Methodological Approaches to Account for Residential Self-Selection and Time-Varying Confounding in the Association Between the Neighborhood Environment and Cardiovascular Disease

by

Yamini Kesavan

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Epidemiological Science)
in the University of Michigan
2013

Doctoral Committee:

Professor Ana V. Diez Roux, Chair
Professor Jeffrey D. Morenoff
Associate Professor Lynda D. Lisabeth
Assistant Professor Brisa N. Sánchez

Dedication

To my parents, for their unwavering love and support,

for instilling in me a respect and love for learning,

and for always believing in me.

I would also like to acknowledge my study group and friends from the Harvard School of Public Health (HSPH): Jaime, Erica, Andrea, Holly, Kathleen, and Jen. It was at HSPH that I really discovered epidemiology, and observing their hard work in the doctoral program inspired me to pursue a doctorate. I will always treasure our friendship.

I thank my family, near and far, for standing with me through this journey. To my parents: you have shown me that with the love and support of family, anything is possible. To my brother: through each step of life, you have been my constant companion and my best friend, challenging me to push my boundaries and live in the moment. To my aunt: you inspire me every day to be strong and resilient. Finally, to Tushar: You give me the strength to face each day with confidence and without fear, and you inspire me to follow my dreams. Words cannot express the light and joy you have brought into my life.

Table of Contents

List of Figures

List of Tables

ABSTRACT

Despite the growing body of research investigating the relationship between the neighborhood environment and cardiovascular disease-related outcomes, many studies have methodological limitations. A major challenge in the study of neighborhood exposures and health-related outcomes arises due to the possibility that individuals may select where to live based on individual behaviors or preferences (i.e. residential self-selection). Another challenge is the need to account for time-varying confounding, which may result when neighborhood exposures and individual-level factors vary over time, possibly acting as both confounders and mediators. This dissertation used longitudinal data from the Multi-Ethnic Study of Atherosclerosis (MESA) to examine: 1) whether individual physical activity and diet quality were related to selection of neighborhoods based on availability of physical activity and favorable food resources; 2) whether changes in the availability of physical activity resources was related to changes in individual physical activity; and 3) whether neighborhood socio-economic composition was related to incident cardiovascular disease after accounting for time-varying confounding. The findings show that among individuals who moved, those who were more physically active moved to neighborhoods with greater availability of physical activity resources. The results also show that changes in physical activity resources over time were related to changes in physical activity over time, with a stronger association observed among older adults. Overall, these results suggest that individuals select their neighborhoods based on physical activity behaviors, but at the same time changes in physical activity resources over time are associated with

simultaneous changes in the physical activity of residents. The estimation of the relationship between neighborhood socio-economic composition and incident cardiovascular disease revealed that increasing neighborhood deprivation was associated with a greater risk of incident cardiovascular disease. This association was similar after using methods to account for time-variation in both the exposure and covariates. Using longitudinal methods to account for residential selection and time-varying confounding in the context of neighborhoods and health provides more robust evidence that can inform policies aimed at improving residential environments to help promote healthy living.

CHAPTER 1

Introduction

Over the past several years, many studies have investigated the role of neighborhoods and residential environments in health (1-4). The impetus for this research stems from growing evidence that suggests that in addition to considering individual risk factors with respect to individual health outcomes, considering the factors related to the environments in which individuals live may also be important in predicting health outcomes. The theory that residential environments may be related to health outcomes has been explored in the context of cardiovascular disease and its risk factors (5-7), and several mechanisms related to the physical and social environments of neighborhoods have been suggested to be involved in this relationship (5). Various measures of neighborhood socio-economic composition have been found to be associated with cardiovascular disease and its risk factors, independent of individual-level socio-economic characteristics (6, 8-12). In addition, features of the neighborhood physical environment, such as access to healthy foods and recreational facilities, have been shown to be associated with individual diet, physical activity, BMI, and cardiovascular events (13-19).

Despite the growing body of research investigating the relationship between the neighborhood environment and cardiovascular disease (CVD) and its risk factors, many studies are cross-sectional and/or have methodological limitations. A major challenge associated with conducting cross-sectional and some longitudinal studies in the context of the neighborhood environment

and CVD-related outcomes arises due to the possibility of residential self-selection. This is because individuals may self-select where they live based on individual characteristics or preferences which, if related to the both the neighborhood construct and the outcome of interest, may lead to a bias in the resulting effect estimates. For example, several studies have found that a better walking environment is associated with higher levels of walking (20-23), that a better physical environment (access to parks, recreational facilities) is associated with higher levels of physical activity (14, 15, 24, 25), and that better access to healthy foods is associated with a better diet (19, 26). However, the statistical associations reported in these studies may not reflect causal relationships if there are reciprocal relationships between the built environment and the health-related outcomes. It is possible that individuals that walk more frequently, are more physically active, and have better diets choose to live in neighborhoods that have better walking environments, better access to resources that support physical activity, and better access to healthy foods, respectively. Although selection into neighborhoods based on certain health-related attributes is often noted as a major methodological challenge and has been raised as an argument for the limited causal inference that can be made from studies on the built environment and health-related behaviors, whether these individual behaviors predict residential mobility (whether individuals move) and residential selection (to which types of neighborhoods individuals move) has not been investigated empirically. Addressing this issue of residential self-selection necessitates quantifying the predictors of residential mobility and selection and accounting for these processes in longitudinal analyses to better understand if specific personal characteristics, behaviors, or preferences are related to the types of neighborhoods in which people live or to which people move. A few recent studies have tried to address the self-selection issue by measuring and adjusting for individuals' preferences for healthy behaviors and

for living in areas that support these behaviors (27-31), finding mixed results for the influence of self-selection on the observed estimates. However, thus far no studies have examined whether the actual health-related behaviors of individuals, such as physical activity and diet, predict whether individuals select neighborhoods that support these behaviors.

Investigating the presence of self-selection into particular neighborhoods due to CVD-related health behaviors is just one step to furthering our understanding of the relationship between neighborhoods and CVD-related outcomes. Designing longitudinal studies that are able to account for changes in the neighborhood environment and covariates in addition to changes in the outcome are also needed to obtain estimates that are less likely to be biased due to self-selection. A specific area of research in which more robust longitudinal studies are needed is in the study of the neighborhood environment and individual physical activity. Physical activity is an important established behavioral risk factor for CVD (32, 33), and several cross-sectional studies and a few recent longitudinal studies have investigated the role of the neighborhood physical environment on individual physical activity levels (15, 21, 23, 30, 31, 34-40). Many of these studies have found positive associations between perceived or objective measures of the neighborhood physical environment and a few studies have found weak or null associations. Limitations of these studies include 1) incorrectly defining or measuring the neighborhood exposure in relation to the type of physical activity studied, and 2) not accounting for changes in the outcome, exposure, and covariates over time, which can adversely impact the causal interpretation of the findings. For this reason, longitudinal analyses are needed to investigate whether individuals who are exposed to better neighborhood environments experience favorable *changes* in physical activity over time. However, in addition to individual behavioral outcomes

3

changing over time, neighborhoods themselves change over time and evolve based on economic cycles, demographic shifts, migration, and other factors. Therefore, longitudinal analyses that follow changes in neighborhoods and changes in individuals over time are needed (36, 41). In particular, longitudinal analyses which investigate how changes in specific neighborhood environmental attributes are associated with changes in related physical activity outcomes is necessary to capture if the neighborhood environment is influencing individual behavior. This type of *change versus change* analysis is uncommon in neighborhood health effects research, and is needed to address limitations that arise in cross-sectional analyses and longitudinal analyses with only single point in time measures of the exposure and/or covariates.

An additional challenge related to residential self-selection arises in the study of the neighborhood environment with respect to CVD event outcomes due to the possibility of residual confounding and time-varying confounding when conducting longitudinal analyses. This additional complexity arises if the exposure and/or covariate values change over time and if neighborhood exposures are simultaneously consequences and causes of time-varying intermediates (Figure 3.1). For example, individual physical activity may affect where a person lives (i.e. selection into neighborhoods as previously described) but may also be a consequence of prior neighborhood exposures. Under these circumstances, adjustment for these confounders using traditional methods may yield biased estimates of neighborhood effects. Special methods such as marginal structural models (MSMs) are needed to account for these processes, but these methods have rarely been used to investigate neighborhood health effects. Sampson et al (42) used MSMs to investigate the relationship between living in a disadvantaged neighborhood and verbal ability in African American children, and Cerda et al (43) used MSMs to estimate the

4

association between neighborhood poverty and alcohol use in young adults. However, the application of MSMs has not yet been explored in the context of neighborhoods and CVD outcomes. The consideration of cumulative effects, rather than a point in time exposure, is also of interest. For example, the effect of the neighborhood environment on CVD risk may only be observable after exposure to an unfavorable neighborhood environment over a longer period of time. However, few longitudinal datasets have time-varying neighborhood data available to contrast results using cumulative versus single point in time measures.

Summary and Specific Aims

Though recent research has begun to investigate the associations discussed above using longitudinal designs, many of these studies have faced challenges related to accounting for self-selection of individuals into particular neighborhoods. Among the reasons for these challenges include incomplete or unavailable data on time-varying measures on one or more of the following: 1) neighborhood exposures, 2) individual-level socio-economic factors, 3) individual-level behavioral and CVD risk factors, and 4) individual-level residential mobility and selection (when and to where individuals moved). Challenges to previous work related to self-selection and time-varying confounding have limited the ability to establish causal relationships between the neighborhood environment and CVD and its risk factors. Hence, further research in this area, addressing these methodological limitations, is warranted.

This dissertation uses longitudinal data from the Multi-Ethnic Study of Atherosclerosis (MESA), with available time-varying measures on residential mobility, the neighborhood environment, and individual-level socio-economic, behavioral, and biomedical factors, to investigate the

relationship between neighborhood of residence and CVD-related outcomes while addressing the limitations outlined above. Specifically, the aims of this dissertation are to 1) investigate and quantify the magnitude of the association of health-related behaviors, specifically physical activity and diet, with residential mobility and residential selection (Aim 1), 2) investigate the relationship between change in an objective measure of neighborhood physical environment and change physical activity using appropriate longitudinal methods (Aim 2), and 3) investigate associations of long-term cumulative and recent exposure measures of neighborhood environments with CVD outcomes while accounting for time-varying confounding (Aim 3). The 3 aims of the dissertation and associated hypotheses are detailed below.

SPECIFIC AIM 1

To examine whether selected health behaviors at baseline (specifically individual physical activity and diet) are predictors of residential mobility among MESA participants by a) investigating whether physical activity and diet predict whether or not participants move during follow-up, b) investigating, among movers, whether physical activity and diet predict the type of neighborhood environment to which participants move and c) investigating, among non-movers, whether physical activity and diet predict whether and how the participants' neighborhood environments change during follow-up.

Hypotheses

1) Individuals' physical activity and diet at baseline will be associated with whether participants move during follow-up.

a. Individuals who are less physically active, as measured by self-reported physical activity, will be more likely to move than individuals who are more physically active.

b. Individuals with poorer diets will be more likely to move than individuals with better diets.

2) Among movers, individuals' physical activity and diet at baseline will be associated with recreational facility and food store availability in the neighborhood environments to which they move.

a. Movers who are more physically active will be more likely to move to neighborhoods with better availability of recreational facilities than movers who are less physically active.

b. Movers with better diets will be more likely to move to neighborhoods with better availability of healthy food stores than movers with poorer diets.

3) Among non-movers, individuals' physical activity and diet at baseline will be associated with a change in recreational facility and food store availability over time in the neighborhood environments in which they stay.

a. The change in availability of recreational facilities over time will be more favorable among non-movers who are more physically active as compared to non-movers who are less physically active.

b. The change in neighborhood availability of healthy food stores over time will be more favorable among non-movers with better diets as compared to non-movers with poorer diets.

SPECIFIC AIM 2

To investigate the relationship between changes in built environment measures of neighborhood physical activity resources, specifically access to recreational facilities, and changes in individual physical activity longitudinally in MESA.

Hypotheses

1) Better availability of recreational facilities at baseline will be associated with a change in physical activity over time.

   a. Individuals living in neighborhoods with better availability of recreational facilities at baseline will experience a more favorable change in physical activity over time as compared to individuals living in neighborhoods with poorer availability recreational facilities.

2) Changes in availability of recreational facilities over time will be associated with changes in individual physical activity over time.

   a. Individuals living in neighborhoods with greater improvements in availability of recreational facilities over time will experience greater increases in physical activity over time as compared to individuals living in neighborhoods with lesser improvements in availability of recreational facilities over time.

SPECIFIC AIM 3

To investigate the relationship between neighborhood environment and incident CVD and to further our understanding of this association by a) using cumulative measures to characterize neighborhood environment and b) using more complex modeling techniques that account for

time-varying confounding, specifically marginal structural models (MSMs), to determine the importance of time-varying confounding in comparison to using traditional regression methods.

Hypotheses

1) Living in a neighborhood of greater disadvantage will be associated with a higher risk of incident CVD.

   a. Participants living in neighborhoods of greater disadvantage, as measured by a composite measure of neighborhood socioeconomic composition, will have a higher risk of CVD.

2) The strength of the association between neighborhood environment and incident CVD will differ depending on the timing and duration of exposure.

   a. The association between cumulative exposure to neighborhood environment (between baseline and time t) will be stronger than the association between recent exposure (exposure over only the last one year of follow-up) and incident CVD.

3) The inverse association between neighborhood environment and incident CVD will be attenuated after adjusting for time-varying covariates, including individual socio-economic factors and CVD risk factors, using traditional methods.

4) If time-varying confounding is present, using marginal structural models that account for time-varying confounding will result in estimates that differ meaningfully from those obtained from traditional methods.

   a. In the presence of time-varying confounding, marginal structural models will result in stronger estimates of effect.

b. In the absence of time-varying confounding, there will be no meaningful difference between estimates obtained from marginal structural models and those obtained from traditional methods.

Conceptual Model

To better understand and navigate the complex relationships between the variables involved when examining the relationship between the neighborhood environment and CVD, we used the following conceptual model. This conceptual model helped to guide the development of hypotheses and models for the aims of this dissertation.

Figure 1.1 – Conceptual model: Hypothesized causal relationships (solid lines) and possible bi-directional associations between the neighborhood environment, individual-level factors, and cardiovascular disease.
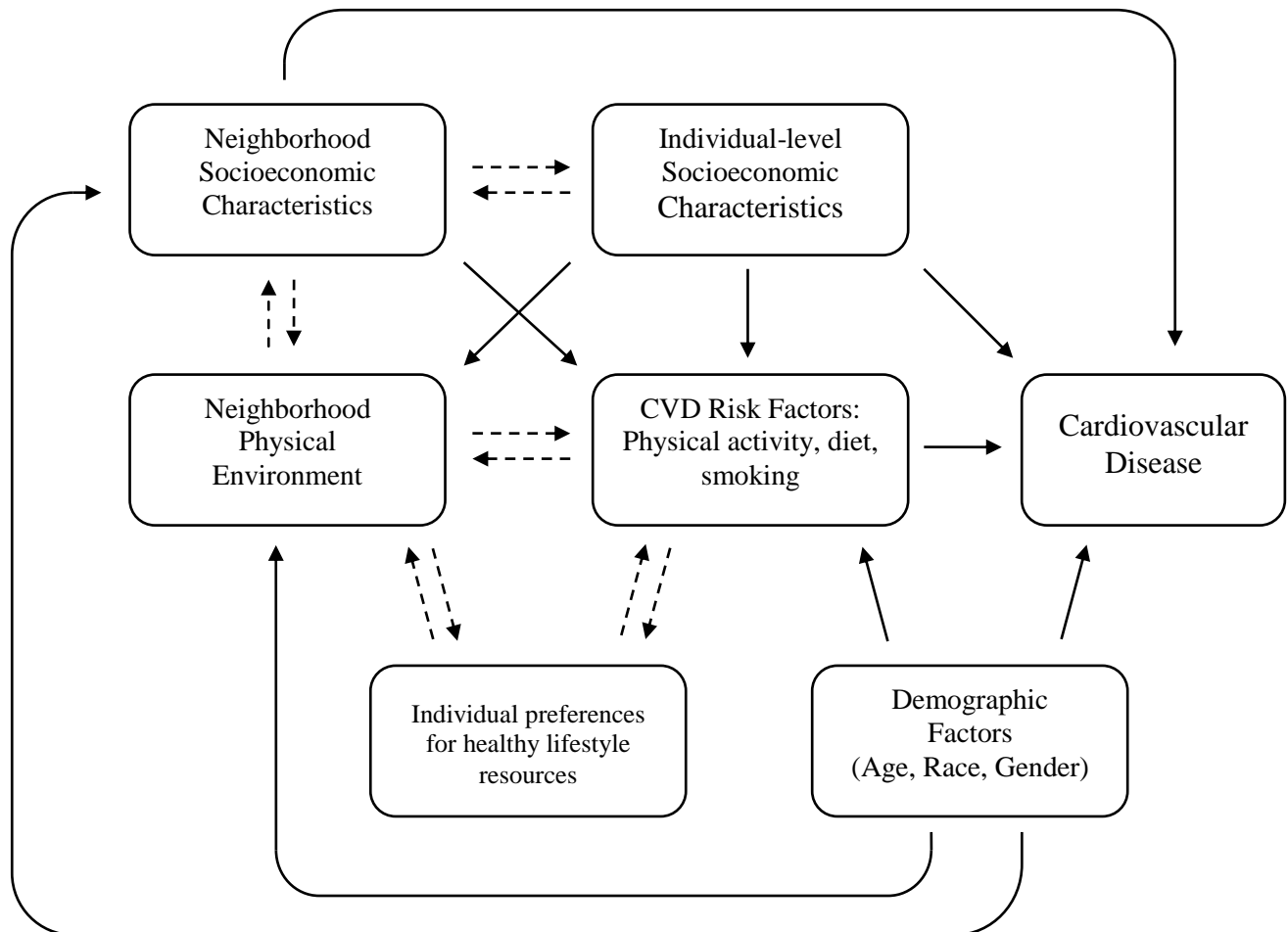
Complexities related to residential selection may arise due to the possibility of reciprocal relationships between neighborhood exposures and individual-level factors, as indicated by the bi-directional arrows between variables such as neighborhood environment and physical activity. For example, a measure of the physical environment, such as availability of recreational facilities, may influence individual physical activity either directly or by changing an individual's preference to engage in physical activity, perhaps because the exposure to facilities and other individuals who are using these facilities increases the motivation to exercise. However, individual physical activity may influence the neighborhood physical environment either directly (physically active individuals may demand better access to resources in their neighborhoods, or businesses may be more likely to build facilities in places where individuals are more likely to use them) or through influencing individuals' preferences to live in neighborhoods with better availability of facilities. These relationships may be confounded by individual-level socio-economic or demographic factors, or possibly by neighborhood-level socio-economic factors.

In addition, when examining the relationship between the neighborhood environment and CVD, it is important to note that though the stable demographic covariates, such as age, gender, and race/ethnicity can be conceptualized as classic confounders, individual-level socio-economic factors (income, education), behavioral factors (physical, activity, diet, smoking), and biomedical risk factors for CVD (diabetes, hypertension) may have potential reciprocal relationships with the neighborhood environment. Extending this concept to longitudinal data, if these variables change over time, then time-varying confounding is a possibility (Figure 3.1). These variables may be influenced by prior exposure, but may in turn influence future exposure.

