnancy, 2013).

Gestational hypertension is elevated BP > 20 weeks of gestation with no proteinuria or other signs of pre-eclampsia, while chronic hypertension is elevated BP diagnosed prior to pregnancy (Task Force on Hypertension in Pregnancy, 2013). The term "pregnancy-induced hypertension" is no longer commonly used among clinicians or scientists, because it has been deemed too non-specific and was used to signify different conditions in different countries (e.g., pre-eclampsia in the U.S., gestational hypertension without proteinuria in the UK) (Dadelszen et al., 2016; Hutcheon et al., 2011). About one-third of women (35%) diagnosed with gestational hypertension < 34 weeks into their pregnancy will develop pre-eclampsia, which takes an average of 5 additional weeks to manifest (Dadelszen et al., 2016). About 20% of women with chronic hypertension will develop pre-eclampsia (Dadelszen et al., 2016).

Burden of disease

Globally, HDP complicates 5-10% of pregnancies and is a direct cause of 14% of maternal deaths, making it the second-leading cause of maternal mortality worldwide (Hutcheon et al., 2011; Say et al., 2014). It is the top cause of maternal mortality in industrialized nations, accounting for 16% of maternal deaths (Hutcheon et al., 2011). The prevalence of HDP is estimated at 10% in the United States, but a smaller proportion of cases result in death in the U.S. (7.4% in 2011-2013), placing HDP at No. 7 on the list of causes of maternal mortality (Centers for Disease Control and Prevention, 2017; Wagner, Barac, & Garovic, 2007). In Florida, HDP prevalence among women who

16

experienced a live birth was 12.6% in 2011, down 33.6% from 10-year peak of 19.0% in 2002 (Florida PRAMS, n.d.). However in 2012, 15.9% of maternal deaths in Florida were attributible to HDP — twice the proportion for the U.S. (7.4%) (Florida Perinatal Quality Collaborative, 2015; Centers for Disease Control and Prevention, 2017).

Between 2005 and 2014, the proportion of inpatient deliveries that included a diagnosis of pre-eclampsia/eclampsia rose nearly 21 percent (Fingar et al., 2017). Among the almost 177,000 deliveries complicated by HDP in the U.S. in 2014, 84% involved pre-eclampsia, 15% chronic hypertension with superimposed pre-eclampsia, and 1% eclampsia (Fingar et al., 2017). Mean length and cost of hospital stays for deliveries involving pre-eclampsia/eclampsia were 70 percent higher than other deliveries (Fingar et al., 2017). Other studies found that HDP adds an average of $6,152 to $14,458 to a woman's hospital expenses in the U.S. and an additional $2,483 to $2,969 to her newborn care costs (Huynh et al., 2013; Law, McCoy, Lynen, Curkendall, Gatwood, Juneau, & Landsman-Blumberg, 2015b; 2015a). Estimates of U.S. annual expenditures related to this group of illnesses range from $1.6 billion to $2.2 billion (Agency for Healthcare Research and Quality, n.d.; Stevens et al., 2017).

Adverse maternal and child health outcomes

For each woman who dies from HDP, an estimated 20 to 30 others experience substantial morbidities (Dadelszen et al., 2016). Pre-eclampsia is estimated to increase a woman's probability of having at least one adverse outcome from 4.6% to 10.1% (Stevens et al., 2017). One of the most serious conditions associated with pre-eclampsia is HELLP syndrome. The "H" stands for hemolysis (breakdown of red blood cells), "EL"

for elevated liver enzymes), and "LP" for low platelet count. There are no reliable estimates of HELLP prevalence, but the syndrome's case-fatality rate is estimated at 25% for mothers and 7.4% to 34% for fetuses (Haram, Svendsen, & Abildgaard, 2009). Other associated morbidities include placental abruption/hemorrhage, pulmonary edema, renal failure, blindness and stroke (Cunningham et al., 2010; Dadelszen et al., 2016; Hutcheon et al., 2011; Stevens et al., 2017). The rate of pregnancy-related stroke cases attributable to HDP increased an estimated 102.6% (from 0.8 to 1.6 per 10,000 pregnancy hospitalizations) in the U.S. between 1994 and 2011 (Leffert, Clancy, Bateman, Bryant, & Kuklina, 2015). In 2011, women who experienced a stroke were 5.2 times as likely to have been hospitalized with HDP (Leffert et al., 2015).

Fetal health risks associated with HDP include intrauterine growth restriction, chronic hypoxia/acidosis, preterm birth, low birth weight, respiratory distress syndrome, sepsis, and stillbirth (Duley, 2009; Fingar et al., 2017; Flenady et al., 2011; Say et al., 2014; Stevens et al., 2017; Task Force on Hypertension in Pregnancy, 2013; Wolf et al., 2004). Being born to a mother with pre-eclampsia is estimated to almost double an infant's probability of having at least one adverse outcome — from 7.9% to 14.2% (Stevens et al., 2017). In a meta-analysis of studies from high-income countries, among women who experienced a stillbirth, the adjusted odds of pregnancy-induced hypertension, pre-eclampsia, and eclampsia were 1.3, 1.6, and 2.2, respectively (Flenady et al., 2011). Furthermore, HDP can have consequences later in life: There is mounting evidence that HDP increases the risk of subsequent metabolic syndrome and circulatory problems in mothers and their offspring (Bellamy, Casas, Hingorani, & Williams, 2007; Duley, 2009; Flenady et al., 2011; Lykke et al., 2009; Task Force on Hypertension in Pregnancy,

18

2013). Recent studies have found that HDP might be a risk factor for poor motor development, attention-deficit/hyperactivity disorder (ADHD) and autism spectrum disorder in children, as well (Böhm et al., 2016; Curran, Khashan, O'keeffe, & Kenny, 2016; Grace, Bulsara, Pennell, & Hands, 2014).

Racial/ethnic disparities

In the United States, there are racial/ethnic HDP disparities in both morbidity and mortality. Gaps between non-Hispanic White and Black women are particularly stark. Compared with Whites, Black women have higher odds of all HDP conditions, and this disparity appears to increase with age (Fingar et al., 2017; R. Gold, Gold, Schilling, & Modilevsky, 2014; Lo, Mission, & Caughey, 2013; Miranda et al., 2010; Shen, Tymkow, & MacMullen, 2005). Black women's rate of preeclampsia/eclampsia was 60% higher than that of White women in 2014 (Fingar et al., 2017). Black women are also about 3 times as likely to die from pre-eclampsia or eclampsia (Tucker, Berg, Callaghan, & Hsia, 2007; Zhang, Meikle, & Trumble, 2003).

Estimates of Hispanic women's HDP risk have varied depending on the specific type or types of HDP examined. In a large sample of non-Hispanic White and Hispanic women in Massachusetts (n=3,200), Hispanic women's relative risk of pre-eclampsia increased from 1.0 to 1.9 once the following covariates were added to the model: age, baseline BP and body mass index (BMI), smoking, gestational age at the first prenatal visit, multiple gestation, and preexisting diabetes mellitus or gestational diabetes mellitis (Wolf et al., 2004). An analysis of four years of National Center for Health Statistics data found that HDP incidence was significantly lower for Hispanic women compared with

non-Hispanic Whites for all but the youngest and oldest maternal age groups (<15 years

and 45-54 years) (R. Gold et al., 2014). The association was the same in those two age

groups as well, but low birth numbers in those strata likely reduced statistical power. In a

2011 report from the National Center for Health Statistics, Hispanic women's rate of

chronic hypertension during pregnancy (6.8 per 100,000 births) was about half that of

non-Hispanic Whites (12.7 per 100,000) and one-fourth that of non-Hispanic Blacks

(25.7 per 100,000) (Martin, Hamilton, Ventura, et al., 2011). Hispanic women also had a

much lower rate of gestational hypertension (28.9 per 100,000 births) compared with

non-Hispanic Whites (46.1 per 100,000) and one-fourth that of non-Hispanic Blacks

(50.2 per 100,000). Still, the rate of chronic hypertension among Hispanic women has

steadily increased since 2000, albeit not as rapidly as the rates of non-Hispanic Whites or

Blacks (Martin et al., 2011). An analysis of National (Nationwide) Inpatient Sample data

found that Hispanic women's odds of pre-eclampsia were the same as those of non-

Hispanic Whites, but they were 21% less likely to develop gestational hypertension (Shen

et al., 2005). Yet in the aforementioned Massachusetts study, compared with non-

Hispanic White women, Hispanic women who initially presented with gestational

hypertension were 3.3 times as likely to develop pre-eclampsia (Wolf et al., 2004).

Individual-level risk factors

Non-modifiable HDP risk factors include extremes of maternal age (≤ 18 years and

≥ 35 years), nulliparity, and multifetal gestation (Cunningham et al., 2010; Hutcheon et

al., 2011; Stevens et al., 2017). Medical risk factors include all types of diabetes. In a

2010 review, type 1 diabetes was associated with a two-fold increase in gestational

hypertension risk and a 5-fold to 6-fold increase in pre-eclampsia incidence (Colatrella et

al., 2010). Gestational diabetes is a risk factor for HDP and vice versa (Colatrella et al., 2010). Hypertensive disorders of pregnancy (HDP) are between 2 and 4 times more common in women with type 2 diabetes (Colatrella et al., 2010). The adjusted odds ratio (aOR) of pre-eclampsia among women with chronic hypertension (aOR=13.5), pre-existing diabetes prior to pregnancy (aOR=3.4), or both conditions (aOR=12.5) was much higher than that of women without either condition (Yanit, Snowden, Cheng, & Caughey, 2012). Other medical risk factors include chronic kidney disease, hyperlipidemia, autoimmune disorders, diabetic neuropathy, and thrombophilia (abnormal blood-clotting) (Dadelszen et al., 2016).

Among non-modifiable risk factors, both overweight/obesity and excessive gestational weight gain have been associated with HDP (Cunningham et al., 2010; Hutcheon et al., 2011). Moreover, higher BMI during a previous pre-eclampsia pregnancy is a risk factor for pre-eclampsia in a subsequent pregnancy (Dadelszen et al., 2016). Surprisingly, smoking is a protective factor for pre-eclampsia: A systematic review of 48 epidemiological studies conducted over a period of nearly 50 years found that smoking during pregnancy reduced a woman's pre-eclampsia risk by as much as half, and studies suggest there is a dose-response relationship between smoking and mild forms of pre-eclampsia (Hackshaw et al., 2010, Karamuchi & Levine, 2010). However, a secondary analysis of a European randomized controlled trial found a significant positive association between smoking and chronic hypertension with superimposed pre-eclampsia (aOR=1.79) (Chappell et al., 2008). Other pre-eclampsia risk factors include bacterial infection, conception within a short time frame of beginning a sexual relationship,

21

interval between pregnancies, and depression (Cunningham et al., 2010; Dadelszen et al., 2016).

Hypertensive disorders of pregnancy (HDP) are also associated with individual-level economic deprivation. In 2014, rates of pre-eclampsia/eclampsia were higher among women receiving Medicaid (41.2 per 1,000 births) and Medicare (53.4 per 1,000 births) compared with the privately insured (35.8 per 1,000 births) (Fingar et al., 2017).

ii.   **Neighborhood deprivation**

Increasingly, public health researchers and government officials are emphasizing the importance of identifying upstream factors that contribute to health outcomes as a necessary step toward more successful public policy and sustainable, effective health interventions (Braveman, Cubbin, Egerter, & Pedregon, 2011; MacDonald, 2004; Sampson, 2012; Stock & Ellaway, 2013; Wallerstein, Yen, & Syme, 2011). Upstream factors — also known as "the causes of the causes" — are social/ecological determinants of health influenced by the distribution of economic resources and power (Commission on Social Determinants of Health, 2008; Gehlert et al., 2008; Marmot & Wilkinson, 2009). Examples include racial or gender inequity, job insecurity, and barriers to education. One key upstream factor is neighborhood deprivation, which is theorized to negatively influence health outcomes independently of individual-level poverty — whose impact it may also amplify (Cohen et al., 2000; Macintyre, Maciver, & Sooman, 1993; Murray et al., 2006; Pickett, 2001).

Association with risk behaviors and chronic diseases

Neighborhood deprivation, which encompasses area-level socioeconomic factors and access to resources, has been associated with unhealthy behavior, racial/ethnic health disparities, increased risk of both infectious and chronic diseases, and all-cause mortality (Acevedo-Garcia, 2005; Andrews et al., 2014; Ford & Browning, 2011; Frank & Bjornstrom, 2011; Jackson, Smith, Tabnak, & Vugia, 2015; Karriker-Jaffe, 2013; Slopen, Non, Williams, Roberts, & Albert, 2014). In a study of >200,000 diabetics in California, deprivation — measured by the eight-item Neighborhood Deprivation Index (NDI) (Messer et al., 2006a) — was positively associated with high cholesterol, obesity, poor glycemic control, and hypertension (Laraia et al., 2006). Relationships between NDI and obesity, poor glycemic control persisted even after adjustment for individual-level covariates, including income and education. Living in an area of high deprivation in Sweden was associated with the cardiovascular disease risk factors of obesity, physical inactivity, and smoking (Cubbin et al., 2006). Yet that study, which assessed deprivation using a seven-item Care-Need Index, found no relationship between chronic hypertension and deprivation. Multi-item indices captured four aspects of neighborhood context — disadvantage, affluence and gentrification, racial/ethnic/immigrant composition, and elderly composition — in a study on hypertension in Chicago, which found that hypertension was significantly associated with low education (odds ratio [OR]=1.5) and Black race (OR=1.8) while controlling for individual-level socioecnomic variables (Morenoff et al., 2007). But adjusting for the four neighborhood measures reduced the educational disparity by 10-15% (depending on whether the model included covariates).

Moreover, neighborhood factors completely accounted for the hypertension disparity found between Whites and Blacks.

Association with maternal child health outcomes

Neighborhood deprivation has been linked to poor maternal and child outcomes. In a study of data from the South Carolina Pregnancy Risk Assessment and Monitoring System (PRAMS), low birth weight was associated with neighborhood-level high poverty and low education (Nkansah-Amankra, 2010). When disadvantage was measured using an eight-item Neighborhood Deprivation Index (NDI) (Messer et al., 2006a), preterm birth was associated with neighborhood deprivation for Black mothers but not White mothers in North Carolina (Messer, Laraia, Savitz, Kaufman, & Dole, 2006b). Preterm birth (aOR=1.24) and low birth weight (aOR=1.19) were associated with living in neighborhoods in the highest NDI quartile in New York City (Janevic et al., 2010). In a study of deprivation and maternal child outcomes in Chicago, a four-item index (poverty, education, public housing, and unemployment) was used to measure neighborhood economic disadvantage. In stratified analyses, there was a significant relationship between disadvantage and preterm birth only for Black women (Masi, Hawkley, Harry Piotrowski, & Pickett, 2007). Disadvantage was significantly associated with small for gestational age only within Black and Hispanic subgroups and low birth weight within each of the three racial/ethnic subgroups.

Neighborhood deprivation might also partially explain the racial disparity in low birth weight in the U.S. An analysis of 95,711 births in Chicago found that infants born to Black mothers weighed an average of 297 grams less than infants born to White mothers

(Buka, 2003). But adjusting for individual-level risk factors reduced the racial gap in birth weight by 143 grams, and additional adjustment for neighborhood-level economic disadvantage reduced the gap by another 30 grams. Moreover, neighborhood economic disadvantage — measured by a three-item index of poverty, public assistance and unemployment — accounted for the majority of between-neighborhood variance for Black (80.8%) and White (76.3%) mothers.

Initial findings concerning HDP-related morbidity and area-level disadvantage were mixed. No association was found between pregnancy-induced hypertension and neighborhood income or unemployment in the Netherlands, while pre-eclampsia was more prevalent in high-income areas of Sweden (Agyemang et al., 2009; Gudmundsson, Björgvinsdóttir, Molin, Gunnarsson, & Marsal, 1997). Yet a prospective study of Norwegian neighborhoods found dose-response relationships between area-level disadvantage and gestational hypertension and pre-eclampsia, and researchers found a correlation between neighborhood poverty and pre-eclampsia hospitalization rates in New York but only among Hispanic women (Clausen, Øyen, & Henriksen, 2006; Tanaka et al., 2007). However, a 2014 national-level analysis of HDP-related hospitalizations in the U.S. found that pre-eclampsia/eclampsia rates were 26% higher in the poorest ZIP codes than in the wealthiest ZIP codes (Fingar et al., 2017). There are several potential explanations for these inconsistent findings: Many of these studies used crude measures of area-level deprivation (e.g., median income, percentage of immigrants) or simplistic statistical methods (e.g., t-tests, single-level regressions). It is also possible that neighborhood effects may not be as pronounced in Europe, where many of the

aforementioned early studies were conducted (Body-Gendrot, 2011; Dreier, Mollenkopf, & Swanstrom, 2014; Wacquant, 1993).

Two recent North Carolina studies indicated connections between area-level variables and HDP. The first found that relationships between maternal health behaviors and neighborhood contextual factors differed by race/ethnicity and that pregnancy-induced hypertension was significantly associated with neighborhood factors but only among non-Hispanic White women (Vinikoor-Imler, Messer, Evenson, & Laraia, 2011b). Specifically, physical incivilities (a six-item index measuring degradation of neighborhood structures and spaces) were associated with higher odds of smoking and inadequate weight gain for both non-Hispanic Black and non-Hispanic White women — as well as excessive weight gain for non-Hispanic White women. There was also a significant inverse relationship between pregnancy-induced hypertension and neighborhood walkability but only for Non-Hispanic white women. The second study found positive associations between neighborhood deprivation, particulate matter exposure, and gestational hypertension even after controlling for race/ethnicity, age, smoking status, and parity (Vinikoor-Imler, Gray, Edwards, & Miranda, 2011a). As with most previous research in this area, both of these studies adjusted for individual-level socioeconomic — measured in these studies by maternal education — rather than exploring potential interactions. To date, there has been a dearth of research on the extent to which area-level characteristics such as neighborhood deprivation might interact with individual-level factors to influence risk of maternal outcomes such as HDP — i.e., whether deprivation amplification is a risk factor for HDP.

iii.   Methodological challenges in neighborhood health research

Lack of a gold standard unit of analysis

Despite the recent focus on neighborhoods and health, because there is no set unit of

analysis, findings cannot be synthesized to guide health interventions and policies.

Instead, the choice of neighborhood boundary type varies widely, even among studies of

the same outcome or exposure variable. This is particularly problematic because both the

size of areal unit chosen (scale effect) and how units are aggregated (zoning effect) can

influence the magnitude and direction of a measure of association (Briant, Combes, &

Lafourcade, 2010; Openshaw, 1984). This phenomenon is known as the Modifiable Areal

Unit Problem (MAUP) (Openshaw, 1984). Two components comprise the MAUP: the

scale effect and the zoning effect. Scale effect pertains to the size of the chosen unit of

analysis — e.g., ZIP Code Tabulation Areas (ZCTAs) vs. census tracts or block groups.

The zoning effect pertains to boundaries, the different ways units can be aggregated at a

particular scale — without changing their size (Schuurman, Bell, Dunn, & Oliver, 2007).

Census units

The vast majority of studies have used administrative boundaries such as census tracts

to approximate neighborhoods (Yen, Michael, & Perdue, 2009). The reasons for this

reliance on census units are fairly straightforward: Such data are free, publicly accessible

online and updated at three-, five-, and 10-year intervals, depending on which variables

one needs. However, there is concern that census units might not accurately reflect the

boundaries of locally meaningful neighborhoods in terms of social, economic, historical,

and cultural factors (Dunn, 2009; Guo & Bhat, 2007; Sampson, Morenoff, & Gannon-

Rowley, 2002). As previously described in the geography, criminology, and public health literature, census units are relatively arbitrarily delinated boundaries of questionable real-world relevance beyond administrative purposes (Diez Roux, 2007; Hipp, 2007; Hogue, Kramer, Cooper, Drews-botsch, & Waller, 2010; Ross, Tremblay, & Graham, 2004; Wei, Cabrera-Barona, & Blaschke, 2016). According to Sabel, Kihal, Bard & Weber (2013), when defining and studying neighborhoods, "we are not just interested in capturing sterile spaces of habituation, but rather places where … social interactions and relations occur, where people have emotional attachment and, moreover, a sense of place" (p. 111).

While some census boundaries — specifically tracts — are delineated to be relatively homogenous in terms of sociodemographics, heterogeneity can arise over the 10-year census period, particularly in areas that experience rapid changes in population size (Messer & Kaufman, 2006). (See Appendix A: Glossary for definitions of commonly used census units.) Moreover, census boundaries can vary drastically in size and shape — particularly in suburban and rural areas (United States Census Bureau, n.d). This increases the likelihood of low-income areas being administratively "lumped in" with wealthy enclaves to make one very heterogeneous "neighborhood" — a form of misclassification bias that could hinder researchers' ability to accurately measure relationships between health outcomes and neighborhood-level factors (Sabel, Kihal, Bard, & Weber, 2013). Nonetheless, census units' accessibility, broad coverage, and consistent reporting schedule are unmatched, making them a practical choice (Messer & Kaufman, 2006; Sabel, Kihal, Bard, & Weber, 2013).

Alternative neighborhood boundaries

Some neighborhood researchers have explored methods to construct novel boundaries in hopes of more closely approximating actual communities. A 2007 study attempted to define New York City's neighborhoods through reviews of census data and land use maps and subsequent structured, qualitative street-level observations (Weiss, Ompad, Galea, & Vlahov, 2007). The study prioritized homogeneity within neighborhoods and heterogeneity across neighborhoods as well as physical obstructions between neighborhoods that helped serve as visual lines of demarcation for residents. The researchers stated that this was "a relatively efficient method" for generating neighborhood boundaries. However, since each field observation required 30 minutes to 4 hours to complete, scalability seems somewhat questionable. Among the limitations described was the absence of input from community residents, which was not collected because of time and money barriers.

Researchers in Urbana, Illinois took a different approach — asking residents to draw their own concepts of neighborhood boundaries on a GIS map that contained the following information: regional data (population distribution by census block, roads); land use; landmarks (buildings of importance); neighborhood amenities (parks, stores, churches, health care); accessibility to schools; commuting patterns; housing at block level (density, tenure, value); demographics; social issues by block (education, poverty, unemployment); and crime statistics (Talen & Shah, 2007). Nearly all participants (94%) were able to successfully complete the task, which was intended to bring a "human touch" to defining neighborhoods. Yet the size of the neighborhoods (10-400 acres) and

the attributes on which participants based their decisions varied widely. Furthermore, the sample was quite small (n=18), all participants were White, and the vast majority (88%) held white-collar jobs.

The rise of social media has allowed for another novel way to collect input from community residents. On some photo-sharing websites, users "tag" their pictures with keywords. Those photos are also "geotagged" with latitude and longitude coordinates. A 2008 study collected geotagged photos from publicly available websites such as Flickr, Panoramio, and Locr and used spatial statistics to approximate boundaries of neighborhoods through an iterative process (Wilske, 2008). While this is a relatively low-resource way to incorporate public opinion into neighborhood boundaries, the method has several notable limitations: There is potential for both homonyms (a place that shares a name with a person, object, etc.) and polysemes (>2 places that share the same name). Scale/level of granularity is also of concern: For example, a participant could tag a photo as "Flamingo," "South Beach," or "Miami." Tourists or others less familiar with an area might tag photos incorrectly. Moreover, while the authors did not mention selection bias, it seems to be another potential limitation: For example, people who tag photos might tend to be of similar socioeconomic status.

While the aforementioned novel methods seem to hold promise for future neighborhood research endeavors, resource constraints in the field of public health limit the scalability and feasibility of such ideas (Frieden, 2014; Kingsley, Coulton, & Pettit, 2016; Mujahid, Diez Roux, Morenoff, & Raghunathan, 2007; van Panhuis et al., 2014). Recently, private sector companies have begun to develop innovative ways to define and

identify locally meaningful neighborhoods, particularly in the United States. By combining geospatial and statistical analysis with qualitative methods, they have attempted to develop boundaries infused with "ground truth." The open-source nature of many of these online tools offers private citizens an opportunity to repurpose their data to create additional resources. Yet academic scientists may not be aware of these resources or might not know that many are accessible to public researchers at no charge. Aim 1 of this dissertation is to review neighborhood boundaries developed by private businesses and citizen collaborations, comparing their relative strengths and limitations to explore their utility for future neighborhood health studies.

Natural neighborhoods

Some researchers have explored re-aggregating census data into more socioeconomically homogeneous units called "natural neighborhoods" (NNs) with the goal of more closely approximating "meaningful," "locally relevant" communities (Bissonnette, Wilson, Bell, & Shah, 2012; Parenteau & Sawada, 2011; Pickett, 2001; Ross et al., 2004). Researchers build NNs using GIS software that finds natural statistical patterns or clusters based on values of one or more variables (Esri, n.d.).

Compared with other, more elaborate ways of creating new neighborhood boundaries such as qualitative analysis, NNs require relatively little investment of time or monetary resources. NNs can be generated for any U.S. city or metropolitan area using easily accessible census data using relatively simple GIS techniques. This dissertation includes a sensitivity analysis of three traditional census units and three types of NNs formed by aggregating block groups and census tracts at different scales. Geographically weighted

regression is used to evaluate the strength of association between HDP prevalence and

neighborhood deprivation at these six units of analysis. The unit that produces the best-

performing model (based on proportion of HDP variance explained) is then used in the

subsequent multilevel analysis of individual-level and area-level HDP predictors.

References

Acevedo-Garcia, D. (2005). Zip code-level risk factors for tuberculosis: neighborhood environment and residential segregation in New Jersey, 1985-1992. *American Journal of Public Health*, *91*, 734–741.

