



Neighborhood Greenness and Prostate Cancer: Association and Explanation in Diverse Populations

Citation

Iyer, Hari S. 2020. Neighborhood Greenness and Prostate Cancer: Association and Explanation in Diverse Populations. Doctoral dissertation, Harvard T.H. Chan School of Public Health.

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:42676011>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

Neighborhood Greenness and Prostate Cancer: Association and Explanation in Diverse Populations

Hari S. Iyer

A Dissertation Submitted to the Faculty of
The Harvard T. H. Chan School of Public Health
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Science
in the Department of Epidemiology
Harvard University
Boston, Massachusetts
March, 2020

Neighborhood Greenness and Prostate Cancer: Associations and Explanation in Diverse Populations

Abstract

Prostate cancer (CaP) is the most common non-cutaneous cancer diagnosed among men in the United States. Neighborhood greenness could confer benefits to men at risk of CaP by promoting healthy lifestyles and reducing mortality.

In Chapter 1, we estimated the association between neighborhood greenness and 28-year risk of lethal CaP in the Health Professionals Follow-up Study. We assigned satellite-derived Normalized Difference Vegetation Index (NDVI) with 1 kilometer resolution linked to the participants' home or work address at the start of follow-up. An interquartile range increase in NDVI was associated with 5% lower rate of lethal CaP (aHR: 0.95, 95% CI: 0.88, 1.03), with stronger associations in non-movers (aHR: 0.92, 95% CI: 0.85, 1.01). Inverse associations were observed among men in high (aHR: 0.90, 95% CI: 0.82, 0.99) but not low (aHR: 1.11, 95% CI: 0.95, 1.29, $P_{\text{het}}=0.086$) population density areas. There was no evidence of mediation by vigorous physical activity.

In Chapter 2, we estimated the association between neighborhood greenness and 10-year cause-specific mortality among men with CaP in the Pennsylvania Cancer Registry. NDVI with 250m resolution was assigned to participants' residential address at diagnosis. Comparing men in highest to lowest NDVI quintile, all-cause (aHR: 0.88, 95% CI: 0.84, 0.92, $P_{\text{trend}}<0.0001$), prostate-specific (aHR: 0.88, 95% CI: 0.80, 0.99, $P_{\text{trend}}=0.0021$), and cardiovascular-specific (aHR: 0.82, 95% CI: 0.74, 0.90, $P_{\text{trend}}<0.0001$) mortality were lower. Hypothetical interventions

to increase NDVI led to non-significant reductions in all-cause (-5.3%) and prostate-specific (-23.2%), but not cardiovascular-specific mortality disparities (+50.5%).

In Chapter 3, we estimated the association between neighborhood greenness and cardiometabolic risk factors in a cross-sectional, multi-country study in sub-Saharan Africa. NDVI with 250m resolution was assigned to a geocode corresponding to the center of the school or village where the participant was recruited. A 0.11 unit NDVI increase was associated with lower BMI (β : -1.01, 95% CI: -1.35, -0.67), lower odds of overweight/obesity (aOR: 0.73, 95% CI: 0.62, 0.85), diabetes (aOR: 0.77, 95% CI: 0.62, 0.96), and having ≥ 3 allostatic load components compared to none (aOR: 0.66, 95% CI: 0.52, 0.85). Associations for BMI, overweight/obesity, and allostatic load remained statistically significant after Bonferroni correction.

Table of Contents

Abstract.....	ii
List of Figures with Captions.....	v
List of Tables with Captions.....	vi
Acknowledgements.....	viii
Introduction.....	1
Chapter 1: The Association between Neighborhood Greenness and Incidence of Lethal Prostate Cancer: A Prospective Cohort Study	10
Abstract.....	11
Introduction.....	13
Methods.....	14
Results.....	19
Discussion.....	27
References.....	31
Supplementary Appendix.....	37
Chapter 2: The Contribution of Neighborhood Greenness to Mortality among Men with Prostate Cancer: A Registry-based Cohort Study of Black and White Men	47
Abstract.....	48
Introduction.....	50
Methods.....	51
Results.....	55
Discussion.....	72
References.....	76
Supplementary Appendix.....	83
Chapter 3: Neighborhood Greenness and Burden of Non-communicable Diseases in Sub-Saharan Africa: A Multi-country Cross-sectional Study	90
Abstract.....	92
Introduction.....	94
Methods.....	95
Results.....	101
Discussion.....	109
References.....	113
Supplementary Appendix.....	120