CHAPTER 1 REFERENCES

1. Diez Roux A. Investigating neighborhood and area effects on health. *Am J Public Health* 2001;91(11):1783-9.
2. Diez Roux A, Mair C. Neighborhoods and health. *Ann N Y Acad Sci* 2010;1186:125-45.
3. Pickett KE, Pearl M. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. *J Epidemiol Community Health* 2001;55:111-22.
4. Sampson R. The neighborhood context of well-being. *Perspect Biol Med* 2003;46(3 Suppl):S53-64.
5. Diez Roux A. Residential environments and cardiovascular risk. *J Urban Health* 2003;80(4):569-89.
6. Winkleby M, Sundquist K, Cubbin C. Inequities in CHD incidence and case fatality by neighborhood deprivation. *Am J Prev Med* 2007;32(2):97-106.
7. Cubbin C, Sundquist K, Ahlén H, et al. Neighborhood deprivation and cardiovascular disease risk factors: protective and harmful effects. *Scand J Public Health* 2006;34(3):228-37.
8. Diez Roux A, Merkin S, Arnett D, et al. Neighborhood of residence and incidence of coronary heart disease. *N Engl J Med* 2001;345(2):99-106.
9. Sundquist K, Winkleby M, Ahlén H, et al. Neighborhood socioeconomic environment and incidence of coronary heart disease: a follow-up study of 25,319 women and men in Sweden. *Am J Epidemiol* 2004;159(7):655-62.
10. Chaix B, Rosvall M, Merlo J. Neighborhood socioeconomic deprivation and residential instability: effects on incidence of ischemic heart disease and survival after myocardial infarction. *Epidemiology* 2007;18(1):104-11.
11. Diez-Roux A, Nieto F, Muntaner C, et al. Neighborhood environments and coronary heart disease: a multilevel analysis. *Am J Epidemiol* 1997;146(1):48-63.
12. Borrell L, Diez Roux A, Rose K, et al. Neighbourhood characteristics and mortality in the Atherosclerosis Risk in Communities Study. *Int J Epidemiol* 2004;33(2):398-407.
13. Owen N, Humpel N, Leslie E, et al. Understanding environmental influences on walking; Review and research agenda. *Am J Prev Med* 2004;27(1):67-76.
14. Hoehner C, Brennan Ramirez L, Elliott M, et al. Perceived and objective environmental measures and physical activity among urban adults. *Am J Prev Med* 2005;28(2 Suppl 2):105-16.
15. Diez Roux A, Evenson K, McGinn A, et al. Availability of recreational resources and physical activity in adults. *Am J Public Health* 2007;97(3):493-9.
16. Morland K, Wing S, Diez Roux A. The contextual effect of the local food environment on residents' diets: the atherosclerosis risk in communities study. *Am J Public Health* 2002;92(11):1761-7.
17. Morland K, Diez Roux A, Wing S. Supermarkets, other food stores, and obesity: the atherosclerosis risk in communities study. *Am J Prev Med* 2006;30(4):333-9.
18. Mujahid M, Diez Roux A, Shen M, et al. Relation between neighborhood environments and obesity in the Multi-Ethnic Study of Atherosclerosis. *Am J Epidemiol* 2008;167(11):1349-57.
19. Moore L, Diez Roux A, Nettleton J, et al. Fast-food consumption, diet quality, and neighborhood exposure to fast food: the multi-ethnic study of atherosclerosis. *Am J Epidemiol* 2009;170(1):29-36.

20. Saelens B, Handy S. Built environment correlates of walking: a review. *Med Sci Sports Exerc* 2008;40(7 Suppl):S550-66.

21. Rodríguez D, Evenson K, Diez Roux A, et al. Land use, residential density, and walking. The multi-ethnic study of atherosclerosis. *Am J Prev Med* 2009;37(5):397-404.

22. Rhodes R, Brown S, McIntyre C. Integrating the perceived neighborhood environment and the theory of planned behavior when predicting walking in a Canadian adult sample. *Am J Health Promot* 2006;21(2):110-8.

23. Wendel-Vos W, Droomers M, Kremers S, et al. Potential environmental determinants of physical activity in adults: a systematic review. *Obes Rev* 2007;8(5):425-40.

24. McNeill L, Wyrwich K, Brownson R, et al. Individual, social environmental, and physical environmental influences on physical activity among black and white adults: a structural equation analysis. *Ann Behav Med* 2006;31(1):36-44.

25. Powell L, Chaloupka F, Slater S, et al. The availability of local-area commercial physical activity-related facilities and physical activity among adolescents. *Am J Prev Med* 2007;33(4 Suppl):S292-300.

26. Moore L, Diez Roux A, Nettleton J, et al. Associations of the local food environment with diet quality--a comparison of assessments based on surveys and geographic information systems: the multi-ethnic study of atherosclerosis. *Am J Epidemiol* 2008;167(8):917-24.

27. Cao X, Handy SL, Mokhtarian PL. The influences of the built environment and residential self-selection on pedestrian behavior: evidence from Austin, TX. *Transportation* 2006;33:1-20.

28. Bagley MN, Mokhtarian PL. The impact of residential neighborhood type on travel behavior: A structural equations modeling approach. Annals of Regional Science. *36* 2002;2.

29. Handy S, Cao X, Mokhtarian P. Self-Selection in the Relationship between the Built Environment and Walking: Empirical Evidence from Northern California. *Journal of the American Planning Association* 2006;72(1):55-74.

30. Pinjari AR, Bhat CR, Hensher DA. Residential self-selection effects in an activity time-use behavior model. *Transportation Research Part B* 2009;43:729–48.

31. Schwanen T, Mokhtarian PL. What affects commute mode choice: Neighborhood physical structure or preferences toward neighborhoods? *Journal of Transport Geography* 2005;13(1):83–99.

32. Batty GD. Physical activity and coronary heart disease in older adults. A systematic review of epidemiological studies. *European journal of public health* 2002;12(3):171-6.

33. Batty GD, Lee IM. Physical activity and coronary heart disease. *Bmj* 2004;328(7448):1089-90.

34. Ball K, Bauman A, Leslie E, et al. Perceived environmental aesthetics and convenience and company are associated with walking for exercise among Australian adults. *Prev Med* 2001;33(5):434-40.

35. Brownson R, Kelly C, Eyler A, et al. Environmental and policy approaches for promoting physical activity in the United States: a research agenda. *J Phys Act Health* 2008;5(4):488-503.

36. Cook T, Shagle S, Degirmencioglu S. Capturing social process for testing mediational models of neighborhood effects. In: Brooks-Gunn G DG, Aber J, ed. *Neighborhood*

poverty: context and consequences for children. New York, NY: Russell Sage Foundation Press, 1997:94–119.

37. Handy S, Boarnet M, Ewing R, et al. How the built environment affects physical activity: views from urban planning. *Am J Prev Med* 2002;23(2 Suppl):64-73.

38. Saelens B, Sallis J, Frank L. Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. *Ann Behav Med* 2003;25(2):80-91.

39. Sallis J, Cervero R, Ascher W, et al. An ecological approach to creating active living communities. *Annu Rev Public Health* 2006;27:297-322.

40. Troped P, Saunders R, Pate R, et al. Associations between self-reported and objective physical environmental factors and use of a community rail-trail. *Prev Med* 2001;32(2):191-200.

41. O'Campo P. Invited commentary: Advancing theory and methods for multilevel models of residential neighborhoods and health. *Am J Epidemiol* 2003;157(1):9-13.

42. Sampson RJ, Sharkey P. Neighborhood Selection and the Social Reproduction of Concentrated Racial Inequality. *Demography* 2008;45(1):1-29.

43. Cerda M, Diez-Roux AV, Tchetgen ET, et al. The relationship between neighborhood poverty and alcohol use: estimation by marginal structural models. *Epidemiology* 2010;21(4):482-9.

CHAPTER 2

The association between health behaviors and both residential mobility and residential selection in the Multi-Ethnic Study of Atherosclerosis (MESA)

INTRODUCTION

A large number of observational studies have documented associations between neighborhood factors and health outcomes (1-3), resulting in growing interest in neighborhood-level interventions to improve population health and reduce health disparities. However, important methodological challenges have raised persistent questions regarding whether the observed associations reflect causal processes. Among these methodological challenges is the possibility of residential self-selection (4), whereby individuals select where they live based on personal characteristics or preferences which, if related to the both the neighborhood construct and the outcome of interest, may lead to biased estimates of the true relationship between a neighborhood-level exposure and a related outcome. For example, several studies have found that a better physical environment (access to parks, recreational facilities) is associated with higher levels of physical activity (5-8) and that greater healthy food availability is associated with a better diet (9, 10). However, the statistical associations reported in these studies may not reflect causal relationships. It is possible that individuals who are more physically active and have higher quality diets choose to live in neighborhoods that have greater access to resources that support physical activity and greater access to healthy food stores, respectively.

Although selection into neighborhoods based on health-related attributes (including predispositions to health-related behaviors) is often noted as a major methodological challenge, it is rarely investigated empirically. Quantifying the presence and magnitude of the impact of behaviors on residential mobility, and specifically whether healthy behaviors are associated with mobility into environments that support these behaviors, is important for the interpretation of the estimates of neighborhood effects on health derived from observational studies.

There is a large literature on socio-demographic predictors of residential mobility. Studies have shown that race, income, age of head of household, and life cycle factors are associated with residential mobility (9, 10). Past work often emphasized housing needs at various stages of the life cycle as one of the principal determinants of residential mobility, but placed less emphasis on other individual-level factors, such as behavioral factors, individual preferences, or community contexts (11). Factors related to health status have also been found to be related to residential mobility (12). Overall, existing evidence in the U.S. suggests that older age, higher income, minority race, home ownership, longer duration of residence, and better health are associated with lower residential mobility (13, 14).

A growing body of work has begun to examine the impact of other factors, such as preference for certain kinds of health-related environments, on residential mobility. An early study found that accessibility of resources was not strongly predictive of residential mobility (whether individuals moved) or residential selection (the types of neighborhoods to which individuals moved), but the measures of accessibility were self-reported (15). A few studies have investigated the role of personal lifestyle preferences in relation to residential selection. Using structural equations

16

modeling, Bagley et al (16) found that the cross-sectional relationship between neighborhood environment and walking was largely attributable to individuals with certain lifestyle preferences self-selecting into certain types of neighborhoods. However, more recently, using different cross-sectional data from the same area, Schwanen et al (17) observed that the type of neighborhood which people indicate they prefer does not always reflect the type of neighborhood in which they actually live, and found that neighborhood environment was associated with travel behavior after accounting for individual preferences. In another study, Handy et al (18) found that indicating a preference for physical activity options in the neighborhood and for having stores within walking distance were related to walking, but neighborhood characteristics remained significantly associated with walking behavior after accounting for these preferences, suggesting that the built environment had a direct effect on walking behavior.

Given the uncertainty regarding the role of residential self-selection in the study of neighborhood effects, understanding the predictors of residential mobility and selection, especially in a longitudinal framework, is critical to elucidating if residential self-selection may play a role in the association between the built environment and health-related outcomes. Despite the large number of studies linking neighborhood environments to key health behaviors like diet and physical activity (1-3, 9, 19-21), no studies to our knowledge have examined the effects of physical activity and diet on residential mobility and residential selection.

We used longitudinal data from a large, multi-ethnic study of adults aged 45-84 years old at baseline to gain insight into the impact of residential self-selection on the relationship between the built environment and both physical activity and diet. We examined whether individual

17

physical activity and diet were associated with whether or not individuals moved, and whether movers reporting more physical activity and higher quality diets were more likely to move to neighborhoods with more favorable physical activity and healthy food resources than those with worse physical activity and diets. In addition, because it is plausible that "health conscious" residents effect changes in neighborhood environments and/or choose to continue living in neighborhoods that are changing favorably over time (which could also result in biased estimates of neighborhood health effects from observational studies), we also examined whether non-movers reporting more physical activity and higher quality diets were more likely to live in neighborhoods that experienced favorable changes in physical activity and food environments over time.

METHODS

Study sample

The sample consisted of participants from the Multi-Ethnic Study of Atherosclerosis (MESA), a longitudinal study of cardiovascular disease conducted at six sites (Baltimore, Maryland; Chicago, Illinois; Forsyth County, North Carolina; Los Angeles, California; New York, New York; and St. Paul, Minnesota). MESA comprises a multi-ethnic (approximately 38% white, 28% African-American, 23% Hispanic, 11% Chinese descent) population-based sample of 6,814 men and women aged 45-84 years who were free of clinical cardiovascular disease at baseline. Data was collected at baseline (between 2000 and 2002) with 4 subsequent follow-up visits approximately every two years. We included participants from all six sites who had data available on physical activity and diet at baseline. The mean follow-up time was 9.3 years.

Analysis variables

Residential mobility and selection:

Two distinct but related outcomes were investigated: 1) whether or not an individual moved (henceforth referred to as residential mobility) and 2) the type of neighborhood to which an individual moved (henceforth referred to as residential selection) (15). We evaluated these separately since the factors that may predict whether or not individuals move may be different from the factors that predict to which kinds of neighborhoods individuals move. For each MESA participant, information on the number of moves during follow-up and the new census tract to which the participant moved were available. Residential mobility was categorized as yes/no, based on whether the participant moved at least once during follow-up. Approximately 32% of MESA participants moved at least during once during follow-up. Of those that moved, 66% moved only once, 21% moved twice, 13% moved more than 2 times, and over 90% of movers changed census tracts.

Physical activity: Physical activity was self-reported at baseline via a semi-quantitative questionnaire. Previous work has provided evidence for the validity and reliability of this survey, which was adapted from the Cross-Cultural Activity Participation Study (22). Participants were asked about the types of activities they did "in a typical week in the past month". Some of these activities were sub-categorized by level of intensity (light, moderate, heavy). For each activity for which they marked yes, participants were asked how many days in a week they did this activity and the average time per day they did the activity in hours and minutes. Minutes of activity were summed for each activity type and multiplied by its metabolic equivalent (23). We used a summary measure of physical activity which included recreational

activities corresponding to moderate and vigorous exercise. We calculated recreational physical activity, measured in MET-minutes per week (MET-min/week), to include the following: walking for exercise, dance, team sports (e.g., softball, basketball, soccer, volleyball), dual sports (e.g., tennis, racquetball), individual activities (e.g., golf, bowling, yoga, tai chi) and conditioning activities (e.g., aerobics, running, jogging, swimming, bicycling, rowing, karate). Higher MET-min/week for the activities indicated a higher total volume of recreational physical activity.

Diet: Diet was measured at baseline using a previously validated 120-item food frequency questionnaire based on the Insulin Resistance Atherosclerosis Study (IRAS) (24). Participants reported the serving size of specific foods/beverages (small, medium, large), as well as how often they consumed these items. Dietary quality was characterized by the Alternative Healthy Eating Index (AHEI) (25, 26). The AHEI is a summary index of dietary factors that have been associated with a reduction in major chronic disease risk, and was derived from the intake of 9 components: vegetables, fruits, nuts/soy, ratio of white to read meat, cereal fiber, trans fat, ratio of polyunsaturated to saturated fats, multi-vitamin use, and alcohol. Higher AHEI scores indicate a more favorable dietary quality (higher intake of vegetables, fruits, nuts/soy, higher ratio of white to red meat and polyunsaturated to saturated fats, lower intake of trans fats, moderate alcohol consumption, and multi-vitamin use).

Recreational facilities and food stores: The recreational facility and food store data were derived from the National Establish Time Series (NETS) developed by Walls & Associates, using data from Dun & Bradstreet. Data was purchased for every zip code within a 5 mile radius

of any reported MESA participant address in the United States. Data were purchased for the years 2000-2007 so they could be linked to MESA exams 1-4. Addresses were geocoded using Tele-Atlas EZ-Locate. Data on supermarket availability was supplemented using data from the Nielsen TDLinx Service Supermarket Retail Category Database (TD) (27).

Recreational facilities included facilities for indoor conditioning, bowling, golf, team sports, racquet sports, water activities, and instructional facilities. Food stores included supermarkets, grocery stores, fruit and vegetable markets, convenience stores, bakeries, fast food establishments, health food stores, and drinking places. To determine which facility types to include, Standard Industrial Classification (SIC) codes were selected based on existing lists (8, 28). Recreational facility and food store densities were created for different categories of recreational facilities and food stores for the years 2000-2007. We summed the density measures calculated for recreational physical activity resources (e.g. resources for indoor conditioning activities, biking, hiking, team sports, racquet sports, swimming), instructional facility resources (e.g. facilities that offer instruction in physical activities), and other water activity resources (e.g. surfing) to obtain a summary measure which corresponded to exercise-related physical activities. A measure of favorable food store density was created by summing the densities for supermarket chain and non-chain stores and fruit and vegetable markets. We used this measure of favorable food stores in these analyses since supermarket density has been shown to be a reasonable proxy for healthy food access in prior work (9).

For both recreational facilities and food stores, the data was linked to the MESA addresses each year and densities were calculated for several buffer sizes around each MESA participant's home

address for the years 2000-2007 (overlapping with MESA exams 1-4). The 1-mile buffer was selected for primary analyses based on earlier studies in this cohort which observed that a majority of physically active participants reported that they exercised within one mile of their residence all or at least half the time (5) and that a healthy diet was associated with access to supermarkets within a 1-mile radius of participants' homes (9). Kernel estimation (29) was used to calculate the densities such that resources closer in distance to participants' addresses were given more weight than resources farther in distance from their addresses. Densities are expressed in units per square mile and were created using the Esri (Redmond, CA) ArcGIS software.

Covariates: Individual-level covariates measured at baseline only included race/ethnicity (White, non-Hispanic Black, Hispanic, Chinese), gender, age, education (<high school, high school, some college, college/graduate school). Covariates measured at each visit included household income (<$20,000, $20,000-49,999, $50,000-74,999, and >$75,000), marital status (never-married, married, divorced, separated, widowed), employment status (homemaker, full-time, part-time, unemployed, retired), home type (rent, own, mortgage), and number of years residing in the neighborhood. Neighborhood-level covariates included median household income derived from the 2000 U.S. Census and from the 2005-2009 American Community Survey (ACS).

Statistical analyses

We first examined the distribution of individual-level variables and neighborhood characteristics by (1) whether or not individuals moved during follow-up and (2) by tertiles of physical activity

and AHEI score at baseline. We also calculated the mean annual change in recreational facility densities and food store densities around participants' residences over the course of follow-up for movers and non-movers separately.

Residential mobility

We used Cox proportional hazards models with time to first move as the dependent variable to estimate whether health behaviors (diet and physical activity) were associated with the probability of moving. Participants who did not move were censored at the last date of follow-up. Since socio-demographic characteristics as well as home ownership and residential history may confound the association between individual-level variables and moving, we adjusted for age, race/ethnicity, gender, household income, education, marital status, current employment, home ownership, and number of years lived in the neighborhood.

Residential selection

In order to determine whether movers reporting more physical activity were more likely to move to neighborhoods with more favorable physical activity resources than those reporting less physical activity, we modeled the recreational density of the new location (after the move) as a function of individual-level physical activity before the move, adjusting for the same covariates in the residential mobility models, as well as recreational density prior to the move. Since mobility could be observed across three intervals (exam 1-2, exam 2-3, and exam 3-4) and physical activity was available at 3 exams, each participant had up to three observations. In the case of individuals who moved in multiple intervals, we used marginal models (Generalized

Estimating Equations (GEE)) to account for correlations between observations within an individual.

In addition, in order to determine if physically active residents lived in neighborhoods that were more likely to experience favorable changes in densities over time (when they are not moving), we conducted similar to those detailed above for the moving intervals, this time using the intervals in which no moves occurred (i.e. we modeled densities at the end of the interval adjusted for recreational density at the beginning of the interval in non-moving intervals). The moving interval and non-moving interval analyses detailed above for physical activity and recreational densities were adapted to investigate diet as a predictor and food store densities as the outcome. Since diet was measured only once at baseline, we estimated associations of diet at baseline with food store densities after the first move among movers, adjusted for socio-demographic characteristics and food store densities at baseline. In addition, in order to determine if residents with better diets lived in neighborhoods that were more likely to experience favorable changes in densities over time, we also estimated associations of diet at baseline with densities just prior to exam 4 after adjustment for covariates among persons who did not move between exams 1 and 4. SAS version 9.3 (Cary, NC) was used for all analyses and all p-values reported were two-sided.

RESULTS

Of the 6814 MESA participants enrolled at baseline, 6191 participated in the MESA neighborhood study. The median 1-mile recreational facility density at baseline was 2.1 facilities per square mile for both non-movers and movers (Table 2.1). Recreational facility

density increased over time for both movers and non-movers, with a greater increase observed in non-movers. The median 1-mile favorable food store density was 0.8 and 0.9 facilities per square mile for non-movers and movers, respectively. Favorable food store density decreased over time for movers but did not change appreciably for non-movers.