Agency for Healthcare Research and Quality. (n.d.). 2014 National Diagnoses — Clinical Classification Software (CCS), Principal Diagnosis: #183 Hypertension complicating pregnancy, childbirth and the puerperium — Standard Errors (SE). Retrieved March 1, 2016, from hcupnet.ahrq.gov

Agyemang, C., Vrijkotte, T. G. M., Droomers, M., van der Wal, M. F., Bonsel, G. J., & Stronks, K. (2009). The effect of neighbourhood income and deprivation on pregnancy outcomes in Amsterdam, The Netherlands. *Journal of Epidemiology & Community Health*, *63*, 755–760.

Ananth, C. V., Keyes, K. M., & Wapner, R. J. (2013). Pre-eclampsia rates in the United States, 1980-2010: Age-period-cohort analysis. *BMJ*, *347*, f6564–f6564.

Andrews, J. O., Mueller, M., Newman, S. D., Magwood, G., Ahluwalia, J. S., White, K., & Tingen, M. S. (2014). The Association of individual and neighborhood social cohesion, stressors, and crime on smoking status among African-American women in Southeastern US subsidized housing neighborhoods. *Journal of Urban Health*, *91*, 1158–1174.

Bellamy, L., Casas, J. P., Hingorani, A. D., & Williams, D. J. (2007). Pre-eclampsia and risk of cardiovascular disease and cancer in later life: systematic review and meta-analysis. *BMJ*, *335*, 974–974.

Bissonnette, L., Wilson, K., Bell, S., & Shah, T. I. (2012). Neighbourhoods and potential access to health care: The role of spatial and aspatial factors. *Health & Place*, *18*, 841–853.

Body-Gendrot, S. (2011). *The Social Control of Cities? A Comparative Perspective*. New York: Wiley & Sons.

Böhm, S., Curran, E. C., Kenny, L. C., O'keeffe, G. W., Murray, D., & Khashan, A. S. (2016). The effect of hypertensive disorders of pregnancy on the risk of attention-deficit/hyperactivity disorder in the offspring. *Journal of Neonatal-Perinatal Medicine*, *6*, 169–170.

Braveman, P., Cubbin, C., Egerter, S., & Pedregon, V. (2011). *Neighborhoods and health*. Robert Wood Johnson Foundation Commission to Build a Healthier America.

Briant, A., Combes, P. P., & Lafourcade, M. (2010). Dots to boxes: do the size and shape of spatial units jeopardize economic geography estimations? *Journal of Urban Economics*, *67*, 287–302.

Buka, S. L. (2003). Neighborhood support and the birth weight of urban infants. *American Journal of Epidemiology*, *157*, 1–8.

Centers for Disease Control and Prevention. (2017). Pregnancy Mortality Surveillance System. Retrieved September 2, 2017, from https://www.cdc.gov/reproductivehealth/maternalinfanthealth/pmss.html

Chappell, L. C., Enye, S., Seed, P., Briley, A. L., Poston, L., & Shennan, A. H. (2008). Adverse Perinatal Outcomes and Risk Factors for Preeclampsia in Women With Chronic Hypertension: A Prospective Study. *Hypertension*, *51*, 1002–1009.

Clausen, T., Øyen, N., & Henriksen, T. (2006). Pregnancy complications by overweight and residential area. A prospective study of an urban Norwegian cohort. *Acta Obstetricia Et Gynecologica Scandinavica*, *85*, 526–533.

Cohen, D., Spear, S., Scribner, R., Kissinger, P., Mason, K., & Wildgen, J. (2000). "Broken windows" and the risk of gonorrhea. *American Journal of Public Health*, *90*, 230–236.

Colatrella, A., Loguercio, V., Mattei, L., Trappolini, M., Festa, C., Stoppo, M., & Napoli, A. (2010). Hypertension in diabetic pregnancy: impact and long-term outlook. *Best Practice & Research Clinical Endocrinology & Metabolism*, *24*, 635–651.

Commission on Social Determinants 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. Geneva: World Health Organization.

Cubbin, C., Sundquist, K., Ahlén, H., Johansson, S. E., Winkleby, M., & Sundquist, J. (2006). Neighborhood deprivation and cardiovascular disease risk factors: Protective and harmful effects. *Scandinavian Journal of Public Health*, *34*, 228–237.

Cunningham, F., KJ, L., SL, B., Hauth, J., Rouse, D. J., & Spong, C. Y. (2010). *Williams Obstetrics*. Toronto: McGraw Hill Medical.

Curran, E. A., Khashan, A. S., O'keeffe, G. W., & Kenny, L. C. (2016). Hypertension in pregnancy and autism spectrum disorder in a British cohort. *Journal of Neonatal-Perinatal Medicine*, *6*, 153.

Dadelszen, von, P., Stones, W., & Mathai, M. (2016). *The FIGO Textbook of Pregnancy Hypertension: An Evidence-Based Guide to Monitoring, Prevention and Management*. London: The Foundation for the Global Library of Women's Medicine.

Diez Roux, A. V. (2007). Neighborhoods and health: where are we and were do we go from here? *Revue D'épidémiologie Et De Santé Publique*, *55*, 13–21.

Dreier, P., Mollenkopf, J. H., & Swanstrom, T. (2014). *Place Matters: Metropolitics for the Twenty-First Century* (3rd ed.). Lawrence, Kan.: University Press of Kansas.

Duley, L. (2009). The Global impact of pre-eclampsia and eclampsia. *Seminars in Perinatology*, *33*, 130–137.

Esri. (n.d.). Tool Reference: Spatial Statistics Toolbox — Grouping Analysis. Retrieved March 18, 2016, from http://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/grouping-analysis.htm

Fingar, K. R., Mabry-Hernandez, I., Ngo-Metzger, Q., Wolff, T., Steiner, C. A., & Elixhauser, A. (2017). Delivery hospitalizations involving preeclampsia and eclampsia, 2005–2014. Rockville, MD: Agency for Health Care Research and Quality (AHRQ).

Flenady, V., Koopmans, L., Middleton, P., Frøen, J. F., Smith, G. C., Gibbons, K., et al. (2011). Major risk factors for stillbirth in high-income countries: a systematic review and meta-analysis. *The Lancet*, *377*, 1331–1340.

Florida Perinatal Quality Collaborative. (2015). *Florida Hypertension in Pregnancy Toolkit: A Quality Improvement Initiative*. Retrieved from https://usfhealth.app.box.com/s/7lu25e7i8w7ifyfp0tjo47g5t8txh9nq

Florida PRAMS. (n.d.). Florida PRAMS trend report: Results from the Florida Pregnancy Risk Assessment Monitoring System: 2000-2011. Tallahassee: Florida Department of Health Bureau of Epidemiology. Retrieved from http://www.floridahealth.gov/statistics-and-data/survey-data/pregnancy-risk-assessment-monitoring-system/reports/_documents/TrendReport.pdf

Ford, J. L., & Browning, C. R. (2011). Neighborhood Social Disorganization and the Acquisition of Trichomoniasis Among Young Adults in the United States. *American Journal of Public Health*, *101*, 1696–1703.

Frank, R., & Bjornstrom, E. (2011). A tale of two cities: Residential context and risky behavior among adolescents in Los Angeles and Chicago. *Health & Place*, *17*, 67–77.

Frieden, T. R. (2014). Six Components Necessary for Effective Public Health Program Implementation. *American Journal of Public Health*, *104*, 17–22.

Gehlert, S., Sohmer, D., Sacks, T., Mininger, C., McClintock, M., & Olopade, O. (2008). Targeting Health Disparities: A Model Linking Upstream Determinants To Downstream Interventions. *Health Affairs*, *27*, 339–349.

Gold, R., Gold, K., Schilling, M., & Modilevsky, T. (2014). Effect of age, parity, and race on the incidence of pregnancy associated hypertension and eclampsia in the United States. *Journal of Neonatal-Perinatal Medicine*, *4*, 46–53.

Grace, T., Bulsara, M., Pennell, C., & Hands, B. (2014). Maternal hypertensive diseases negatively affect offspring motor development. *Journal of Neonatal-Perinatal Medicine*, *4*, 209–214.

Gudmundsson, S., Björgvinsdóttir, L., Molin, J., Gunnarsson, G., & Marsal, K. (1997). Socioeconomic status and perinatal outcome according to residence area in the city of Malmö. *Acta Obstetricia Et Gynecologica Scandinavica*, *76*, 318–323.

Hackshaw, L., Hackshaw, L., McEwen, A., McEwen, A., West, R., West, R., et al. (2010). Quit attempts in response to smoke-free legislation in England. *Tobacco Control*, *19*, 160–164.

Haram, K., Svendsen, E., & Abildgaard, U. (2009). The HELLP syndrome: Clinical issues and management. A review. *BMC Pregnancy and Childbirth*, *9*, 89.

Hipp, J. R. (2007). Block, Tract, and Levels of Aggregation: Neighborhood Structure and Crime and Disorder as a Case in Point. *American Sociological Review*, *72*, 659–680.

Hogue, C. R., Kramer, M. R., Cooper, H. L., Drews-botsch, C. D., & Waller, L. A. (2010). Do measures matter ? Comparing measures of racial residential segregation. *International Journal of Health Geographics*, *9*, 1–15.

Hutcheon, J. A., Lisonkova, S., & Joseph, K. S. (2011). Epidemiology of pre-eclampsia and the other hypertensive disorders of pregnancy. *Best Practice & Research Clinical Obstetrics & Gynaecology*, *25*, 391–403.

Huynh, L., McCoy, M., Law, A., Tran, K. N., Knuth, S., Lefebvre, P., et al. (2013). Systematic Literature Review of the Costs of Pregnancy in the US. *PharmacoEconomics*, *31*, 1005–1030.

Jackson, R., Smith, D., Tabnak, F., & Vugia, D. (2015). Disparities of shigellosis rates among California children by race/ethnicity and census tract poverty level, 2000–2010. *The Pediatric Infectious Disease Journal*, *34*, 843–847.

Janevic, T., Savitz, D. A., Stein, C. R., Kaufman, J. S., Mason, S. M., & Herring, A. H. (2010). Neighborhood deprivation and adverse birth outcomes among diverse ethnic groups. *Annals of Epidemiology*, *20*, 445–451.

Karriker-Jaffe, K. J. (2013). Neighborhood socioeconomic status and substance use by U.S. adults. *Drug and Alcohol Dependence*, *133*, 212–221.

Kingsley, G., Coulton, C. J., & Pettit, K. L. (2016). *Strengthening Communities with Neighborhood Data*. Rowman & Littlefield. Retrieved from https://www.neighborhoodindicators.org/sites/default/files/publications/13805-urban_kingsley.pdf

Laraia, B. A., Messer, L., Kaufman, J. S., Dole, N., Caughy, M., O'Campo, P., & Savitz, D. A. (2006). Direct observation of neighborhood attributes in an urban area of the US south: characterizing the social context of pregnancy. *International Journal of Health Geographics*, *5*, 11.

Law, A., McCoy, M., Lynen, R., Curkendall, S. M., Gatwood, J., Juneau, P. L., & Landsman-Blumberg, P. (2015a). Costs of Newborn Care Following Complications During Pregnancy and Delivery. *Maternal and Child Health Journal*, *19*, 2081–2088.

Law, A., McCoy, M., Lynen, R., Curkendall, S. M., Gatwood, J., Juneau, P. L., & Landsman-Blumberg, P. (2015b). The prevalence of complications and healthcare costs during pregnancy. *Journal of Medical Economics*, *18*, 533–541.

Leffert, L. R., Clancy, C. R., Bateman, B. T., Bryant, A. S., & Kuklina, E. V. (2015). Hypertensive Disorders and Pregnancy-Related Stroke. *Obstetrics & Gynecology*, *125*, 124–131.

Lo, J. O., Mission, J. F., & Caughey, A. B. (2013). Hypertensive disease of pregnancy and maternal mortality. *Current Opinion in Obstetrics and Gynecology*, *25*, 124–132.

Lykke, J. A., Lykke, J. A., Langhoff-Roos, J., Langhoff-Roos, J., Sibai, B. M., Sibai, B. M., et al. (2009). Hypertensive Pregnancy Disorders and Subsequent Cardiovascular Morbidity and Type 2 Diabetes Mellitus in the Mother. *Hypertension*, *53*, 944–951.

MacDonald, M. A. (2004). From Miasma to Fractals: The Epidemiology Revolution and Public Health Nursing. *Public Health Nursing*, *21*, 380–391.

Macintyre, S., Maciver, S., & Sooman, A. (1993). Area, class and health: Should we be focusing on places or people? *Journal of Social Policy*, *22*, 213–234.

Marmot, M., & Wilkinson, R. G. (Eds.). (2009). *Social Determinants of Health* (2nd ed.). London: Oxford University Press.

Martin, J. A., Hamilton, B. E., Ventura, S. J., et al. (2011). Births: Final data for 2009 National Vital Statistics Reports; Vol. 60 No. 1. Hyattsville, MD: National Center for Health Statistics.

Masi, C. M., Hawkley, L. C., Harry Piotrowski, Z., & Pickett, K. E. (2007). Neighborhood economic disadvantage, violent crime, group density, and pregnancy outcomes in a diverse, urban population. *Social Science & Medicine*, *65*, 2440–2457.

Messer, L. C., & Kaufman, J. S. (2006). Using Census Data to Approximate Neighborhood Effects. In J. M. Oakes & J. S. Kaufman, *Methods in Social Epidemiology* (pp. 209–236). New York: John Wiley & Sons Incorporated..

Messer, L. C., Laraia, B. A., Kaufman, J. S., Eyster, J., Holzman, C., Culhane, J., et al. (2006a). The development of a standardized Neighborhood Deprivation Index. *Journal of Urban Health*, *83*, 1041–1062.

Messer, L. C., Laraia, B. A., Savitz, D. A., Kaufman, J. S., & Dole, N. (2006b). Neighborhood Crime, Deprivation, and Preterm Birth. *Annals of Epidemiology*, *16*, 455–462.

Miranda, M. L., Swamy, G. K., Maxson, P., James, S., Edwards, S., & Gelfand, A. (2010). Disparities in maternal hypertension and pregnancy outcomes: Evidence from North Carolina, 1994-2003, *Public Health Reports, 125*, 579–587.

Morenoff, J. D., House, J. S., Hansen, B. B., Williams, D. R., Kaplan, G. A., & Hunte, H. E. (2007). Understanding social disparities in hypertension prevalence, awareness, treatment, and control: The role of neighborhood context. *Social Science & Medicine*, *65*, 1853–1866.

Mujahid, M. S., Diez Roux, A. V., Morenoff, J. D., & Raghunathan, T. (2007). Assessing the Measurement Properties of Neighborhood Scales: From Psychometrics to Ecometrics. *American Journal of Epidemiology*, *165*, 858–867.

Murray, C. J. L., Kulkarni, S. C., Michaud, C., Tomijima, N., Bulzacchelli, M. T., Iandiorio, T. J., & Ezzati, M. (2006). Eight Americas: Investigating mortality disparities across races, counties, and race-counties in the United States. *PLoS Medicine*, *3*, e260.

Nkansah-Amankra, S. (2010). Neighborhood contextual factors, maternal smoking, and birth outcomes: Multilevel analysis of the South Carolina PRAMS Survey, 2000-2003, *Journal of Women's Health, 19*, 1543–1552.

Office of Disease Prevention and Health Promotion. (2017, September 6). Healthy People 2020 topics & objectives: Maternal, infant, and child health. Retrieved from https://www.healthypeople.gov/2020/topics-objectives/topic/maternal-infant-and-child-health#seven

Openshaw, S. (1984). Ecological fallacies and the analysis of areal census data, *Environment and Planning A: Economy and Space, 16*, 17–31.

Parenteau, M.-P., & Sawada, M. C. (2011). The modifiable areal unit problem (MAUP) in the relationship between exposure to NO2 and respiratory health. *International Journal of Health Geographics, 10*, 58.

Pickett, K. E. (2001). Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. *Journal of Epidemiology & Community Health, 55*, 111–122.

Ross, N. A., Tremblay, S., & Graham, K. (2004). Neighbourhood influences on health in Montréal, Canada. *Social Science & Medicine, 59*, 1485–1494.

Sabel, C. E., Kihal, W., Bard, D., & Weber, C. (2013). Creation of synthetic homogeneous neighbourhoods using zone design algorithms to explore relationships between asthma and deprivation in Strasbourg, France. *Social Science & Medicine, 91*, 110–121.

Sampson, R. J., Morenoff, J. D., & Gannon-Rowley, T. (2002). Assessing "Neighborhood Effects": Social Processes and New Directions in Research. *Annual Review of Sociology, 28*, 443–478.

Sampson, R. J. (2012). *Great American City: Chicago and the Enduring Neighborhood Effect*. Chicago: University of Chicago Press.

Say, L., Chou, D., Gemmill, A., Tunçalp, Ö., Moller, A.-B., Daniels, J., et al. (2014). Global causes of maternal death: a WHO systematic analysis. *The Lancet Global Health, 2*, e323–e333.

Schuurman, N., Bell, N., Dunn, J. R., & Oliver, L. (2007). Deprivation indices, population health and geography: An evaluation of the spatial effectiveness of indices at multiple scales. *Journal of Urban Health, 84*, 591–603.

Shen, J., Tymkow, C., & MacMullen, N. (2005). Disparities in maternal outcomes among four ethnic populations. *Ethnicity & Disease, 15*, 492–497.

Slopen, N., Non, A., Williams, D. R., Roberts, A. L., & Albert, M. A. (2014). Childhood Adversity, Adult Neighborhood Context, and Cumulative Biological Risk for Chronic Diseases in Adulthood. *Psychosomatic Medicine, 76*, 481–489.

Stevens, W., Shih, T., Incerti, D., Ton, T. G. N., Lee, H. C., Peneva, D., et al. (2017). Short-term costs of preeclampsia to the United States health care system. *American Journal of Obstetrics and Gynecology, 217*, 237–248.e16.

Stock, C., & Ellaway, A. (2013). *Neighbourhood Structure and Health Promotion*. Boston, MA: Springer.

Talen, E., & Shah, S. (2007). Neighborhood Evaluation Using GIS. *Environment and Behavior*, *39*, 583–615.

Tanaka, M., Jaamaa, G., Kaiser, M., Hills, E., Soim, A., Zhu, M., et al. (2007). Racial Disparity in Hypertensive Disorders of Pregnancy in New York State: A 10-Year Longitudinal Population-Based Study. *American Journal of Public Health*, *97*, 163–170.

Task Force on Hypertension in Pregnancy. (2013). *Hypertension in Pregnancy*. Washington, D.C.: American College of Obstetricians and Gynecologists.

Tucker, M. J., Berg, C. J., Callaghan, W. M., & Hsia, J. (2007). The Black–White disparity in pregnancy-related mortality from 5 conditions: Differences in prevalence and case-fatality rates. *American Journal of Public Health*, *97*, 247–251.

United Nations. (2015). The Millennium Development Goals Report 2015. Retrieved September 3, 2017, from http://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20rev%20(July%201).pdf

United States Census Bureau. (n.d.). Geography: 2010 Geographic Terms and Concepts. Retrieved May 22, 2014, from https://www.census.gov/geo/reference/terms.html

van Panhuis, W. G., Paul, P., Emerson, C., Grefenstette, J., Wilder, R., Herbst, A. J., et al. (2014). A systematic review of barriers to data sharing in public health. *BMC Public Health*, *14*, 3.

Vinikoor-Imler, L. C., Gray, S. C., Edwards, S. E., & Miranda, M. L. (2011a). The effects of exposure to particulate matter and neighbourhood deprivation on gestational hypertension. *Paediatric and Perinatal Epidemiology*, *26*, 91–100.

Vinikoor-Imler, L. C., Messer, L. C., Evenson, K. R., & Laraia, B. A. (2011b). Neighborhood conditions are associated with maternal health behaviors and pregnancy outcomes. *Social Science & Medicine*, *73*, 1302–1311.

Wacquant, L. J. (1993). Urban Outcasts: Stigma and Division in the Black American Ghetto and the French Urban Periphery. *International Journal of Urban and Regional Research*, *17*, 366–383.

Wagner, S. J., Barac, S., & Garovic, V. D. (2007). Hypertensive pregnancy disorders: Current concepts. *The Journal of Clinical Hypertension*, *9*, 560–566.

Wallerstein, N. B., Yen, I. H., & Syme, S. L. (2011). Integration of Social Epidemiology and Community-Engaged Interventions to Improve Health Equity. *American Journal of Public Health*, *101*, 822–830.

Wei, C., Cabrera-Barona, P., & Blaschke, T. (2016). Local geographic variation of public services inequality: Does the neighborhood scale matter? *International Journal of Environmental Research and Public Health*, *13*, 981.

Weiss, L., Ompad, D., Galea, S., & Vlahov, D. (2007). Defining Neighborhood Boundaries for Urban Health Research. *American Journal of Preventive Medicine*, *32*, S154–S159.

Wilske, F. (2008). Approximation of Neighborhood Boundaries Using Collaborative Tagging Systems. *GI-Days*, *32*, 179–187.

Wolf, M., Shah, A., Jimenez-Kimble, R., Sauk, J., Ecker, J. L., & Thadhani, R. (2004). Differential risk of hypertensive disorders of pregnancy among Hispanic women. *Journal of the American Society of Nephrology*, *15*, 1330–1338.

Yanit, K. E., Snowden, J. M., Cheng, Y. W., & Caughey, A. B. (2012). The impact of chronic hypertension and pregestational diabetes on pregnancy outcomes. *American Journal of Obstetrics and Gynecology*, *207*, 333.e1–333.e6.

Yen, I. H., Michael, Y. L., & Perdue, L. (2009). Neighborhood environment in studies of health of older adults. *American Journal of Preventive Medicine*, *37*, 455–463.

Zhang, J., Meikle, S., & Trumble, A. (2003). Severe Maternal Morbidity Associated with Hypertensive Disorders in Pregnancy in the United States. *Hypertension in Pregnancy*, *22*, 203–212.

CHAPTER III

METHODOLOGY

i.   Aim 1: Thematic review

Search strategy

The focus of this thematic review is to identify alternative sources of neighborhood

boundaries developed outside of academic or government research that might have

potential future uses in U.S. neighborhood health research. Therefore, this review was

limited to resources developed by the private sector that either: are viable candidates to

someday augment or replace census boundaries in neighborhood health studies; employ

novel methods that could be expanded upon or adapted; or provide necessary context

regarding key themes or trends in this emerging field.

Three researchers conducted data searches independently, then pooled their findings

before collectively deciding on most important and pertinent studies to include in the

review. Relevant resources were identified through keyword searches on Google, Google

Scholar, and an online directory of application programming interfaces (APIs)

(ProgrammableWeb, n.d.). An API is a set of procedural building blocks (protocols,

routines, etc.) that governs how parts of one or more applications relate to and interact

with each other. Some websites, such as Google, offer users access to extensive portions

of their data through an "API key," (Chakraborty, Wilson, Sarraf, & Jana, 2015; Sheehan,

2013). For researchers with basic coding skills, this represents a new opportunity to

obtain large datasets, often at no charge.

Study inclusion criteria were: i) information provided in English, ii) polygons or

centroids developed/designed by private sector and/or private citizens, and iii) resources

accessible to researchers or the general public, whether for a fee or free of charge. There were two exclusion criteria: i) polygons or centroids designed solely by an academic institution or government agency/committee (federal, state, or local) and ii) evidence that a resource/website had not been updated in more than a year. As shown in Table 1, search terms related to neighborhoods and GIS analysis were used, and a Google search was conducted for colloquial phrases a consumer might use when looking for local goods and services. Google uses an algorithm to rank its search results based on relevance (Google, n.d.). For this review, only the first 10 pages (i.e., 100 results) for each search term were reviewed. Two sets of Google search terms and one set of Google Scholar terms were used. Thus, a total of 30 pages or 300 results were reviewed. Finally, websites of 9 companies known for using interactive maps to display information (airbnb, Bing, Esri, Google, RedFin, Trulia, Walk Score, Yelp!, Zillow) were reviewed, and domain-specific searches for the aforementioned GIS-related terms were conducted. These sites' partner, investor, and developer webpages were scrutinized for relevant content.

On 19 occasions, information on one website alluded to similar tools created by another organization (Table 2). These additional resources were researched and included if they met the aforementioned study inclusion criteria. The following details were extracted from each website and entered into an Excel sheet: website URL, organization type, geographic coverage area(s); available data types/file formats; data sources, collection, and analysis; access fees; and any other pertinent information. Had reviewers disagreed on whether any resources met the inclusion criteria, majority vote would have ruled. However, no such disagreements arose. Findings were grouped into four categories based on the type of organization that created the neighborhood resource.