List of Figures with Captions

Figure 1.1. Hazard Ratios and Confidence Intervals for the Association Between Baseline Normalized Difference Vegetation Index (NDVI) and Lethal Prostate Cancer Incidence in the Health Professionals' Follow-up Study, United States, 1986-2014.....	23
Figure S1.1. Health Professionals' Follow-up Study geocoded locations (Home, Work, Other) at baseline (1988).....	41
Figure S1.2. NDVI 1989 values at Health Professionals' Follow-up Study geocoded locations (Home, Work, Other).....	42
Figure 2.1. Participant Residential Address Locations in Pennsylvania Cancer Registry Prostate Cancer Cohort Study from 2000 to 2015	59
Figure 2.2. Normalized Difference Vegetation Index (NDVI) July 2000 values at Pennsylvania Cancer Registry participants' residential address locations from 2000 to 2015	60
Figure 2.3. Standardized 10-year Cause-specific (All-Cause, Cardiovascular-, Prostate-specific) Mortality Risk and 95% Confidence Intervals Under Three Scenarios ^a of Normalized Difference Vegetation Index (NDVI) at Diagnosis among Black and White Men with Prostate Cancer in Pennsylvania)	70

List of Tables with Captions

Table 1.1. Age-standardized Characteristics by Quintile of Normalized Difference Vegetation Index (NDVI) Among Men in the Health Professionals Follow-up Study From 1986 to 2014	20
Table 1.2. Hazard Ratios for the Association Between Baseline Normalized Difference Vegetation Index (NDVI) and Lethal Prostate Cancer Incidence in the Health Professionals Follow-up Study, United States, 1986-2014, Stratified by Population Density (high: ≥ 1000 , low: < 1000 people/mi ²) and Address Type (Work, Home)	24
Table 1.3. Hazards Ratios for the Association Between Baseline Normalized Difference Vegetation Index (NDVI) and Lethal Prostate Cancer Incidence in the Health Professionals Follow-up Study, United States, 1986-2014, Stratified by Address, Prostate-specific Antigen Screening, and Census Region	26
Table S1.1. Odds Ratios for the Association Between Normalized Difference Vegetation Index (NDVI) and Quintiles of Vigorous Physical Activity (PA), Health Professionals Follow-up Study, 1986 (N=47,958).....	43
Table S1.2 Hazard Ratios for the Association Between Baseline Normalized Difference Vegetation Index (NDVI) and Lethal Prostate Cancer Incidence in the Health Professionals Follow-up Study, United States, 1986-2014	44
Table S1.3. Hazard ratios for the Association Between Cumulative Updated Average Normalized Difference Vegetation Index (NDVI) and Lethal Prostate Cancer Incidence in the Health Professionals Follow-up Study, United States, 1986-2014, Stratified by Population Density	45
Table S1.4. E-values for Hazard Ratios for the Association Between Baseline Normalized Difference Vegetation Index (NDVI) and Lethal Prostate Cancer, Total Population and Men who did not Move During Follow-up, Health Professionals Follow-up Study, United States, 1986-2014, stratified by population density.....	46
Table 2.1. Descriptive Characteristics of Pennsylvania Cancer Registry Cohort Stratified by Quintile of Baseline Normalized Difference Vegetation Index (NDVI) Quintile, from 2000 to 2015	57
Table 2.2. Cox Proportional Hazards Models for Association between Normalized Difference Vegetation Index (NDVI) at Diagnosis and Cause-specific Mortality among Pennsylvania Prostate Cancer Patients Diagnosed between 2000 and 2015	62
Table 2.3. Cox Proportional Hazards Models for Association between Normalized Difference Vegetation Index (NDVI) at Diagnosis and Cause-specific Mortality among Pennsylvania Prostate Cancer Patients Diagnosed between 2000 and 2015, Stratified by Race, Stage and Population Density	64
Table 2.4. Cause-specific 10-year Mortality Risks, Disparities, and 95% Confidence Intervals Under Three Levels of Normalized Difference Vegetation Index (NDVI) at Diagnosis among Black and White Men with Prostate Cancer in Pennsylvania, 2000 to 2015.....	68
Table S2.1. Sensitivity Analysis for Competing Risks: Cox Proportional Hazards Model Estimates of Association between Normalized Difference Vegetation Index (NDVI) at Diagnosis and Prostate-specific Mortality and Cardiovascular Mortality using Inverse Probability of Censoring Weights	87
Table S2.2. E-values for Robustness to Unmeasured Confounding for Cox Model Hazard Ratio Estimates of the Association between Normalized Difference Vegetation Index (NDVI) at Diagnosis and Rate of Cause-specific Mortality among men with prostate cancer	88