Non-movers reported a median physical activity level of 851 MET min/week and movers reported a median physical activity level of 761 MET min/week at baseline (Table 2.1). Movers and non-movers reported similar AHEI scores at baseline. Movers were more likely to be under 65 years of age, of Chinese descent, widowed or separated/divorced, employed, in the lower income group, and renting their homes as compared to non-movers. Non-movers were more likely to be non-Hispanic Black, married, retired, and home owners as compared to movers.

The median 1-mile recreational facility density at baseline was higher for participants in the highest tertile of physical activity, but the change in density over time did not differ by physical activity level (Table 2.1). Participants in the highest tertile of physical activity were less likely to move and more likely to be male, white, college-educated, in the highest income group, retired, and home owners, as compared to participants in the lowest tertile of physical activity.

The median 1-mile favorable food store density at baseline was slightly higher for participants in the highest tertile of AHEI score, but the change in density did not differ by AHEI tertile (Table 2.1). Participants in the highest tertile of AHEI score were more likely to be older, female, of Chinese descent, college-educated, in the highest income group, retired, and home owners, as compared to participants in the lowest tertile of AHEI score.

In adjusted Cox proportional hazards models, older age, higher income, home ownership, and longer time living in the neighborhood were associated with a lower hazard of moving (Table 2.2). Chinese participants had a higher hazard of moving and non-Hispanic Black and Hispanic participants had a lower hazard of moving, as compared to non-Hispanic White participants. Never married participants had a lower hazard of moving and divorced/widowed/separated participants had a higher hazard of moving, as compared to married participants. Participants in the higher tertiles of physical activity had a lower hazard of moving as compared to participants in the lowest tertile. Better diet quality, as measured by AHEI score, was not associated with moving.

In movers, we observed a positive, statistically significant association between physical activity level at the start of an exam interval (prior to moving) and recreational facility density at the end of the corresponding interval (after a move has occurred), after adjusting for age, race/ethnicity, gender, education, income, marital status, employment status, home type, and number of years residing in the neighborhood (mean difference in recreational density comparing the highest tertile of physical activity to the lowest tertile = 1.46 facilities per square mile) (Table 2.3). The association in movers was moderately attenuated after adjustment for recreational density at the start of the intervals.

In non-movers, we observed a much weaker association between physical activity at the start of the exam interval and recreational density at the end of the interval (mean difference in recreational density comparing the highest tertile of physical activity to the lowest tertile = 0.08

facilities per square mile) (Table 2.3). This association was no longer statistically significant after adjusting for density at the start of the interval.

We did not observe associations between baseline AHEI score and subsequent food store densities in movers or non-movers. For all models, sensitivity analyses using the ½ mile and 2-mile buffers and analyses excluding outliers for physical activity level yielded results similar to those shown.

DISCUSSION

In this multi-ethnic sample, we observed that individuals with higher physical activity were more likely to move during follow-up. We also observed a positive association between physical activity tertile and subsequent recreational density in movers even after adjusting for prior recreational density, which suggests that movers who are more physically active are more likely to move to neighborhoods with better availability of recreational facilities as compared to movers who are less active. We observed weaker associations in non-movers, suggesting that physical activity level was not strongly associated with a more favorable change in the availability of facilities for non-movers.

We did not observe a relationship between baseline diet and subsequent favorable food store density in this sample. However, we were limited to using the baseline measure of diet since diet was not measured at the subsequent follow-up visits used in these analyses.

In addition to physical activity and diet, we investigated several socio-demographic predictors of residential mobility. We found strong associations when examining age, race, income, marital

status, home ownership, and duration of residence in relation to residential mobility, which were all consistent with previous studies (13, 30-36). To our knowledge, no previous studies have investigated the relationship between healthy behaviors, such as physical activity and diet, and residential selection, specifically moving to or staying in neighborhoods with better availability of related facilities. Librett et al (37) reported a modest association between regular physical activity and a desire to live in an activity-friendly community, suggesting that physical activity may be a contributing factor in a person's choice of neighborhood.

Since selection into neighborhoods based on health-related behaviors may inflate the observed associations between neighborhood characteristics and health behavior outcomes, some studies have attempted to control for residential self-selection. Cao et al (38) found that built environment characteristics were related to walking behavior even after adjusting for preferences to live in walkable neighborhoods. Previous studies that have controlled for residential self-selection in the association between walkability and physical activity have found mixed results, with some finding that adjusting for self-selection has little impact on the associations, others finding an attenuation in the associations, and one reporting a strengthening of the observed association (38-43).

Despite previous findings that associations between the built environment and physical activity persisted after adjustment for self-selection, our finding of a positive association between physical activity and subsequent neighborhood recreational facility density suggests that the role of residential self-selection should be considered when investigating the association between the neighborhood physical environment and physical activity, especially among people who move

residences. Since residential self-selection is especially of concern in cross-sectional studies, these findings suggest that in the absence of longitudinal data on mobility, collecting data on residential history and adjusting for whether individuals recently moved may be one strategy to reduce bias due to possible residential selection. Other strategies for addressing residential self-selection include controlling for individual preferences and using longitudinal models, joint discrete choice models, or structural equations modeling.

Limitations

Measurement error is possible in the physical activity and diet measures since they were self-reported. Recreational facility and food store densities did not account for characteristics of the facilities such as quality, hours, or cost, and we did not have information on utilization of the facilities. In addition, there may be considerable measurement error in the databases used to obtain information on the availability of these facilities (44-46). There is little information on the buffer size that is most relevant for individual use of facilities. We used the 1-mile buffer based on previous work (5, 9). We did not account for perceived measures of physical activity resources or healthy food availability, and it is possible that self-selection based on physical activity or healthy eating behaviors is more strongly related to perceived access to resources. In addition, we did not have information on individual preferences for resources, and it is possible that these preferences are a better predictor of residential selection than the behaviors themselves.

We did not observe an association between diet and residential mobility or residential selection into neighborhoods with better access to favorable food stores in this sample, but we were

limited by the time-invariant measure of diet available, and investigated only one composite measure of a healthy diet. It is possible that specific components of a healthy diet, such as fruit and vegetable intake or characteristics of an unhealthy diet, such a fat intake or fast food consumption may be better predictors of selection into neighborhoods with access to related types of food stores. Future studies with repeated measures on diet and food stores are warranted to investigate whether diet is related to residential selection.

Strengths

Strengths of this study include the availability of longitudinal data on recreational densities and food store densities, physical activity measures, and individual and neighborhood-level covariates. We used objective measures of recreational facility and food store availability and our measure of physical activity accounted for both duration and intensity of activities. We were also able to adjust for several covariates related to residential mobility including key socio-demographic variables as well as housing tenure and duration of residence. Furthermore, we had detailed information on residential mobility, including when individuals moved and to where they moved.

Conclusion

Our findings support the hypothesis that physical activity may be related to residential self-selection into neighborhoods with better availability of physical activity resources, especially for people who move. This research complements the mixed findings in the existing literature that have examined the associations between the physical environment and physical activity after controlling for preferences related to self-selection. It is possible that part of the association

between the environment and physical activity is a direct effect and that part of the association is the result of residential self-selection, highlighting the need for public health strategies that focus on developing activity-friendly neighborhoods as well as targeting individual preferences and attitudes. We did not observe that diet was associated with self-selection into neighborhoods with better healthy food resources, but note that we were limited by our single point in time measure of diet. Further longitudinal studies investigating the role individual physical activity and diet as well as individual preferences in relation to residential self-selection are needed to clarify the impact of the neighborhood environment on physical activity and diet. Continued efforts to account for residential self-selection in these analyses through longitudinal designs and adjusting for individual preferences are warranted.

Further longitudinal studies investigating the role individual physical activity as well as individual preferences in relation to residential self-selection are need to clarify the impact of the neighborhood environment on physical activity, and continued efforts to adjust for residential self-selection in these analyses through longitudinal designs and controlling for individual preferences is warranted.

Table 2.1 – *Baseline characteristics and change in recreational facility and food store densities by move status and tertiles of physical activity and Alternative Healthy Eating Index (AHEI), No. (%), Median (p25-p75) or Mean (SD), the Multi-Ethnic Study of Atherosclerosis*

			Tertile of physical activity		Tertile of AHEI score	
	Non-movers	Movers	1 (low)	3 (high)	1 (low)	3 (high)
1-mile recreational facility density	2.1 (0.8-4.7)	2.1 (0.9-4.2)	1.8 (0.7-3.8)	2.4 (0.9-5.6)	--	--
Change in recreational density	0.5 (0.01)	0.1 (0.04)	0.4 (0.03)	0.4 (0.03)	--	--
1-mile favorable food store density	0.8 (0.2-3.3)	0.9 (0.3-2.5)	--	--	0.8 (0.2-2.4)	1.1 (0.3-3.3)
Change in food store density	-0.004 (0.003)	-0.1 (0.01)	--	--	-0.04 (0.008)	-0.03 (0.009)
Physical activity (MET min/week)	851 (210-2100)	761 (75-1980)	0 (0-157.5)	2835 (2070-4320)	--	--
AHEI score	43.7 (11.2)	44.0 (11.2)	--	--	31.7 (5.0)	56.4 (6.2)
Age (%)						
45-64	54	63	58	57	63	52
65-84	44	37	42	43	37	48
Gender (%)						
Female	52	53	57	46	50	55
Male	48	47	43	54	50	45
Race (%)						
White	40	37	32	44	39	43
Non-Hispanic Black	29	16	14	9	28	16
Hispanic	21	24	28	29	26	24
Chinese	10	23	26	18	7	17
Education (%)						
High school or less	35	35	46	25	42	26
Some college/College	65	65	54	75	57	74
Household Income (%)						
<$ 49,999	58	62	67	52	63	52
>=$50,000	42	38	33	48	37	48
Marital status						
Married	63	59	62	63	61	65
Widowed/Divorced/Separated	28	35	30	29	31	28
Never Married	9	6	8	8	8	7
Employment status						
Homemaker	11	12	13	8	11	12
Working	47	54	50	50	52	46
Unemployed	2	3	3	2	2	2
Retired	40	31	34	40	35	40
Hometype						
Rent	23	38	30	26	29	24
Mortgage	42	36	40	40	42	41
Own	32	20	26	31	26	32
Other	3	6	3	3	3	3
% Moved	--	--	35	30	32	32

[*]Neighborhood-level covariate

Table 2.2 – Adjusted hazard ratios (HR) of moving associated with socio-demographic characteristics, physical activity (MET min/week) and Alternative Healthy Eating Index (AHEI), the Multi-Ethnic Study of Atherosclerosis

	Adjusted HR[1]
Age (%)	
45-54	1.0
55-64	0.49 (0.42-0.58)***
65-74	0.45 (0.38-0.53)***
75-84	0.26 (0.21-0.32)***
Gender (%)	
Male	1.0
Female	0.90 (0.80-1.02)
Race (%)	
White	1.0
Non-Hispanic Black	0.74 (0.62-0.87)**
Hispanic	0.79 (0.66-0.95)**
Chinese	1.29 (1.08-1.55)**
Education (%)	
<High school	1.0
High school	0.87 (0.71-1.07)
Some college	0.98 (0.81-1.18)
College/graduate school	0.94 (0.77-1.16)
Household Income (%)	
<$20,000	1.0
$20,000-$49,999	0.89 (0.76-1.04)
$50,000-$74,999	0.70 (0.56-0.87)**
$75,000	0.66 (0.53-0.83)**
Marital status	
Married	1.0
Divorced /Widowed/Separated	1.24 (1.07-1.41)**
Never Married	0.59 (0.45-0.77)***
Currently working	
No	1.0
Yes	1.35 (1.18-1.54)***
Hometype	
Rent	1.0
Mortgage	0.55 (0.47-0.64)***
Own	0.64 (0.54-0.77)***
Physical activity	
Lowest tertile	1.0
Middle tertile	0.65 (0.57-0.75)***
Highest tertile	0.40 (0.33-0.46)***
AHEI	
Lowest tertile	1.0
Middle tertile	1.07 (0.93-1.24)
Highest tertile	1.06 (0.91-1.24)
Time lived in neighborhood (years)	0.97 (0.97-0.98)***

***p<0.0001 **p<0.01, *p<0.05
[1]*Adjusted for age, gender, race/ethnicity, education, income, marital status, work status, home type, and number of years lived in neighborhood*

Table 2.3 – *Mean differences (estimate (SE)) in recreational facility densities associated with previous physical activity level, and mean differences in food store densities associated with baseline AHEI score, for movers and non-movers, the Multi-Ethnic Study of Atherosclerosis*

	Physical activity tertile			AHEI Tertile		
	1 (low)	2	3 (high)	1 (low)	2	3 (high)
Model 1[a]						
Movers[a]	Ref	1.08 (0.4)*	1.46 (0.4)**	Ref	0.05 (0.2)	0.16 (0.2)
Non-movers[a]	Ref	-0.02 (0.03)	0.08 (0.04)*	Ref	0.05 (0.2)	0.15 (0.2)
Model 2[b]						
Movers[b]	Ref	0.87 (0.4)	1.25 (0.4)**			
Non-movers[b]	Ref	-0.02 (0.02)	0.01 (0.02)			

[a]*Adjusted for age, age*time, race, gender, education, income, marital status, employment status, home type, number of years lived in the neighborhood*
[b]*Additionally adjusted for beginning density (density level at the start of an exam interval)*
****p<0.0001* ***p<0.01* **p<0.05*

CHAPTER 2 REFERENCES

1.	Diez Roux A. Investigating neighborhood and area effects on health. *Am J Public Health* 2001;91(11):1783-9.
2.	Pickett KE, Pearl M. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. *J Epidemiol Community Health* 2001;55:111-22.
3.	Sampson R. The neighborhood context of well-being. *Perspect Biol Med* 2003;46(3 Suppl):S53-64.
4.	Diez Roux A. Neighborhoods and health: where are we and were do we go from here? *Rev Epidemiol Sante Publique* 2007;55(1):13-21.
5.	Diez Roux A, Evenson K, McGinn A, et al. Availability of recreational resources and physical activity in adults. *Am J Public Health* 2007;97(3):493-9.
6.	Hoehner C, Brennan Ramirez L, Elliott M, et al. Perceived and objective environmental measures and physical activity among urban adults. *Am J Prev Med* 2005;28(2 Suppl 2):105-16.
7.	McNeill L, Wyrwich K, Brownson R, et al. Individual, social environmental, and physical environmental influences on physical activity among black and white adults: a structural equation analysis. *Ann Behav Med* 2006;31(1):36-44.
8.	Powell L, Chaloupka F, Slater S, et al. The availability of local-area commercial physical activity-related facilities and physical activity among adolescents. *Am J Prev Med* 2007;33(4 Suppl):S292-300.
9.	Moore L, Diez Roux A, Nettleton J, et al. Associations of the local food environment with diet quality--a comparison of assessments based on surveys and geographic information systems: the multi-ethnic study of atherosclerosis. *Am J Epidemiol* 2008;167(8):917-24.
10.	Moore L, Diez Roux A, Nettleton J, et al. Fast-food consumption, diet quality, and neighborhood exposure to fast food: the multi-ethnic study of atherosclerosis. *Am J Epidemiol* 2009;170(1):29-36.
11.	Lee BA, Oropesa RS, Kanan JW. Neighborhood Context and Residential Mobility. *Demography* 1992;31(2):249-70.
12.	Larson A, Bell M, Young A. Clarifying the relationships between health and residential mobility. *Soc Sci Med* 2004;59(10):2149-60.
13.	Clark WA, Onaka JL. Life Cycle and Housing Adjustment as Explanations of Residential Mobility. *Urban Studies* 1983;20(1):47 — 57.
14.	Halliday T, Kimmitt M. Selective migration and health in the USA, 1984-93. *Popul Stud (Camb)* 2008;62(3):321-34.
15.	Butler EW, Kaiser EJ. Prediction of Residential Movement and Spatial Allocation. *Urban Affairs Review* 1971;6:477-93.
16.	Bagley MN, Mokhtarian PL. The impact of residential neighborhood type on travel behavior: A structural equations modeling approach. Annals of Regional Science. *36* 2002;2.
17.	Schwanen T, Mokhtarian PL. What affects commute mode choice: Neighborhood physical structure or preferences toward neighborhoods? *Journal of Transport Geography* 2005;13(1):83–99.

18. Handy S, Cao X, Mokhtarian P. Self-Selection in the Relationship between the Built Environment and Walking: Empirical Evidence from Northern California. *Journal of the American Planning Association* 2006;72(1):55-74.

19. A OF, Sen B, Rahurkar S, et al. The relationship between built environments and physical activity: a systematic review. *Am J Public Health* 2012;102(10):e7-e13.

20. Addy CL, Wilson DK, Kirtland KA, et al. Associations of perceived social and physical environmental supports with physical activity and walking behavior. *Am J Public Health* 2004;94(3):440-3.

21. Morland K, Wing S, Diez Roux A. The contextual effect of the local food environment on residents' diets: the atherosclerosis risk in communities study. *Am J Public Health* 2002;92(11):1761-7.

22. LaMonte M, Durstine J, Addy C, et al. Physical activity, physical fitness, and Framingham 10-year risk score: the cross-cultural activity participation study. *J Cardiopulm Rehabil* 2001;21(2):63-70.

23. Ainsworth BE, Haskell WL, Whitt MC, et al. Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc* 2000;32(9 Suppl):S498-504.

24. Mayer-Davis E, Vitolins M, Carmichael S, et al. Validity and reproducibility of a food frequency interview in a Multi-Cultural Epidemiology Study. *Ann Epidemiol* 1999;9(5):314-24.

25. McCullough ML, Willett WC. Evaluating adherence to recommended diets in adults: the Alternate Healthy Eating Index. *Public Health Nutr* 2006;9(1A):152-7.

26. Nettleton JA, Steffen LM, Mayer-Davis EJ, et al. Dietary patterns are associated with biochemical markers of inflammation and endothelial activation in the Multi-Ethnic Study of Atherosclerosis (MESA). *Am J Clin Nutr* 2006;83(6):1369-79.

27. Auchincloss AH, Moore KA, Moore LV, et al. Improving retrospective characterization of the food environment for a large region in the United States during a historic time period. *Health Place* 2012;18(6):1341-7.

28. Gordon-Larsen P, Nelson MC, Page P, et al. Inequality in the built environment underlies key health disparities in physical activity and obesity. *Pediatrics* 2006;117(2):417-24.

29. Silverman BW. *Density estimation for statistics and data analysis*. Chapman & Hall/CRC; 1986.

30. Speare AJ, Goldstein S, Frey WH. *Residential Mobility, Migration and Metropolitan Change*. Cambridge, Massachusetts: Ballinger Publishing Company; 1975.

31. McCarthy KF. *The Household Life Cycle and Housing Choices*. Santa Monica, California: The Rand Corporation; 1976.

32. Varady DP. Determinants of Residential Mobility Decisions The Role of Government Services in Relation to Other Factors. *Journal of the American Planning Association* 1983;49(2):184-99.

33. Pickvance CG. Life cycle, housing tenure and intra-urban residential mobilty: a causal model. *Sociological Review* 1973;21(2):279-97.

34. Newman SJ, Duncan GJ. Residential problems, dissatisfaction and mobility. *Journal of the American Planning Association* 1978;45(2):154-62.

35. Feijten P, van Ham M. Residential mobility and migration of the divorced and separated. *Demographic Research* 2007;17.

36.	South SJ, Deane GD. Race and Residential Mobility: Individual Determinants and Structural Constraints. *Social Forces* 1993;21(1):147.

37.	Librett JJ, Yore MM, Schmid TL, et al. Are self-reported physical activity levels associated with perceived desirability of activity-friendly communities? *Health Place* 2007;13(3):767-73.