Table 1: Search string for private sector sources of novel neighborhood boundaries

Website	Search Terminology
Google Scholar* restricted to 2013-2016	(neighborhood OR neighborhoods OR 'hood) AND (define OR definition OR "informal space" OR "location based analytics" OR "location intelligence tool" OR metric OR microtarget OR tool)
Google*	(neighborhood OR neighborhoods OR 'hood) AND (centroid OR boundary OR polygon OR shapefile) AND (app OR API OR "location based analytics" OR "location intelligence tool" OR gazetteer)
Google*	(best OR good OR low-crime OR safe) AND (neighborhood OR neighborhoods OR 'hood) AND (centroid OR boundary OR polygon OR shapefile)
Programmable Web API directory	neighborhood

*First 10 pages reviewed

API – application program interface

Table 2: Websites from original search that produced secondary sources for review

URL
1. http://web.lotadata.com/blog/neighborhood-data-landscape-part-one-geometry/
2. https://www.quora.com/How-do-you-get-boundaries-of-neighborhoods-via-the-Google-Maps-api
3. https://gis.stackexchange.com/questions/29067/openstreetmap-neighborhood-boundaries
4. https://www.walkscore.com/professional/travel-time-api.php
5. https://www.opendataphilly.org/dataset/philadelphia-neighborhoods
6. http://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/modeling-spatial-relationships.htm
7. http://geolode.org/?q=tag%3Dneighborhoods
8. https://github.com/DNAinfoData/Draw-Your-Neighborhood
9. http://code.flickr.net/2011/01/08/flickr-shapefiles-public-dataset-2-0/
10. https://www.trulia.com/voices/In_My_Neighborhood/I_was_wondering_if_you_had_neighborhood_boundary_m-844418
11. https://www.citylab.com/life/2015/09/how-many-neighborhoods-is-too-many-for-one-map/403474/
12. https://www.fastcodesign.com/1669554/a-map-of-your-city-s-invisible-neighborhoods-according-to-foursquare
13. http://www.marketwired.com/press-release/urban-mapping-enhances-industry-leading-neighborhood-boundary-data-technology-expands-1424697.htm
14. http://ncase.me/polygons/
15. http://wiki.openstreetmap.org/wiki/Wikimapia
16. https://www.azavea.com/
17. https://medium.com/airbnb-engineering/behind-the-scenes-airbnb-neighborhoods-cef63242eab7
18. https://www.programmableweb.com/api/philly-hoods
19. https://www.directionsmag.com/article/2554

ii.   Aims 2 and 3: Sensitivity analysis and multilevel analysis

Data sources

Area-level demographic data were obtained from the U.S. Census Bureau American Community Survey (U.S. Census Bureau, n.d.). Data to acertain hypertensive disorders of pregnancy (HDP) status and all individual-level variables were obtained from the 2008–2012 Florida Department of Health (DOH) birth records for Miami-Dade County (n = 159,069). The study protocol was approved by the Institutional Review Boards at Florida DOH and Florida International University. A total of 11.6% of birth records were excluded because of missing or misspelled addresses, leaving 140,551 records. For continuous variables, missing data ($< 5\%$) was imputed using the multiple imputation package in SPSS 24.0. Participants with missing data on ≥ 1 categorical variable were excluded. After these exclusions for missing variables, the final sample was 121,421 for the Aim 2 and Aim 3 analysis. When Cohen's d statistics were calculated, there were no significant differences between the missing and non-missing cases on any study variables.

Neighborhood units

Geographic boundaries were obtained from the U.S. Census Bureau (U.S. Census Bureau, n.d.). To identify the residential location of each mother in terms of census tracts and ZCTAs, maternal addresses were geocoded using ArcGIS 10.5 software (Esri, 2017). Birth record data were linked to census data using the unique geographic identifiers.

Natural neighborhoods (NNs) were constructed using the ArcGIS grouping analysis tool. The attributes used to distinguish clusters from each other are referred to as

"analysis fields." When a spatial constraint is specified, the tool uses a connectivity graph (minimum spanning tree) to identify natural groupings (Esri, n.d.). I used this tool to create three different alternative areal units that were equivalent in scale to existing census boundaries: BG small NNs – 1,560 block groups clustered into 507 NNs; BG large NNs – 1,560 block groups clustered into 78 NNs; and CT large NNs – 507 census tracts clustered into 78 NNs. These three types of NN represented block groups reaggregated at the scale of census tracts (BG small NNs), block groups reaggregated at the scale of ZCTAs (BG large NNs), and census tracts reaggregated at the scale of ZCTAs (CT large NNs). To increase socioeconomic homogeneity within each cluster, the eight factors of the neighborhood deprivation index were entered as analysis fields. To make the scale of the NNs equivalent to that of either census tracts (n=507) or ZCTAs (n=78), I requested 507 or 78 groups depending on the type of NN being generated. I set the spatial constraint to "contiguous edges or corners" and set k to 1 to ensure that each micro unit (e.g., block group or census tract) within a given NN shared an edge with at least 1 other micro unit.

Variables included in the analysis

Outcome of interest

Hypertensive Disorders of Pregnancy: First, a composite HDP variable was created by combining responses to three questions from the birth record: "Was mother diagnosed with gestational hypertension (pregnancy-induced hypertension, pre-eclampsia, etc.) during this pregnancy?", "Was mother diagnosed with gestational hypertension (eclampsia) during this pregnancy?", and "Did the mother have a history of chronic

hypertension prior to this pregnancy?". Available answer choices were yes, no, and unknown. The outcome variable was dichotomized as yes/no, and participants who answered "unknown" were excluded from the sample.

For Aim 2, the outcome variable was HDP prevalence in Miami-Dade County neighborhoods — aggregated to each of the six areal units used in this study. The standard prevalence definition was used: percentage of each neighborhood's population who have the disease of interest. The number of HDP cases in a neighborhood was divided by the total number of births in that neighborhood. The quotient was then multiplied by 100 to obtain a percentage. All women who reported having gestational hypertension, pre-eclampsia, eclampsia, or chronic hypertension during their pregnancy were considered HDP cases. For Aim 3, the outcome of interest was individual-level HDP status, using the composite, dichotomous HDP variable described above.

Area-level exposure

Neighborhood deprivation: The area-level exposure of interest, neighborhood deprivation, was measured with an eight-item index (Messer et al., 2006) used in previous area-level health studies (Elo et al., 2009; Gustafson, Lewis, Wilson, & Jilcott-Pitts, 2012; Janevic et al., 2010; Laraia et al., 2006). It comprised eight variables that represented five domains: *poverty* (% households below the poverty line; % female-headed households with dependent children; % residents receiving public assistance; % residents earning < $30,000 annually); *under-education* (% residents ≥ age 25 without a high school diploma); *unemployment* (% unemployed residents;); *occupation* (% adult males who do not hold managerial/professional jobs); and *crowded households*

47

(% households with > 1 person per room). A neighborhood deprivation index (NDI)

score was calculated for each neighborhood using principal components analysis (PCA)

with varimax rotation to weight each variable's relative contribution to the score.

Neighborhood deprivation index scores were standardized to have a mean of 0 and a

standard deviation of 1, with higher scores reflecting higher deprivation. For Aim 3, NDI

scores were divided into tertiles, which were categorized as high (T3), moderate (T2),

and low neighborhood deprivation (T1).

Aim 3 individual-level variables

Educational attainment: Individual-level educational attainment was used a proxy for

individual-level socio-economic deprivation. It was categorized as < high school

diploma / General Equivalency Diploma (GED) age < 18 years, and < high school

diploma / GED age ≥ 18 years, and ≥ high school diploma / GED, with the latter serving

as the reference group.

Insurance status: Insurance status was also used as a proxy for individual-level socio-

economic deprivation. It was categorized as self-pay (a proxy for lack of insurance),

Medicaid/Medicare, and private insurance/other, with the latter serving as the reference

group.

Race/ethnicity: Individual-level race/ethnicity was categorized as non-Hispanic

White, non-Hispanic Black, Hispanic and Other race/ethnicity.

Pre-pregnancy body mass index (BMI): This behavioral variable was calculated from

self-reported height and weight using the standard BMI formula of kilograms of body

weight divided by height in meters squared. Body mass index was categorized into six groups: underweight (<18.5 kg/m^2), normal weight (18.5-24.9 kg/m^2), overweight (25.0-29.9 kg/m^2), category I obese (30-34.9 kg/m^2), category II obese (35.0-39.9 kg/m^2), and category III obese (≥40.0 kg/m^2).

Gestational weight gain: This behavioral variable was calculated by subtracting self-reported pre-pregnancy weight from the mother's self-reported weight at the time of delivery. Responses were dichotomized as excessive and not excessive based on pre-pregnancy BMI status and the Institute of Medicine criteria for gestational weight gain. Institute of Medicine guidelines stipulate that during pregnancy, obese women should gain 11-20 lbs., overweight 15-25 lbs., normal weight 25-35 lbs., and underweight 28-40 lbs. (Rasmussen, Yaktine, Institute of Medicine Committee to Reexamine IOM Pregnancy Weight Guidelines, Food and Nutrition Board and Board on Children, Youth, and Families, 2009).

Smoking during pregnancy: This behavioral variable was dichotomized as smoked cigarettes during pregnancy vs. did not smoke cigarettes during pregnancy.

Number of prenatal care visits: This behavioral variable was entered as a grand mean-centered continuous variable.

Nulliparity, multifetal gestation, gestational diabetes, pre-existing diabetes (Type 1 or 2): These four medical variables were dichotomized as yes/no. The birth record did not distinguish between Type 1 and 2 diabetes, so a single dichotomous variable for "pre-existing diabetes (Type 1 or 2)" was used.

<u>Maternal age</u>: This variable was entered as a grand mean-centered, continuous variable.

Statistical analysis

Aim 2 analysis

Aim 2 of this dissertation was to conduct a sensitivity analysis to compare the strength of association between HDP prevalence and neighborhood deprivation at six different units of analysis — block group, census tract, ZIP Code Tabulation Area (ZCTAs), and three types of NN formed by aggregating block groups and census tracts at three different scales — using structural equation modeling and geographically weighted regression. By comparing the R^2 values for the six geographic units, I determined the unit of analysis to use for Aim 3.

For each of the six areal units, NDI and HDP prevalence maps were generated in ArcGIS 10.5, for a total of 12 maps. Values for both variables were expressed as quintiles to foster easier visual comparison between maps. Additional descriptive statistics were calculated in SPSS 24.0 (IBM Corp, Armonk, NY, 2016). Because the distributions of some neighborhood deprivation variables were skewed, nonparametric Kruskal-Wallis H tests were used to compare median values across areal units. Considered a nonparametric alternative to the one-way ANOVA, the Kruskal-Wallis H test's null hypothesis is that the groups' medians are equal (Laerd Statistics, 2013). The skewedness of the neighborhood variables did not affect my regression analyses because the variables were entered into the regression models as a single, standardized index variable.

After confirming the presence of a linear relationship between HDP prevalence and

neighborhood deprivation using linearity tests in SPSS 24.0, a two-step regression

process was used to evaluate the relationship between HDP prevalence and neighborhood

deprivation at the six areal units under investigation: block groups, census tracts, ZCTAs,

and the three types of NN. First, separate ordinary least squares (OLS) regression models

were constructed for each areal unit, with HDP as the outcome variable and

neighborhood deprivation as the independent variable. Unadjusted regression

coefficients, 95% confidence intervals, p-values, and R^2 values were recorded. Models

found to have statistically significant regression coefficients in the OLS analysis were

then analyzed using geographically weighted regression (GWR). This type of regression

has the benefit of accounting for spatial variation in independent variable (e.g.,

neighborhood deprivation as measured by NDI score) and dependent variables (e.g., HDP

prevalence) (Fotheringham, Charlton, & Brunsdon, 1998). Unlike traditional forms of

regression analysis, GWR produces regression coefficients for each individual

geographic unit in the data set instead of a regression coefficient for the overall model

(Fotheringham et al., 1998). Thus GWR regression coefficients could not be included in

the Aim 2 results. A map displaying areas of under- and over-prediction in each GWR

model and the R^2 value were generated for each GWR model. Also known as the

coefficient of determination, R^2 is a model fit statistic that measures the proportion of

variance in a dependent variable explained by a model's independent variable(s) (Brown,

2006). To determine the optimal areal unit for studying the relationship between HDP

and neighborhood deprivation, GWR R^2 values from all six areal units were compared.

R^2 values also were used to explore the MAUP's potential influences on the relationships

between HDP and neighborhood deprivation. I compared the three census unit models to

51

evaluate the scale effect and compared models of the same scale that were aggregated differently to evaluate the zoning effect.

The differences in the models for Aims 2 and 3 are intentional: The initial Aim 2 models are simple with one dependent variable (HDP prevalence) and one indepent variable (neighborhood deprivation score) per unit of analysis. This is to avoid infringing on the uniqueness of the more complex Aim 3 multilevel model, which is intended to test hypotheses that could help inform future maternal health interventions and policies. It would stand to reason that if neighborhood deprivation influences HDP prevalence, any potential relationship between neighborhood deprivation and individual-level HDP odds would occur through the same mechanism(s). Thus, the appropriate scale/unit of analysis should be the same for both.

Aim 3 analysis

Aim 3 of this dissertaton was to quantify the relationship between neighborhood deprivation and individual-level HDP status in Miami-Dade County, Florida using multilevel logistic regression. The cross-level interactions between neighborhood deprivation and two measures of individual-level deprivation — low educational attainment (< high school diploma/GED) and lack of access to health care (uninsured/Medicaid recipient) were also examined.

- *Hypothesis 1a: Women living in neighborhoods with higher levels of deprivation will have higher odds of HDP compared with those in low-deprivation neighborhoods.*

- *Hypothesis 1b: Women with low educational attainment who live in deprived neighborhoods will have higher odds of HDP compared with women with low educational attainment in low-deprivation neighborhoods.*

- *Hypothesis 1c: Women who lack access to health care who live in deprived neighborhoods will have higher odds of HDP compared with women who lack access to health care in low-deprivation neighborhoods.*

Individual-level descriptive statistics were calculated and compared using ANOVA and Chi-square tests. Traditional, individual-level binary logistic regression was conducted to determine which variables would be included in the multilevel analysis. Only variables significant at the 0.05 level were retained for multilevel modeling.

A two-level logistic regression model was constructed in SPSS 24.0 using the "General Linear Mixed" procedure with random intercepts to allow the outcome variable, HDP, to vary randomly across natural neighborhoods (Heck, Thomas, & Tabata, 2012). Neighborhood deprivation and individual-level variables and covariates were entered as fixed parameters. "Neighborhood deprivation x educational attainment" "neighborhood deprivation x insurance status" were entered as cross-level interaction terms.

References

Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York: Guilford Press.

Chakraborty, A., Wilson, B., Sarraf, S., & Jana, A. (2015). Open data for informal settlements: Toward a user's guide for urban managers and planners. *Journal of Urban Management*, *4*, 74–91.

Elo, I. T., Culhane, J. F., Kohler, I. V., O'Campo, P., Burke, J. G., Messer, L. C., et al. (2009). Neighbourhood deprivation and small-for-gestational-age term births in the United States. *Paediatric and Perinatal Epidemiology*, *23*, 87–96.

Esri. (2017). ArcGIS Desktop: ArcMap. Retrieved from
http://desktop.arcgis.com/en/arcmap/

Esri. (n.d.). Tool Reference: Spatial Statistics Toolbox — Grouping Analysis. Retrieved
March 18, 2016, from http://pro.arcgis.com/en/pro-app/tool-reference/spatial-
statistics/grouping-analysis.htm

Fotheringham, A. S., Charlton, M. E., & Brunsdon, C. (1998). Geographically Weighted
Regression: A natural evolution of the expansion method for spatial data analysis.
Environment and Planning A, *30*, 1905–1927.

Google. (n.d.). How Google Search Works: Search Algorithms. Retrieved March 18,
2017, from https://www.google.com/search/howsearchworks/algorithms/

Gustafson, A. A., Lewis, S., Wilson, C., & Jilcott-Pitts, S. (2012). Validation of food
store environment secondary data source and the role of neighborhood deprivation in
Appalachia, Kentucky. *BMC Public Health*, *12*, 688.

Heck, R. H., Thomas, S., & Tabata, L. (2012). *Multilevel Modeling of Categorical
Outcomes Using IBM SPSS*. Routledge: Taylor and Francis.

Janevic, T., Savitz, D. A., Stein, C. R., Kaufman, J. S., Mason, S. M., & Herring, A. H.
(2010). Neighborhood deprivation and adverse birth outcomes among diverse ethnic
groups. *Annals of Epidemiology*, *20*, 445–451.

IBM Corp. Released 2016. IBM SPSS Statistics for Windows, Version 24.0. Armonk,
NY: IBM Corp.

Laerd Statistics. (2013). Kruskal-Wallis H Test using SPSS Statistics. Retrieved July 13,
2017, from https://statistics.laerd.com/spss-tutorials/kruskal-wallis-h-test-using-spss-
statistics.php

Laraia, B. A., Messer, L., Kaufman, J. S., Dole, N., Caughy, M., O'Campo, P., & Savitz,
D. A. (2006). Direct observation of neighborhood attributes in an urban area of the US
south: characterizing the social context of pregnancy. *International Journal of Health
Geographics*, *5*, 11.

Messer, L. C., Laraia, B. A., Kaufman, J. S., Eyster, J., Holzman, C., Culhane, J., et al.
(2006). The development of a standardized Neighborhood Deprivation Index. *Journal of
Urban Health*, *83*, 1041–1062.

ProgrammableWeb. (n.d.). ProgrammableWeb. Retrieved October 15, 2016, from
http://www.programmableweb.com

Rasmussen, K. M., Yaktine, A. L., Institute of Medicine Committee to Reexamine IOM Pregnancy Weight Guidelines, Food and Nutrition Board and Board on Children, Youth, and Families. (2009). Weight gain during pregnancy: Reexamining the guidelines. Washington D.C.: National Academies Press.

Sheehan, J. (2013, March 12). APIs are dead, Long live APIs. Retrieved September 4, 2017, from https://thenextweb.com/dd/2013/03/12/apis-are-dead-long-live-apis/#.tnw_2fGlh12E

U.S. Census Bureau. (n.d.). American FactFinder. Retrieved September 8, 2017, from https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml

U.S. Census Bureau. (n.d.). Geography: TIGER Products. Retrieved September 8, 2017, from https://www.census.gov/geo/maps-data/data/tiger.html

CHAPTER IV

RESULTS

i.   Aim 1: Thematic review

Aim 1 of this dissertation was to conduct a thematic review of neighborhood

boundaries developed by private sector organizations to explore their potential utility in

public health research. I identified 22 organizations that have collected geospatial data or

developed community-based tools that could potentially advance neighborhood health

research (Table 3). Four companies' resources facilitate searches for local goods and

services, and two companies design tools for real estate searches. Seven companies offer

location-based intelligence tools, and collaborations among private citizens have

generated nine other viable resources. Key advantages of tools developed by each type of

organization are summarized in Table 4.

Local search

When looking for local business or services, consumers no longer have to limit their

search options to ZIP codes or street addresses: Companies have designed interactive

maps that allow users to browse within a specific neighborhood (Biancalana et al., 2011;

DeMers, 2014; Smith, 2017). Google Maps has data on more than 100,000,000 places

worldwide, including many U.S. neighborhood boundaries. In developing these

boundaries, Google Maps was able to leverage the sizeable resources of its parent

company, Alphabet. Data sources for its maps include Google directory and search

engine data, Google Map Maker user feedback, information gathered by its Street View

vehicles, and traditional government data (Capps, n.d.). Application program interface

(API) usage is free for < 2,500 data requests and < 25,000 map loads per day, and non-

profits can apply for a free premium usage license (Google Maps API, n.d.). Any maps produced using these data must be available to the public for free.

Similarly, Yelp! offers boundaries for neighborhoods across the globe. These data are free to the public via the company's API. Yelp! also posts an academic dataset online, which is free to anyone willing to complete a request form and adhere to the usage agreement (Yelp!, n.d.). On Yelp's website, one can find links to published papers that have used the company's data (Google Scholar, n.d.; Pranata & Susilo, 2016; Schomberg, Haimson, Hayes, & Anton-Culver, 2016; Sussman et al., 2014).

Real estate

It is often said that real estate is all about "location, location, location." When potential customers can quickly and precisely identify properties in the specific area where they wish to live, they are more likely to rent or buy property. In July 2017, the top two real estate websites, Zillow and Trulia, garnered 59,000,000 visits in a single month (Statista, n.d.). Companies in this industry have developed user-friendly tools that allow customers to take a virtual stroll through a community and view key details, such as school rankings and crime statistics. Many of these same indicators are germane to neighborhood health studies.

One real estate company whose resources might benefit public health researchers is Zillow. Its website includes 7,000 neighborhood polygons from 41 states and Washington, D.C. To develop these boundaries, Zillow took a truly mixed methods approach: "various tactics, including calling individual chambers of commerce, tourism and convention boards, speaking with real estate agents and community members in these

areas, as well as using available online local sources" (Zillow, n.d.). The polygons are posted on the company's website, in the form of ArcGIS shapefiles, the industry standard format for GIS analysis (GISGeography, 2015). The polygons are available for use under a Creative Commons license.

By augmenting high-tech and traditional data sources with realtor feedback, Home Junction has developed a "proprietary algorithm" to construct three types of local boundaries: districts, neighborhoods, and subdivisions ("Boundaries For Developing IDX Search & Filter Applications," n.d.). The company also sells raw data to complement these polygons. Colleges, schools, and government agencies can apply for "courtesy use" privileges. Researchers should be aware that the company prioritizes bulk data retrieval for paying customers — meaning courtesy users' bulk requests could be unexpectedly delayed. However, those with coding skills can retrieve their own data via the company's API as soon as an agreement has been reached and paperwork completed.

Location-based intelligence

Experts have estimated that 80% of business data have a spatial component (Forbes, 2015). The relatively new discipline of "location intelligence" or "location-based intelligence" allows companies to make better use of this location-based data by blending traditional data-driven decision-making (business intelligence) with geospatial analysis (Panian, 2012). Proponents say that location-based intelligence enhances companies' profit-making potential by revealing previously unnoticed relationships between variables and new markets for growth (Forbes, 2015).

One specific type of location-based intelligence is market segmentation. Clustering people based on shared characteristics allows companies to tailor advertising and marketing strategies in hopes of increasing sales. One can categorize a population by location (geographic segmentation), gender, race/ethnicity, income, education (demographic segmentation), or even personality traits, motivation, and lifestyle (psychographic segmentation) (Cant, Strydom, Jooste, & Plessis, 2009). In recent years, companies have begun combining multiple types of segmentation in order to "micromarket" products to specific communities or neighborhoods. The data and boundary files that inform these micromarketing campaigns are available to anyone who can afford the required access fees. Of the many types of resources described in this review, it is likely that this is the one with which some health researchers are already familiar. Still, the depth and breadth of data offered by these resources warrant their inclusion in this review.

Location-based intelligence firm Maponics sells traditional geospatial data and market segmentation datasets. It purports to have "the largest database ever compiled for neighborhood boundaries" (Maponics, n.d.). After acquiring Urban Mapping's data in 2015, Maponics' neighborhood repository now totals nearly 200,000 polygons from 68 countries, including the United States. Feedback from realtors — which the company calls "expert sourcing" — along with traditional data sources serve as the basis for these boundaries, which are used by industries ranging from real estate, local search and direct marketing to social media and mobile apps (Schutzberg, 2008). Clients include eHarmony, Redfin, and YikYak. Maponics estimates that 95% of all social media users

encounter its data (Maponics, n.d.). Its parent company, Pitney Bowes, also sells neighborhood boundaries and segmentation datasets.

Esri Tapestry targets clients in both private industry and the scientific research community (Esri, 2017). Tapestry, which requires a paid subscription, can subdivide locations into units commonly used in business — Congressional Districts, Core Based Statistical Areas, or Designated Market Areas — as well as census units ranging from states to block groups (Esri, 2017). The platform's instructional materials use the terms "block group" and "neighborhood" interchangeably.

Developed through a combination of cluster analysis and data mining, Tapestry has three classification schemes: Urbanization Groups, Life Modes, and Segments (Esri, 2017). The six Urbanization Groups are based on locales: Principal Urban Centers, Urban Periphery, Metro Cities, Suburban Periphery, Semirural, and Rural. The 14 Life Modes classify groups based on shared experience (e.g., generational cohort) or an essential demographic trait (e.g., wealth). These Life Mode groups can be subdivided into 67 Segments. For example, the Life Mode group called "Ethnic Enclaves" consists of Hispanic people who immigrated to the U.S. One of the group's six segments is "Southwestern Families," who are described as having a median age of 33.8 years, a median income of $27,000, an unemployment rate of 12% (Esri Demographics, 2014). While 30% of people in this segment are college-educated, 30% lack a high school diploma. They are described as budget-conscious and inclined to base consumer decisions on how a product might improve or organize their life. In addition to these demographic and consumer data, Tapestry can be used in tandem with Esri's Community

Analyst software, which contains crime statistics, health data, and additional census variables. Using the company's "Neighborhood toolset" available in its ArcMap software, researchers can generate their own novel neighborhood boundaries (Esri, n.d.).

Citizen collaboration

The widespread availability of high-quality open-source data and the popularity of crowd-sourcing reflect the online community's emphasis on transparency and cooperation. Mappers have capitalized on these trends — finding novel ways to combine publicly available data from unexpected and traditional sources and tasking online contributors with further expanding upon and/or refining the results (Crooks et al., 2015; Padmanabhan et al., 2014). On most projects of this nature, anyone willing to follow the established protocol, from seasoned GIS researchers to hobbyists to first-time mappers, is welcome to participate. Various monikers have been used to describe this phenomenon, including "collaborative mapping" and "participatory GIS," and the resulting data are often referred to as "volunteered geographic information" (VGI) (Goodchild, 2007).