Table S2.3. Cox Proportional Hazards Models for Association between Cumulative Updated Average NDVI over follow-up and Cause-specific Mortality among Pennsylvania Prostate Cancer Patients Diagnosed between 2000 and 2015	89
Table 3.1. Descriptive Characteristics of Study Participants from four Sites in Three sub-Saharan African Countries	102
Table 3.2. Cross-sectional associations between Normalized Difference Vegetation Index (0.11 unit increase) and components of allostatic load using Multiple Imputation among Participants from four Sites in Three sub-Saharan African Countries.....	105
Table 3.3. Cross-sectional associations between Normalized Difference Vegetation Index (0.11 unit increase) and Non-Communicable Disease Prevalence Using Log-Linear Models using Multiple Imputation to Evaluate Mediation by Body Mass Index among Participants from four Sites in Three sub-Saharan African Countries.....	108
Table S3.1. Cross-sectional associations between Normalized Difference Vegetation Index (0.11 unit increase) and components of allostatic load among Participants from four Sites in Three sub-Saharan African Countries (complete case).....	120
Table S3.2. Sensitivity Analysis for Unmeasured Confounding using E-values for Associations between Normalized Difference Vegetation Index (0.11 unit increase) and allostatic load components	122

Acknowledgements

I wish to begin by thanking each of my thesis committee members for their mentorship during my time at Harvard Chan. Dr. Michelle Holmes welcomed me to the school, and I will be forever grateful for her taking a chance on me, an applicant without any previous cancer research experience. Dr. Linda Valeri's clear explanations and practical examples in her causal mediation class helped me recognize how to apply these methods in my own work. Anyone who has had the privilege of working with Dr. Peter James knows that his passion for geospatial methods is contagious. The projects presented in this thesis would have been impossible without the analytic foundation he laid across multiple datasets, software environments, and publications. Lastly, Dr. Timothy Rebbeck served as a compassionate, patient, and generous dissertation chair and advisor. His gentle nudges over many conversations about thesis projects helped whittle down what seemed to be an impossibly broad set of hypotheses into manageable chunks. The committee meetings where I had the opportunity to receive constructive feedback from these world-leading experts in cancer epidemiology, causal inference methods, and environmental epidemiology will remain among my fondest memories at the school.

In addition to committee members, I must acknowledge support from additional faculty and staff whose contributions strengthened the quality and conduct of this work. Dr. Francine Laden and Dr. Jaime Hart helped me think carefully about exposure assessment for these studies, and generously shared a wealth of institutional knowledge on exposure data and linkage to the Harvard cohorts and other large population-based datasets. Researchers within the Africa/Harvard Partnership for Cohort Research and Training provided important insight and interpretation of results from my third thesis aim. I am particularly grateful for the generosity of Dr. Francis Bajunirwe and Dr. Joan Mutyoba, who hosted me at Mbarara University and

Makerere University during a trip to Uganda. Dr. Lorelei Mucci made invaluable contributions as an expert on prostate cancer epidemiology in the Health Professionals cohort. Jeffrey Blossom at the Harvard Center for Geographic Analysis provided essential technical support for Geographic Information Systems on all of these thesis projects. Lastly, I wish to thank Ellen Furxhi, Eric DiGiovanni, Caroline Huntington, and Warisha Amin for administrative support throughout my time in the program.

This work would not have been possible without financial support. I am grateful to Dr. Meir Stampfer for providing support during my first four years of the program through the National Institute of Health T32 in Cancer Epidemiology training grant, and to the Harvey Fineberg Fellowship in Cancer Prevention for providing support during my final year.