38.	Cao X, Handy SL, Mokhtarian PL. The influences of the built environment and residential self-selection on pedestrian behavior: evidence from Austin, TX. *Transportation* 2006;33:1-20.

39.	Owen N, Cerin E, Leslie E, et al. Neighborhood walkability and the walking behavior of Australian adults. *Am J Prev Med* 2007;33(5):387-95.

40.	Pinjari AR, Bhat CR, Hensher DA. Residential self-selection effects in an activity time-use behavior model. *Transportation Research Part B* 2009;43:729–48.

41.	Sallis JF, King AC, Sirard JR, et al. Perceived environmental predictors of physical activity over 6 months in adults: activity counseling trial. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association* 2007;26(6):701-9.

42.	Van Dyck D, Cardon G, Deforche B, et al. Neighborhood SES and walkability are related to physical activity behavior in Belgian adults. *Prev Med* 2010;50 Suppl 1:S74-9.

43.	Van Dyck D, Cardon G, Deforche B, et al. Relationships between neighborhood walkability and adults' physical activity: How important is residential self-selection? *Health & Place* 2011;17:1011-4.

44.	Boone JE, Gordon-Larsen P, Stewart JD, et al. Validation of a GIS facilities database: quantification and implications of error. *Ann Epidemiol* 2008;18(5):371-7.

45.	Han E, Powell L, Slater S. Validation of secondary commerical data sources for physical activity facilities in urban and nonurban settings. *Journal of physical activity & health* 2012;9(8):1080-8.

46.	Hoehle F, Schootman M. Concordance of commercial data sources for neighborhood-effects studies. *Journal of urban health: bulletin of the New York Academy of Medicine* 2010;87(4):713-25.

CHAPTER 3

Longitudinal associations between neighborhood recreational facilities and change in recreational physical activity in the Multi-Ethnic Study of Atherosclerosis

INTRODUCTION

After smoking, physical inactivity is the second leading modifiable risk factor for chronic disease and has a significant impact on mortality in western countries (1). Previous research examining the worldwide burden of chronic disease attributable to physical inactivity found that eliminating physical inactivity would increase the life expectancy of the world's population (2), and a recent longitudinal study found that reducing physical activity over time was related to increased functional disability (3). Individual-level factors, such as lower socioeconomic status, are associated with decreased leisure-time physical activity (4, 5). However, recent work has suggested that contextual factors may also be important in influencing physical activity. Creating social and physical environments that are conducive to physical activity may be especially important in motivating individuals to engage in physical activity and in enhancing the impact of individual-targeted interventions (6).

Mixed land use and residential density have been shown to be associated with activities such as walking or biking for transportation (7-11). In addition, previous studies have found that the availability of physical activity equipment was associated with vigorous physical activity/sports and that recreational facility availability and neighborhood aesthetics were related to recreational

38

physical activity (12-19). Conversely, some studies have found weak associations between neighborhood environment and physical activity (20-22). A systematic review found that several studies investigating the relationship between neighborhood environment and physical activity had null/inconclusive findings, noting that the environmental attributes investigated in many studies may be defined incorrectly in relation to the type of physical activity studied (23). This lack of specificity may weaken the relationship between the predictor and outcome variables (24). Stronger associations have been found in studies that investigated hypotheses linking more narrowly defined combinations of environmental attributes with the specific activities to which they may be related (23).

One limitation to existing research is that most analyses are cross-sectional, and due to the possibility of residential self-selection, causal inferences are limited. Longitudinal data that track both persons and neighborhoods over time are needed to examine whether changes in environments are related to changes in physical activity (25, 26). This type of *change versus change* analysis is uncommon in neighborhood health research, and is needed to address limitations that arise in cross-sectional analyses. A second limitation is the absence of large population studies with time-varying measures of neighborhood attributes.

We used longitudinal data from a large, multi-ethnic study of adults to investigate the association between baseline recreational facility density and change in physical activity over time, and the association between change in density and change in physical activity. We hypothesized that individuals living in neighborhoods with better availability of recreational facilities at baseline would experience more favorable changes in physical activity as compared to individuals living

in neighborhoods with poorer availability of facilities. We also hypothesized that individuals living in neighborhoods with greater increases in the availability of recreational facilities would experience more favorable changes in physical activity as compared to individuals living in neighborhoods with smaller increases (or declines) in facilities.

MATERIALS AND METHODS

Study sample

The sample consisted of participants from the Multi-Ethnic Study of Atherosclerosis (MESA), a longitudinal study of cardiovascular disease which comprises a population-based, multi-site (Baltimore, Maryland; Chicago, Illinois; Forsyth County, North Carolina; Los Angeles, California; New York, New York; and St. Paul, Minnesota), multi-ethnic (approximately 38% white, 28% African-American, 23% Hispanic, 11% Chinese descent) sample of 6,814 men and women aged 45-84 years who were free of clinical cardiovascular disease at baseline. We based these analyses on data from the baseline visit (July 2000-September 2002) and two follow-up visits (approximately 1.6 and 3.2 years later). Participants were observed on at least 2 occasions (94% had 3 observations).

Recreational facilities

Recreational facility data was purchased from the National Establish Time Series database from Walls & Associates. Home addresses were geocoded using Tele-Atlas EZ-Locate. Data was purchased for 2000-2007, spanning the time range for the MESA exams used in our analyses. 133 Standard Industrial Classification (SIC) codes were selected based on existing lists (27, 28).

Densities were created for different categories of facilities for the years 2000-2007. We summed the densities calculated for facilities offering indoor conditioning, biking, hiking, team sports, racquet sports, swimming, physical activity instruction, and water activities to obtain a summary measure which corresponded to exercise-related activities.

Densities were calculated for several buffer sizes around each MESA participant's home address for each year. Densities were matched to each MESA participant at each time point, such that densities changed over time if the neighborhood resources changed or if the participant moved. The 1-mile buffer was selected for analyses based on a study in a sub-sample of this cohort which observed that a majority of physically active participants reported that they exercised within one mile of their residence all or at least half the time (19). Kernel estimation (29) was used to calculate the densities such that facilities closer in distance to participants' addresses were given more weight than facilities farther in distance. Densities are expressed in units per square mile (henceforth, densities indicates densities/mi^2) and were created using the Esri (Redmond, CA) ArcGIS software.

Physical activity assessment

Physical activity was assessed at baseline and two subsequent follow-up exams via a semi-quantitative questionnaire adapted from the Cross-Cultural Activity Participation Study. Previous work has provided evidence for the validity and reliability of this survey (30). Participants were asked how many days a week and the average time per day they engaged in activities during "a typical week in the past month". Minutes of activity were summed for each activity and multiplied by its metabolic equivalent (31). We used a summary measure of

physical activity which included the activities we hypothesized would be related to recreational facilities. Recreational physical activity (RPA) was calculated to include walking for exercise, dance, team sports (e.g., softball, basketball), dual sports (e.g., tennis), individual activities (e.g., golf, yoga) and conditioning activities (e.g., running, swimming, bicycling), and was measured in metabolic equivalent-minutes per week (MET-min/week). Higher MET-min/week for the activities indicated a higher total volume of physical activity. Individual records indicating total physical activity of more than 24 hours per day were excluded.

Measurement of covariates

Covariate data were obtained at baseline and follow-up visits and included age, race/ethnicity (White, non-Hispanic Black, Hispanic, Chinese), education (<high school, high school, some college, college/graduate school), household income (<$20,000, $20,000-49,999, $50,000-74,999, and >$75,000), season of exam visit (December-February, March-May, June-August, September-November), employment, and move status (never moved, moved at least once). Neighborhood-level covariates obtained from census data included % unemployed, % poverty, and median household income.

Statistical analyses

We examined the distribution of individual-level variables across the three exams and estimated the mean annual change in recreational density and RPA. We categorized recreational density and RPA into tertiles and examined the distribution of individual-level variables across the highest and lowest tertiles. Finally, we estimated unadjusted associations of the individual-level variables with baseline levels and changes in RPA.

After conducting preliminary analyses using categories and smoothed lines to check the linearity assumption, RPA and recreational density were included as continuous variables in longitudinal models. We used linear mixed effects models to estimate the association between recreational density and change in RPA over time, adjusting for age, gender, race/ethnicity, education, household income (time-varying), and season of exam (time-varying) in final models. To estimate the association between baseline density and change in RPA (model A), we used linear mixed effects models to model repeated physical activity measures on each participant as a function of recreational density at baseline, time since baseline, an interaction between density and time, and time-invariant and time-varying covariates. The interaction between time and baseline recreational density was used to estimate the impact of baseline recreational density on changes over time in RPA.

To assess the association between change in density and change in RPA (model B), we used similar mixed effects models, with the addition of change in density since baseline as a time-varying covariate and an interaction between change in density and time to the models. This interaction term was used to assess whether change in recreational density over time modified the change over time in RPA.

All models included a random intercept and a random time slope for each participant, to allow the baseline responses as well as the time slope to vary between individuals. Coefficients from the final model were used to compare the physical activity trajectories over time for different levels of baseline densities and change in densities.

We also stratified by age at baseline, whether participants moved (32% of participants moved during follow-up), income, and race/ethnicity to assess effect measure modification, or whether the association between recreational density and RPA response trajectories differed by these characteristics. We investigated each interaction in separate models and retained those with $p<0.05$ in the final model.

Sensitivity analyses were conducted using the ½ mile and 2-mile buffers for recreational density. SAS version 9.3 (Cary, NC) was used for all analyses and all p-values reported are two-sided.

RESULTS

Of the 6814 MESA participants, 6191 participated in the MESA neighborhood study. After excluding visits at which participants reported greater than 24 hours of total physical activity per day and records with missing data on the exposure and outcome, 23 participants were excluded, leaving 6168 for analysis. The excluded participants did not differ from the analytic sample with respect to socio-demographic factors.

At baseline, 58% of the participants were younger than 65 years and 48% were male; 40% were white, 26% were non-Hispanic black, 22% were Hispanic, and 12% were Chinese (Table 3.1). Individuals reported a median RPA level of 840 MET min/week at baseline, which declined to 825 MET min/week at exam 3 (mean follow-up=3.2 years). The mean annual change in RPA during the study period was -17.7 MET min/week. Overall, recreational facility density

increased during the study period from a mean density of 2.1 facilities/mi^2 to a mean density of 2.7 (mean annual change = 0.4).

Participants in the highest tertile of recreational density were less likely to be non-Hispanic Black or Hispanic, and more likely to have higher levels of RPA and live in neighborhoods of greater socioeconomic disadvantage (Table 3.2). Participants in the highest tertile of RPA were more likely to be male, white, college-educated, earn greater than $75,000, and have greater access to recreational facilities.

In unadjusted models, male gender, higher education and higher income were associated with higher baseline RPA (Table 3.3). Hispanic and Chinese race/ethnicity and lower recreational density were associated with lower baseline RPA. Older age at baseline was associated with a greater decline in RPA over time.

In models estimating the association between baseline recreational density and change in RPA (model A), we observed a positive, statistically significant association between recreational facility density and baseline RPA, after adjusting for covariates (mean difference in baseline RPA associated with a 1-unit change in recreational density (95% confidence interval (CI) = 18.7 (13.1, 24.4) MET min/wk) (Table 3.4). Overall, RPA decreased over time as the cohort aged (mean annual change in RPA for individuals living in areas with 0 resource density was -31.6 MET min/wk). The interaction between baseline recreational density and time was positive, suggesting that higher baseline density was associated with a less pronounced decline in RPA

over time, but the estimate was not statistically significant (mean difference in annual change in RPA for each 1-unit increase in density at baseline (CI) = 1.5 (-0.3-3.4)).

In models estimating the association between change in recreational density since baseline and change in RPA (model B), the interaction between change in density and time was positive and statistically significant in the unadjusted model (mean difference in annual change in RPA for each 1-unit increase in the change in density since baseline (CI) = 11.6 (1.9-21.1)) (Table 3.4), indicating that a greater increase in recreational density over time was associated with a less pronounced decline in RPA over time, even after adjustment for the effect of baseline density on change in RPA. Estimates were slightly attenuated after adjusting for individual-level covariates (mean difference in annual change (CI) = 10.3 (0.7-19.9)), but the association remained statistically significant. Adjusting for employment, move status, neighborhood-level covariates, and site did not meaningfully alter the results.

Figure 3.1 shows estimated changes in RPA over time for the 90th and the 10th percentiles of baseline levels and changes in densities as estimated from the fully adjusted model in Table 3.4 (model B). Overall, persons who lived in areas with higher recreational densities at baseline had higher levels of RPA than those who lived in areas with lower densities at baseline. In addition, within each set, persons who experienced greater increases in recreational densities over time showed less pronounced declines or even small increases in RPA over time. Overall, we observed the most favorable change in RPA over time for the group in the 90th percentile of baseline density and the 90th percentile of change in density, and the least favorable change in

RPA for the group in the 10th percentile of baseline density and 10th percentile of change in density.

In age-stratified models, the associations between baseline recreational density and change in RPA was stronger among adults age 65 and over (mean annual change (CI) = 4.6 (1.7-7.6)) than among those under 65 (mean annual change (CI) = -1.0 (-3.4-1.4)), and the test for heterogeneity was statistically significant ($P<0.01$). Similar patterns were observed when analyses relating changes in densities to changes in RPA were stratified by age: stronger associations were observed among older adults (mean difference in annual change (CI) = 16.5 (0.9-32.1)) than among those under 65(mean difference in annual change (CI) = 6.2 (-5.9-18.2)), but the test for heterogeneity was not statistically significant ($P>0.1$). There was no significant heterogeneity by move status, income, or race/ethnicity for either model. Analyses using buffer sizes of ½ mile and 2-mile yielded results similar to those shown.

DISCUSSION

In this multi-ethnic sample, we observed a positive association between recreational facility density and recreational physical activity (RPA) at baseline. We also observed a decline in RPA over time. We did not observe statistically significant associations between baseline recreational density and change in RPA in the overall sample. However, while on average RPA declined in the sample as a whole, individuals who experienced greater increases in the densities of recreational resources around their homes experienced a less pronounced decline in RPA over

47

time. Among individuals who experienced the largest increases in density of resources, the decline in RPA was practically eliminated.

We used mixed effects models to estimate associations of recreational resources with RPA change. Model A revealed cross-sectional associations between baseline density levels and baseline RPA but showed no statistically significant associations between baseline density and change in RPA over time. An important limitation of this model is that it does not account for changes in densities over time, and assumes that densities at a single time point impact future trajectories, which may be unrealistic if densities change over time. However, Model A allows for comparisons with other work that may only have measurement at a single time point available. In contrast, Model B showed that greater increases in densities over time were associated with a statistically significant less pronounced decline in RPA over time. The comparison between these models highlights the impact of accounting for changes in the exposure over time.

Recent studies have suggested that age may modify the association between the neighborhood environment and physical activity (32, 33). In this study, the observed associations between baseline recreational density and change in RPA were stronger among older adults, but the difference between the age groups was not as strong for the change in density models. There is limited previous research investigating the environmental correlates of physical activity in adults of different age groups. One study reported positive associations between perceived access to recreational facilities and physical activity among adults age 80 and over (34). Longitudinal studies investigating the neighborhood environment and changes in walking have reported that

48

proximity to parks/trails was associated with maintaining/increasing walking in older men (35),

and that proximity to parks/gyms was associated with smaller declines in walking (36). A study

examining interactions of physical environment with age suggest that older adults may be more

influenced by environmental attributes (37), perhaps because they may spend more time in their

neighborhoods. Our results suggest that further research is warranted to understand the impact

of recreational facilities on RPA in older adults.

Thus far, few studies have examined the association between changes in neighborhood

environment and changes physical activity in adults. A recent study found that increases in

access to pay facilities around the home were associated with increases in physical activity in

young males (38). A secondary analysis of physical activity intervention trials found that

increases in perceived access to facility/home equipment was associated with increases in

moderate/vigorous physical activity (39), and a study in California found that changes in

neighborhood attributes were associated with self-reported changes in physical activity within

the neighborhood (40). To our knowledge, ours is the first study to show an association between

change in recreational facility density and change in recreational physical activity in adults.

Limitations

Recreational facility densities in this study were based on commercial data and did not take into

account quality, hours, or cost, nor do we have information on whether the participants were

utilizing the resources. In addition, we did not include resources such as parks or other physical

environment features that may be related to physical activity. Previous studies examining the

quality of the commercial databases suggest that there may be measurement error in the facility

measures (41-43). However, it is unlikely that this error is patterned in such a way that it would create the associations that we observed.

There is little information on the relevant buffer size for use of recreational facilities. We used the 1-mile buffer since most of the active individuals in this sample reported that they use facilities within 1-mile of their residences, and found that results did not vary in analyses using the ½ mile and 2-mile buffers. We used the same buffer size for the whole sample, though the relevant buffer may vary by site. Adjusting for site in our models slightly attenuated the observed estimates. Furthermore, since the 1-mile buffer was measured around residential addresses, we did not account for facilities near the workplace.

We chose to present estimates unadjusted for population density because our exposure of interest was the presence of facilities within a certain radius. On one hand, more densely populated areas may have more resources available. On the other hand, there is no specific reason that availability of facilities should affect physical activity conditional on the number of persons who live in the area. In addition, population density and urban development, while relevant for walking for transportation, are unlikely to be related to RPA and hence unlikely to be confounders in our analyses. Sensitivity analyses adjusting for population density and %retail/commercial facilities yielded results similar to those presented.

Since physical activity was self-reported, measurement error is possible. The form of the questionnaire likely results in overestimates of time spent engaging in activities; thus absolute values may not be interpretable. We excluded visits at which participants reported greater than

24 hours of total activity per day to exclude inflated measures of self-reported activity from influencing the results. Despite the limitations of the absolute values, given that the questionnaire was similar across time points, relative positions and trends over time remain meaningful.

In using mixed effects models, we cannot rule out residual confounding due to unmeasured covariates, especially variables related to residential self-selection, such as individual preference for physical activity resources, which are difficult to measure.

Strengths

Our study is unique in having longitudinal data on recreational densities, physical activity measures, and individual and neighborhood-level covariates. We used objective measures of recreational densities and our physical activity measure accounted for both duration and intensity of a range of moderate/vigorous activities. In addition, our study pairs the neighborhood construct of interest, recreational facilities, with a relevant physical activity measure (RPA), addressing the limitation in previous studies of the environmental construct of interest not corresponding closely enough to the outcome measure of interest.

With available time-varying measures, we were able to estimate the association between change in an objective measure of resource availability and change in RPA. Although we cannot completely rule out confounding or bi-directional causation (changes in RPA of residents influencing changes in densities), this type of change versus change analysis is a major step

forward as compared to cross-sectional analyses, which may be hampered by the possibility that individuals with certain activity preferences choose neighborhoods with better resources.

Conclusion

In summary, the results from this longitudinal study suggest that greater access to recreational facilities may benefit middle-aged and older adults by enabling them to maintain activity levels as they age and experience slower declines in physical activity over time, which may lead to reduced morbidity and increased functional ability. Data on quality of physical activity resources, individual preferences for living in a neighborhood with physical activity resources, and whether individuals are using available resources would allow for better estimation of the causal relationship between availability of recreational resources and physical activity.