One of the most extensive collaborative GIS projects is OpenStreetMap. Founded in 2004, OpenStreetMap stresses the value of local knowledge (OpenStreetMap, n.d.). The website accepts original data collected via manual surveying, GPS tracking, and methods that leverage smartphone technology and thus might be more accessible to novice mappers (e.g., audio-mapping, photo-mapping, etc.). Contributors can also import existing public data, combine datasets, or refine others' work. Freely available under a Creative Commons license, OpenStreetMap data are used by the general public, researchers, and also commercial organizations, including Apple, Craigslist, Flickr,

61

Foursquare, Twitter, and The Wall Street Journal. Information is available for every region of the globe, including neighborhood centroids for some areas of the United States. Completeness and level of granularity for a given location depend on contributor input. According to a 2013 report, half of the 10 most densely mapped locations on OpenStreetMap were in Cameroon, three were in France and the others were in Martinique and Brazil (OpenStreetMap, n.d.). OpenStreetMap surpassed > 3 million registered in 2016, but just 700,000 have ever contributed to the site and only 1.5% of users contributed at least one edit in 2015 (OpenStreetMap, n.d.).

Social media has also helped spur new spatial analysis methods. Posting photos online is a national pastime, and many people use keyword "tags" to help ensure that friends notice their posts. Some websites "geotag" these keywords with latitude and longitude coordinates. Quattroshapes, "the global polygon gazetteer," uses a larger pool of information, combining geotagged Flickr photos, FourSquare check-ins, Natural Earth data, census data and other publicly accessible sources to generate shapefiles, which it offers for free under a Creative Commons license (Quattroshapes, n.d.). Another gazetteer, Who's on First, further expands upon this method ("Who's On First," n.d.). As an initial step, information from Quattroshapes, Natural Earth, and other sources is aggregated. Then, for any given point, Who's on First displays the degree of concordance between the aforementioned data sources. Site visitors can further tweak the boundaries.

Who's on First also uses data from a now defunct website called Zetashapes, which billed itself as "an experiment in crowd-sourced U.S. neighborhood polygons." While the Zetashapes website has been disabled, the data are still available through Github.

Zetashapes relied on geotagged Flickr photos and 2010 Census data to construct its boundaries. Contributors edited the free, publicly available shapefiles and uploaded the modified versions for others' use.

On the opposite end of the spectrum, some small-scale projects can still offer useful data or methodological ideas for researchers. The People Organizing Place Neighborhood Map project focuses solely on the city of Dallas, Texas, yet its approach is quite ambitious (bc workshop, n.d.). Founded in 2011 by the nonprofit Building Community Workshop (bc workshop, n.d.), the project's first map was based on archival information gathered from planning initiatives, homeowners associations, crime watch groups, and other publicly available sources (bc workshop, 2015). Organizers then took the map on a "road show," gathering community input at festivals, town hall meetings, etc., and using that qualitative data to add, subtract, refine, and rename neighborhoods. A series of community initiatives followed, including recording neighborhood stories, converting unused spaces into art projects and mapping city council agenda items to spur citizen involvement. Know Your Neighborhood is an online directory of neighborhood-specific resources while Draw Your Neighborhood allows users to add their opinions to the ever-evolving Dallas map (bc workshop, 2015).

One might assume this project is of little relevance to researchers based outside Texas. On the contrary, it may hint at a feasible, scalable way for public health researchers to use a mixed methods approach when identifying neighborhood boundaries. Instead of all qualitative data collection taking place in person, one could move a large proportion — or even all of it — onto a website.

63

Table 3: Alternative sources of neighborhood centroids or polygons

Company	Geographic Coverage	Output formats	Methods	Access Fee	Other information
Local search					
1. DineHere	Neighborhood boundaries hundreds of U.S. cities	SHP	Not available	Free	
2. Google Maps	>100,000 places across the world, including neighborhood boundaries for some U.S. cities	JavaScript, VML	Not available	< $2,500	Free online course on how to use Google Maps APIs available via Udacity [a]
3. Microsoft Bing Maps	Global geospatial data; U.S. neighborhood centroids	GEORSS, JavaScript	Not available		
4. Yelp!	Neighborhood boundaries for many midsize/large U.S. and international cities; academic dataset available upon request	API: Serialized PHP, Serialized Python; Academic dataset: JSON	Not available	Free	Dataset Challenge open to students; Yelp! website links to publications based on academic dataset [b]
Real estate					
1. Home Junction	Boundaries for U.S. neighborhoods, subdivisions, districts, school attendance zones, property parcels, buildings, MLS, ZIP codes, other municipal areas; demographics, market trends, other variables	JSON, XML, EWKT; delivered in bulk or API	Mixed methods, including realtor feedback and data from Google and census	Varies; "courtesy use" for colleges, schools, government agencies under certain conditions	Unified dataset of multiple boundaries; company will fulfill custom data requests for additional fee; courtesy requests via API fulfilled more quickly than those bulk courtesy requests
2. Zillow	Boundaries for 7,000 U.S. neighborhoods (in 41 states and Washington, D.C.)	SHP; API [c], REST, XML		Shapefiles are free; daily limit on free API usage	API returns demographic and housing market data at neighborhood level

Company	Geographic Coverage	Output formats	Methods	Access Fee	Other information
Location-based intelligence					
1. Azavea	United States community/neighborhood boundaries	Unknown	Based on market segmentation, demographics, other factors as requested	Varies; some data available for free	Data analysis, software development available; non-profits can apply to host a summer intern [d]
2. DistrictBuilder	U.S. community/neighborhood boundaries drawn by user	SHP	Data sources: administrative boundaries, GoogleMaps, Esri ArcGIS Online, OpenStreetMap, Bing maps	Data is free and required software is open-source, but hosting / installation costs vary	Contiguity, compactness, population statistics calculated as user draws boundaries.
3. Claritas Segmentation & Market Solutions	U.S. neighborhood boundaries and segmentation data	Unknown	Not available	Varies	Recently purchased by Carlyle from Nielsen
4. Esri Tapestry/ Spatial Anaylst	U.S. neighborhood boundaries, segmentation data, ability to make novel boundaries using	CSV, DBF, GDB, SDC, SHP, TXT, XLS	Mixed methods, including cluster analysis and data mining; user input	Varies	Can be combined with Esri Community Analyst (crime statistics, health data, additional census variables)
5. Pitney Bowes	Global polygons and centroids at various scales, including U.S. neighborhoods; segmentation data; other geospatial analysis options	ASCII, BMP, DBF, EMF, GeoTiff, TAB, JPG, TIFF, WMF, others	Data from government, home owners' associations, property records, "customer input"	Varies	Allows import of SHP and other common geospatial and relational database formats; Pitney Bowes also owns Maponics
6. Maponics	~200,000 neighborhood boundaries across 68 countries, including U.S.; Nielsen PRIZM lifestyle segmentation, MicroBuild household segmentation, crime, walkability,	KML, MySQL, PostGIS, SHP, TAB, WKT	Data from government, home owners' associations, public service groups, property	Varies	Updated quarterly; other boundary types (e.g., social, metro) available); Maponics is owned by Pitney Bowes

Company	Geographic Coverage	Output formats	Methods	Access Fee	Other information
	census demographics, other factors upon request		records, "customer input"		
7. StatSilk	Software platform that includes global geospatial data of varying scales, including some neighborhood boundaries; additional data on other topics including health, economics, crime, and environment	CSV, JPG, PDF, SVG, XLS	Wide range of open data sources; users can import additional online data sources or their own data	Varies; discount on certain packages for governments, schools, non-profits	
Citizen collaboration					
1. MapIt	Global administrative boundaries and post codes; more extensive information for United Kingdom	GeoJSON files	Not available	Free for non-profits	
2. Neighborland	Select U.S. neighborhood centroids	JSON	Not available	Free	
3. OpenStreet Map	Global data at various spatial scales, including U.S. neighborhood centroids	GeoJSON, OSM, SHP, SVG	Combination of Bing, MapQuest, various other private and public data, and public's ongoing input via website	Free	
4. POP Neighborhood Map	Neighborhoods and "super neighborhoods" (i.e., macro areas) of Dallas, Texas	GeoJSON	Base map that included boundaries from a variety of municipal sources, refined by "citizen experts" in series of public events. Maps still evolving based on public's input via website.	Free	Site includes details on land use and development history, directory of neighborhood groups, and personal stories from local residents

66

Company	Geographic Coverage	Output formats	Methods	Access Fee	Other information
5. Quattroshapes	Some data from each of the seven continents; scale and level of detail vary from region to region	SHP	"Global polygon gazetteer" uses FourSquare check-ins, Flickr geotags, Natural Earth data, government data to determine "dominant place ID."	Free	
6. The Neighborhood Project	San Francisco neighborhood boundaries	JSON	Combination of OpenStreetMap, Craigslist housing posts and public's ongoing online input	Free	
7. Who's On First	Boundaries (of varying scales) from across the world	GeoJSON	Gazetteer combines data from Natural Earth, Quattroshapes, zetashapes and other sources; determines concordance between those sources; the public can further refine neighborhood boundaries.	Free	Methods and rationale documented in lengthy detail on website
8. Wikimapia	>200,000 neighborhood boundaries from around the world	JSON, KML, XML	Base map derived from GoogleMaps; crowd-sourced refinement of boundaries is ongoing	Free	
9. Zolk	Chicago, Illinois	Google Earth file or KML	GIS data from City of Chicago	Free	

URLs for Table 3 Resources

Local Search

1. DineHere: http://dinehere.us/neighborhoods.html

2. Google Maps APIs: https://developers.google.com/maps/get-started/

3. Microsoft Bing Maps: https://www.microsoft.com/maps/choose-your-bing-maps-API.aspx

4. Yelp!: https://www.yelp.com/developers/documentation/v2/neighborhood_list

Real Estate

1. Home Junction: https://www.homejunction.com/boundaries/#neighborhoods

2. Zillow: www.zillowgroup.com/news/7000-neighborhood-boundary-files-in-shapefile-format/

Location Based Intelligence

1. Azavea: https://www.azavea.com/about/

2. DistrictBuilder: http://www.districtbuilder.org/

3. Claritas: https://segmentationsolutions.nielsen.com/mybestsegments/

4. Esri: http://doc.arcgis.com/en/esri-demographics/data/tapestry-segmentation.htm; http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/an-overview-of-the-neighborhood-tools.htm

5. Pitney Bowes: http://www.pitneybowes.com/us/data/boundary-data/neighborhood-boundaries.html

6. Maponics: http://www.maponics.com/products/communities/neighborhood-boundaries

7. StatSilk: https://www.statsilk.com/software/

Citizen Collaboration

1. MapIt: http://global.mapit.mysociety.org/

2. Neighborland: https://neighborland.com/docs

3. OpenStreetMap: https://www.openstreetmap.org/about

4. Know Your Neighborhood: http://peopleorganizingplace.com/know/

5. Quattroshapes: http://quattroshapes.com/

6. The Neighborhood Project: https://hood.theory.org/

7. Who's On First: https://whosonfirst.mapzen.com/

8. Wikimapia: http://wikimapia.org/api

9. Zolk: http://chicagomap.zolk.com/about.html

Supplemental links referenced in table

a. Yelp! Dataset Challenge: https://www.yelp.com/dataset/challenge

b. Udacity course: https://www.udacity.com/course/google-maps-apis--ud864

c. Zillow API: https://www.zillow.com/howto/api/APIOverview.htm

d. Azavea summer internship: http://www.summerofmaps.com/

Table 4: Typical* advantages of tools designed by particular types of private sector organizations

	Local Search	Real Estate	Location-based Intelligence	Citizen Collaboration
Cost: Most resources are free	X	X		X
Longevity: Resources updated regularly and likely to be available for foreseeable future		X	X	
Transparency: Clear documentation of data sources, statistical analysis, sponsors, etc.				X
Coverage: Data for wide geographic area available, with similar level of detail throughout		X	X	
Enrichment: Boundaries often augmented by information on demographics, schools, crime statistics, neighborhood quality measures		X	X	
Customization: Group accepts requests for additional data or combinations of datasets			X	
Freedom: Few if any limitations on how researchers use tools or combine them with other resources			X	X

*Intended to serve as a rule of thumb for researchers: There are exceptions to this pattern

ii.   Aim 2: Sensitivity analysis

Aim 2 of this dissertation was to conduct a sensitivity analysis to compare the strength of association between HDP prevalence and neighborhood deprivation at six different units of analysis — block group, census tract, ZIP Code Tabulation Area (ZCTAs), and three types of natural neighborhood formed by aggregating block groups and census tracts at three different scales — using ordinary least squares (OLS) regression and geographically weighted regression (GWR). By comparing the R^2 values for the six geographic units, I determined the unit of analysis to use for Aim 3.

The effects of the MAUP are evident in the varying patterns of HDP prevalence in Miami-Dade County, Florida in 2008-2012 (Figure 2) and neighborhood deprivation (Figure 3) at different scales and aggregations. Median HDP prevalence (Table 5) differed significantly between types of areal unit. ZCTAs had the highest median HDP at 5.1%, and block groups had the lowest at 3.3%. The standard deviation was even more inconsistent. Block groups (SD=6.5%) had the largest overall HDP standard deviation of any areal unit. Among NNs, HDP prevalence varied the most within CT large NNs (SD=5.3%). When NDI variables were examined, median values of all eight factors differed significantly between areal units.

Principal component factor loadings (Table 6) ranged from 0.33 to 0.94. While no statistical test was used to compare the loadings, differences between areal units were evident: PCA values for block groups and the two types of NNs created from those block groups appeared to be more inconsistent than those for other census tracts, ZCTAs, and CT large NNs. However, the range of Cronbach's alpha values was relatively narrow,

70

with a low of 0.88 (BG small NNs) and a high of 0.94 (census tracts), suggesting high internal validity for the indices generated for each areal unit.

Based on R^2 values (Table 7), the proportion of variance in HDP prevalence explained by neighborhood deprivation was highest at the level of CT large NN (OLS R^2=0.14; GWR R^2=0. 27), followed by ZCTAs (OLS R^2=0.09; GWR R^2=0.14). The map of under- and over-prediction (Figure 3) provides a visual representation of the extent to which the CT large NN model out-performed the other models — particularly block groups (OLS R^2=0.01; GWR^2=0.02). The BG large NN model was not significant in OLS regression and thus was not tested with GWR. In the CT large NN model, a one-unit increase in deprivation was associated with a 1.94% increase in HDP prevalence.

In addition to the HDP and NDI maps, R^2 values were used to evaluate the MAUP. In terms of scale effects, the smallest areal units – block groups – produced the least acceptable model. The ZCTA model fit the data substantially better than the census tract model in OLS. However when spatial variation was accounted for in GWR analysis, the census tract model fit produced a larger R^2 value. In terms of the zoning effect, CT large NNs and BG large NNs — the two types of NN comparable to ZCTAs — performed as well or better than the census unit whose scale they approximated. However, BG small NNs did not explain as much HDP variance as their comparable unit, census tracts.

Based on the R^2 values from the GWR analysis, the CT large NN model fit the data best. Thus, this area-level unit of analysis was used for the multilevel regression in Aim 3: Neighborhood deprivation, aggregated by CT large NNs and individual-level characteristics were examined as predictors of individual-level HDP.

Figure 2: Prevalence of hypertensive disorders of pregnancy in Miami-Dade County, Florida, 2008–2012, at six areal units
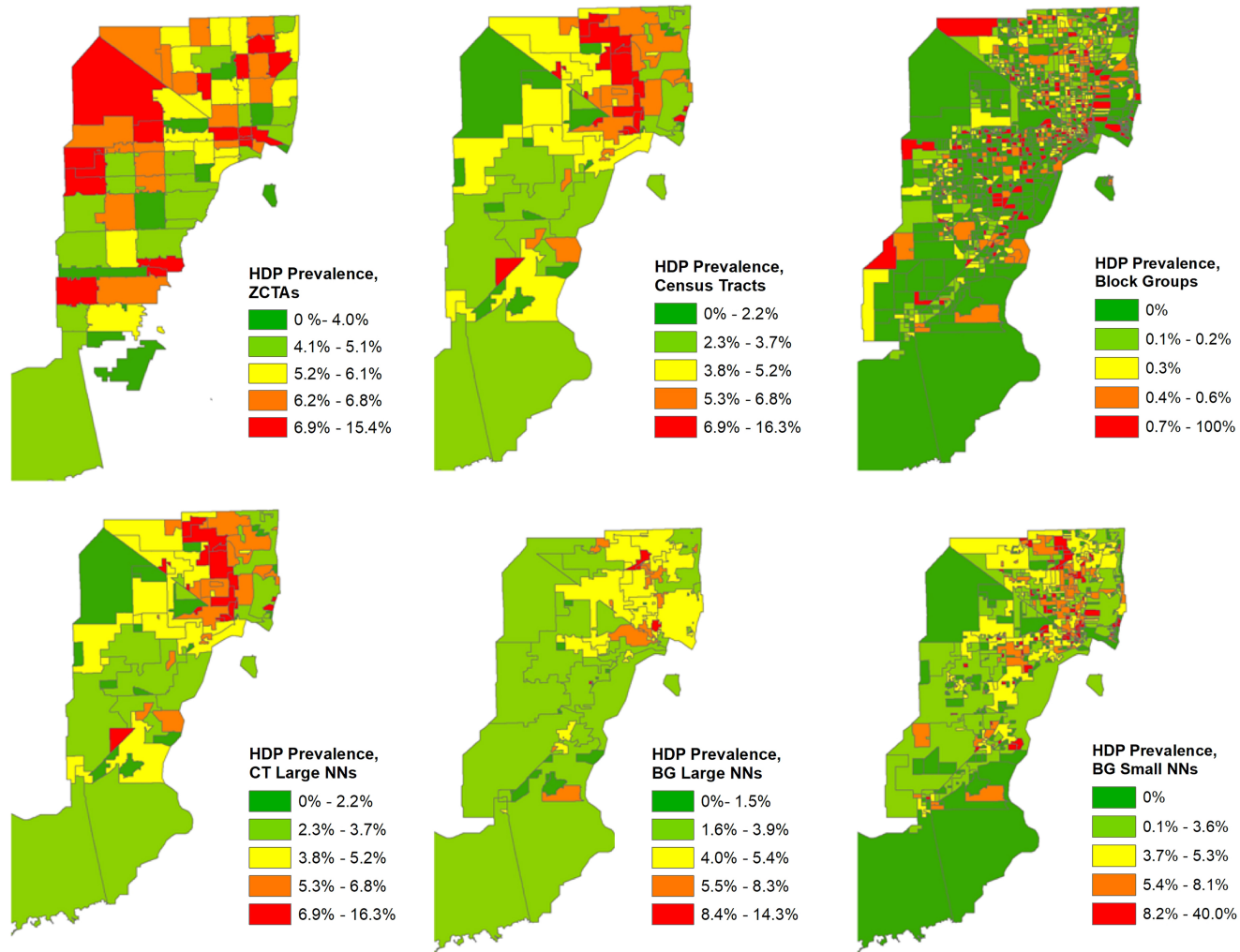
ZCTA – ZIP code tabulation area; NN – natural neighborhood; CT – census tract; BG – block group

Figure 3: Neighborhood deprivation quintiles in Miami-Dade County, Florida, 2008–2012, at six areal units
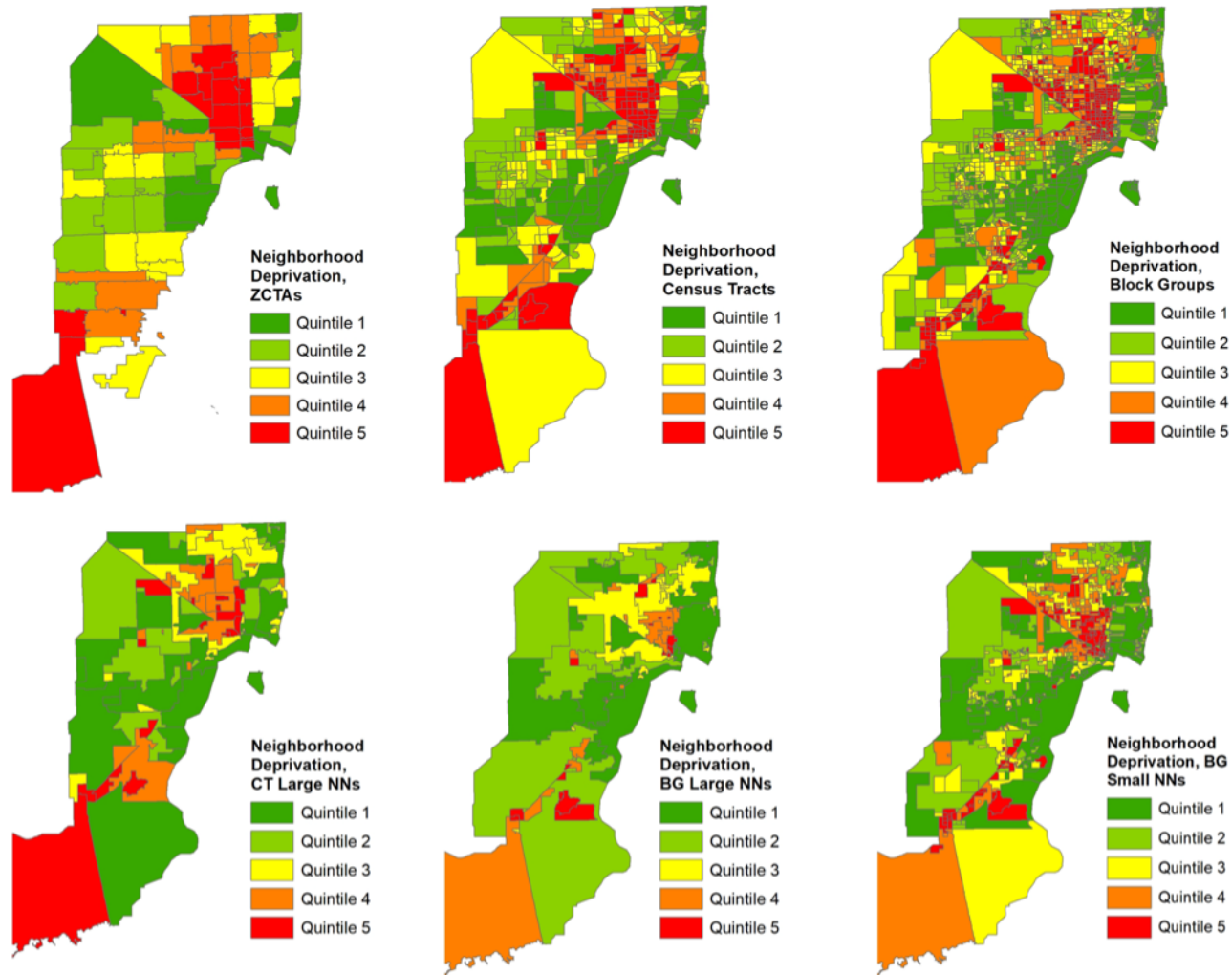
Neighborhood deprivation measured by eight-item index; ZCTA – ZIP code tabulation area; NN – natural neighborhood; CT – census tract; BG – block group

Table 5: Kruskal-Wallis H test and Levine nonparametric test comparisons of neighborhood variables measured at six different areal units in a sample of women who gave birth in Miami-Dade County Florida, 2008–2012

	Median (Standard Deviation)					
	Block Groups (n=1,560)	Census Tracts (n=507)	ZIP Code Tabulation Areas (n=78)	Small NNs from Block Groups (n=507)	Large NNs from Block Groups (n=78)	Large NNs from Census Tracts (n=78)
Prevalence of hypertensive disorders of pregnancy***	3.3% (6.5%)	4.9% (2.7%)	4.9% (1.4%)	4.9% (3.9%)	5.0% (1.8%)	5.1% (5.3%)
Neighborhood deprivation						
Household poverty [1]*	16.7% (15.7%)	17.1% (12.9%)	17.5% (10.5%)	22.5% (17.9%)	27.3% (20.9%)	24.3% (16.7%)
Female-headed households [2]*	10.0% (12.9%)	12.1% (9.1%)	11.7% (6.2%)	11.9% (16.2%)	13.9% (21.3%)	13.7% (12.7%)
Low-income residents [3]**	33.3% (20.6%)	40.9% (18.9%)	32.0% (15.0%)	40.4% (21.3%)	47.8% (21.7%)	52.1% (21.7%)
Public assistance [4]*	19.7% (17.5%)	22.4% (14.3%)	20.0% (12.7%)	24.9% (19.0%)	33.4% (20.3%)	26.4% (14.3%)
Blue-collar employment [5]*	76.5% (22.6%)	75.7% (18.9%)	72.3% (17.7%)	80.7% (21.8%)	84.2% (20.4%)	81.2% (18.0%)
Unemployment [6]***	6.5% (5.6%)	6.9% (3.8%)	10.8% (4.6%)	7.5% (6.9%)	8.4% (8.4%)	8.1% (4.7%)
Under-educated adults [7]*	18.1% (14.6%)	20.0% (12.4%)	32.0% (15.0%)	21.8% (15.6%)	25.8% (17.1%)	23.5% (13.4%)
Crowded households [8]*	3.4% (6.7%)	4.6% (5.7%)	4.9% (2.9%)	4.7% (8.5%)	7.1% (12.8%)	6.0% (10.6%)

*p < 0.05; **p < 0.01; ***p < 0.001

[1] Proportion of households below the poverty line

[2] Proportion of female-headed households with dependent children

[3] Proportion of residents earning < $30,000 annually

[4] Proportion of residents receiving public assistance

[5] Proportion of adult males who do not hold managerial/professional jobs

[6] Proportion of unemployed residents

[7] Proportion of residents ≥ age 25 without a high school diploma / GED

[8] Proportion of households with > 1 person per room

Table 6: Principal component loadings and Cronbach's alpha, Neighborhood Deprivation Index at six different areal units

	Block Groups (n=1,560)	Census Tracts (n=507)	ZIP Code Tabulation Areas (n=78)	Small NNs from Block Groups (n=507)	Large NNs from Block Groups (n=78)	Large NNs from Census Tracts (n=78)
Neighborhood deprivation						
Household poverty	0.85	0.90	0.94	0.63	0.92	0.63
Female-headed households	0.58	0.69	0.78	0.76	0.69	0.85
Low-income residents	0.86	0.89	0.91	0.70	0.89	0.77
Public assistance	0.86	0.88	0.93	0.79	0.82	0.87
Blue-collar employment	0.73	0.82	0.89	0.73	0.81	0.88
Crowded households	0.55	0.61	0.82	0.33	0.45	0.80
Unemployment	0.50	0.64	0.83	0.73	0.48	0.73
Under-educated adults	0.77	0.85	0.87	0.87	0.59	0.93
Cronbach's alpha	0.90	0.95	0.91	0.88	0.89	0.94

Table 7: Ordinary least squares and geographically weighted regression analysis of prevalence of hypertensive disorders of pregnancy and neighborhood deprivation measured at six different areal units in a sample of women who gave birth in Miami-Dade County Florida, 2008–2012

	Ordinary Least Squares Regression		Geographically Weighted Regression
Unit of Analysis	**β (95% CI)**	**R^2**	**R^2**
Block Groups (n=1,560)	0.28** (0.07, 0.49)	0.01	0.02
Census Tracts (n=507)	0.47*** (0.24, 0.70)	0.03	0.07
ZIP Code Tabulation Areas (n=78)	0.35*** (0.18, 0.53)	0.18	0.13
Natural neighborhoods (NNs)			
Small NNs from Block Groups (n=507)	0.45** (0.11, 0.79)	0.01	0.05
Large NNs from Block Groups (n=78)	0.08 (-0.33, 0.50)	0.00	N/A [a]
Large NNs from Census Tracts (n=78)	1.94** (0.82, 3.06)	0.14	0.27

a: GWR is not performed if OLS regression does not yield significant results.