Table 3.1. *Socio-demographic Characteristics of Study Participants at Baseline and 2 Follow-up Exams, No. (%), Mean (SD), or Median (p25-p75) and Mean Annual Change in Recreational Facility Density and Recreational Physical Activity (RPA), the Multi-Ethnic Study of Atherosclerosis*

	Baseline	Exam 2	Exam 3	Mean annual change (SE)
Individual-level covariates				
Mean follow-up (years)	--	1.6	3.2	
1-mile recreational facility density	2.1 (0.8-2.5)	2.3 (1.0-5.1)	2.7 (1.0-5.6)	0.4 (0.01)**
RPA (MET min/wk)	840 (150-2070)	735 (0-1755)	825 (150-1890)	-17.7 (7.4)*
Age (%)				
45-54	30	23	19	
55-64	28	29	30	
65-74	29	31	31	
75-84	13	17	20	
Gender (%)				
Female	52			
Male	48			
Race/ethnicity (%)				
White	40			
Non-Hispanic Black	26			
Hispanic	22			
Chinese	12			
Education (%)				
<High school	18			
High school	19			
Some college	24			
College	39			
Household income (%)				
<$20,000	22	23	23	
$20,000-$49,999	37	37	35	
$50,000-$74,999	17	16	17	
$75,000	24	24	25	
Neighborhood-level covariates				
%Unemployed	0.08	0.08	0.08	
%Poverty	0.16	0.15	0.15	
Median household income ($)	44, 549	44,922	48,316	

*$P<0.05$, **$P<0.001$

Table 3.2. *Baseline Characteristics According to Tertiles of Recreational Facility Density and Recreational Physical Activity (RPA), No. (%), Mean (SD), or Median (p25-p75), the Multi-Ethnic Study of Atherosclerosis*

	Tertile of recreational density		Tertile of RPA	
	1 (low)	3 (high)	1 (low)	3 (high)
Recreational facility density	0.5 (0.03-0.83)	6.4 (4.5-15.8)	1.8 (0.7-3.8)	2.5 (0.9-5.6)
RPA (MET min/wk)	735 (52.5-1890)	1103 (300-2377.5)	0 (0-157.5)	2835 (2070-4590)
Age (%)				
45-54	29	28	30	30
55-64	29	27	28	27
65-74	30	29	29	31
75-84	12	16	13	12
Gender (%)				
Female	50	54	57	53
Male	50	46	43	47
Race (%)				
White	41	39	32	45
Non-Hispanic Black	16	10	28	28
Hispanic	35	24	26	18
Chinese	8	27	14	9
Education (%)				
<High school	15	19	26	10
High school	22	15	23	16
Some college	27	23	23	27
College/graduate school	36	43	28	47
Household Income (%)				
<$20,000	21	24	29	16
$20,000-$49,999	38	35	38	36
$50,000-$74,999	19	15	17	18
$75,000	22	26	16	30
%Unemployed[*]	0.06	0.10	0.08	0.07
%Poverty[*]	0.12	0.20	0.17	0.14
Median household income ($)[*]	47,187	41,915	41,634	47,068

Neighborhood-level covariate

54

Table 3.3. *Unadjusted Mean Differences in Recreational Physical Activity (RPA) (MET min/week) at Baseline and Mean Differences in Annual Change in RPA Over Follow-up Related to Baseline Covariates, the Multi-Ethnic Study of Atherosclerosis*

	Baseline RPA	Change in RPA
Age (%)		
45-54	Reference	Reference
55-64	21.5	-5.6
65-74	17.7	-19.3
75-84	-29.6	-57.4[*]
Gender (%)		
Female	Reference	Reference
Male	432.0[***]	2.2
Race (%)		
White	Reference	Reference
Non-Hispanic Black	-35.2	-8.6
Hispanic	-326.1[***]	0.1
Chinese	-538.5[***]	-21.5
Education (%)		
<High school	Reference	Reference
High school	385.4[***]	-20.5
Some college	718.9[***]	-15.0
College or graduate school	745.0[***]	-9.6
Household income (%)		
<$20,000	Reference	Reference
$20,000-$49,999	387.0[***]	-18.2
$50,000-$74,999	431.4[***]	22.6
$75,000	737.6[***]	36.7
Move Status		
Never moved	Reference	Reference
Moved at least once	-96.6	22.8
Baseline recreational density		
Highest tertile	Reference	Reference
Middle tertile	-286.6[***]	-12.8
Lowest tertile	-327.2[***]	-5.6
Study Site		
New York	Reference	Reference
Illinois	11.8	49.5
Maryland	-113.9	32.6
Minnesota	-238.3*	-19.6
California	-682.9***	-12.3
North Carolina	-363.0**	-33.9

[***] $P<0.0001$ [**] $P<0.01$ [*] $P<0.05$

Table 3.4. *Mean Differences in Recreational Physical Activity (RPA)(MET min/wk) at Baseline and Mean Differences in Annual Changes in RPA Associated With Baseline Recreational Densities and Changes in Recreational Densities Over Time, the Multi-Ethnic Study of Atherosclerosis*

	Model 1[a] Estimate (SE)	Model 2[b] Estimate (SE)
Model A: Baseline recreational density + interaction of baseline recreational density with time		
Mean difference in baseline RPA per 1-unit increase in rec density at baseline	24.1 (2.8)***	18.7 (2.9)***
Mean annual change in RPA for individuals with density of 0 at baseline	-26.3 (9.2)**	-31.6 (9.8)**
Mean difference in annual change in RPA per 1-unit increase in baseline density	1.3 (0.9)	1.5 (0.9)
Model B: Baseline recreational density + interaction of change in recreational density with time		
Mean difference in baseline RPA per 1-unit increase in rec density at baseline	23.6 (2.9)***	19.2 (2.9)***
Mean annual change in RPA for individuals with density of 0 at baseline	-27.8 (9.8)**	-32.7 (9.9)*
Mean difference in annual change in RPA per 1-unit increase in change in density	11.6 (4.9)*	10.3 (4.9)*

***$P<0.0001$, **$P<0.01$, *$P<0.05$

[a]*Unadjusted model*

[b]*Model adjusted for age, age*time, race, gender, education, income, season*

Figure 3.1 - Change in recreational physical activity (RPA) over time, comparing 4 different groups, categorized as either the 90th or 10th percentile of baseline (BL) recreational density, and either the 90th or 10th percentile of change (Δ) in recreational density. [*]
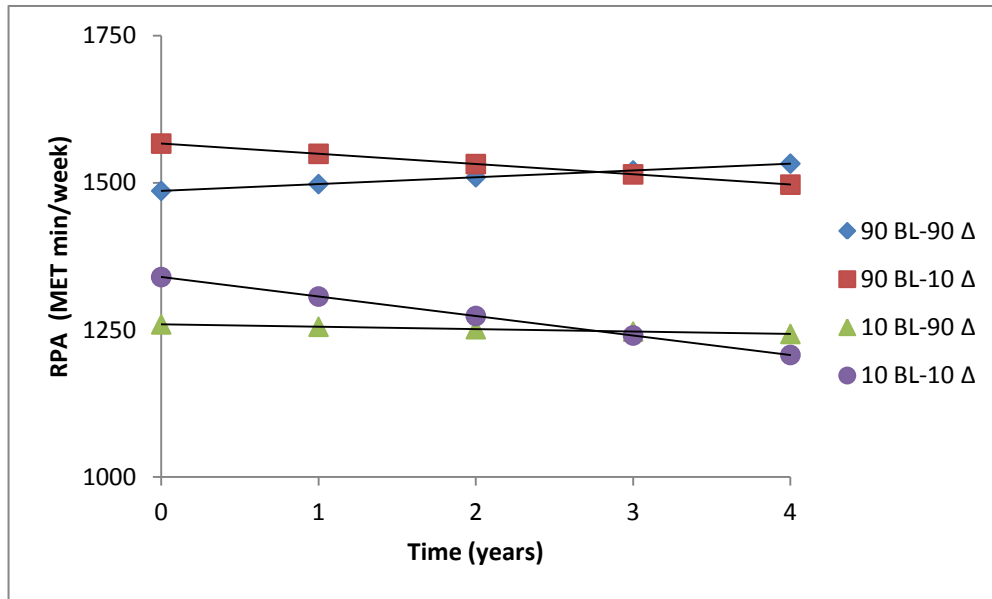
[*]*Estimated from Model 2B in Table 3.4 (adjusted for all covariates)*

CHAPTER 3 REFERENCES

1.   Services USDoHaH. Physical Activity and Health: A Report of the Surgeon General. . Atlanta, GA: US Departmentof Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, 1996.
2.   Lee IM, Shiroma EJ, Lobelo F, et al. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet* 2012;380(9838):219-29.
3.   Gretebeck RJ, Ferraro KF, Black DR, et al. Longitudinal change in physical activity and disability in adults. *American journal of health behavior* 2012;36(3):385-94.
4.   Brug J, Oenema A, Ferreira I. Theory, evidence and Intervention Mapping to improve behavior nutrition and physical activity interventions. *Int J Behav Nutr Phys Act* 2005;2(1):2.
5.   Trost S, Owen N, Bauman A, et al. Correlates of adults' participation in physical activity: review and update. *Med Sci Sports Exerc* 2002;34(12):1996-2001.
6.   Sallis J, Cervero R, Ascher W, et al. An ecological approach to creating active living communities. *Annu Rev Public Health* 2006;27:297-322.
7.   Rodríguez D, Evenson K, Diez Roux A, et al. Land use, residential density, and walking. The multi-ethnic study of atherosclerosis. *Am J Prev Med* 2009;37(5):397-404.
8.   Handy S, Boarnet M, Ewing R, et al. How the built environment affects physical activity: views from urban planning. *Am J Prev Med* 2002;23(2 Suppl):64-73.
9.   Schwanen T, Mokhtarian PL. What affects commute mode choice: Neighborhood physical structure or preferences toward neighborhoods? *Journal of Transport Geography* 2005;13(1):83–99.
10.   Pinjari AR, Bhat CR, Hensher DA. Residential self-selection effects in an activity time-use behavior model. *Transportation Research Part B* 2009;43:729–48.
11.   Saelens B, Sallis J, Frank L. Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. *Ann Behav Med* 2003;25(2):80-91.
12.   Sallis JF, Hovell MF, Hofstetter CR, et al. A multivariate study of determinants of vigorous exercise in a community sample. *Prev Med* 1989;18(1):20-34.
13.   Sallis JF, Johnson MF, Calfas KJ, et al. Assessing perceived physical environmental variables that may influence physical activity. *Research quarterly for exercise and sport* 1997;68(4):345-51.
14.   Brownson RC, Baker EA, Housemann RA, et al. Environmental and policy determinants of physical activity in the United States. *Am J Public Health* 2001;91(12):1995-2003.
15.   De Bourdeaudhuij I, Sallis JF, Saelens BE. Environmental correlates of physical activity in a sample of Belgian adults. *Am J Health Promot* 2003;18(1):83-92.
16.   Giles-Corti B, Donovan RJ. Socioeconomic status differences in recreational physical activity levels and real and perceived access to a supportive physical environment. *Prev Med* 2002;35(6):601-11.
17.   Addy CL, Wilson DK, Kirtland KA, et al. Associations of perceived social and physical environmental supports with physical activity and walking behavior. *Am J Public Health* 2004;94(3):440-3.

18. Powell KE, Martin LM, Chowdhury PP. Places to walk: convenience and regular physical activity. *Am J Public Health* 2003;93(9):1519-21.
19. Diez Roux A, Evenson K, McGinn A, et al. Availability of recreational resources and physical activity in adults. *Am J Public Health* 2007;97(3):493-9.
20. Troped P, Saunders R, Pate R, et al. Associations between self-reported and objective physical environmental factors and use of a community rail-trail. *Prev Med* 2001;32(2):191-200.
21. Ball K, Bauman A, Leslie E, et al. Perceived environmental aesthetics and convenience and company are associated with walking for exercise among Australian adults. *Prev Med* 2001;33(5):434-40.
22. Brownson R, Kelly C, Eyler A, et al. Environmental and policy approaches for promoting physical activity in the United States: a research agenda. *J Phys Act Health* 2008;5(4):488-503.
23. Wendel-Vos W, Droomers M, Kremers S, et al. Potential environmental determinants of physical activity in adults: a systematic review. *Obes Rev* 2007;8(5):425-40.
24. Giles-Corti B, Timperio A, Bull F, et al. Understanding physical activity environmental correlates: increased specificity for ecological models. *Exercise and sport sciences reviews* 2005;33(4):175-81.
25. Cook T, Shagle S, Degirmencioglu S. Capturing social process for testing mediational models of neighborhood effects. In: Brooks-Gunn G DG, Aber J, ed. *Neighborhood poverty: context and consequences for children*. New York, NY: Russell Sage Foundation Press, 1997:94–119.
26. O'Campo P. Invited commentary: Advancing theory and methods for multilevel models of residential neighborhoods and health. *Am J Epidemiol* 2003;157(1):9-13.
27. Powell L, Chaloupka F, Slater S, et al. The availability of local-area commercial physical activity-related facilities and physical activity among adolescents. *Am J Prev Med* 2007;33(4 Suppl):S292-300.
28. Gordon-Larsen P, Nelson M, Page P, et al. Inequality in the built environment underlies key health disparities in physical activity and obesity. *Pediatrics* 2006;117(2):417-24.
29. Silverman BW. *Density estimation for statistics and data analysis*. Chapman & Hall/CRC; 1986.
30. LaMonte M, Durstine J, Addy C, et al. Physical activity, physical fitness, and Framingham 10-year risk score: the cross-cultural activity participation study. *J Cardiopulm Rehabil* 2001;21(2):63-70.
31. Ainsworth BE, Haskell WL, Whitt MC, et al. Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc* 2000;32(9 Suppl):S498-504.
32. Van Cauwenberg J, De Bourdeaudhuij I, De Meester F, et al. Relationship between the physical environment and physical activity in older adults: a systematic review. *Health Place* 2011;17(2):458-69.
33. Ding D, Gebel K. Built environment, physical activity, and obesity: what have we learned from reviewing the literature? *Health Place* 2012;18(1):100-5.
34. Chad KE, Reeder BA, Harrison EL, et al. Profile of physical activity levels in community-dwelling older adults. *Med Sci Sports Exerc* 2005;37(10):1774-84.
35. Michael YL, Perdue LA, Orwoll ES, et al. Physical activity resources and changes in walking in a cohort of older men. *Am J Public Health* 2010;100(4):654-60.

36. Li F, Fisher J, Brownson RC. A multilevel analysis of change in neighborhood walking activity in older adults. *Journal of aging and physical activity* 2005;13(2):145-59.

37. Sallis JF, King AC, Sirard JR, et al. Perceived environmental predictors of physical activity over 6 months in adults: activity counseling trial. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association* 2007;26(6):701-9.

38. Boone-Heinonen J, Guilkey DK, Evenson KR, et al. Residential self-selection bias in the estimation of built environment effects on physical activity between adolescence and young adulthood. *Int J Behav Nutr Phys Act* 2010;7:70.

39. Ries AV, Dunsiger S, Marcus BH. Physical activity interventions and changes in perceived home and facility environments. *Prev Med* 2009;49(6):515-7.

40. Handy SL, Cao X, Mokhtarian PL. The causal influence of neighborhood design on physical activity within the neighborhood: evidence from Northern California. *Am J Health Promot* 2008;22(5):350-8.

41. Boone JE, Gordon-Larsen P, Stewart JD, et al. Validation of a GIS facilities database: quantification and implications of error. *Ann Epidemiol* 2008;18(5):371-7.

42. Han E, Powell L, Slater S, et al. Validation of secondary commercial data sources for physical activity facilities in urban and nonurban settings. *Journal of physical activity & health* 2012;9(8):1080-8.

43. Hoehner CM, Schootman M. Concordance of commercial data sources for neighborhood-effects studies. *Journal of urban health : bulletin of the New York Academy of Medicine* 2010;87(4):713-25.

The association between neighborhood socio-economic composition and incident cardiovascular disease: methodological approaches to account for time-varying confounding, The Multi-Ethnic Study of Atherosclerosis (MESA)

Introduction

Within the body of work on neighborhoods and health outcomes, a large number of studies have focused specifically on the relationship between neighborhood environment and cardiovascular disease (CVD) and CVD risk factors (1-3). Several studies have found that lower neighborhood socioeconomic status (4) is associated with increased risk of cardiovascular disease, after adjusting for individual-level factors (3, 5-9). Despite recent studies documenting associations between neighborhood environments and coronary heart disease (CHD), there is a need for continued research to address the limitations of this work (8, 10-12) in order to determine whether the observed associations reflect causal processes. In particular, longitudinal studies that follow subjects and neighborhoods over time are necessary to improve causal inferences. A few studies have investigated the association between neighborhood characteristics and *incidence* of cardiovascular disease. One prospective study found that lower neighborhood SES, as measured by a summary score constructed based on wealth and income, education, and occupation, was associated with increased risk of CHD, after adjustment for individual-level socioeconomic characteristics (7). Similarly, two separate prospective studies conducted in Sweden found that lower neighborhood income and education and a measure of neighborhood

deprivation were associated with increased risk of CHD, though these studies had little or no information on individual-level risk factors for CHD (i.e. smoking, physical activity, diet) (13).

A major methodological challenge in these analyses is the need to account for exposures and/or covariates that vary over time. Time-varying covariates may simultaneously confound and mediate neighborhood effects. For example, individual physical activity may affect where a person lives (i.e. more physically active individuals may select to live in neighborhoods with more physical activity resources) but may also be a consequence of prior neighborhood exposures (i.e. prior exposure to an environment with higher resources may increase individuals' future physical activity). Time-varying confounding results when 1) time-varying covariates are associated with subsequent exposure, 2) time-varying covariates are associated with the outcome, independent of exposure, and 3) exposure predicts subsequent time-varying covariates (Figure 1). Consider the variable physical activity in the context of studies investigating whether the neighborhood environment affects CVD outcomes. Individual physical activity (PA) may be 1) related to the type of neighborhood an individual chooses to live in (PA at time t is associated with neighborhood exposure at time $t+1$), 2) related to future development of CVD (PA at time t is associated with having a CVD event at time $t+1$), and 3) the consequence of prior neighborhood exposures (neighborhood exposure at time t-1 predicts PA at time t). Under these circumstances, both adjusted and unadjusted analyses yield biased estimates of the causal effect of the exposure on the outcome. In these contexts, more sophisticated methodologies are needed to account for time-varying confounding.

One approach that has been used to account for time-varying confounding is the use of marginal structural models (MSMs). MSMs are specifically advantageous in the situation where time-dependent covariates may be both confounders and mediators with respect to the exposure-outcome association of interest (14, 15). To adjust for time-varying confounding using this approach, a pseudo-population is created by weighting each individual by the inverse of the conditional probability of receiving the exposure he/she actually received. This pseudo-population is balanced with respect to the distribution of possible confounders across levels of the exposure. Using this weighting approach, the time-varying confounders that are in the pathway between the exposure and outcome are now equally distributed across levels of exposure and hence conditioning on the time-varying confounder (which would be problematic if the confounder is also a mediator) is no longer necessary (see Figure 1). The final exposure-outcome model is then fit to the pseudo-population data, which is unconfounded by the time-varying covariates. A few studies have applied MSMs in the context of either investigating CVD outcomes *or* neighborhood exposures. Ilomaki et al (16) used MSMs to investigate the association between alcohol and myocardial infarction and Nandi et al (17) recently applied MSMs to investigate the relationship between childhood socio-economic status and CVD. In the context of neighborhood exposures, Sampson et al (18) used MSMs to estimate the association between neighborhood disadvantage and verbal ability in African-American children and Cerda et al (19) applied MSMs to investigate the association between neighborhood poverty and alcohol use. However, this analytic approach has not yet been explored in the context of estimating the association between neighborhood exposures and cardiovascular outcomes, as most studies lack time-varying measures on both the neighborhood exposures and biomedical risk factors for CVD.

A second methodological challenge is the need to consider cumulative effects of exposure. Most longitudinal studies to date focus on exposure measured at a single point in time. However, the effect of neighborhood environment on CVD risk may only be observable after exposure to an unfavorable neighborhood environment over a longer period of time (20). With the availability of time-varying exposure information, investigating measures of cumulative exposure may better capture the neighborhood exposure that is associated with CVD. In addition, comparing the long-term cumulative exposure to the recent exposure may also reflect the underlying biological mechanisms involved in the pathway between neighborhood socioeconomic composition and incident CVD. A stronger association between long-term cumulative exposure to neighborhood deprivation and incident CVD would suggest that the pathway is more likely mediated by biological changes that accumulate more slowly over time, such as atherosclerosis. In contrast, a stronger association between recent exposure and incident CVD would suggest a pathway mediated by acute biological responses, such as stress or inflammatory pathways.

We used longitudinal data from a large, multi-ethnic study of adults to investigate the relationship between neighborhood SES and incident CVD by 1) contrasting results from using long-term cumulative exposure with results from using recent exposure measures to characterize neighborhood SES and 2) contrasting results from traditional regression models including time-varying covariates with marginal structural models which account for time-varying confounding.

METHODS

*Sample***:** The sample consisted of participants from the Multi-Ethnic Study of Atherosclerosis (MESA), a longitudinal study of cardiovascular disease conducted at six sites (Baltimore, Maryland; Chicago, Illinois; Forsyth County, North Carolina; Los Angeles, California; New York, New York; and St. Paul, Minnesota). MESA comprises a multi-ethnic (approximately 38% white, 28% African-American, 23% Hispanic, 11% Chinese descent) population-based sample of 6,814 men and women aged 45-84 years who were free of clinical cardiovascular disease at baseline. Data was collected at baseline (between 2000 and 2002) with 4 subsequent follow-up visits approximately every two years. Approximately 28% of participants were lost to follow-up by exam 5 (this includes missing an exam and deaths).

*Neighborhood socio-economic composition***:** We obtained census tract-level data on neighborhood demographics from the 2000 U.S. census and from the American Community Survey (ACS, 2005-2009). These two datasets were used to create factor scales of neighborhood SES determinates using principal components analysis (PCA). For our analyses, we used an aggregate measure of neighborhood socio-economic composition as measured by a factor score which included: 1) % individuals 25+ with at least a Bachelor's degree, 2) % individuals with managerial/professional occupation, 3) % individuals 25+ with at least a high school diploma, 4) Median owner value of households, 5) median household income, and 6) % of household income >$50,000. A higher factor score indicated worse neighborhood SES. The neighborhood factor score was linked to the MESA addresses for each participant based on the year of the address (either the 2000 Census for years 2000-2004 or the 2005-2009 ACS for years 2005-2012). Two

measures of neighborhood SES were created to reflect exposure history over time: 1) long-term cumulative exposure (the average of the neighborhood factor score from baseline through the month prior to a CVD event or end of follow-up (t-1)) and 2) recent cumulative exposure (the average exposure over the one year prior to an event or end of follow-up.

Cardiovascular disease (CVD): In the MESA study, participants were included only if they were free of clinical CVD at baseline. During follow-up, participants were contacted by a telephone interviewer at intervals of 9-12 months to inquire about any interim hospital admission, cardiovascular outpatient diagnoses/procedures, and deaths. Diagnoses were verified by obtaining death certificates, medical records, and next-of-kin interviews. Hospital records were obtained successfully on approximately 98% of hospitalized CVD events and other information was obtained on 96% of outpatient diagnoses. CVD events were only included if they were determined to be definite or probable by trained reviewers. Events classified as incident CVD for these analyses included myocardial infarction, resuscitated cardiac arrest, definite angina, probable angina (if followed by revascularization), stroke, stroke death, coronary heart disease death, other atherosclerotic death, and other CVD death.

Covariates: Baseline covariates included age (continuous, centered at the mean), race/ethnicity (White, non-Hispanic Black, Hispanic, Chinese), gender, and education (<high school, high school, some college, college/graduate school. Time-varying covariates included household income (<$20,000, $20,000-49,999, $50,000-74,999, and >$75,000), employment status, smoking (never, former, current), physical activity (self-reported intensity and duration, measured in MET min/week), BMI (kg/m2), and history of diabetes. Additional covariates that

were included as part of model building included self-reported diet, number of years lived in the neighborhood, marital status, home type (own or rent), recently moved (in the past 6 months), hypertension, LDL, and current alcohol use. These covariates were not included in final models as the addition of these covariates did not alter the results.

Statistical analyses: In descriptive analyses, we examined the distribution of CVD events and baseline covariates by quintile of neighborhood SES at baseline. We also examined the distribution of the time-varying covariates across the 5 MESA exams. For any exposure and covariate data that was missing for an individual at a particular time point, the past or future non-missing value closest in time was assigned.

Prior to conducting regression analyses, we explored whether the time-varying covariates in these data acted as both confounders and mediators. To assess the presence of time-varying confounding, we 1) used repeated measures linear models to estimate the associations between lagged measures of time-varying covariates (t-12) and subsequent neighborhood SES (to assess any confounding effects) and 2) used logistic models (for dichotomized covariates) and linear models (for continuous covariates) to estimate associations between a lagged measure of neighborhood SES (t-12) and subsequent time-varying covariates (to assess any mediating effects).

To estimate the association between neighborhood SES and incident CVD, we first used Cox proportional hazards models with baseline and time-varying covariates, which has been the traditional approach to model the probability of disease as a function of past exposure and

confounder history. Models were adjusted for age, race/ethnicity, gender, education, income, currently working, BMI, physical activity, smoking status, and presence of diabetes. Cox models were estimated using the SAS procedure PROC PHREG.

We also estimated marginal structural models (MSMs) using inverse-probability-of-treatment-and-censoring weighting to account for the possible presence of time-varying confounding. The parameters of a MSM can be consistently estimated using a class of estimators called the inverse-probability-of-treatment-and-censoring weighted estimators, assuming exchangeability, consistency, positivity, and correct specification of the models (14, 15, 21). This model is fit in a two-stage process: 1) estimate each subject's probability of having his/her own exposure history and the probability of remaining uncensored, and use these predicted probabilities to derive inverse-probability-of-treatment-and-censoring weights (IPTCW), and 2) estimate the exposure-outcome association in a regression model that is weighted using the IPTWCs.

We followed the approach of Hernan and Robins (14, 15) to create both the treatment weights and the censoring weights for each participant. In our example, the "treatment" is the continuous exposure neighborhood SES. The treatment weights refer to the probability of having the exposure history the participant actually experienced and the censoring weights refer to the probability of remaining uncensored. A participant was censored if he/she was lost to follow-up (includes deaths) at time t. The final weight (IPTCW) is calculated as the product of the treatment and censoring weights.

To estimate the treatment weights, logistic regression models are traditionally used for dichotomous treatments. Since our measure of neighborhood SES was continuous, we used a normal exposure probability density function and linear regression models to estimate the treatment weights as per Cerda, et al (19). Probability densities were conditional on neighborhood SES exposure history, baseline covariates (age, gender, race, education), time-varying covariates (household income, employment status, BMI, physical activity, smoking, and history of diabetes), and remaining uncensored up to time t.

The general form of the treatment weight up to time t was defined as (adapted from Cerda, et al):

$$sw_i(t) = \prod_{t=0}^{t} \frac{f(neigh_i(t)|\overline{C_i}(t) = 0, \ \overline{neigh}_i(t-1), V_i}{f(neigh_i(t)|\overline{C_i}(t) = 0, \ \overline{neigh}_i(t-1), \ \overline{L}_i(t-1), V_i}$$

In the equation above, $neigh_i$ is the neighborhood SES exposure at time t and $\overline{neigh}_i(t-1)$ represents prior exposure history up to t-1. $\overline{C_i}(t)$ is equal to 1 if a respondent was lost to follow-up by time t. V_i represents a vector of baseline covariates, and $\overline{L}_i(t-1)$ represents a vector of time-varying covariates. The denominator represents the conditional probability density of neighborhood SES exposure received by person i at time t with baseline covariates V_i, prior exposure, and time-varying covariates until t-1, given that they remained uncensored at time t. The numerator was defined similarly to stabilize the weight, and included only covariates measured at baseline, including the baseline values of the time-varying covariates.

To estimate the numerator of the treatment weight, we used a linear regression model to predict the exposure as a function of prior exposure and baseline covariates. To estimate the denominator of the treatment weight, we used a similar model with the addition of time-varying covariates. We used the SAS procedure PROC GLM to obtain regression estimates and their predicted values. The predicted values were used to create the treatment weights based on the normal density assumption.

We used logistic regression models to estimate the censoring weights, modeling the probability of remaining uncensored up to time t as per Hernan, et al (14). These weights were combined with the treatment weights to create final stabilized weights. We considered several different model specifications to determine the final weighting model based on the distribution of the final stabilized weights, with the ideal distribution having a mean near 1 and few extreme values (15, 22). We observed some very large stabilized weights, so we opted to trim the weights. Since a very small percentage of the weights were extreme, we trimmed the top and bottom 0.4%, assigning the top and bottom 0.4^{th} percentile to the value of the 99.6^{th} and 0.4^{th} percentile, respectively. The final stabilized weights had mean of 1.06 and a maximum value of 11.6.

The final weighted regression models were estimated using pooled logistic regression using SAS procedure PROC GENMOD, which closely estimates a Cox proportional hazards model. This is because PROC PHREG is unable to handle the time-dependent weights that are required for the MSM. The final models included the exposure (either long-term cumulative exposure or recent exposure), all baseline covariates (age, race, gender, education), and time. In this analysis weighted by the IPTCW, each participant's probability of having a particular exposure history at

each time point is unrelated to the time-varying covariates that were included when creating the

weights. The structure of the final weighted MSM was defined as:

$$\text{logit}\,\Pr\{D(i)=1|D(t-1)=0,\overline{Neigh}(t),V_i\}=\beta_0+\beta_1*\overline{Neigh}+C'V_i+\beta_3*month+C'splinemonth$$

To allow for time to be modeled flexibly, the weighting models and the final MSM included time

as a continuous variable and a transformation of time specified as a cubic spline with 5 knots.

All analyses were conducted using SAS 9.3 (Cary, NC) and all reported *p*-values are 2-sided.

RESULTS

Of the 6814 MESA participants enrolled at baseline, 6191 gave consent to participate in the

MESA neighborhood study. After excluding participants missing data on the exposure,

covariates, or outcome, 6138 participants remained for analysis. Participants were followed

from 2000-2012) and 458 cases of incident CVD were documented over 669,635 person-months

of follow-up. Average follow-up time was approximately 108 months.

The mean neighborhood factor score at baseline was -2.4 for participants in the lowest quintile of

neighborhood SES and 1.3 for participants in the highest quintile of neighborhood SES (Table

4.1). The mean neighborhood factor score declined during between baseline and the end of

follow-up from -0.3 to -0.7, suggesting that the neighborhoods were improving slightly over time

(a higher score indicates greater neighborhood deprivation). Participants in the highest quintile

of neighborhood factor score were more likely to be non-Hispanic Black and Hispanic and more

likely to be in the lowest education and income groups as compared to participants in the lowest

quintile of neighborhood factor score. These participants were also more likely to be current

smokers, report lower levels of physical activity and higher BMI, have diabetes or hypertension at baseline, and develop CVD in the future.

The results of investigating the assumptions for time-varying confounding by estimating the relationships between time-lagged covariates and future exposure and time-lagged exposure and future covariate values are shown in Table 4.2 and 4.3. Household income, working status, smoking, physical activity, and history of diabetes, all measured at t-12, were associated with neighborhood SES at time (Table 4.2). Examining the possible mediating role of the time-varying covariates shows that living in a worse neighborhood at t-12 is associated with being in the lowest income group, working, and having diabetes at time t and associated with being less physically active at time t (Table 4.3).

Using traditional methods to estimate Cox proportional hazards models (unweighted regression), we observed a 18% increase in the risk of CVD for every 1-unit increase in the long-term cumulative average of the neighborhood factor score up to t-1 (adjusted hazard ratio (HR) (95% CI) for a 1-unit increase in neighborhood factor score = 1.18 (1.08-1.28)), after adjusting for the time-variant covariates age, race, gender, income, and education (Table 4.2, Model 1). The results were slighted attenuated after additionally adjusting for potential time-varying confounders/mediators (income, currently working, physical activity, BMI, smoking, and history of diabetes) (Table 4.2, Model 2a). Estimates obtained from modeling recent exposure (cumulative average from t-12 to t-1) yielded estimates similar to those obtained when using long-term cumulative exposure.

Using MSMs (IPTCW weighted models), we observed a 7% increase in the risk of CVD for each 1-unit increase in the long-term cumulative average of the neighborhood factor score up to time t-1, but the estimate was no longer statistically significant (adjusted OR (CI) = 1.07 (0.97-1.17) (Table 2, Model 3b). Estimates obtained from modeling recent exposure using MSMs yielded estimates similar to those obtained when using long-term cumulative exposure.

DISCUSSION

In this multi-ethnic sample of men and women, we observed a greater incidence of CVD with increasing neighborhood deprivation using traditional time-to-event models, even after adjustment for individual-level income, education, and occupation. Adjusting for individual-level biomedical risk factors, which may be mediators in the pathway between neighborhood SES and incident CVD, attenuated the estimates slightly. Using MSMs to adjust for time-varying confounding, we observed a weaker association between increasing neighborhood deprivation and incident CVD, and the estimates were no longer statistically significant.

We estimated both traditional regression models and MSMs to compare the results from both approaches to determine whether accounting for time-varying confounding resulted in estimates that varied greatly from the models that did not account for time-varying confounding. If there were a significant amount of time-varying confounding in these data, we would expect results from the MSMs to differ significantly from those obtained using traditional methods. Adjusting for time-varying confounders in the MSMs may have accounted for the slight attenuation of the observed association as compared to the traditional model estimates. We expected the results

73

from the traditional analyses to be biased towards the null, since including time-varying

covariates in these models is analogous to "adjusting out" the effect of an intermediate, and

would result in an underestimate of the true association (though only a small amount of

attenuation was observed with the addition of potential mediators to the traditional models. The

weaker associations observed in the MSM models may reflect the true direction of bias due to

time-varying confounding in the relationship between neighborhood deprivation and incident

CVD, which suggests an overestimate of the relationship when using traditional models. The

unexpected attenuation of the results after accounting for time-varying confounding could also

be partially related to the time lag selected for the time-varying covariates (t-12), and it is

possible that defining the time-varying covariates at a time point more proximal to the time of

event would yield different results. However, despite the slight attenuation in the MSM

estimates, the magnitudes of the estimates from the two sets of models were not greatly different,

with largely overlapping confidence intervals. Furthermore, it has been noted that estimates

from standard and MSM approaches cannot be compared directly, since the exposure effect in

MSMs is marginal with respect to time-dependent covariates but conditional with respect to

time-invariant covariates (16, 23).

Our finding that greater neighborhood disadvantage was associated with increased incidence of

CVD is similar to results from other studies. Several studies have found that living in

disadvantaged neighborhoods was associated with various CVD outcomes, including CHD (7, 9,

subclinical CVD (Nordstrom, 2004 #1362, 13, 24-28), stroke (25, 29, 30), and prognosis,

adverse events and survival after myocardial infarction (6, 31-33). Similar to findings from

Diez-Roux, et al (7) and Sundquist, et al (9), as well as other studies, we found that adjusting for

individual-level socio-economic covariates, such as income and education, did not change the observed associations between neighborhood disadvantage and incident CVD, and that the addition of potential biomedical risk factors/mediators attenuated the associations only slightly.

Despite the large body of work documenting associations between neighborhood disadvantage and CVD outcomes, a few studies have found weak associations, or associations that are greatly attenuated after adjusting for individual-level covariates. In a cohort of middle-aged and older British men, Morris et al (34) found that men living in the most socially deprived neighborhoods had a higher incidence of CHD than men in the least deprived neighborhoods, but this association did not persist after adjustment for individual-level socio-demographic factors. In Sweden, Merlo et al (35) observed an association between low neighborhood income and risk of ischemic heart disease, which was consistent with the results from several previous Swedish studies (9, 27, 36). However, the association was attenuated after adjusting for individual socio-economic characteristics and adjusting for other confounders using a quasi-experimental family-based design. The authors concluded, in contrast to several previous Swedish studies that the causal association between neighborhood disadvantage and heart disease risk was unclear and that neighborhoods in Sweden have a small effect on heart disease risk (35). To our knowledge, ours is one of the first studies in the U.S. to investigate the association between neighborhood SES and incident CVD using time-varying measures on both the exposure and individual-level covariates, and the first study to use marginal structural models to evaluate the influence of time-varying confounding in this context.

Several mechanisms have been suggested as playing role in the pathway between neighborhood SES and cardiovascular disease (2). Individuals residing in disadvantaged neighborhoods may be more exposed to tobacco advertising and resources such as fast-food restaurants, and have less availability of healthy foods and physical activity resources (7, 26). The lack of access to health-promoting resources and the greater exposure to health-damaging resources may in turn affect lifestyle factors such as smoking, diet, and physical activity, which have been shown to be related to neighborhood disadvantage (1, 37). Other mechanisms that may explain the association between neighborhood disadvantage and CVD include lower access to health care resources and increased stress caused by increased psychosocial hazards in the neighborhood such as violent crime and abandoned buildings (9, 28, 38-40). Chronic stress has been shown to be associated with CVD risk factors (41-44).

Limitations

Though we adjusted for several covariates in the both the traditional models and the weighting models for the MSMs, some residual confounding is possible. The summary score used to characterize neighborhood SES may not fully capture the exposure that is most relevant to CVD risk and does not include any specific objective features of the environment or individual perceptions of the environment. It has been argued that more detailed measures of the neighborhood environment are needed to understand the specific mechanisms involved in the relationship between neighborhood SES and CVD (38). Few studies have considered specific features of the neighborhood environment in relation to CVD, and future studies should

investigate specific features of the physical and social environments of neighborhoods (both objective and perceived) to better understand the mechanisms involved.

In addition, our exposure does not take into account non-residential exposures, and it is possible that exposures to workplace or other neighborhood environments may modify or confound the relationship between neighborhood of residence and CVD (45). Though we used a time-varying measure of neighborhood SES, we only had Census data at 2 time points. Having only two points of variation for the exposure may have resulted in some measurement error in the exposure. However, this may have been slightly mitigated due to the available data on residential mobility of individuals. The movers in this sample contributed to additional variability in the exposure, since the neighborhood exposure changed when an individual moved. Those who moved during follow-up (approximately 30%) had more than 2 points of change in the neighborhood exposure provided they moved out of their census tracts (over 90% of movers changed census tracts).

Strengths

Strengths of this study included 1) the large, population-based sample, 2) the longitudinal prospective design, and 2) time-varying data on the neighborhood exposure, individual-level socio-economic factors, individual-level behavioral covariates, and biomedical CVD risk factors. The extensive information available on individual-level covariates enabled us to reduce confounding from unmeasured covariates. We also had information on residential mobility and length of time lived in the neighborhood, which has been noted as variables that have been

unavailable in past studies . In addition, our measure of incident CVD was validated using several sources, including medical records and death certificates.

Another strength of our study is that we investigated a long-term cumulative measure of neighborhood exposure, and compared this to a recent measure of neighborhood exposure. Past studies have relied on baseline measures of the neighborhood exposure, which may not reflect the exposure of interest if the exposure changes over time (due to changes in neighborhood composition or the individuals moving to different neighborhoods) or if an accumulation of the exposure over time is the relevant construct. Our findings suggest that there is no difference between long-term cumulative exposure and more recent exposure with respect to risk of CVD. However, our cumulative exposure spanned an average of 9 years of follow-up, and it is possible that the accumulation of exposure that may be relevant is over a longer time span (over the life course). Ours is the first study to examine longer-term cumulative exposure with respect to CVD incidence, which addresses limitations of previous work that did not consider any time lags or accumulation of exposure. Future work examining cumulative versus recent neighborhood exposures may help to elucidate the biological pathway involved in the association between neighborhood deprivation and incident CVD.

We also used MSMs to adjust for time-varying confounding in the context of the neighborhood environment and CVD outcomes. This approach has been warranted since most studies over the past decade have been adjusting for individual-level behavioral factors and biomedical CVD risk factors, which may be confounders but may also lie in the pathway between the neighborhood environment and CVD. Though adjustment for these covariates is appropriate to control for

confounding, bias may result if these individual behaviors and risk factors are a consequence of prior exposure and a risk factor for future exposure. By comparing the estimates from the MSMs to traditional models, we were able to assess if time-varying confounding was causing any bias in the traditional estimates. The results we observed from the MSMs were slightly weaker than the results obtained from traditional models, suggesting that there was not an appreciable amount of time-varying confounding in our study, and that any small amount of time-varying confounding biased traditional estimates slightly away from the null. These results may have implications for interpreting previous studies examining the neighborhood environment and CVD, in that previous findings may be slightly overestimating associations in the presence of time-varying confounding. However, these results should be interpreted with caution since it is possible that we were not able to capture the true amount of time-varying confounding in this study given the limited time points for which the exposure was available. Nevertheless, this study contributes to the body of research on neighborhoods and CVD by paving the way for future studies to analyze longer-term, time-varying measures of exposure and to employ stronger methodological tools to address the complex relationships that may arise between neighborhood exposures and time-varying covariates. Further studies using exposure measures that capture neighborhood SES variation over multiple time points is warranted.