*p < 0.05, **p < 0.01, ***p < 0.001

Figure 4: Clusters of under- and over-prediction in geographically weighted regression analysis at five areal units in models of prevalence of hypertensive disorders of pregnancy and neighborhood deprivation in Miami-Dade County, Florida, 2008–2012
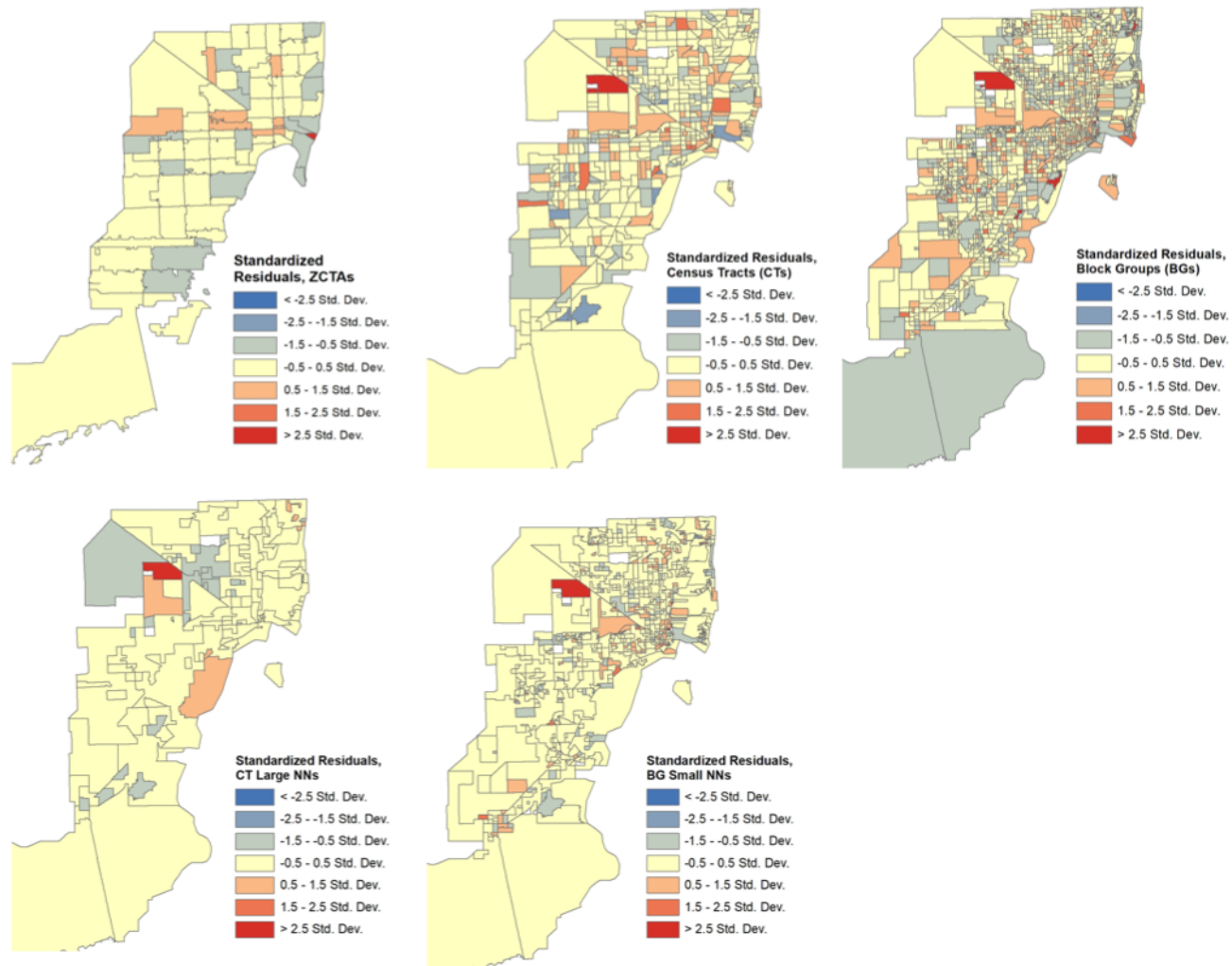
iii.   Aim 3: Multilevel Analysis

Descriptive statistics

Aim 3 of this dissertation was to quantify the relationship between neighborhood deprivation and individual-level HDP status in Miami-Dade County, Florida using multilevel logistic regression. I also examined cross-level interactions between neighborhood deprivation and two measures of individual-level deprivation — low educational attainment and lack of access to health care. Individual-level descriptive statistics are shown in Table 8. Compared with Hispanic and non-Hispanic White women, significantly larger proportions of non-Hispanic Black women had less than a high school education ($\geq$ age 18: 16.6%; < age 18: 3.7%), were Medicaid recipients (67.0%) and had fewer prenatal care visits (Mean=10.6, SD=3.5). Larger proportions of Hispanic women were uninsured (16.9%) and had pre-existing diabetes (1.1%).

The macro unit of analysis was large census tract natural neighborhoods (large CT NNs), created by aggregating census tracts to the scale of ZCTAs based on an eight-item neighborhood deprivation index. In neighborhoods with higher deprivation, the overall proportion of women $\geq$ age 18 years with low educational attainment was greater, peaking at 15.8% in the high deprivation tertile (Figure 5). The difference was most prominent within the Hispanic subgroup. In neighborhoods with high deprivation, 17.1% of adult Hispanic mothers had low educational attainment, compared with 11.1% in low deprivation neighborhoods — a difference of 35.1%. Among women < age 18 years, the proportion of low educational attainment differed significantly between neighborhood deprivation tertiles for the overall sample and all racial/ethnic subgroups (Figure 6). However, among women in the Other race/ethnicity group, the proportion of women

< age 18 years with low educational attainment was greatest not in the high deprivation tertile (1.0%) but instead in the moderate deprivation tertile (1.2%).

The proportion of uninsured women was largest in the high deprivation tertile, both in the overall sample and within each racial/ethnic subgroup (Figure 7). Among Hispanic mothers, 20.5% of those in high deprivation neighborhoods were uninsured, compared with 16.9% in moderate and 13.7% in low deprivation neighborhoods. The proportion of Medicaid recipicients differed significantly as well (Figure 8). The pattern was most marked among non-Hispanic Blacks. In high deprivation neighborhoods, 69.9% of non-Hispanic Black mothers were Medicaid recipients, compared with 64.1% in low deprivation neighborhoods a difference of 5.8 percentage points.

Adult women with low-educational attainment accounted for a larger proportion of HDP cases in high deprivation neighborhoods (16.4%) compared with moderate (13.4%) or low deprivation (12.2%) neighborhoods (Figure 9). Hypertensive disorders of pregnancy (HDP) cases were also classified by insurance status and neighborhood deprivation (Figure 10). Medicaid recipients accounted for the majority of HDP cases overall (46.8%) as well as in each neighborhood subcategory. Uninsured women accounted for nearly one-third more HDP cases in high deprivation neighborhoods (23.4%) compared with low deprivation neighborhoods (15.9%).

Single-level logistic regression analysis

While controlling for maternal age and number of prenatal care visits, HDP was significantly associated with all 10 categorical variables examined in the single-level logistic regression analysis (Table 9), so all were retained for the multilevel logistic

regression analysis. In the full single-level model, HDP was most strongly linked to pre-existing diabetes (aOR=6.09, CI: 5.12, 7.25) and Category III obesity (aOR=5.30, CI: 4.76, 5.92). Non-Hispanic Black race/ethnicity (aOR=1.62, CI: 1.44, 1.82) and lack of health insurance (aOR=1.63, CI: 1.51, 1.76) were the sociodemographic factors with the largest measures of association.

Multilevel logistic regression analysis

An unconditional, or null, model consisting of only a randomly varying intercept was constructed to determine whether odds of HDP varied between neighborhoods (Table 10). Had this model not reached statistical significance, multilevel modeling would not have proceeded. The Level 2 variance with no predictors in the model was 0.015 (CI=0.007, 0.030), which translates to an interclass correlation (ICC) of 0.005 and a median odds ratio of 1.23. This suggests that: 1.) 0.5% of the total variation in HDP prevalence is due to differences between neighborhoods and 2.) the median case residual heterogeneity is 1.23. In common language, this means that if the median case in the sample lived a neighborhood with higher probability of HDP, her likelihood of having HDP would be 1.23 times higher than it was in her original, less-deprived neighborhood. While the proportion of HDP variance between neighborhoods is low, similar studies of neighborhood factors and non-communicable diseases have found Level-2 variability between 1% and 3% (Ford & Browning, 2011; Merlo, Wagner, Ghith, & Leckie, 2016; Mujahid et al., 2008; Ross, Tremblay, & Graham, 2004; Slopen, Non, Williams, Roberts, & Albert, 2014).

In the unadjusted multilevel model (Table 10), women living in areas of both moderate (OR=1.20, CI: 1.09, 1.32) and high (OR=1.15, CI: 1.07, 1.24) neighborhood deprivation had higher odds of HDP than women in living in areas of low neighborhood deprivation. This relationship remained statistically significant even after all interaction terms and demographic covariates and number of prenatal care visits were added. In the final model, compared with women living in low deprivation neighborhoods, odds of HDP were 1.16 times as high among those in high deprivation neighborhoods and 1.13 times as high among those in moderate deprivation neighborhoods.

Individual-level indicators of socioeconomic deprivation were also linked to HDP (Table 8). Compared with mothers with private insurance, Medicaid recipients (aOR=1.12, CI: 1.05, 1.18) and those who paid out of pocket (aOR=1.69, CI: 1.56, 1.84) were more likely to report HDP. There was also a significant relationship between a lack of a high school diploma/GED and HDP among teenage mothers (aOR=1.34, CI: 1.10, 1.63) but not for mothers ≥ age 18 (aOR=1.05, CI: 0.99, 1.12). Compared with non-Hispanic White women, non-Hispanic Black women were about 58% more likely to have HDP. Membership in the Other race/ethnicity category was a protective factor (aOR=0.72, CI: 0.59, 0.89), while there was no association between HDP and Hispanic ethnicity in the final multilevel model. The cross-level interactions between neighborhood deprivation and insurance status and between neighborhood deprivation and educational atainment (not shown) did not reach statistical significance and were removed from the model.

Among modifiable individual-level factors, overweight/obesity was strongly associated with HDP. Compared with their underweight and normal weight counterparts, odds of HDP for overweight women were 2.58 times as high (CI: 1.98, 3.37). Depending on the category of severity, obesity was associated with nearly double to more than five times the odds of HDP. Excessive gestational weight gain (aOR=1.44, CI: 1.34, 1.54) and smoking (aOR=1.46, CI: 1.09, 1.96) were also associated with higher likelihood of HDP. Pre-existing diabetes (aOR=5.52, CI: 4.71, 6.48) and gestational diabetes (aOR=3.64, CI: 3.13, 3.99) were the non-modifiable risk factors with the largest measures of association.

Adding variables to the multilevel model in stepwise fashion allowed for preliminary investigations of potential pathways through which deprivation influences HDP risk. When modifiable factors such as gestational weight gain and smoking status were added in Model 3, measures of assocation for educational attainment and insurance increased. Conversely, the association between HDP and race/ethnicity diminished. With the introduction of non-modifiable medical factors including gestational diabetes and pre-existing diabetes in Model 4, the opposite pattern was observed for education, but all variables remained statistically significant.

Table 8: Results of ANOVA and Chi-square analyses: individual-level characteristics by race/ethnicity among women in Miami-Dade County, Florida who had a live birth, 2008–2012

	Total N = 121,421	Non-Hispanic Black (22.6%)	Hispanic (62.6%)	Non-Hispanic White (5.2%)	Other race / ethnicity (9.5%)
Outcome variable					
Hypertensive disorders of pregnancy***	5.4%	7.5%	4.9%	6.1%	2.7%
Sociodemographic factors					
Educational attainment***					
< HS diploma/GED, age <18	1.9%	3.7%	1.4%	1.0%	1.0%
< HS diploma/GED, age ≥18	13.1%	16.6%	13.8%	5.6%	4.4%
HS diploma/GED	85.0%	79.7%	84.8%	93.4%	94.6%
Insurance status***					
Self-pay	14.8%	12.8%	16.9%	7.1%	10.6%
Medicaid	47.7%	67.0%	45.8%	32.8%	22.4%
Private insurance/other	37.5%	20.2%	37.3%	60.8%	67.0%
Maternal age***					
Mean years	28.6 (6.3)	26.6 (6.5)	29.0 (6.1)	29.8 (6.1)	30.6 (5.8)
Modifiable factors					
BMI status***					
Obese III	1.6%	7.9%	3.2%	1.5%	0.0%
Obese II	2.7%	10.1%	6.7%	2.9%	0.1%
Obese I	7.2%	21.2%	20.0%	7.7%	0.2%
Overweight	16.4%	38.0%	49.5%	19.9%	0.7%
Normal / Underweight	62.8%	5.7%	10.5%	66.9%	98.7%
Gestational weight gain***					
Excessive	55.3%	51.6%	56.7%	82.3%	40.2%
Not excessive	44.7%	48.4%	43.3%	17.7%	59.8%
Smoked during pregnancy***					
Yes	0.6%	0.7%	0.5%	1.7%	1.0%
No	99.4%	99.3%	99.5%	98.3%	99.0%
Prenatal care***					
Mean visits	11.8 (3.4)	10.6 (3.5)	12.0 (3.3)	12.7 (3.6)	11.9 (3.4)
Non-modifiable factors					
Nulliparity***					
Yes	56.4%	61.4%	55.5%	55.1%	50.9%
No	43.6%	24.6%	44.5%	44.9%	49.1%
Multifetal gestation***					
Yes	3.3%	3.3%	3.1%	4.4%	4.0%
No	96.7%	96.7%	96.9%	95.6%	96.0%
Gestational diabetes***					
Yes	2.8%	2.1%	3.1%	3.2%	2.4%
No	97.2%	97.9%	96.9%	96.8%	97.6%
Pre-existing diabetes***					
Yes	0.5%	0.6%	1.1%	0.7%	0.3%
No	99.5%	99.4%	98.9%	99.3%	99.57%

*** *p<0.001*

Figure 5: Low educational attainment by race/ethnicity and level of neighborhood deprivation among women ≥18 years in Miami-Dade County, Florida who had a live birth, 2008–2012
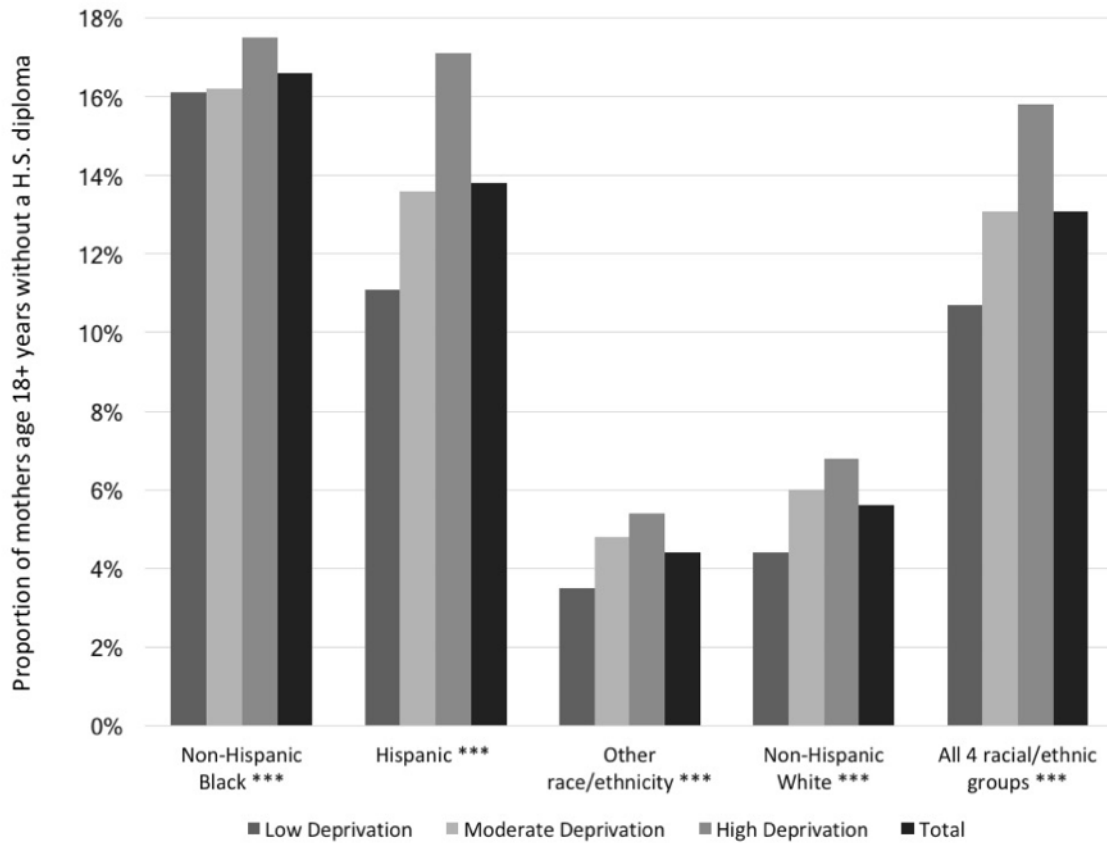
*** $p<0.001$, ** $p<0.01$, * $p<0.05$

Figure 6: Low educational attainment by race/ethnicity and level of neighborhood deprivation among women <18 years in Miami-Dade County, Florida who had a live birth, 2008–2012
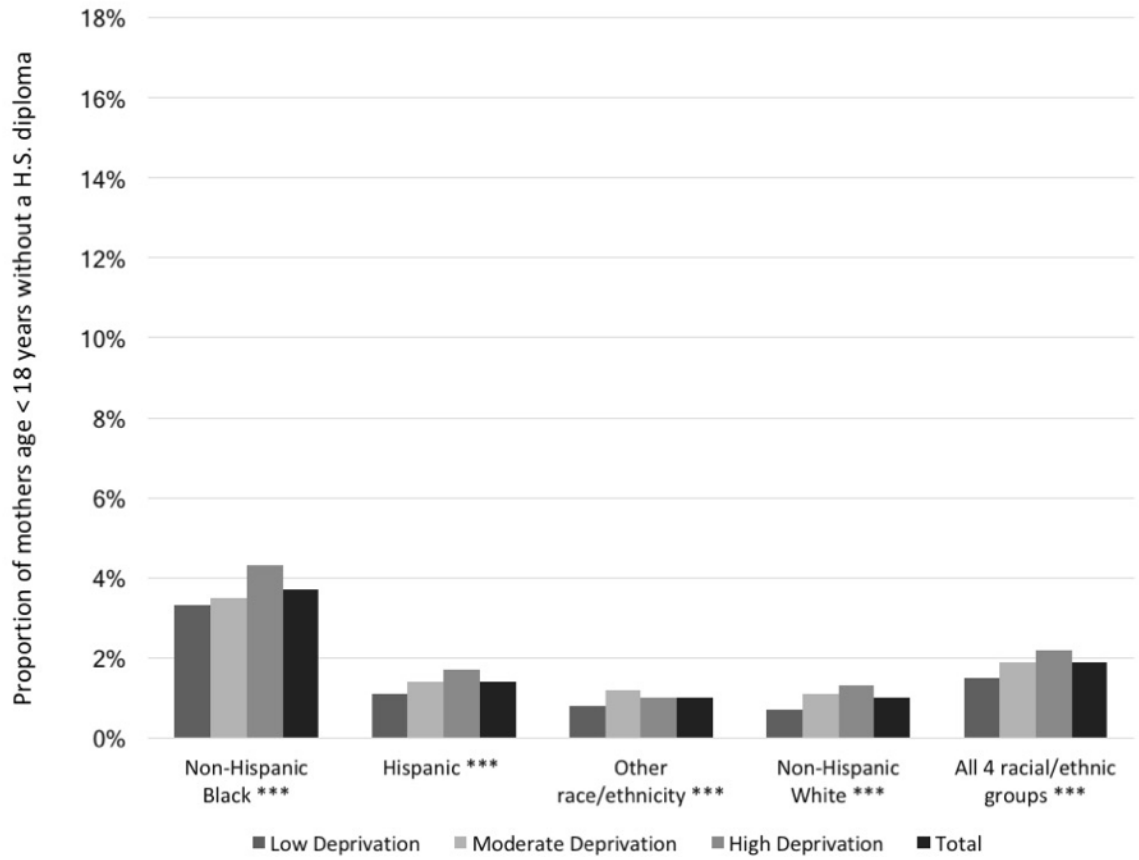
*** $p<0.001$, ** $p<0.01$, * $p<0.05$