Conclusion

Using time-varying measures on neighborhood exposure and covariates, our findings confirm results from previous studies which report mostly moderate associations between neighborhood SES and risk of CVD, though our MSM estimates suggested weaker associations. Adjusting for

time-varying confounding and using a long-term cumulative measure of neighborhood exposure

versus a recent measure of exposure do not appear to alter these estimates appreciably. Policies

and resources targeting prevention of CVD should include neighborhood-level approaches in

addition to individual-level interventions to help individuals to attain and sustain healthy

behaviors that may reduce the overall burden of CVD.

Figure 4.1: Directed acyclic graphs (DAGs) illustrating time-varying confounding (study population) and the removal of time-varying confounding using inverse probability weighting (pseudopopulation).

Study Population

Pseudopopulation
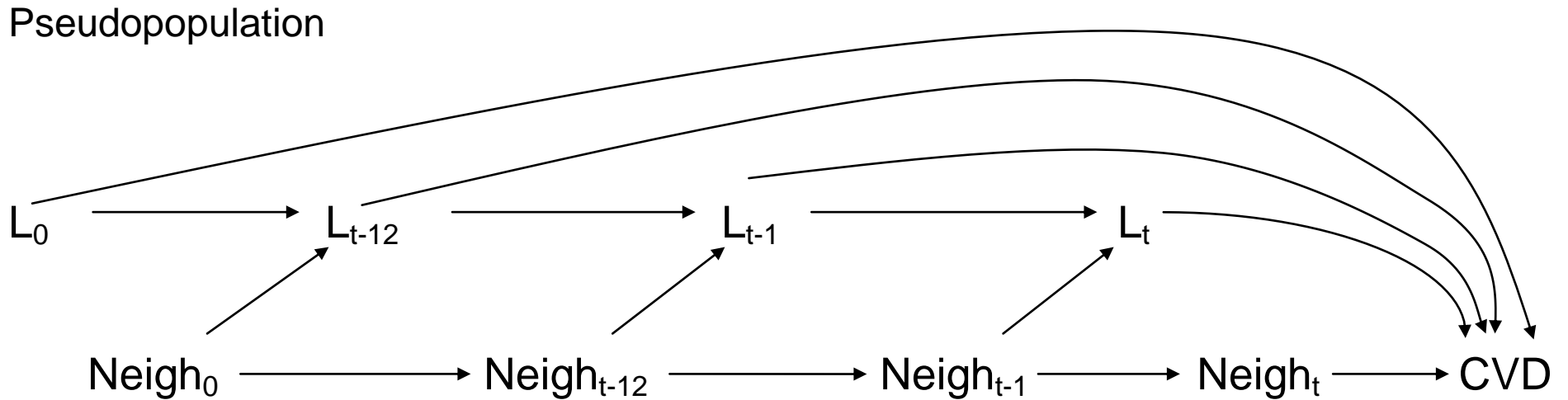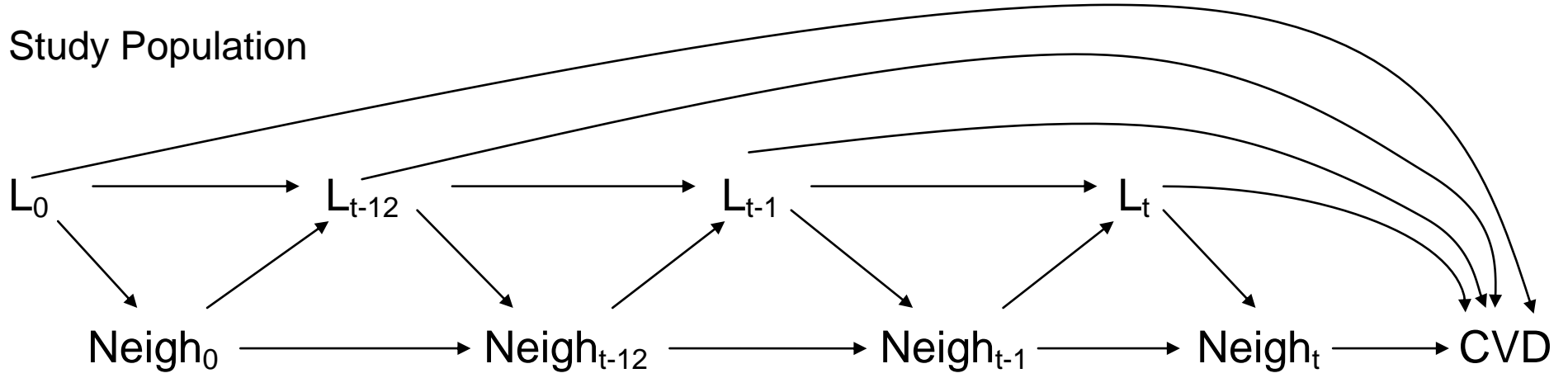
$Neigh_x$ = Neighborhood exposure at time x L_0 = baseline covariates (i.e., age, gender)
L_x = time-varying covariates at time x (i.e., income, physical activity)

Table 4.1 – *Baseline characteristics by quintile of SES factor score, No. (%) or Mean value, the Multi-Ethnic Study of Atherosclerosis*

	Quintile of Neighborhood Factor Score			Exam Number				
	1(better)	3	5 (worse)	1	2	3	4	5
Neighborhood factor score	-2.4	-0.04	1.3	-0.3	-0.3	-0.3	-0.7	-0.7
Incident CVD after baseline (%)	5.8	7.8	7.5					
Age (years)	62.2	61.2	60.7	61.9	63.5	64.8	66.2	69.7
Gender (%)								
Female	51	53	54					
Male	49	47	44					
Race (%)								
White	68	36	23					
Non-Hispanic Black	12	35	20					
Hispanic	7	15	50					
Chinese	13	13	7					
Education (%)								
<High school	4	15	35					
High school	6	19	24					
Some college	21	33	27					
College/graduate school	69	32	14					
Household Income (%)								
<$20,000	10	22	38	23	23	23	22	20
$20,000-$49,999	21	40	44	37	37	36	36	35
$50,000-$74,999	15	19	13	17	16	17	17	17
$75,000	54	19	5	23	23	24	25	28
Smoking								
Never	49	53	50	50	47	45	45	45
Former	43	34	34	37	42	44	45	47
Current	8	13	17	13	11	11	10	8
Current alcohol use	73	54	47	56	51	50	45	44
Physical activity (MET min/wk)	1312	825	525	840	743	840	840	1650
BMI (kg/m2)	26.8	28.6	29.1	28.3	28.4	28.3	28.4	28.4
History of diabetes (%)	7	12	15	12	15	17	19	23
History of hypertension (%)	36	46	43	44	46	48	52	59
SBP	121.7	126.8	126.4	125.9	124.2	123.2	123.5	123.8
LDL	115.7	117.6	118.6	117.2	113.6	117.8	111.2	106.2

Table 4.2 – Parameter estimates and standard errors from repeated measures linear regression models investigating the relationship between selected lagged measures of time-varying covariates (at t-12) and neighborhood factor score over time, the Multi-Ethnic Study of Atherosclerosis

Time-varying covariates	Estimate	SE
Household Income (%)		
<$20,000	Ref	
$20,000-$49,999	0.0009	0.002
$50,000-$74,999	-0.02[***]	0.003
$75,000	-0.03[***]	0.003
Working		
No	Ref	
Yes	-0.02[***]	0.002
Smoking		
Never	Ref	
Former	-0.1[***]	0.005
Current	-0.1[***]	0.004
Physical activity (MET min/wk)	0.0004[***]	0.0001
History of diabetes (%)		
No	Ref	
Yes	0.02[***]	0.003

[***] $P < 0.0001$
Adjusted for age, gender, race/ethnicity, education, household income, working, smoking, physical activity, and diabetes

Table 4.3 – Odds ratios and 95% confidence intervals (for dichotomous covariates) and parameter estimate and standard error (for continuous physical activity), estimating the association between lagged neighborhood SES (at *t*-12) and selected time-varying covariates, the Multi-Ethnic Study of Atherosclerosis

Time-varying covariates	OR or Parameter Estimate	95%CI or SE
Lowest household income (<$20,000)	1.28	1.20, 1.35
Working (Yes)	1.07	1.02, 1.13
Smoking (Never)	1.03	0.99, 1.08
History of diabetes (Yes)	1.13	1.07, 1.19
Physical activity (MET min/wk)	-0.06[***]	0.01

[***]$P < 0.0001$
Adjusted for age, gender, race/ethnicity, education, household income, working, smoking, physical activity, and diabetes

Table 4.4 – Hazard ratios or odds ratios of CVD per 1-unit increase in neighborhood SES score from traditional models and marginal structure models (MSM)

	Traditional Models (unweighted)		MSM
Neighborhood SES Factor	Model 1	Model 2a	Model 2b
Long-term cumulative avg[¥]	1.18 (1.08-1.28)	1.13 (1.03-1.24)	1.07 (0.97-1.17)
Recent exposure[§]	1.18 (1.07-1.29)	1.15 (1.04-1.27)	1.08 (1.00-1.19)

[¥]Long-term cumulative average, defined as the mean exposure from baseline through t-1
[§]Recent average exposure, defined as the mean exposure from t-12 through t-1

Model 1 – Adjusted for age, gender, education (time-invariant covariates)
Model 2 – Additionally adjusted for income, currently working, log physical activity, bmi, smoking status, diabetes (time-varying covariates)

CHAPTER 4 REFERENCES

1. Cubbin C, Sundquist K, Ahlén H, et al. Neighborhood deprivation and cardiovascular disease risk factors: protective and harmful effects. *Scand J Public Health* 2006;34(3):228-37.
2. Diez Roux A. Residential environments and cardiovascular risk. *J Urban Health* 2003;80(4):569-89.
3. Winkleby M, Sundquist K, Cubbin C. Inequities in CHD incidence and case fatality by neighborhood deprivation. *Am J Prev Med* 2007;32(2):97-106.
4. Lee IM, Ewing R, Sesso HD. The built environment and physical activity levels: the Harvard Alumni Health Study. *Am J Prev Med* 2009;37(4):293-8.
5. Borrell L, Diez Roux A, Rose K, et al. Neighbourhood characteristics and mortality in the Atherosclerosis Risk in Communities Study. *Int J Epidemiol* 2004;33(2):398-407.
6. Chaix B, Rosvall M, Merlo J. Recent increase of neighborhood socioeconomic effects on ischemic heart disease mortality: a multilevel survival analysis of two large Swedish cohorts. *Am J Epidemiol* 2007;165(1):22-6.
7. Diez Roux A, Merkin S, Arnett D, et al. Neighborhood of residence and incidence of coronary heart disease. *N Engl J Med* 2001;345(2):99-106.
8. Diez-Roux A, Nieto F, Muntaner C, et al. Neighborhood environments and coronary heart disease: a multilevel analysis. *Am J Epidemiol* 1997;146(1):48-63.
9. Sundquist K, Winkleby M, Ahlén H, et al. Neighborhood socioeconomic environment and incidence of coronary heart disease: a follow-up study of 25,319 women and men in Sweden. *Am J Epidemiol* 2004;159(7):655-62.
10. Brownson R, Kelly C, Eyler A, et al. Environmental and policy approaches for promoting physical activity in the United States: a research agenda. *J Phys Act Health* 2008;5(4):488-503.
11. Hart C, Ecob R, Smith G. People, places and coronary heart disease risk factors: a multilevel analysis of the Scottish Heart Health Study archive. *Soc Sci Med* 1997;45(6):893-902.
12. Smith G, Hart C, Watt G, et al. Individual social class, area-based deprivation, cardiovascular disease risk factors, and mortality: the Renfrew and Paisley Study. *J Epidemiol Community Health* 1998;52(6):399-405.
13. Sundquist K, Malmström M, Johansson S. Neighbourhood deprivation and incidence of coronary heart disease: a multilevel study of 2.6 million women and men in Sweden. *J Epidemiol Community Health* 2004;58(1):71-7.
14. Hernán M, Brumback B, Robins J. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 2000;11(5):561-70.
15. Robins J, Hernán M, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;11(5):550-60.
16. Ilomaki J, Hajat A, Kauhanen J, et al. Relationship between alcohol consumption and myocardial infarction among ageing men using a marginal structural model. *European journal of public health* 2012;22(6):825-30.
17. Nandi A, Glymour MM, Kawachi I, et al. Using marginal structural models to estimate the direct effect of adverse childhood social conditions on onset of heart disease, diabetes, and stroke. *Epidemiology* 2012;23(2):223-32.

18. Sampson RJ, Sharkey P, Raudenbush SW. Durable effects of concentrated disadvantage on verbal ability among African-American children. *Proceedings of the National Academy of Sciences of the United States of America* 2008;105(3):845-52.

19. Cerda M, Diez-Roux AV, Tchetgen ET, et al. The relationship between neighborhood poverty and alcohol use: estimation by marginal structural models. *Epidemiology* 2010;21(4):482-9.

20. Diez Roux A. Neighborhoods and health: where are we and were do we go from here? *Rev Epidemiol Sante Publique* 2007;55(1):13-21.

21. Hernan MA, Robins JM. Estimating causal effects from epidemiological data. *J Epidemiol Community Health* 2006;60(7):578-86.

22. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol* 2008;168(6):656-64.

23. Kaufman JS. Marginalia: comparing adjusted effect measures. *Epidemiology* 2010;21(4):490-3.

24. Foraker RE, Rose KM, Kucharska-Newton AM, et al. Variation in rates of fatal coronary heart disease by neighborhood socioeconomic status: the atherosclerosis risk in communities surveillance (1992-2002). *Ann Epidemiol* 2011;21(8):580-8.

25. Hamano T, Kawakami N, Li X, et al. Neighbourhood environment and stroke: a follow-up study in Sweden. *PloS one* 2013;8(2):e56680.

26. Kawakami N, Li X, Sundquist K. Health-promoting and health-damaging neighbourhood resources and coronary heart disease: a follow-up study of 2 165 000 people. *J Epidemiol Community Health* 2011;65(10):866-72.

27. Pickett KE, Pearl M. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. *J Epidemiol Community Health* 2001;55:111-22.

28. Pollack CE, Slaughter ME, Griffin BA, et al. Neighborhood socioeconomic status and coronary heart disease risk prediction in a nationally representative sample. *Public health* 2012;126(10):827-35.

29. Brown AF, Liang LJ, Vassar SD, et al. Neighborhood disadvantage and ischemic stroke: the Cardiovascular Health Study (CHS). *Stroke; a journal of cerebral circulation* 2011;42(12):3363-8.

30. Neser WB, Tyroler HA, Cassel JC. Social disorganization and stroke mortality in the black population of North Carolina. *Am J Epidemiol* 1971;93(3):166-75.

31. Gerber Y, Koton S, Goldbourt U, et al. Poor neighborhood socioeconomic status and risk of ischemic stroke after myocardial infarction. *Epidemiology* 2011;22(2):162-9.

32. Tonne C, Schwartz J, Mittleman M, et al. Long-term survival after acute myocardial infarction is lower in more deprived neighborhoods. *Circulation* 2005;111(23):3063-70.

33. Villanueva C, Aggarwal B. The association between neighborhood socioeconomic status and clinical outcomes among patients 1 year after hospitalization for cardiovascular disease. *Journal of community health* 2013;38(4):690-7.

34. Morris RW, Wannamethee G, Lennon LT, et al. Do socioeconomic characteristics of neighbourhood of residence independently influence incidence of coronary heart disease and all-cause mortality in older British men? *European journal of cardiovascular prevention and rehabilitation : official journal of the European Society of Cardiology, Working Groups on Epidemiology & Prevention and Cardiac Rehabilitation and Exercise Physiology* 2008;15(1):19-25.

35.	Merlo J, Ohlsson H, Chaix B, et al. Revisiting causal neighborhood effects on individual ischemic heart disease risk: a quasi-experimental multilevel analysis among Swedish siblings. *Soc Sci Med* 2013;76(1):39-46.

36.	Chaix B. Geographic life environments and coronary heart disease: a literature review, theoretical contributions, methodological updates, and a research agenda. *Annu Rev Public Health* 2009;30:81-105.

37.	Macintyre S. Deprivation amplification revisited; or, is it always true that poorer places have poorer access to resources for healthy diets and physical activity? *Int J Behav Nutr Phys Act* 2007;4:32.

38.	Augustin T, Glass TA, James BD, et al. Neighborhood psychosocial hazards and cardiovascular disease: the Baltimore Memory Study. *Am J Public Health* 2008;98(9):1664-70.

39.	Elliott M. The stress process in neighborhood context. *Health Place* 2000;6(4):287-99.

40.	Hill TD, Ross CE, Angel RJ. Neighborhood disorder, psychophysiological distress, and health. *Journal of health and social behavior* 2005;46(2):170-86.

41.	Black PH, Garbutt LD. Stress, inflammation and cardiovascular disease. *Journal of psychosomatic research* 2002;52(1):1-23.

42.	Kaplan JR, Pettersson K, Manuck SB, et al. Role of sympathoadrenal medullary activation in the initiation and progression of atherosclerosis. *Circulation* 1991;84(6 Suppl):VI23-32.

43.	Merkin SS, Basurto-Davila R, Karlamangla A, et al. Neighborhoods and cumulative biological risk profiles by race/ethnicity in a national sample of U.S. adults: NHANES III. *Ann Epidemiol* 2009;19(3):194-201.

44.	Moyer AE, Rodin J, Grilo CM, et al. Stress-induced cortisol response and fat distribution in women. *Obes Res* 1994;2(3):255-62.

45.	Inagami S, Cohen DA, Finch BK. Non-residential neighborhood exposures suppress neighborhood effects on self-rated health. *Soc Sci Med* 2007;65(8):1779-91.

CHAPTER 5

DISCUSSION

Overall summary of research and findings

Though there is a sizeable body of evidence indicating that there is an association between the neighborhood environment and cardiovascular disease and its risk factors (specifically physical activity and diet), methodological limitations related to residential self-selection and the absence of prospective studies with time-varying data on neighborhoods and individuals have warranted further examination of these associations using 1) time-varying information on neighborhoods and a range of individual-level covariates and 2) more sophisticated longitudinal methods. Much of the research in this area over the past decade has been cross-sectional, and only a few recent longitudinal studies have attempted to account for potential reciprocal relationships between neighborhood exposures and individual-level outcomes, usually by adjusting for individual self-reported preferences measured at one point in time (1-3), and few studies have addressed the possible presence of time-varying confounding (4). Although previous studies have provided evidence of a true association between the neighborhood environment and CVD-related outcomes, the debate in the literature regarding whether these associations are causal has continued. This dissertation contributes to the literature by using longitudinal data to 1) investigate if individual physical activity and diet are related to moving to or staying in neighborhoods with better access to recreational facilities and favorable foods stores, 2)

investigate whether changes in neighborhood recreational facilities over time are related to changes in physical activity over time, and 3) investigate whether neighborhood socio-economic composition is associated with incident CVD after accounting for possible time-varying confounding by individual-level factors.