Figure 7: Lack of health insurance by race/ethnicity and level of neighborhood deprivation among women in Miami-Dade County, Florida who had a live birth, 2008–2012
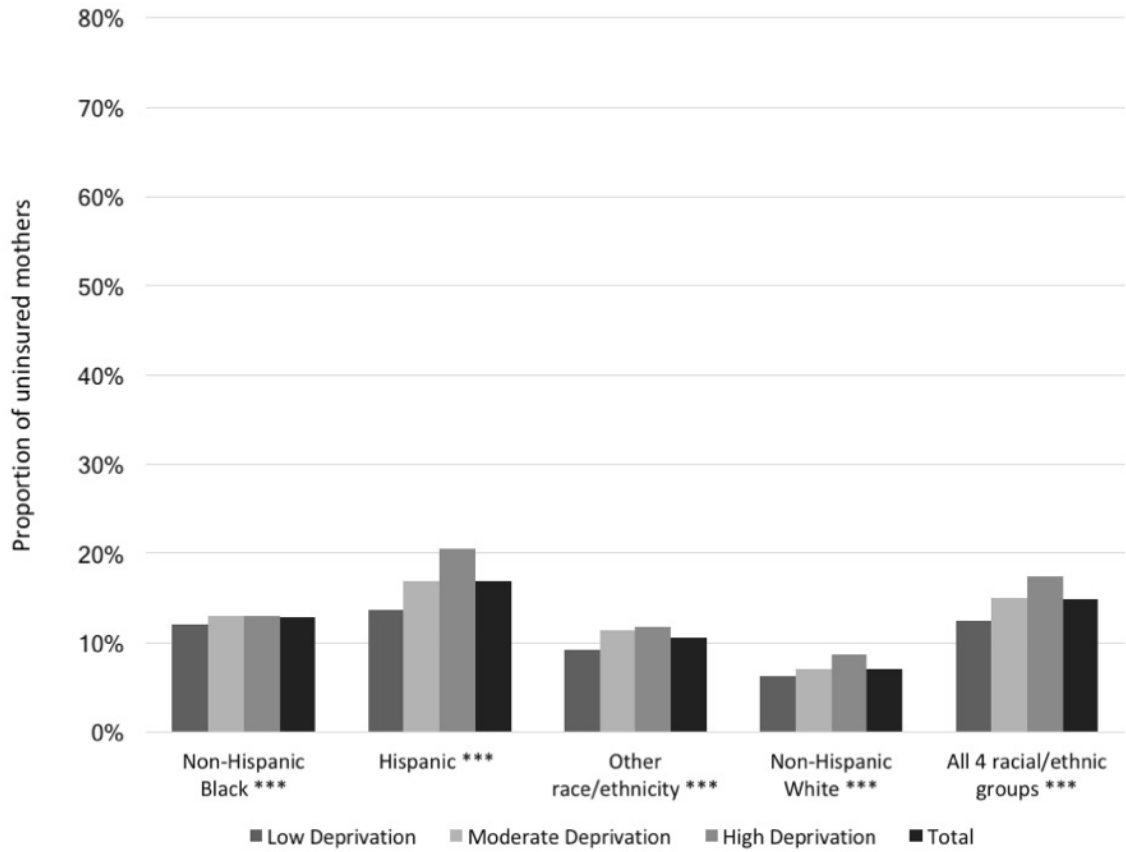
Figure 8: Proportion of Medicaid recipients by race/ethnicity among women in Miami-Dade County, Florida who had a live birth, 2008–2012
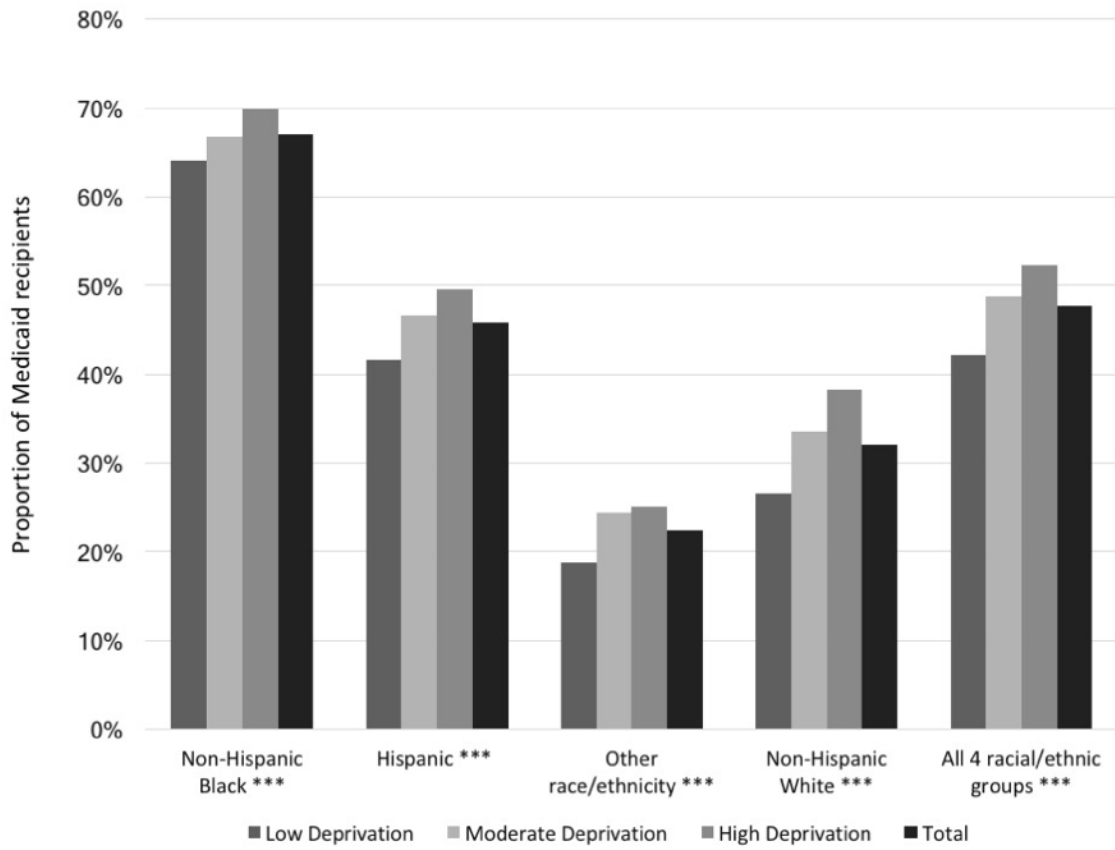
*** $p<0.001$, ** $p<0.01$, * $p<0.05$

Figure 9: Proportion of hypertensive disorders of pregnancy cases by educational attainment and neighborhood deprivation in Miami-Dade County, Florida, 2008–2012
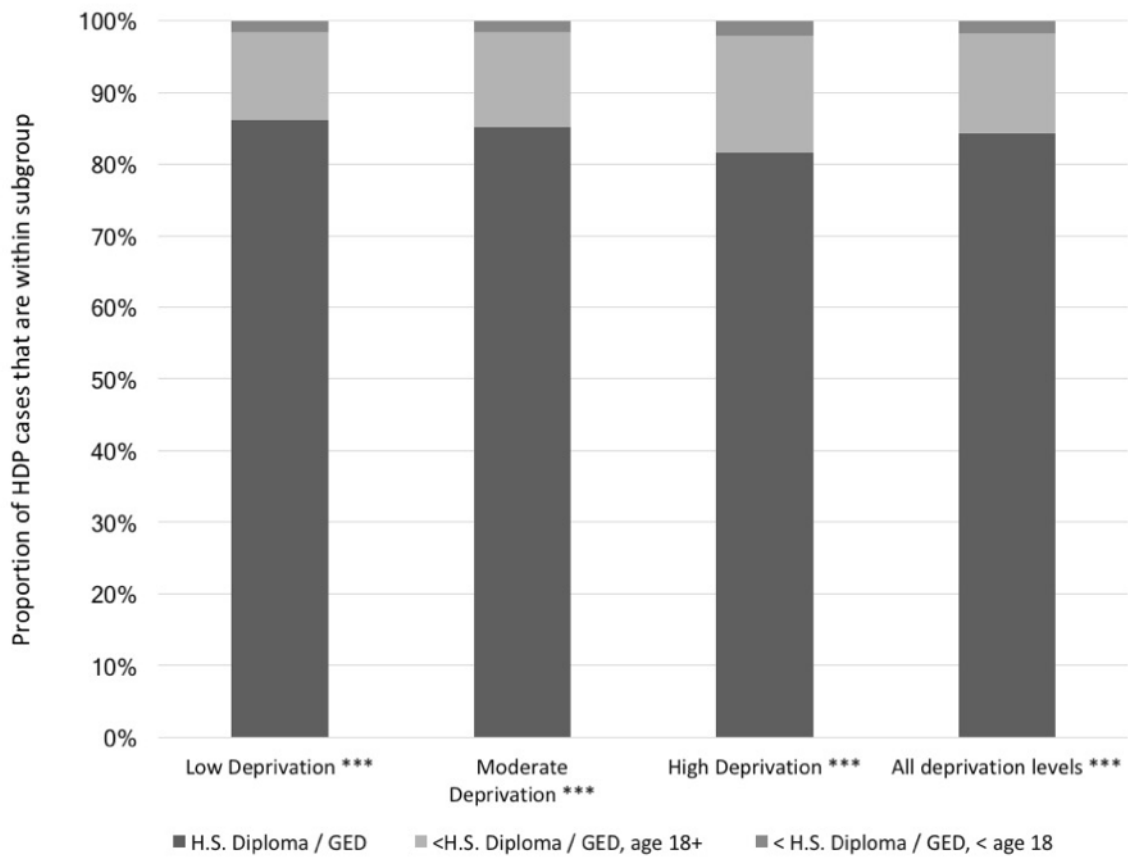
*** $p<0.001$, ** $p<0.01$, * $p<0.05$

Figure 10: Proportion of hypertensive disorders of pregnancy cases by insurance status and neighborhood deprivation in Miami-Dade County, Florida, 2008-2012
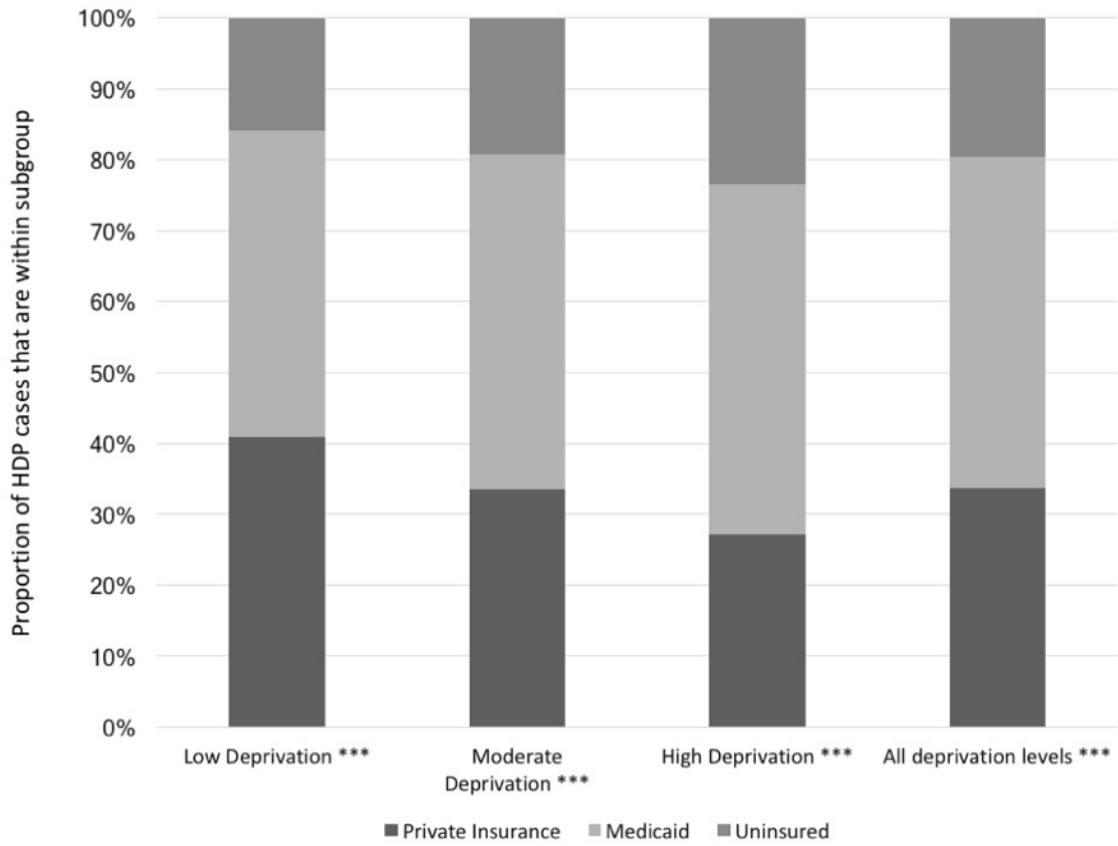
*** $p<0.001$, ** $p<0.01$, * $p<0.05$

Table 9: Results of single-level logistic regression: Unadjusted odds ratios (ORs) and adjusted odds ratios (aORs)[a] and 95% confidence intervals (CIs) for hypertensive disorders of pregnancy among women in Miami-Dade County, Florida who had a live birth, 2008–2012

	OR (95% CI) N = 121,421	aOR (95% CI) N = 121,421
Race/ethnicity		
Non-Hispanic Black	1.25 (1.12, 1.39)	1.62 (1.44, 1.82)
Hispanic	0.80 (0.72, 0.89)	1.11 (1.00, 1.25)
Other race / ethnicity	0.42 (0.36, 0.49)	0.86 (0.73, 1.01)
Non-Hispanic White	Ref	Ref
Educational attainment		
< High school diploma/GED, age <18	0.91 (0.75, 1.10)	0.97 (0.91, 1.03)
< High school diploma/GED, age ≥18	1.07 (1.00, 1.15)	0.92 (0.76, 1.12)
High school diploma/GED	Ref	Ref
Insurance status		
Self-pay	1.47 (1.37, 1.57)	1.63 (1.51, 1.76)
Medicaid	1.08 (1.02, 1.14)	0.97 (0.91, 1.03)
Private insurance/other	Ref	Ref
Modifiable characteristics		
Pre-pregnancy BMI status		
Obese III	6.36 (5.75, 7.03)	5.30 (4.76, 5.92)
Obese II	3.92 (3.57, 4.30)	3.36 (3.04, 3.71)
Obese I	2.79 (2.60, 3.00)	2.40 (2.22, 2.60)
Overweight	1.89 (1.77, 2.01)	1.65 (1.54, 1.77)
Normal weight / underweight	Ref	Ref
Excessive gestational weight gain		
Yes	1.62 (1.54, 1.70)	1.38 (1.30, 1.46)
No	Ref	Ref
Smoking during pregnancy		
Yes	1.53 (1.18, 1.98)	1.57 (1.19, 2.06)
No	Ref	Ref
Medical characteristics		
Pre-existing diabetes		
Yes	7.56 (6.43, 8.91)	6.09 (5.12, 7.25)
No	Ref	Ref
Gestational diabetes		
Yes	4.62 (4.24, 5.05)	4.03 (3.68, 4.43)
No	Ref	Ref
Multifetal gestation		
Yes	2.66 (2.42, 2.93)	2.82 (2.54, 3.12)
No	Ref	Ref
Nulliparity		
Yes	1.26 (1.20, 1.33)	1.58 (1.50, 1.67)
No	Ref	Ref

[a]: Grand mean-centered variables entered as continuous covariates: maternal age and prenatal care visits

Table 10: Results of multilevel logistic regression: Unadjusted and adjusted odds ratios (aORs)[a] and 95% confidence intervals (CIs) for hypertensive disorders of pregnancy among women in Miami-Dade County, Florida who had a live birth, 2008–2012 (N = 121,421)

	Null Model	Model1 aOR (95% CI)	Model 2 aOR (95% CI)
Fixed effects			
Neighborhood deprivation			
High	———	1.20 (1.09, 1.32)	1.17 (1.08, 1.27)
Moderate	———	1.15 (1.07, 1.24)	1.12 (1.05, 1.21)
Low	———	Ref	Ref
Educational attainment			
< HS diploma, age <18	———	———	1.39 (1.13, 1.69)
< HS diploma, age ≥18	———	———	0.95 (0.89, 1.02)
High school diploma/GED	———	———	Ref
Insurance status			
Self-pay	———	———	1.57 (1.45, 1.70)
Medicaid	———	———	1.07 (1.01, 1.13)
Private insurance/other	———	———	Ref
Race/ethnicity			
Non-Hispanic Black	———	———	1.62 (1.39, 1.89)
Hispanic	———	———	1.14 (1.01, 1.30)
Other race/ethnicity			0.58 (0.52, 0.66)
Non-Hispanic White	———	———	Ref
Random effects			
Level-2 variance	0.015 (0.007, 0.030)	0.010 (0.004, 0.024)	0.008 (0.003 0.020)
Intraclass correlation (ICC)	0.005	0.003	0.002
Median odds ratio (MOR)	1.23	1.10	1.09

[a]*: Grand mean-centered variables entered as continuous covariates: maternal age and prenatal care visits*

Table 10: Results of a multilevel logistic regression: Adjusted odds ratios (aORs)[a] and 95% confidence intervals (CIs) for hypertensive disorders of pregnancy among women in Miami-Dade County, Florida who had a live birth, 2008–2012 (N = 121,421)

	Model 3 aOR (95% CI)	Model 4 aOR (95% CI)
Fixed effects		
Neighborhood deprivation		
High	1.16 (1.06, 1.26)	1.16 (1.07, 1.26)
Moderate	1.12 (1.04, 1.21)	1.13 (1.05, 1.21)
Low	Ref	Ref
Educational attainment		
< HS diploma, age <18	1.51 (1.24, 1.85)	1.34 (1.10, 1.63)
< HS diploma, age ≥18	0.96 (0.90, 1.02)	1.05 (0.99, 1.12)
High school diploma/GED	Ref	Ref
Insurance status		
Self-pay	1.65 (1.53, 1.79)	1.69 (1.56, 1.84)
Medicaid	1.06 (1.00, 1.12)	1.12 (1.05, 1.18)
Private insurance/other	Ref	Ref
Race/ethnicity		
Non-Hispanic Black	1.49 (1.28, 1.73)	1.58 (1.37, 1.84)
Hispanic	1.11 (0.98, 1.25)	1.11 (0.98, 1.27)
Other race/ethnicity	0.78 (0.62, 0.98)	0.72 (0.59, 0.89)
Non-Hispanic White	Ref	Ref
Modifiable characteristics		
Pre-pregnancy BMI status		
Category III obese	5.50 (4.39, 6.90)	5.34 (4.27, 6.68)
Category II obese	3.78 (3.06, 4.60)	3.74 (3.05, 4.58)
Category I obese	2.47 (2.02, 3.02)	2.52 (2.07, 3.07)
Overweight	2.35 (1.77, 3.13)	2.58 (1.98, 3.37)
Normal / underweight	Ref	Ref
Excessive gestational weight gain		
Yes	1.49 (1.39, 1.60)	1.44 (1.34, 1.54)
No	Ref	Ref
Smoking during pregnancy		
Yes	1.52 (1.19, 1.96)	1.46 (1.09, 1.96)
No	Ref	Ref
Medical characteristics		
Pre-existing diabetes		
Yes	———	5.52 (4.71, 6.48)
No	———	Ref
Gestational diabetes		
Yes	———	3.64 (3.13, 3.99)
No	———	Ref
Multifetal gestation		
Yes	———	2.82 (2.50, 3.18)
No	———	Ref
Nulliparity		
Yes	———	1.77 (1.67, 1.88)
No	———	Ref
Random effects		
Level-2 variance	0.008 (0.003 0.021)	0.007 (0.002, 0.020)
Intraclass correlation (ICC)	0.002	0.002
Median odds ratio (MOR)	1.09	1.08

[a]*: Grand mean-centered variables entered as continuous covariates: maternal age and prenatal care visits*

References

Azavea. (n.d.). About Azavea - beyond dots on a map. Retrieved July 13, 2016, from https://www.azavea.com/about/

bc workshop. (2015, August 14). The History of the POP Neighborhood Map. Retrieved September 4, 2017, from http://www.bcworkshop.org/posts/the-history-of-the-pop-neighborhood-map

bc workshop. (n.d.). Dallas: Know Your Neighborhood. Retrieved September 7, 2017, from http://peopleorganizingplace.com/know/

bc workshop. (n.d.). Draw Your Neighborhood. Retrieved September 4, 2017, from http://peopleorganizingplace.com/draw/

Biancalana, C., Flamini, A., Gasparetti, F., Micarelli, A., Millevolte, S., & Sansonetti, G. (2011). Enhancing traditional local search recommendations with context-awareness. *Umap*, 6787, 335–340.

Boundaries For Developing IDX Search & Filter Applications. (n.d.). Retrieved March 17, 2016, from https://www.homejunction.com/boundaries/

Cant, M. C., Strydom, J. W., Jooste, C. J., & Plessis, P. J. (Eds.). (2009). *Marketing Management* (5 ed.). Cape Town, South Africa: Juta Academic.

Capps, K. (n.d.). How many neighborhoods is too many for one map? Retrieved September 9, 2016, from http://www.citylab.com/tech/2015/09/how-many-neighborhoods-is-too-many-for-one-map/403474/

Crooks, A., Pfoser, D., Jenkins, A., Croitoru, A., Stefanidis, A., Smith, D., et al. (2015). Crowdsourcing urban form and function. *International Journal of Geographical Information Science*, 29(5), 720–741.

DeMers, J. (2014, December 2). How to own your neighborhood in google local search rankings. Retrieved September 4, 2017, from https://www.forbes.com/sites/jaysondemers/2014/12/02/how-to-own-your-neighborhood-in-google-local-search-rankings/#e986a6a1d8ba

Esri. (2017, May). Tapestry Segmentation: Methodology. An Esri White Paper. Retrieved November 3, 2016, from http://downloads.esri.com/esri_content_doc/dbl/us/J9941_Tapestry_Segmentation_%20Methodology_2017.pdf

Esri. (n.d.). ArcMap: An overview of the Neighborhood toolset. Retrieved September 8, 2017, from http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/an-overview-of-the-neighborhood-tools.htm

Esri Demographics. (2014). LifeMode Group 7F: Ethnic Enclaves. Southwestern Families. Retrieved September 4, 2017, from http://downloads.esri.com/esri_content_doc/dbl/us/tapestry/segment32.pdf

Forbes. (2015, March 12). Location Intelligence Leads to Business Insights and Creates Competitive Advantage. Retrieved October 23, 2016, from http://www.forbes.com/sites/forbespr/2015/03/12/location-intelligence-leads-to-business-insights-and-creates-competitive-advantage/#7f0c6bc81f12

Ford, J. L., & Browning, C. R. (2011). Neighborhood social disorganization and the acquisition of trichomoniasis among young adults in the United States. *American Journal of Public Health*, 101, 1696–1703.

GISGeography. (2015, November 30). The ultimate list of GIS formats - Geospatial file extensions. Retrieved September 8, 2017, from http://gisgeography.com/gis-formats/

GitHub. (n.d.). Azavea. Retrieved April 20, 2016, from https://github.com/azavea

Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69(4), 211–221.

Google Maps API. (n.d.). Google Maps API. Retrieved November 2, 2016, from https://developers.google.com

Google Scholar. (n.d.). Academic papers written with Yelp! academic dataset. Retrieved September 4, 2017, from https://scholar.google.com/scholar?q=citation%3A+Yelp+Dataset&btnG=&hl=en&as_sdt=0%2C5

Kim, L. (2014, October 14). 9 Places You Can Learn How to Code (for Free). Retrieved September 4, 2017, from https://www.inc.com/larry-kim/9-places-you-can-learn-how-to-code-for-free.html

Maponics. (n.d.). Key Industries. Retrieved September 13, 2016, from http://www.maponics.com/solutions/key-industries

Maponics. (n.d.). Neighborhood Boundaries. Retrieved March 1, 2016, from http://www.maponics.com/products/communities/neighborhood-boundaries

Merlo, J., Wagner, P., Ghith, N., & Leckie, G. (2016). An original stepwise multilevel logistic regression analysis of discriminatory accuracy: The case of neighbourhoods and health. *PLoS ONE*, 11, e0153778.

Mujahid, M. S., Diez Roux, A. V., Morenoff, J. D., Raghunathan, T. E., Cooper, R. S., Ni, H., & Shea, S. (2008). Neighborhood characteristics and hypertension. *Epidemiology, 19,* 590–598.

Neis, P., & Zipf, A. (2012). Analyzing the contributor activity of a volunteered geographic information project — the case of OpenStreetMap. *ISPRS International Journal of Geo-Information*, 1(3), 146–165.

OpenStreetMap. (n.d.). The most densely mapped locations in OSM. Retrieved October 29, 2016, from http://fred.dev.openstreetmap.org/density/

OpenStreetMap. (n.d.). Wiki: Stats. Retrieved October 28, 2016, from http://wiki.openstreetmap.org/wiki/Stats

Padmanabhan, A., Wang, S., Cao, G., Hwang, M., Zhang, Z., Gao, Y., et al. (2014). FluMapper: A cyberGIS application for interactive analysis of massive location-based social media. *Concurrency and Computation: Practice and Experience, 26*(13), 2253–2265.

Panian, Z. (2012). A New Dimension of Business Intelligence: Location-based Intelligence. *World Academy of Science, Engineering and Technology, International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, 6(3), 338–343. Retrieved from http://waset.org/Publication/a-new-dimension-of-business-intelligence-location-based-intelligence/9520

Pranata, I., & Susilo, W. (2016). Are the most popular users always trustworthy? The case of Yelp. *Electronic Commerce Research and Applications*, 20, 30–41.

Public Mapping Project. (n.d.). Software. Retrieved April 20, 2017, from http://www.publicmapping.org/resources/software

Quattroshapes. (n.d.). The Global Polygon Gazetteer. Retrieved October 9, 2016, from http://quattroshapes.com/

Redfin. (n.d.). Opportunity Score: Powered by Walk Score. Retrieved September 4, 2017, from https://labs.redfin.com/about-opportunity-score

Redfin. (n.d.). Real-Time Data Center: Downloadable Housing Market Data from Redfin. Retrieved September 4, 2017, from https://www.redfin.com/blog/data-center

Redfin. (n.d.). Terms of Use. Retrieved October 28, 2016, from https://www.redfin.com/about/terms-of-use

Ross, N. A., Tremblay, S., & Graham, K. (2004). Neighbourhood influences on health in Montréal, Canada. *Social Science & Medicine*, 59, 1485–1494.

Rummo, P. E., Guilkey, D. K., Shikany, J. M., Reis, J. P., & Gordon-Larsen, P. (2017). How do individual-level sociodemographics and neighbourhood-level characteristics influence residential location behaviour in the context of the food and built environment? Findings from 25 years of follow-up in the CARDIA Study. *Journal of Epidemiology & Community Health, 71*(3), 261–268.

Schomberg, J. P., Haimson, O. L., Hayes, G. R., & Anton-Culver, H. (2016). Supplementing public health inspection via social media. *PLoS ONE*, 11(3), e0152117.

Schutzberg, A. (2008, February 13). Neighborhood boundaries: the next big thing in geographic data. Retrieved October 10, 2016, from http://www.directionsmag.com/entry/neighborhood-boundaries-the-next-big-thing-in-geographic-data/122761

See, L., Fritz, S., & Leeuw, J. (2013). The rise of collaborative mapping: trends and future directions. *ISPRS International Journal of Geo-Information, 2*(4), 955–958.

Slopen, N., Non, A., Williams, D. R., Roberts, A. L., & Albert, M. A. (2014). Childhood adversity, adult neighborhood context, and cumulative biological risk for chronic diseases in adulthood. *Psychosomatic Medicine, 76*, 481–489.

Smith, C. R. (2017, May 12). How to pay to play if you want to win local search results. Retrieved September 4, 2017, from https://www.forbes.com/sites/forbesagencycouncil/2017/05/12/how-to-pay-to-play-if-you-want-to-win-local-search-results/#78c84c7810c7

Statista. (n.d.). Statista. Retrieved September 3, 2017, from https://www.statista.com

Sussman, S., Garcia, R., Cruz, T. B., Baezconde-Garbanati, L., Pentz, M. A., & Unger, J. B. (2014). Consumers' perceptions of vape shops in Southern California: An analysis of online Yelp reviews. *Tobacco Induced Diseases, 12*(1), 1972.