Chapter 2 examined whether higher physical activity predicted whether individuals moved to neighborhoods with better availability of resources that supported physical activity, and also examined whether better diet predicted whether individuals moved to neighborhoods with greater availability of favorable food stores. Similar analyses were conducted among individuals who did not move between exams to examine whether changes to the neighborhood recreational facilities or favorable food stores around stable residents were predicted by their physical activity and diet. Finding evidence for these associations would suggest that self selection (persons with more favorable behaviors moving on average to areas with better resources) or endogeneity (the behavior of residents affecting subsequent changes in their neighborhoods) need to be accounted for in studies of neighborhood health effects. Movers who were more physically active prior to moving were more likely to move to neighborhoods with greater availability of recreational facilities than movers who were less active. This association was slightly weakened after adjusting for recreational density prior to the move, suggesting that some of the association observed prior to adjustment was due to the cross-sectional association between recreational density and physical activity prior to moving and the correlation between recreational densities in the neighborhood of origin and the new neighborhood. Among non-movers, physical activity was not strongly associated with a more favorable change in the availability of facilities, suggesting that more physically active stable residents were not effecting favorable changes in

their neighborhoods. For both movers and non-movers, diet did not predict moving to or staying in neighborhoods with better availability of favorable food stores. However, the analyses were limited to using a single time-point measure of diet, since diet was only measured at baseline.

By examining "reciprocal relationships", these analyses further the understanding of the role of these specific health behaviors in residential selection, which allows for more meaningful interpretations of previous and future findings related to the influence of the neighborhood environment on individual physical activity and diet. In other words, the findings in Chapter 2 that residential selection does occur based on physical activity behavior in this sample of multi-ethnic adults suggests that residential self-selection may play a role in the association between the neighborhood environment and physical activity, especially among adults who move residences. Based on these results, attempting to account for self-selection in future studies, either by using robust longitudinal designs, adjustment for individual preferences, or both, continues to be of importance.

Chapter 3 examined the association between neighborhood availability of recreational facilities and changes in physical activity and found that individuals who experienced greater increases in the densities of recreational resources around their homes experienced a less pronounced decline in recreational physical activity over time. Although on average physical activity declined in this sample of middle-aged and older adults, the decline was almost eliminated for individuals who experienced the greatest increases in the density of facilities. The association between availability of facilities and physical activity change was stronger among older adults, which is consistent with the limited previous research investigating whether older adults may be more

influenced by their environments (5, 6). Additional research in older adults is needed to further

corroborate this finding. Although Chapter 2 found some evidence for the selection of

individuals into neighborhoods with better physical activity resources based on previous physical

activity level, the longitudinal change versus change analyses from Chapter 3 suggest that there

is an association between the change in objective availability of physical activity resources and

change in physical activity. Since the analyses from Chapter 3 did not adjust for any measures

of self-selection, completely ruling out the possibility that some of the observed association may

be due to reciprocal relationships is not possible. However, the change versus change analysis is

a major improvement over cross-sectional analyses and longitudinal analyses that use only a

single measure of the neighborhood exposure or physical activity. Taking the results from

chapter 2 and 3 together, is likely that there is some degree of residential selection that is

involved in the relationship between the neighborhood physical environment and physical

activity. Therefore, analyses like those conducted in chapter 3 of this dissertation, examining

change over time in both the neighborhood exposure and physical activity move us one step

closer to estimating the causal association between the neighborhood environment and physical

activity.

Chapter 4 used longitudinal data to investigate the relationship between a composite measure of

neighborhood socio-economic composition and incident CVD by 1) contrasting estimates from

using long-term cumulative and recent measures of exposure and 2) contrasting estimates from

traditional regression models, adjusting for time-varying covariates, with marginal structural

models (MSM) which account for time-varying confounding. Increasing neighborhood

deprivation was associated with a greater incidence of CVD using both traditional models and

MSMs, even after adjusting for individual-level socio-economic factors and time-varying CVD risk factors. These associations did not differ by exposure accumulation and timing; the results for cumulative exposure were similar to those for recent exposure. Adjusting for individual-level biomedical risk factors, which may be mediators in the pathway between neighborhood SES and incident CVD, attenuated the estimates from the traditional models slightly. Estimates from MSMs yielded slightly weaker associations between increasing neighborhood deprivation and incident CVD. Since estimates from the MSMs and unweighted models cannot be directly compared (7) and the confidence intervals between the two sets of estimate overlap greatly, the comparison of the results suggests that there is very little difference between the MSM, which adjusts for time-varying confounding, and traditional models that did not. Since the exposure was measured at only two time points, it may not have been possible to capture the true presence of time-varying confounding in these data. Therefore, the conservative conclusion is that, though it does not appear that time-varying confounding yields biased estimates in the association between neighborhood socio-economic composition and incident CVD using these data, investigating the presence of time-varying confounding empirically when studying this relationship in the future may be of value, especially if the exposure has greater variation over time.

This dissertation moves forward the quest to determine whether there is a causal relationship between the neighborhood environment and physical activity and the neighborhood environment and incident CVD, though the implications of this work can be extended to and examined in other neighborhood exposure-health outcome relationships. If there is a causal association between specific features of the built environment and physical activity, for example, then the

implication is that increasing the availability of recreational facilities should result in improvements in individual physical activity. However, since experimental studies may not be feasible or may be difficult to implement in the context of neighborhoods and health, longitudinal analyses like those investigated in this dissertation move us closer to understanding if these relationships may be causal, though even well-designed longitudinal studies using rigorous methods may be limited in their ability to make causal inferences (8).

Strengths and Limitations

Measurement error is possible in the physical activity and diet measures since they were self-reported. Though the form of the MESA physical activity questionnaire likely results in overestimates of self-reported physical activity, these overestimates are unlikely to result in any systematic bias. Diet, as measured by the Alternative Healthy Eating Index, was measured only at baseline, and examining changes over time in diet in relation to residential selection may be needed to better estimate if diet is associated with selection into neighborhoods based on availability of healthy foods. It is also possible that more specific measures of diet, such as fast food consumption or fruit/vegetable consumption, which were not investigated in this dissertation, are better predictors of selection into neighborhoods based on healthy food availability.

The measures of recreational facility and food store densities that were used in these analyses lacked specific information about the facilities such as quality of facilities, hours of operation, and cost. Although the density measures reflected the presence of facilities within a 1-mile buffer around participants' homes, information on whether participants were aware of these

facilities or using these facilities was not available. Another limitation with using commercial databases to obtain the data on recreational facilities and food stores is that there may be considerable measurement error in the databases (9-11). Since the 1-mile buffer was measured around residential addresses, the density measures did not account for facilities and food stores near the workplace. With respect to estimating the impact of the objective physical environment on physical activity, measures on resources such as parks were not available, and these other features of the physical environment may be related to physical activity.

These analyses also did not account for perceived measures of physical activity resources or healthy food availability, and it is possible that self-selection based on physical activity or healthy eating behaviors is more strongly related to perceived access to resources. It is also possible that individual preferences for availability of resources, which was not available for these analyses, better predicts residential selection than health behaviors do. Though we adjusted for several time-invariant and time-varying characteristics, including neighborhood-level characteristics where appropriate, it is not possible to rule out residual confounding due to unmeasured covariates, especially unmeasured covariates that may be related to selection into neighborhoods, such as individuals' preferences, which are difficult to measure.

Chapter 4 uses a composite measure of neighborhood socio-economic composition as the exposure, calculated using Census data from only 2 time-points. Though this is consistent with past studies that have used various Census measures of neighborhood socio-economic status to investigate possible relationships with CVD outcomes and allows direct comparison with past work, more specific measures of the physical and social environments of neighborhoods,

measured at multiple time points, may better elucidate the mechanisms involved between the neighborhood environment and CVD. Since the exposure was only available at 2 time points for these analyses, we interpolated the available data to the other time points, which may have resulted in some measurement error.

The research conducted for this dissertation used data from the Multi-Ethnic Study of Atherosclerosis, which follows adults who were aged 45-84 at baseline. The generalizability of the results from this work may not extend to younger adults, especially because factors that may impact residential selection may differ greatly for younger adults, for whom residential selection based on health behaviors and/or individual preferences may be overshadowed by selection based on life stage developments, such as job changes, marriage, good school districts, etc. Given the large, multi-ethnic, multi-site sample, the results of this work are likely generalizable to middle-aged and older adults in the U.S., and possibly other developed countries with similar distributions of race and socio-economic status.

Strengths of this dissertation include the availability of longitudinal data on 1) objective measures of recreational densities and food store densities; 2) physical activity measures that accounted for both duration and intensity of activities; 3) individual and neighborhood-level socio-economic characteristics; 4) individual-level behavioral and biomedical CVD risk factors; and 5) residential mobility and housing information, including data on when individuals moved, to where they moved, duration of residence, and whether they owned or rented their homes; 6) cardiovascular outcomes that were validated using medical records and death certificates. A major strength of this dissertation is the use of time-varying measures on all of the above

variables, since the majority of studies in the past that have investigated the association between the neighborhood environment and physical activity or CVD lack time-varying measures on one or more of these variables. The second major strength is the use of longitudinal models that estimate change versus change, and longitudinal models which control for time-varying confounding. Both of these methodological approaches are useful for estimating associations between the neighborhood and health outcomes in the presence of residential selection into neighborhoods based on individual-level factors, and have been implemented in few studies.

Public health significance, policy implications and future research directions

Physical inactivity and poor diet have made the U.S. a nation with a less-than-ideal proportion of overweight and out-of-shape individuals in just a few generations. As individuals age, declining physical inactivity and increasing obesity may increase the risk of developing chronic diseases, such as CVD, diabetes, certain cancers, and depression, among others. The benefits of physical activity on cardiovascular health specifically have been documented extensively (12, 13), and CVD is the leading cause of death for men and women worldwide (14). To target changes in individual-level CVD risk factors to reduce the burden of CVD, intervening on multiple-levels, as posited by the socio-ecological framework (15), may be needed. Studies documenting the associations between neighborhood-level exposures and individual-level CVD risk factors as well as incident CVD can play a significant role in providing evidence for community-based interventions and neighborhood-level changes to be coupled with individual-level interventions.

The current evidence is strengthened by the work of this dissertation, which highlights the application of longitudinal methods to investigate the relationship between the neighborhood

97

environment and physical activity and incident CVD, while empirically investigating issues

related to residential selection as well as time-varying confounding. Though many studies have

investigated these relationships, there remain questions as to how much of the observed

associations are due to the possibility of reciprocal relationships. Understanding the true

magnitude of the association between neighborhood exposures and these health outcomes has

implications for the prevention of CVD and other physical activity-related health outcomes,

since it is difficult to argue that neighborhoods should increase their availability of resources that

support healthy behaviors without understanding if these changes would actually impact

individuals' behaviors.

To inform policy applications, there is a need for this type of longitudinal evidence to explain

whether and to what degree residential self-selection plays a role in the relationship between

neighborhoods and health. Do community design and neighborhood resources play a causal role

in influencing individual behavior, or do individuals with certain lifestyles or preferences choose

environments with particular community designs or availability of health-promoting resources?

Recent and past work has supported both sides of the debate, in that self-selection does appear to

play a role, but the neighborhood environment appears to influence individual outcomes after

accounting for self-selection. Empirically investigating the presence of residential selection, as

in Chapter 2 of this dissertation, is an important contribution to understanding the role

neighborhood resource availability may play in influencing individual behavior and health. It is

likely that both the neighborhood environment and individual preferences are involved, and

empirical evidence supporting this can inform the development of policies that aim to improve

the neighborhoods in which individuals live.

Though self-selection has been cited as a major issue in neighborhood health research, the presence of selection into neighborhoods based on individual factors does not necessarily decrease the importance of the influence of the neighborhood environment on health. Understanding the influence of the neighborhood environment on health is complicated because there are complex relationships between the neighborhood environment, individuals' preferences for particular behaviors, individuals' preferences to live in particular types of neighborhoods (which may or may not be related to behavioral preferences), individuals' *perceptions* of their neighborhood environments, and the objective presence of resources in their neighborhoods. Furthermore, not all individuals will be affected by their environments in the same way, or exert influences on their environments in the same way. The most health-conscious individuals may find a way to keep up their healthy behaviors regardless of their environments, and the least health-conscious individuals may not change their behaviors even if they live in health-promoting environments. It is likely that most individuals fall somewhere the middle, where they are impacted by their environments to varying degrees. Using the neighborhood physical environment and physical activity as an example, improvements to the neighborhood environment can affect individuals by 1) facilitating physical activity for those who prefer to engage in healthy lifestyles and 2) influencing those who may not be driven by preferences to engage in physical activity due to the ease of access to resources that support and encourage a physically active lifestyle. It is likely that part of the association between the environment and physical activity is a direct effect and that part of the association is the result of residential selection, highlighting the need for public health strategies that focus on developing activity-friendly neighborhoods as well as targeting individual preferences and attitudes.

Another reason the presence of self-selection should not diminish the importance of the influence of the neighborhood environment on health is because many individuals cannot actually select where they live based on their health behaviors or preferences due to socio-economic constraints or other factors. For this reason, even if these individuals prefer to engage in healthy behaviors and live in areas with health-promoting resources, they may move to or stay in environments that do not support their preferred behaviors because they have no other options. Creating neighborhood environments that support and encourage healthy behaviors and minimize psychosocial stress, which has been shown to be related to CVD outcomes, continues to be of importance for the health of individuals who are not able to select environments based on their preferences. Future studies examining residential selection and preferences should attempt to collect information not only preferences for health-promoting resources, but also on individuals' perceived abilities to select neighborhoods based on these preferences. Though several studies have collected data on neighborhood preferences, none have published results using data on changing preferences over time. Future longitudinal studies should collect data on how preferences are changing over time, in addition to collecting data on changing neighborhood factors and individual behaviors.

Future studies should also consider the interaction between preferences and objective measures of the neighborhood environment and/or the interaction between the perceived environment and the objective environment. Using physical environment and physical activity as an example once again, if selection into neighborhoods based on individual factors plays a large role in the association between the neighborhood and health-related outcomes, then results investigating the

association between the physical environment and physical activity should find weaker associations among those whose preferences indicate that they do not prefer to live in areas with greater availability of facilities or who perceive that their environment does not have many facilities to support physical activity. One recent study that has examined this type of interaction found that those individuals who did not have a preference for exercising or living in a neighborhood with greater availability of facilities were *more* likely to be influenced by their environments (16). This study showed that the association between neighborhood physical environment and leisure-walking was stronger among those with more perceived barriers to exercise, less perceived psychosocial support, and less perceived benefit of exercise, suggesting that the physical environment benefited those individuals that had a lower preference to exercise and a lower preference for exercise-related facilities. More studies should consider stratifying by individual preferences and perceptions of the environment to better understand how much of an impact the presence of self-selection may have on the association between the neighborhood and health-related outcomes. Future work in this area could be further enhanced by including information on residential history and other variables related to the lifecourse to better capture preferences and behaviors that may be related to residential selection over a longer period time. Especially for individuals who are not able to select their neighborhoods based on preferences, understanding which subgroups of these individuals may benefit more from living in environments that support healthy behaviors may target neighborhood-level changes to specific areas. This may help to reduce racial and socio-economic disparities in physical activity and other health outcomes.

Since true experimental studies are not likely feasible in the context of neighborhoods and health, longitudinal observational studies continue to be of great importance. In addition to longitudinal observation studies, another direction for future research is the design of quasi-experimental studies and natural experiments. These types of studies can, for example, directly measure changes in individual-level outcomes that may result after re-location to a particular type of neighborhood or measures changes in individual outcomes in a neighborhood after some major changes to the built environment are made. A few of these studies have been undertaken in the past few years (6), but further carefully designed studies, perhaps in collaboration with urban planners and community organizations may yield more insight into the direct effect neighborhood-level changes may have on individual outcomes.

Following from chapter 4 of this dissertation, since selection into neighborhoods due to individual-level factors can result in time-varying confounding, future studies investigating the association between the neighborhood environment and incident CVD should consider using methods to adjust for time-varying confounding. The few studies that have adjusted for time-varying confounding in the context of neighborhoods and health-related outcomes have found that adjusting for time-varying confounding yielded less biased estimates than using traditional methods (4). Although the results from Chapter 4 of this dissertation did not observe a difference between models adjusting for time-varying confounding and traditional models, future studies should examine the association between the neighborhood environment and incident CVD with the exposure measured at more than one time point. The absence of strong time-varying confounding in the data used for the Chapter 4 analyses does not necessarily mean that

time-varying confounding is not present, only that it was not apparent using the available measures of exposure and covariates in MESA.

Future studies should also consider more specific neighborhood constructs, such as objective and perceived features of the neighborhood physical and social environments, which are more closely related to the mechanisms that may be involved in the pathway between neighborhood socio-economic composition and CVD. Future studies that investigate the association between objective features of the physical environment with incident CVD outcomes (or other health-related outcomes) should still consider using approaches that adjust for time-varying confounding, including the possibility of time-varying confounding by neighborhood SES.

Conclusion

Overall, this dissertation addresses important methodological limitations to provide evidence with which to make more robust inferences in the study of the neighborhood environment and CVD-related outcomes. The results of this dissertation showed that there was a longitudinal relationship between the neighborhood environment and physical activity that may be partially explained by residential selection. This work also showed that the application of more advanced methodologies to control for time-varying confounding in the study of neighborhood environment and cardiovascular outcomes led to conclusions similar to what has been found in previous studies. The longitudinal approaches undertaken as part of this dissertation should motivate additional research which continues to address these methodological issues in the study of neighborhoods and health. This is important because more robust evidence in this area of research may inform policies that address the residential environment, such as facilitating access

to recreational facilities and access to healthy foods, and considering the physical environment

when designing or re-building neighborhoods. Ultimately, these policies, in addition to

individual-level interventions, may help promote healthy lifestyles and prevent cardiovascular

disease.

CHAPTER 5 REFERENCES

1. Boone-Heinonen J, Guilkey DK, Evenson KR, et al. Residential self-selection bias in the estimation of built environment effects on physical activity between adolescence and young adulthood. *Int J Behav Nutr Phys Act* 2010;7:70.
2. Pinjari AR, Bhat CR, Hensher DA. Residential self-selection effects in an activity time-use behavior model. *Transportation Research Part B* 2009;43:729–48.
3. Handy SL, Cao X, Mokhtarian PL. The causal influence of neighborhood design on physical activity within the neighborhood: evidence from Northern California. *Am J Health Promot* 2008;22(5):350-8.
4. Cerda M, Diez-Roux AV, Tchetgen ET, et al. The relationship between neighborhood poverty and alcohol use: estimation by marginal structural models. *Epidemiology* 2010;21(4):482-9.
5. Van Cauwenberg J, De Bourdeaudhuij I, De Meester F, et al. Relationship between the physical environment and physical activity in older adults: a systematic review. *Health Place* 2011;17(2):458-69.
6. Ding D, Gebel K. Built environment, physical activity, and obesity: what have we learned from reviewing the literature? *Health Place* 2012;18(1):100-5.
7. Kaufman JS. Marginalia: comparing adjusted effect measures. *Epidemiology* 2010;21(4):490-3.
8. Diez Roux A, Mair C. Neighborhoods and health. *Ann N Y Acad Sci* 2010;1186:125-45.
9. Boone JE, Gordon-Larsen P, Stewart JD, et al. Validation of a GIS facilities database: quantification and implications of error. *Ann Epidemiol* 2008;18(5):371-7.
10. Han E, Powell L, Slater S. Validation of secondary commerical data sources for physical activity facilities in urban and nonurban settings. *Journal of physical activity & health* 2012;9(8):1080-8.
11. Hoehle F, Schootman M. Concordance of commercial data sources for neighborhood-effects studies. *Journal of urban health: bulletin of the New York Academy of Medicine* 2010;87(4):713-25.
12. Batty GD. Physical activity and coronary heart disease in older adults. A systematic review of epidemiological studies. *European journal of public health* 2002;12(3):171-6.
13. Batty GD, Lee IM. Physical activity and coronary heart disease. *Bmj* 2004;328(7448):1089-90.
14. Global status report on noncommunicable disaeses 2010. Geneva, 2011, (Organization WH
15. Duncan C, Jones K, Moon G. Health-related behaviour in context: a multilevel modelling approach. *Soc Sci Med* 1996;42(6):817-30.
16. Ding D, Sallis JF, Conway TL, et al. Interactive effects of built environment and psychosocial attributes on physical activity: a test of ecological models. *Ann Behav Med* 2012;44(3):365-74.