Trulia. (n.d.). Trulia Local: live well. Retrieved September 4, 2017, from https://www.trulia.com/local/miami-fl/type:care_essentials_live

Trulia. (n.d.). Trulia Maps. Retrieved September 4, 2017, from https://www.trulia.com/local/

Walk Score. (n.d.). Cities and neighborhoods. Retrieved September 8, 2017, from https://www.walkscore.com/cities-and-neighborhoods/

Walk Score. (n.d.). Data Services: walkability, real estate, and public health data.

WalkScore. (n.d.). Public health and the built environment. Retrieved October 1, 2016, from https://www.walkscore.com/professional/public-health-research.php

Who's On First. (n.d.). Who's On First. Retrieved August 6, 2016, from https://whosonfirst.mapzen.com/

Yelp! (n.d.). Yelp open dataset. Retrieved October 12, 2016, from https://www.yelp.com/dataset

Zillow. (n.d.). 7000+ neighborhood boundary files in shapefile Format. Retrieved February 3, 2016, from http://www.zillowgroup.com/news/7000-neighborhood-boundary-files-in-shapefile-format

CHAPTER V

DISCUSSION

i.   **Neighborhood tools developed by the private sector**

A thematic review serves both as an introduction for readers who are new to a specific topic and as a vehicle for advancing a specific research area. The purpose of this review was to identify alternative sources of neighborhood boundaries generated outside of academia that might be applicable and useful to neighborhood health research. Promoting awareness of such tools has the potential to advance neighborhood health research methodology. The review's chief strengths are the timeliness and utility of the topic. A few recent studies have summarized emerging mapping tools or technology in health research (e.g., Google Maps, drones, etc.) (Schootman et al., 2016; Vandeviver, 2014). However, I believe this to be the first thematic review of nontraditional sources for U.S. neighborhood polygons and centroids. Public health is rich in talent and commitment but lacking in money and other material resources (Frieden, 2014; van Panhuis et al., 2014). For public health researchers who are seeking replacements for census boundaries in studies where sociocultural relevance is an important element of the neighborhood proxy, this review identifies avenues for potentially leveraging private sector innovation.

Local search: advantages and disadvantages

Local search datasets tend to cover large geographic areas but do not necessarily contain demographics or other variables related to health research. Often, they consist of only boundaries or centroids and must be combined with other datasets to be of use to researchers. The companies that design these boundaries have a monetary stake in

ensuring that they closely correspond to actual neighborhoods. This increases their potential utility for health researchers. However, few companies divulge methodological details. Further, because these data are typically posted on websites intended for web developers, it is prudent to learn some basic coding skills and terminology before attempting to acquire boundaries from these sites. Free training materials can be found online (Kim, 2014).

Real estate: advantages and disadvantages

Neighborhood resources created by the real estate industry tend to include numerous variables (i.e., sociodemographic information, crime statistics, and neighborhood quality indicators such as walkability). Given what is known about the purposes and the methods behind these tools, there is reason to believe they might also include at least a modicum of "ground truth." Because Zillow's boundaries can be viewed via its search engine, this dissertation's author and co-reviewers were able to conduct a cursory search for neighborhoods familiar to them (e.g., Miami's Little Haiti neighborhood, Atlanta's Old Forth Ward). Visually, there was strong concordance between Zillow polygons and the reviewers' *a priori* assumptions of where the boundaries should be. Formal geospatial analysis would be needed to test the veracity of these casual observations.

The primary disadvantage of real estate's neighborhood tools is geographic limitations, such as limiting data sets to only select states or to cities with high real estate demand. Nevertheless, this should not hinder researchers whose projects focus on major U.S. urban centers: In most instances, areas omitted from data sets are those that are sparsely populated and/or less likely to generate substantial real estate sales (e.g.,

Wyoming). This would be problematic for those studying rural areas or the entire nation. At least one epidemiological study has used Zillow boundaries already: A 2016 analysis of area-level and individual-level predictors of residential location behavior combined participant and census data with Zillow neighborhood boundaries in Chicago, Minneapolis, and Oakland (Rummo, Guilkey, Shikany, Reis, & Gordon-Larsen, 2017). For the one study location in which Zillow boundaries were not yet available (Birmingham), the researchers used regional planning commission boundaries.

Another common challenge is identifying the origins of real estate tools. Often, sites borrow or purchase nontraditional neighborhood boundaries from other organizations but do not prominently display attribution information. For example, Trulia's website includes a series of interactive maps that allow users to quickly assess crime, commute, and school conditions in a particular community (Trulia, n.d.). One can zoom in all the way to the neighborhood level on these maps, which also display demographics, affordability, and natural hazards. The site even published rankings for a series of indicators the company calls "Live Well," consisting of hospitals, urgent care clinics, pediatricians, pharmacies, and day care locations (Trulia, n.d.). Close scrutiny of the company's website, coupled with extensive web searches, revealed that these maps leverage Google Maps API combined with data from Maponics, CrimeReports, SpotCrime, and Yelp!, among others (Trulia, n.d.). That is why Trulia was not included in the final list of boundary sources in Table 3. Such tools have clear public health applications. However, given the number of parties involved in their creation, acquiring permission to use these tools for research purposes could be a lengthy process (e.g., determining access fees, terms of use agreements, etc.). Still, that does not necessarily

mean one should not pursue this possibility. As always, investigators should base decisions on study aims, data needs, resources, timelines, etc.

Researchers should also note that companies that sell geocoded, neighborhood-level data do not necessarily provide any actual boundary information. Even advertisements for interactive neighborhood maps coupled with key terms such as "shapefiles" or "API" should not be viewed as confirmation that centroids or polygons are available. Several potential websites were excluded from this review on this basis: For example, the website Walk Score rates a neighborhood's "walkability" on a scale of 0 (Car Dependent) to 100 (Walker's Paradise) (Walk Score, n.d.). These walk scores, along with ratings for biking, transit, and pedestrian friendliness, have been incorporated into numerous real estate search engines and apps. Public health researchers can also purchase this information in several formats, including API, spreadsheets and shapefiles (Walk Score, n.d.). In fact, the website — which was designed by Redfin in consultation with the Rockefeller Foundation and Robert Wood Johnson Foundation — links to numerous published studies that have used these data ("Public Health and the Built Environment," n.d.). Walk Score advertises multiple area-level data packages and services, ArcGIS shapefiles, and an API called "Neighborhood Map." But via e-mail, a company representative clarified that none of these options include neighborhood centroids or polygons. Thus, Walk Score was excluded from the list of boundary resources in this review.

Similarly, brokerage firm Redfin offers a wealth of U.S. area-level information in numerous formats. The company has a dedicated webpage providing point-and-click access to data on housing market trends and other variables (Redfin, n.d.). These data are

free and can be used for research purposes so long as attribution is given and certain other guidelines are followed. The company also publishes weekly data visualizations and periodic reports on the "hottest neighborhoods." In partnership with Walk Score — which it acquired in 2014 — Redfin created a tool called "Opportunity Score" that has potential public health implications (Redfin, n.d.). A user enters an address and the tool evaluates the ease or difficulty of commuting to work without a car from that location in 30 minutes or less. However, boundary files are not available, because Redfin uses polygons supplied by Maponics (Redfin, n.d.). Hence, Redfin was also excluded from the table of resources listed in the Aim 2 review.

Location-based intelligence: advantages and disadvantages

Location-based intelligence data sets tend to include a large variety of variables and offer the option to request additional variables for a fee. Some location-based intelligence projects and companies even tailor products specifically to the needs of government, non-profit, and research organizations. Designed by the Public Mapping Project, the District Builder tool allows users to combine traditional and modern data sources and construct new district boundaries (Public Mapping Project, n.d.). The program then calculates district statistics (e.g., population, contiguity, etc.) to help guide users' decisions about whether to adjust the boundaries they have drawn. The software is open-source and thus free. Still, there are costs associated with the installation and hosting necessary to run the software. Azavea, the company that designed District Builder, has a larger neighborhood toolkit (Azavea, n.d.). In addition to U.S. polygons based on demographics and market segmentation, the company accepts custom data requests. Additionally, some of its data are open-source and available through GitHub (GitHub, n.d.). Location-based

intelligence resources tend to have longevity and are typically updated on a consistent basis because they are owned by large companies with a vested interest in keeping their data current. However, few if any of these products are free, even for researchers affiliated with the government or public educational institutions.

Citizen collaboration: Advantages and disadvantages

An obvious strength of using neighborhood boundaries created through citizen collaboration is their accessibility. In addition to being open-source and thus free, these data typically include thorough documentation and a user-friendly interface. The methods used by these sites have another, more distinct advantage: Incorporating photo-taggers' and citizen mappers' opinions about neighborhood names and boundaries injects "ground truth" into these resources. Even so, there are several inherent limitations. Scale or level of granularity is determined by the photo-tagger or citizen mapper. These individuals also might accidentally introduce error into the sample: Someone not familiar with a neighborhood (e.g., a tourist) might incorrectly identify it. Another potential source of error with geotagging stems from the fact that places could share names with a person, object, etc. (i.e., homonyms), or two places could share the same name (polysemes). Lastly, taggers or citizen mappers might be fairly homogenous groups in terms of demographics or other characteristics. For example, a 2012 study found that nearly three-fourths (72%) of OSM contributors lived in Europe (Neis & Zipf, 2012).

While these citizen collaborations have yielded some extensive neighborhood resources, it is important to note that few of these projects endure over long term. During the discovery phase of this paper, I found a consistent pattern: Many sites had intriguing

data sources and analysis plans. But a flurry of initial activity was followed by years of "radio silence," sometimes without anyone ever officially acknowledging that the project had been abandoned. For example, zetashapes was functional when this review began but was shuttered while this dissertation was being revised and was excluded from the final resources list. Also, because participants are not required to have professional training or credentials, data quality is a concern (See, Fritz, & Leeuw, 2013). Because such sites typically document their user protocols in lengthy detail, researchers can weigh the potential risks and benefits of using these data and make a reasonably informed decision.

Considerstations for future research with private sector tools

Regardless of the online source, safeguarding participants' personally identifiable information must remain a key concern for researchers. These new technologies "create new ways to violate human participant protections" (Bader, Mooney, & Rundle, 2016). For example, "direct geocoding" — entering participants' addresses into a site such as Google Maps or Walk Score is unethical. So is what is known as "the needle in a haystack method" — where a researcher attempts to obscure a participant's address when geocoding it on a public website by entering many "fake" addresses along with the real one. Ethical methods include geographic sampling — passive, anonymous data collection through tools such as Google Street View — and geographic imputation — geocoding the address of a non-participant who resides on a nearby/similar street (Bader, Mooney, & Rundle, 2016). Downloading boundary files and geocoding participant information using a non-Web-based software platform such as ArcGIS Desktop avoids the aforementioned ethical concerns.

By seeking community members' input in answering the question of "What is a neighborhood?", entities outside of academia have created tools that might include the ground truth that administrative boundaries such as census tracts and ZIP codes lack (Kingsley, Coulton, & Pettit, 2016; Wahl, 2008). Combining GIS analysis, data mining, and "boots on the ground" qualitative research has allowed the private sector to refine market segmentation and improve user experience and profits (Capps, n.d.; Carroll et al., n.d.; Forbes, 2015; Hayden, 2014; Rosoff, 2015). Thanks in part to the rising popularity of open-source data, private citizens have also contributed to this pool of resources (Crooks et al., 2015; Neis & Zipf, 2012; See, Fritz, & Leeuw, 2013; Van Exel, Dias, & Fruijtier, 2010). Though their motivations differ, businesses and citizens both have a great deal of knowledge about and a vested interest in neighborhoods. Thus, it stands to reason that these tools might more accurately represent real communities than do census boundaries. This is an area with substantial research potential.

Academic researchers also have explored mixed methods approaches to constructing neighborhood boundaries (Cranshaw, Schwartz, Hong, & Sadeh, 2012; Weiss, Ompad, Galea, & Vlahov, 2007; Wilske, 2008). However, many of these studies have had substantial limitations in terms of geographic coverage, sample size, scalability, and long-term feasibility. Such challenges are common in public health because of resource limitations (Frieden, 2014; van Panhuis et al., 2014). This stands in stark contrast to the private sector: Motivated by potential profits, these businesses are willing and able to devote vast monetary and technological resources to neighborhood research and online tools. Many of these resources are available for immediate download at no cost and some are updated quite frequently and consistently. This creates an opportunity to leverage

private investment to benefit public health and invigorate neighborhood health research with relatively little investment of time or funds.

Given their respective strengths and limitations, resources designed by particular groups may be better suited to specific purposes. For example, with a large geographic coverage area and low costs, local search tools would likely be most appropriate for studies that include the entire country as a sampling frame. This also leads the author of this dissertation to suggest that, at present, Google or Yelp! boundary files might be the most viable alternative to census units. For studies that focus solely on one or more urban or populous areas, real estate boundaries are a viable option and tend to be accompanied by richer information on sociodemographics and other key variables than local search boundaries —Yelp!'s additional data options notwithstanding. Zillow's free polygons are available online as ArcGIS shapefiles, the most common format used in GIS analysis (GISGeography, 2015).

Researchers with more substantial monetary resources can explore marketing segmentation tools. These data include rich detail about subtypes of residents within neighborhoods — in terms of not only sociodemographics but also lifestyle and attitudes. This wealth of information could help researchers better understand the pathways through which neighborhood factors influence health behaviors and, in turn, health outcomes. These data could also help inform the planning and implementation of health interventions and policy changes, such as seeking to lower body mass index (BMI) by improving built environment factors such as walkability, aiming to reduce consumption of cigarettes, alcohol, and soda by changing zoning laws to restrict the location of

convenience stores, or seeking to lower overall morbidity and mortality by improving

housing conditions and area resources in deprived neighborhoods (Gibson et al., 2011,

Sanders-Jackson, Parikh, Schleicher, Fortmann, & Henriksen, 2015, Sarkar, Gallacher, &

Webster, 2013).

Since many community collaboration projects are eventually abandoned, they might

not seem very useful at first glance. However, preliminary and short-term studies could

benefit from these often innovative tools while minimizing the risk of losing one's data

mid-project. Further, since these sites are open-source and document their protocols in

great detail, an industrious researcher might be able to take an active role in a

collaborative mapping project and essentially "keep it going" as long as needed.

There is one particularly noteworthy obstacle surrounding resources developed by

most private companies: Details of how they were developed may be considered trade

secrets. Still, one should not assume that pursuing such public-private partnerships would

be futile: It might be possible for companies to create "workarounds" (e.g., signing a

nondisclosure agreement regarding methodological details) — especially if researchers

convey the potential benefits to the community, which would in turn reflect well on the

company who lent its tools.

It is likely that some researchers will balk at the notion of using tools without

knowing their underlying methodology. However, public health researchers should be

accustomed to acting despite uncertainty: After all, a key principle of epidemiology is

that preventing the spread of a disease does not require fully understanding its causes

(Friis & Sellers, 2013). A famous example is John Snow and the London cholera

epidemic of 1853. A year before *Vibrio cholerae* was first isolated and in an era when germ theory had yet to gain traction, Snow needed only a hand-drawn map and his powers of deduction to help stop a deadly outbreak (Friis & Sellers, 2013). If neighborhood health researchers are willing to accept a bit of uncertainty and harness the private sector's powerful resources, imagine how many people might benefit.

ii.   **Aim 1 limitations**

A thematic review is meant to provide the reader with a holistic overview and thus does not cite every available piece of literature on the topic at hand. Given that there is no central repository for information on this subject matter, this dissertation's author and co-reviewers relied almost exclusively on nontraditional sources (i.e., websites, app store results) as opposed to peer-reviewed, academic literature. There was a concerted effort made to apply the same principles of impartiality and scientific rigor to this review as would be required in any other academic paper. Additionally, detailed descriptions of the methods underlying many of the private industry tools were unavailable. Such information is considered proprietary.

iii.   **Sensitivity analysis of census units and natural neighborhoods**

Aim 2 is intended to serve as a case study for an empirical approach to choosing the optimal areal unit for studying the relationship between a specific health outcome and its theorized area-level exposures. The other chief goal is to extensively evaluate the utility of socioeconomically homogeneous natural neighborhoods (NNs) compared with traditional census units. One type of NN performed substantially better than all other

areal units: In the GWR analysis, the CT large NN model explained more than twice as much HDP variance as any traditional census unit.

In a Canadian study whose method of generating NNs was most similar to that used in this dissertation, nitrogen dioxide explained the largest proportion of variance in asthma when measured at the level of NN (Parenteau, Sawada, & Sawada, 2011). A French study found that the relationship between asthma and neighborhood deprivation was stronger when measured within socioeconomically homogeneous neighborhoods compared to census tracts (Sabel, Kihal, Bard, & Weber, 2013). However, one of the earliest studies of NNs found "remarkably similar results" when Canadian census tracts were compared to NNs formed through a mixed methods approach (Ross, Tremblay, & Graham, 2004). Further research is needed to verify whether aggregating census units based on neighborhood deprivation variables might serve as a feasible neighborhood proxy in future studies of HDP and, potentially, other health outcomes that have social determinants.

In general, the strength of association between HDP and neighborhood deprivation as well as the proportion of variance in HDP prevalence explained differed dramatically depending on the chosen geographic unit of analysis. Larger areal units tended to produce better statistical models for my variables of interest and, in GWR analysis, two types of NNs performed better than traditional units of the equivalent scale. This suggests that both aspects of the MAUP, scale and zoning, affect the relationship between HDP and neighborhood deprivation — a finding that corresponds with previous research on area-level variables and health (Gale, Magzamen, Radke, & Tager, 2011; Hipp, 2007; Ortega

Hinojosa et al., 2014; Schuurman, Bell, Dunn, & Oliver, 2007; Tian, Goovaerts, Zhan, & Wilson, 2010). Studies that have reported little or no MAUP effects (Diez Roux et al., 2001; Zhang et al., 2014) typically have focused solely on scale.

The poor performance of block groups in the present study — both in their usual form and as an aggregate building block for NNs — was somewhat surprising. Many area-level health studies have used block group data, often with the assumption that such small units might represent the most realistic available proxies of locally meaningful neighborhoods using census data (Cabrera-Barona, Wei, & Hagenlocher, 2016; Frank et al., 2006; Pearl, Braveman, & Abrams, 2001). However, census analyses suggest that inflated measurement error in block group data negates any potential advantages (Bazuin & Fraser, 2013; Spielman, Folch, & Nagle, 2014).

There is a potential alternative interpretation of my block group findings. It is possible that, compared with census tracts (n=507) and ZCTAs (n=78) the large number of block groups (n=1,560) provided the software with too many potential combinations to generate such relatively small aggregate clusters for the NNs. However, this does not explain the extremely low R^2 value for the original OLS block group model.

Because variance in NDI was standardized within each model and the largest distinctions between R^2 occurred in the GWR analysis, it seems plausible that the superior performance of the CT large NN model was due to larger spatial variation in HDP prevalence. While some will likely question whether ZCTAs are too large to be considered "neighborhoods," area-level exposures can operate at different scales. For

example, schools typically serve larger geographic areas but are still an important health indicator.

iv.   Neighborhood deprivation and hypertensive disorders of pregnancy

Neighborhood-level deprivation and a lack of individual-level resources were both associated with HDP. However, contrary to this dissertation's hypotheses, the two levels of deprivation did not have a significant multiplicative impact on a woman's HDP odds. This stands in contrast to previous research on deprivation amplification concerning other health outcomes and behaviors. A Canadian longitudinal analysis found stark survival differences between poor individuals living in the most and least deprived neighborhoods (Ross, Oliver, & Villeneuve, 2013). When deprivation was defined in terms of material goods in that study, the all-cause mortality gap was 10%. When defined in terms of social resources, the gap was 7%. A recent U.K. study found that individuals with low education who lived in areas of high deprivation were significantly less likely to use weight management strategies (Green et al., 2014).

While my analysis found no significant association between deprivation amplification and HDP, further research is still needed. Education and insurance status were the only available measures of individual-level socioeconomic deprivation, because the birth records lacked income data. Also, it is possible that other variables not included in the eight-item neighborhood deprivation index might better encompass aspects of area-level deprivation that interact with individual-level deprivation to influence health. Neighborhood health researchers have emphasized that conceptualizing scarcity as merely the presence or absence of goods and services is insufficient (Macintyre, 2007;

Stock & Ellaway, 2013). Instead, one must seek to capture multiple aspects of each neighborhood factor, such as type, quality, and spatial distribution. This would require primary data collection and likely qualitative interviews with community residents.

In this dissertation, compared with non-Hispanic Whites, non-Hispanic Black women had higher odds of HDP — a finding consistent with previous studies (R. Gold, Gold, Schilling, & Modilevsky, 2014; Lo, Mission, & Caughey, 2013; Miranda et al., 2010; Shen, Tymkow, & MacMullen, 2005). There was also a racial/ethnic disparity in exposure to neighborhood deprivation: Compared with both non-Hispanic Whites and Hispanics, a significantly larger proportion of non-Hispanic Black women lived in high deprivation neighborhoods in Miami-Dade County, Florida. This aligns with previous findings that Blacks are disproportionately exposed to neighborhood disadvantage (Collins & Williams, 2001; Laraia et al., 2006; Sampson, 2012).

My analysis also confirmed that previously identified individual-level factors play a substantial role in women's HDP risk even after accounting for between-neighborhood variance and upstream factors. There was a clear dose-response relationship between HDP and BMI, which is a modifiable risk factor. This may be of particularly importance for combatting racial disparities since nearly one-third of the Hispanic women and nearly 40% of the non-Hispanic Black women in my sample were obese. Half of Hispanic women and more than a third of non-Hispanic Black women were overweight. Of particular note, 7.9% of Black women were category III obese, which was associated with a more than five-fold increase in HDP risk.

Excessive gestational weight gain was associated with a 44% increase in HDP odds. This finding may be clinically significant for two reasons: 1) Given that the adjustment for numerous other variables — including pre-pregnancy BMI status — this measure of association is relatively large, and 2) The time window of risk for this factor is only about 40 weeks and might be a matter of a relatively small amount of weight.

In a secondary analysis of data from a large randomized controlled trial of vitamin use and HDP complications (n = 9,543), women with a normal pre-pregnancy BMI who gained above the Institute of Medicine guidelines were at increased risk of developing gestational hypertension (odds ratio [OR]=1.5) or pre-eclampsia (OR=2.5) (Johnson, et al., 2013). Similarly, women who were overweight prior to pregnancy and gained more than recommended weight were more likely to develop pre-eclampsia (OR=4.2). Women who were obese prior to pregnancy and gained more than recommended weight had an increased risk of pre-eclampsia (OR=1.9).

In a post-hoc analysis, the mean difference between the two weight gain groups was 18.0 lbs. The largest difference was among obese women: 24.7 lbs. A recent clinical opinion essay referred to pregnancy as a "teachable moment" for weight gain, citing women's concern for the health of their baby and increased contact with health professionals as key lynchpins (Phelan, 2010). Additionally, interventions to prevent gestational weight gain — including a few targeting special populations such as obese and low-income women — have shown some success (Olson, Strawderman, & Reed, 2004; Phelan et al., 2011; Wolff, Legarth, Vangsgaard, Toubro, & Astrup, 2008).

v. Aim 2 and 3 limitations

Despite this study's large sample size, results might not be generalizable to the entire

United States because all participants resided in Miami-Dade County, which is one of the

most racially and ethnically diverse areas of the U.S. (Eitle & Taylor, 2008) with a large

proprortion of foreign-born residents and neighborhoods with relatively short histories

compared with other areas of the U.S. Further, because portions of the birth records data

were self-reported and a large proportion of cases (11.6%) were lost because of

missing/misspelled addresses, both under-reporting and over-reporting are possible.

However, my quality assurance analysis suggested that dropped cases were statistically

similar to the remaining sample. Additionally, some key variables were not available in

this data set – chiefly, individual-level income and mother's country of origin.

Ideally, one would use prospective data in a study of relationships between

neighborhood deprivation and health, because participants might move numerous times

and their socioeconomic circumstances might change substantially during the decades

that it takes most non-communicable diseases to develop. Nonetheless, HDP's natural

history is shorter than that of many other non-communicable diseases. And while it has

been estimated that one-fourth to one-third of U.S. women change residences during

pregnancy, nearly half of them move < 8 miles away from their previous residence

(Miller, Siffel, & Correa, 2009). Further, in the U.S., a person's socioeconomic status

does not tend to change drastically within a short time frame, making it improbable that a

significant proportion of relocations occur between neighborhoods with markedly

different deprivation characteristics. This suggests that the neighborhood deprivation

categories assigned to most study participants reflect the conditions in which they were susceptible to developing HDP.

In the sensitivity analysis, the largest coefficient of variation (as measured by R^2) was 0.27. While this would be considered a small R^2 in some fields of study, it is acceptable by social science research standards (Cohen, 1988; Falk, 1992). In particular, cross-sectional studies tend to underestimate R^2 when the independent variable is theorized to influence the outcome over a long time period (Ableson, 1985). Neighborhood deprivation is an upstream independent variable theorized to have a cumulative impact on health outcomes. Thus, the proportion of variation in HDP prevalence explained by neighborhood deprivation in the census tract large NN model was actually larger than anticipated. Further, the Aim 2 models — used in a sensitivity analysis to choose the unit of analysis for Aim 3 — were intentionally limited to only one independent variable and thus would not be expected to explain a robust proportion of variance in HDP prevalence. By contrast, the multilevel model of individual-level HDP odds examined in Aim 3 included one area-level independent variable as well as eight individual-level independent variables.

References

Ableson, R. (1985). A variance explanation paradox: when a little is a lot. *Psychological Bulletin, 97(*1), 129-133.

Azavea. (n.d.). About Azavea - beyond dots on a map. Retrieved July 13, 2016, from https://www.azavea.com/about/

Bader, M. D. M., Mooney, S. J., & Rundle, A. G. (2016). Protecting personally identifiable information when using online geographic tools for public health research. *American Journal of Public Health*, *106*(2), 206–208.

Bazuin, J. T., & Fraser, J. C. (2013). How the ACS gets it wrong: The story of the American Community Survey and a small, inner city neighborhood. *Applied Geography, 45*, 292–302.

Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York: Guilford Press.

Cabrera-Barona, P., Wei, C., & Hagenlocher, M. (2016). Multiscale evaluation of an urban deprivation index: Implications for quality of life and healthcare accessibility planning. *Applied Geography*, 70, 1–10.

Capps, K. (n.d.). How many neighborhoods is too many for one map? Retrieved September 9, 2016, from http://www.citylab.com/tech/2015/09/how-many-neighborhoods-is-too-many-for-one-map/403474/

Carroll, J. M., Shih, P. C., Kropczynski, J., Cai, G., Rosson, M. B., & Han, K. (n.d.). The Internet of Places at Community-Scale. In *Enriching Urban Spaces with Ambient Computing, the Internet of Things, and Smart City Design* (pp. 1–24). Hershey, PA: IGI Global.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, N.J.: L. Erlbaum Associates.

Collins, C., & Williams, D. R. (2001). Racial residential segregation: a fundamental cause of racial disparities in health. *Public Health Reports, 116*(5), 404–416.

Cranshaw, J., Schwartz, R., Hong, J. I., & Sadeh, N. (2012). The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City. Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2168428

Crooks, A., Pfoser, D., Jenkins, A., Croitoru, A., Stefanidis, A., Smith, D., et al. (2015). Crowdsourcing urban form and function. *International Journal of Geographical Information Science, 29*(5), 720–741.

Diez Roux, A. V., Kiefe, C. I., Jacobs, D. R., Jr, Haan, M., Jackson, S. A., Nieto, F. J., et al. (2001). Area characteristics and individual-level socioeconomic position indicators in three population-based epidemiologic studies. *Annals of Epidemiology, 11*(6), 395–405.

Eitle, D., & Taylor, J. (2008). Are Hispanics the new "Threat?" Minority group threat and fear of crime in Miami-Dade County. *Social Science Research, 37*(4), 1102–1115.

Forbes. (2015, March 12). Location intelligence leads to business insights and creates competitive advantage. Retrieved October 23, 2016, from http://www.forbes.com/sites/forbespr/2015/03/12/location-intelligence-leads-to-business-insights-and-creates-competitive-advantage/#7f0c6bc81f12

Frank, L. D., Sallis, J. F., Conway, T. L., Chapman, J. E., Saelens, B. E., & Bachman, W. (2006). Many Pathways from Land Use to Health: associations between neighborhood walkability and active transportation, body mass index, and air quality. *Journal of the American Planning Association, 72*(1), 75–87.

Frieden, T. R. (2014). Six components necessary for effective public health program implementation. *American Journal of Public Health, 104*(1), 17–22.

Friis, R. H., & Sellers, T. A. (2013). *Epidemiology for Public Health Practice* (5 ed.). Burlington, MA: Jones & Bartlett Learning.

Gale, S. L., Magzamen, S. L., Radke, J. D., & Tager, I. B. (2011). Crime, neighborhood deprivation, and asthma: A GIS approach to define and assess neighborhoods. *Spatial and Spatio-Temporal Epidemiology, 2*(2), 59–67.

Gibson, M., Petticrew, M., Bambra, C., Sowden, A. J., Wright, K. E., & Whitehead, M. (2011). Housing and health inequalities: A synthesis of systematic reviews of interventions aimed at different pathways linking housing and health. *Health & Place*, *17*(1), 175–184.

GISGeography. (2015, November 30). The ultimate list of GIS formats - geospatial file extensions. Retrieved September 8, 2017, from http://gisgeography.com/gis-formats/

Gold, R., Gold, K., Schilling, M., & Modilevsky, T. (2014). Effect of age, parity, and race on the incidence of pregnancy associated hypertension and eclampsia in the United States. *Journal of Neonatal-Perinatal Medicine, 4*(1), 46–53.

Green, M. A., Subramanian, S. V., Strong, M., Cooper, C. L., Loban, A., & Bissell, P. (2014). "Fish out of water": a cross-sectional study on the interaction between social and neighbourhood effects on weight management behaviours. *International Journal of Obesity, 39*(3), 535–541.

Hayden, T. (2014). *The Mobile Commerce Revolution: Business Success in a Wireless World*. Indianapolis, IN: Que Publishing.

Hipp, J. R. (2007). Block, tract, and levels of aggregation: neighborhood structure and crime and disorder as a case in point. *American Sociological Review, 72*(October), 659–680.

Johnson, J., Clifton, R. G., Roberts, J. M., Myatt, L., Hauth, J. C., Spong, C. Y., et al. (2013). Pregnancy outcomes with weight gain above or below the 2009 Institute of Medicine Guidelines. *Obstetrics & Gynecology, 121*(5), 969–975.

Kingsley, G., Coulton, C. J., & Pettit, K. L. (2016). Strengthening communities with data. Rowman & Littlefield. Retrieved from https://www.neighborhoodindicators.org/sites/default/files/publications/13805-urban_kingsley.pdf

Laraia, B. A., Messer, L., Kaufman, J. S., Dole, N., Caughy, M., O'Campo, P., & Savitz, D. A. (2006). Direct observation of neighborhood attributes in an urban area of the US south: characterizing the social context of pregnancy. *International Journal of Health Geographics, 5*(1), 11.

Lo, J. O., Mission, J. F., & Caughey, A. B. (2013). Hypertensive disease of pregnancy and maternal mortality. *Current Opinion in Obstetrics and Gynecology*, 25(2), 124–132.

Macintyre, S. (2007). Deprivation amplification revisited; or, is it always true that poorer places have poorer access to resources for healthy diets and physical activity? *International Journal of Behavioral Nutrition and Physical Activity, 7*, 1–7.

Miller, A., Siffel, C., & Correa, A. (2009). Residential mobility during pregnancy: Patterns and correlates. *Maternal and Child Health Journal, 14*(4), 625–634.

Miranda, M. L., Swamy, G. K., Edwards, S., Maxson, P., Gelfand, A., & James, S. (2010). Disparities in maternal hypertension and pregnancy outcomes: Evidence from North Carolina, 1994–2003. *Public Health Reports*, 125(4), 579–587.

Neis, P., & Zipf, A. (2012). Analyzing the contributor activity of a volunteered geographic information project — The case of OpenStreetMap. *ISPRS International Journal of Geo-Information, 1*(3), 146–165.

Olson, C. M., Strawderman, M. S., & Reed, R. G. (2004). Efficacy of an intervention to prevent excessive gestational weight gain. *American Journal of Obstetrics and Gynecology, 191*(2), 530–536.

Ortega Hinojosa, A. M., Davies, M. M., Jarjour, S., Burnett, R. T., Mann, J. K., Hughes, E., et al. (2014). Developing small-area predictions for smoking and obesity prevalence in the United States for use in Environmental Public Health Tracking. *Environmental Research, 134*, 435–452.

Parenteau, M.-P., & Sawada, M. C. (2011). The modifiable areal unit problem (MAUP) in the relationship between exposure to NO2 and respiratory health. *International Journal of Health Geographics, 10*(1), 58.

Parenteau, M.-P., Sawada, M. C., & Sawada, M. C. (2011). The modifiable areal unit problem (MAUP) in the relationship between exposure to NO2 and respiratory health. *International Journal of Health Geographics, 10*(1), 58.

Pearl, M., Braveman, P., & Abrams, B. (2001). The relationship of neighborhood socioeconomic characteristics to birthweight among 5 ethnic groups in California. *American Journal of Public Health, 91*(11), 1808–1814.

Phelan, S. (2010). Pregnancy: a "teachable moment" for weight control and obesity prevention. *American Journal of Obstetrics and Gynecology, 202*(2), 135.e1–135.e8.

Phelan, S., Phipps, M. G., Abrams, B., Darroch, F., Schaffner, A., & Wing, R. R. (2011). Randomized trial of a behavioral intervention to prevent excessive gestational weight gain: the Fit for Delivery Study. *American Journal of Clinical Nutrition, 93*(4), 772–779.

Rosoff, M. (2015, March 3). What I found when I tried Nextdoor, the neighborhood-based social network that just raised $110 million. Retrieved September 4, 2017, from http://www.businessinsider.com/nextdoor-raises-110-million-at-11-billion-valuation-why-2015-3

Ross, N. A., Tremblay, S., & Graham, K. (2004). Neighbourhood influences on health in Montréal, Canada. *Social Science & Medicine, 59*(7), 1485–1494.

Ross, N., Oliver, L., & Villeneuve, P. (2013). The contribution of neighbourhood material and social deprivation to survival: A 22-year follow-up of more than 500,000 Canadians. *International Journal of Environmental Research and Public Health*, 10(4), 1378–1391.

Sabel, C. E., Kihal, W., Bard, D., & Weber, C. (2013). Creation of synthetic homogeneous neighbourhoods using zone design algorithms to explore relationships between asthma and deprivation in Strasbourg, France. *Social Science & Medicine, 91*, 110–121.

Sampson, R. J. (2012). *Great American City: Chicago and the Enduring Neighborhood Effect*. Chicago: University of Chicago Press.

Sanders-Jackson, A., Parikh, N. M., Schleicher, N. C., Fortmann, S. P., & Henriksen, L. (2015). Convenience store visits by US adolescents: Rationale for healthier retail environments. *Health & Place, 34*, 63–66.

Sarkar, C., Gallacher, J., & Webster, C. (2013). Built environment configuration and change in body mass index: The Caerphilly Prospective Study (CaPS). *Health & Place, 19*, 33–44

Schootman, M., Nelson, E. J., Werner, K., Shacham, E., Elliott, M., Ratnapradipa, K., et al. (2016). Emerging technologies to measure neighborhood conditions in public health: implications for interventions and next steps. *International Journal of Health Geographics, 15*(1), 1.

Schuurman, N., Bell, N., Dunn, J. R., & Oliver, L. (2007). Deprivation indices, population health and geography: An evaluation of the spatial effectiveness of indices at multiple scales. *Journal of Urban Health, 84*(4), 591–603.

See, L., Fritz, S., & Leeuw, J. (2013). The rise of collaborative mapping: trends and future directions. *ISPRS International Journal of Geo-Information, 2*(4), 955–958.

Shen, J., Tymkow, C., & MacMullen, N. (2005). Disparities in maternal outcomes among four ethnic populations. *Ethnicity & Disease, 15*, 492–497.

Spielman, S. E., Folch, D., & Nagle, N. (2014). Patterns and causes of uncertainty in the American Community Survey. *Applied Geography, 46*, 147–157.

Stock, C., & Ellaway, A. (2013). *Neighbourhood Structure and Health Promotion*. (C. Stock & A. Ellaway, Eds.). Boston, MA: Springer.

Tian, N., Goovaerts, P., Zhan, F. B., & Wilson, J. G. (2010). Identification of racial disparities in breast cancer mortality: does scale matter? *International Journal of Health Geographics, 9*(1), 35.

Van Exel, M., Dias, E., & Fruijtier, S. (2010). The impact of crowdsourcing on spatial data quality indicators. *Proceedings of the 6th GIScience International Conference on Geographic Information Science*, 213.

van Panhuis, W. G., Paul, P., Emerson, C., Grefenstette, J., Wilder, R., Herbst, A. J., et al. (2014). A systematic review of barriers to data sharing in public health. *BMC Public Health, 14*(1), 3.

Vandeviver, C. (2014). Applying Google Maps and Google Street View in criminological research. *Crime Science, 3*(1), 13.

Wahl, B. (2008, December 4). Mapping the world ... one neighborhood at a time - Part Two. Retrieved September 4, 2017, from https://www.directionsmag.com/article/2425#v

Weiss, L., Ompad, D., Galea, S., & Vlahov, D. (2007). Defining neighborhood boundaries for urban health research. *American Journal of Preventive Medicine, 32*(6), S154–S159.

Wilske, F. (2008). Approximation of neighborhood boundaries using collaborative tagging systems. *GI-Days, 32*, 179–187.

Wolff, S., Legarth, J., Vangsgaard, K., Toubro, S., & Astrup, A. (2008). A randomized trial of the effects of dietary counseling on gestational weight gain and glucose metabolism in obese pregnant women. *International Journal of Obesity, 32*(3), 495–501.

Zhang, X., Holt, J. B., Lu, H., Wheaton, A. G., Ford, E. S., Greenlund, K. J., & Croft, J. B. (2014). Multilevel regression and poststratification for small-area estimation of population health outcomes: A case study of chronic obstructive pulmonary disease prevalence using the Behavioral Risk Factor Surveillance System. *American Journal of Epidemiology, 179*(8), 1025–1033.

CHAPTER VI

CONCLUSIONS

The thematic review is intended to provide an orientation to alternative sources of neighborhood boundaries developed outside of academic or government research that might have potential future uses in U.S. neighborhood health research. The ultimate goal of producing this review is to empower researchers to seek out resources of this nature for future studies that might someday revitalize neighborhood health methodology. In order for such research to generate data that can guide effective policy changes and public health interventions, the unit of analysis selected must represent ground truth. Real-world validity must be prioritized over convenience and remaining in one's methodological comfort zone. Gaining access to the private sector's robust toolkit could be a key step toward achieving these objectives in an efficient, expedient manner.

In addition to exploring the aforementioned resources individually, I suggest comparing their performance against traditional units of analysis. For example, one could run two separate linear regressions examining risk factors for lowbirthweight within the same population — aggregated by census tracts and by Zillow or Yelp! boundaries — and compare the respective R^2 values. One could try similar comparisons of neighborhood polygons created by different groups within the same category (e.g., Tapestry vs. Maponics); by groups in different categories (Who's on First vs. Google); or the same tool with different populations or health outcomes. Perhaps one type of polygon might be more appropriate for chronic disease studies while another might be more appropriate for maternal morbidity research. The ultimate goal would be for researchers

to reach consensus on which units of analysis are most appropriate for neighborhood health research (or at least specific subfields) so that future studies can be more consistent and thus comparable.

This is quite possibly the first neighborhood health study to examine the effects of the Modifiable Areal Unit Problem (MAUP) and maternal health to this extent: The sensitivity analysis of three census units commonly used in health research along with three alternative areal units formed by re-aggregating those same census units included an examination of both the scale and the zoning effect. Most neighborhood health studies have evaluated only two or three types of areal unit and/or examined only one of the MAUP components. The process described here can easily be replicated by other researchers and serve as an objective guide for selecting the most appropriate areal unit for a particular neighborhood health study.

Methods used to generate natural neighborhoods (NNs) have varied substantially between studies. Some researchers have used qualitative or mixed methods approaches which almost assuredly fulfill the criteria of being relevant or meaningful to community residents. However, reproducing such methods over large areas of geography would require substantial investments of time, money, and staff. This study adapted a method used by Parenteau & Sawada (2011) that utilizes publicly available data and relatively simple GIS methods. Since the previous study was based in Canada, a neighborhood deprivation index widely used in the U.S. was substituted. This process represents a potential path to generating more realistic neighborhood proxies with relatively little investment of resources.

122

In my regression analyses of prevalence of hypertensive disorders of pregnancy (HDP) and neighborhood deprivation, NNs formed by aggregating census tracts (CTs) into larger, socioeconomically consistent units (CT large NNs) — performed better than all traditional census units. It is important to bear in mind that the analysis encompassed only one major metropolitan area and future studies should seek to validate these findings elsewhere in the United States, particularly in rural areas. Still, this dissertation's findings suggest that census boundaries – which are much more commonly used in health research – might not be ideal neighborhood boundaries in all situations. While units such as the tract and ZIP Code Tabulation Area (ZCTA) are certainly more convenient to obtain, they are not inherently related to a community's inner workings. When analyzing relationships between neighborhood deprivation and health outcomes, researchers should consider using NNs based on socioeconomic variables, which could represent a compromise between qualitative and census-based delineations of neighborhoods. More importantly, sensitivity analyses such as the one conducted in Aim 2 should become common practice so that researchers can begin to use empirical data to select a standard unit of analysis for particular combinations of exposure and outcome to ultimately facilitate synthesis of key findings and the advancement of policy and health interventions.

This dissertation also represents one of the first multilevel analyses of deprivation amplification and HDP. In addition to the well-documented statistical advantages of multilevel analysis (T. A. Brown, 2006), examining area- and individual-level factors simultaneously allows researchers to gain a more complete understanding of complex relationships among multiple risk factors. My area-level unit of analysis, the "natural

neighborhood" — specifically the CT large NN — was chosen to increase socioeconomic homogeneity of clusters with the aim of more closely approximating actual Miami-Dade County neighborhoods.

While I did not find evidence to support the notion of deprivation amplification as an upstream risk factor for HDP, area-level deprivation and two individual-level socioeconomic factors (educational attainment and insurance status) were each individually associated with HDP even after including race/ethnicity and numerous behavioral and medical variables in the model. Moreover, there were two significant antenatal modifiable risk factors: excessive gestational weight gain and smoking during pregnancy — each of which represented an increase in HDP odds of nearly 1.5.

Public health interventions in deprived neighborhoods that target expectant mothers who lack personal resources — particularly non-Hispanic Black and Hispanic women — might help reduce HDP prevalence and disparities. Counseling women on the potential HDP-related consequences of smoking during pregnancy and gaining excessive weight — no matter how low or high one's pre-pregnancy body mass index (BMI) — should also be explored as an intervention strategy. Future studies should seek to replicate these findings with national-level data and more robust measures of deprivation at both the area and individual levels. In-depth exploration of the specific linkages between modifiable factors and deprivation in potentially increasing HDP risk is also paramount.

APPENDIX A

GLOSSARY

Automated Program Interface (API): A set of procedural building blocks (protocols, routines, etc.) that governs how parts of one or more applications relate to and interact with each other. API is the technology that enables integration between websites and apps (e.g., Twitter.com, Twitter feeds on other sites, Twitter app, TweetDeck app, etc.) as well as websites that aggregate data from multiple other sites to create search sites (e.g., Hotels.com or Travelocity).

Boundary: A line that separates adjacent political entities (e.g., districts, counties) or geographic zones (e.g., ecosystems). A boundary may or may not follow physical features such as rivers or mountains.

Census Units

Delineated by the United States Census Bureau

> **Block:** The basis for all tabulated U.S. data. Blocks nest within all other census geographic entities and are bounded visible features (e.g., streets, railroad tracks) and invisible boundaries (e.g., administrative boundaries, selected property lines, short line-of-sight extensions of streets). Generally, census blocks are small (e.g., a city block) but can be irregular and large in suburban or rural areas, sometimes encompassing hundreds of square miles. To protect confidentiality, socioeconomic data are not available at the block level.

> **Block group (BG):** A cluster of blocks within a census tract that typically covers a contiguous area. Block groups usually have a population of 600-3,000 people and

never cross state, county, or census tract lines but can cross the boundaries of any other geographic entity. This is the smallest census unit at which area-level socioeconomic data are provided.

Census tract (CT): Relatively permanent subdivision of a U.S. county or equivalent entity that typically follows a visible, identifiable geographic feature and usually covers a contiguous area. Because census tracts are delineated to contain a specific population size (1,200-8,000 people, optimally 4,000 people) their size varies widely depending on population density of an area. Occasionally, tracts are combined because of substantial population decline or split because of population growth.

ZIP Code Tabulation Area (ZCTA): Approximation of U.S. Postal Service (USPS) five-digit ZIP Code service area created from census blocks. Each census block is assigned to a single ZCTA, typically the one that reflects the most frequently occurring ZIP Code for addresses within that block.

Centroid: A feature's geometric center.

Feature: A representation of a real-world object (e.g., building, ocean, street) on a map.

Gazetteer: A list of geographic place names along with their coordinates. Gazetteer entries can also include other information as well, such as area, population, or cultural statistics.

Geocoding: A GIS operation for converting a location description (e.g., address, place name, coordinates) into spatial data that can be displayed as features on a map and used in spatial analysis.

Geographically Weighted Regression (GWR): A local version of spatial regression that provides a local model of a relationship between 2 or more variables by fitting a regression equation to every feature in the dataset. This allows assessment of the spatial heterogeneity in the estimated relationships between variables.

Geographic Information System (GIS): Computer-based mapping and analysis tool that combines common database operations (e.g., query, statistical analysis) with data visualization and geographic analysis unique to maps.

Ground truth: The accuracy of mathematically calculated or remotely sensed data based on information gathered "on the ground" (in the field).

Line: A shape defined by a series of unique, connected x,y coordinate pairs. A line can be straight or curved.

Modifiable Areal Unit Problem (MAUP): A type of spatial misclassification bias that arises from imposing artificial boundaries onto an area. Two components comprise the MAUP, the scale effect and the zoning effect (See definitions elsewhere in this glossary). The MAUP can influence the magnitude and direction of a measure of association.

Natural neighborhood: Groups of census units (e.g., block groups, census tracts, etc.) that have been re-aggregated into more socioeconomically homogeneous clusters with the

goal of more closely approximating actual communities. In this instance, the word "natural" is not used in the typical, vernacular sense, and the authors are in no way suggesting that neighborhood deprivation is natural. Instead, "natural" in this instance denotes "meaningful" or "locally relevant." It also refers to the process by which these neighborhoods are formed: ArcGIS software searches for natural statistical patterns or clusters based on values of one or more variables, such as neighborhood deprivation. The software seeks a solution that maximizes within-group similarity and maximizes between group variability.

Point: A geographic element defined by a pair of x,y coordinates.

Polygon: A closed shape defined by a connected sequence of x,y coordinate pairs in which the first and last pair of coordinates are the same and all others are unique.

Scale effect: One of the two aspects of the Modifiable Areal Unit Problem (MAUP), it is the extent to which the size of areal units influences research findings.

Shapefile: The standard GIS file format for storing information on the location, shape, and attributes of geographic features. A shapefile is stored in a set of related files and contains one feature class.

Zoning effect: One of the two aspects of the Modifiable Areal Unit Problem (MAUP), it is the extent to which the aggregation of areal units influences research findings.

VITA

KELLY MICHELLE WINTER

Born, Spartanburg, South Carolina

2000–2008	Copy Editor and Page Designer
	Tallahassee Democrat and *The Palm Beach Post*
	Tallahassee, Florida and West Palm Beach Florida
2006–2007	Associate of Arts (A.A.)
	Tallahassee Community College
	Tallahassee, Florida
2006–2007	Bachelor of Arts, Psychology (B.A.**)**
	Florida Atlantic University
	Boca Raton, Florida
2009–2012	Master of Public Health (M.P.H.)
	Florida International University
	Miami, Florida
2010–2013	Graduate Assistant and Instructor
	Florida International University
	Miami, Florida
2010–present	Doctoral Candidate
	Florida International University
	Miami, Florida
2013-present	Health Scientist
	Centers for Disease Control and Prevention (CDC)
	Atlanta, Georgia

PUBLICATIONS AND PRESENTATIONS (Selected)

Winter, K., Lee, C.V., O'Sullivan, M. Henry, R., Holton, K., Brunette, G. Late to the party: Online health guidance for mass gatherings draws more traffic near start of events. Poster presentation at 15th Annual Conference of the International Society of Travel Medicine. May 16, 2017, Barcelona, Spain.

Winter, K., Li, T., Maziak, W., Trepka, M., Onsted, J., Madhivanan, P. The weight of a place: Comparing the relationship between hypertension in pregnancy and neighborhood deprivation at three geographical scales Poster presentation at Epidemiology Congress of the Americas. June 22, 2016, Miami, Florida.

Joseph, H., Wojno, A.E., Grady-Erickson, O., Winter, K., et al The Check and Report Ebola Plus (CARE+) Program: The power of interpersonal communication in reaching mobile populations. Oral presentation at National Conference on Health Communication, Marketing, and Media, Aug. 12, 2015, Atlanta, Georgia.

Winter, K. Mapping the relationship between low birth weight and neighborhood deprivation in Miami-Dade County, Florida. Oral presentation at APHA Meeting, Nov. 18, 2014, New Orleans, La.

Winter, K. & Onsted, J. Diabetes mortality is associated with residential segregation in Miami, Florida. Oral presentation at Esri International Users Conference, July 11, 2013, San Diego, Calif.

Winter, K. (2014). Hardcore smoking does not necessarily indicate hardening. *Addiction, 109*(4), 681-681.

Ahmed, N.U., Pelletier, V., Winter, K., Albatineh, A. (2013). Factors explaining racial/ethnic disparities in rates of physician recommendation for colorectal cancer screening. *American Journal of Public Health, 103*(7), e91-e99.

Ahmed, N.U., Winter, K., Albatineh, A., & Haber, G. (2012). Clustering very low-income, insured women's mammography screening barriers into potentially functional subgroups. *Women's Health Issues, 22*(3), 259-266.

Shehadeh, N., McCoy, H.V., Rubens, M.B., Batra, A., Renfrew, R., & Winter K. (2012). The impact of ethnic identity on changes in high risk HIV behaviors in sexually active migrant workers. *Journal of Immigrant and Minority Health, 14*(1), 100-106.

Winter, K. (2012). Impulsivity. In Gellman, M., & Turner, J., (Eds.), *Encyclopedia of Behavioral Medicine*. Springer Reference. Available at www.springerreference.com/docs/html/chapterdbid/345551.html