I must also thank the study participants who graciously offered deeply personal information regarding their cancer diagnoses, geographic locations, and clinical histories to be shared with public health researchers through the Harvard-based Health Professionals Follow-up Study, the Pennsylvania Cancer Registry, and the Africa/Harvard Partnership for Cohort Research and Training. They have placed their faith and trust in the scientific community, faith that useful knowledge that will benefit all of us will be gleaned from their data, and trust that we will use this for the benefit of society at large.

Finally, I wish to thank my family, friends and teachers who have nurtured, encouraged, and inspired me throughout my life and allowed me to reach this point. Your unconditional support over the years has made me confident that I am on the right path, and has carried me through the inevitable ups and downs that arise during academic training.

Introduction

This thesis presents results from three epidemiologic investigations concerning the role of neighborhood greenness, or the natural green vegetation in our environments, in prostate cancer prevention, risk and mortality. While all studies relied on the same satellite-derived exposure to neighborhood greenness, the Normalized Difference Vegetation Index, they differ in several important respects. These studies were conducted in diverse populations from the United States and sub-Saharan Africa, using different epidemiologic study designs, attempting to answer different questions about the relationships between neighborhood greenness and prostate cancer risk factors, incidence and mortality. Scientific contributions of the thesis include empirical assessments of the relationships between neighborhood greenness and several important prostate cancer outcomes and risk factors, as well as use of epidemiologic methods for causal inference to explain the relationships we found. Furthermore, these studies were among the first to leverage satellite image data and geospatial analysis to create exposure metrics and visualize relationships between neighborhood environments and prostate cancer outcomes. Upon completion of this thesis, we not only shed light on whether associations existed between neighborhood greenness and the cancer outcomes under study, but also revealed clues about mediating factors that could drive these associations. Conducting our studies in different populations helped us understand whether these relationships between neighborhood greenness and prostate cancer outcomes and cancer risk factors are generalizable across populations and countries.

Improving our understanding of the role of neighborhood environments and cancer is important for many reasons. Research in this area touches on one of the core debates in public health that began over a century ago, during the first conference in the United States to establish formal public health training programs (1). During this meeting, the attendees debated whether

public health practitioners should focus their efforts on addressing individual lifestyle and biologic causes of disease, or whether their focus should be on addressing social, environmental, and political causes of ill health in the population. This tension between biology and social and environmental context continues to the present day (2-4). With respect to cancer, risk does arise from genetic factors that can accelerate or decelerate progression of tumor-forming mutations – clearly an individual-level contributor (2). However, these genetic factors within the individual exert their effects within a complex social, political and environmental milieu (3-6). Working in sub-Saharan Africa prior to joining Harvard brought this message home very clearly – a cancer patient seeking care in rural Rwanda would have a very different prognosis compared to one living in Boston with the same diagnosis – for reasons that could be far more readily explained by the limited availability of high quality oncology care than differences in biological predisposition for cancer among Americans compared to Rwandese.

The fields that study health effects of neighborhood context - environmental and social science - face unique challenges towards achieving the ultimate goal of epidemiology – making causal inferences (7-9). Though measures of socioeconomic status are readily available, applying those measures within a causal framework becomes challenging (8). For example, simply running correlations between socioeconomic status and cancer incidence rates using cross-sectional cannot be interpreted as an effect in part because the temporal ordering of the two variables is unclear. Is the hospital located in a poor area, and did a rich person move there to be close to their treatment? What exactly is it about living in a rich neighborhood that leads to lower risk? Do benefits arise due to collective sharing of wealth within a community of rich individuals, through better public spending on roads, schools and green spaces and stronger environmental regulation? Or is it simply that at an individual level, the rich are engaging in

healthy lifestyles, have incomes that allow them to access timely cancer care, and are highly educated? None of this nuance can be easily disentangled without rich, longitudinal data that explicitly demarcates the temporal relationships between neighborhood characteristics, behaviors, and the outcomes of interest. These questions also assume that precise measurements of a specific component of that neighborhood is necessary to be able to inform a public health policy, which is often not available (9).

Recognizing these limitations within many studies of neighborhood environment and cancer, I realized that these two challenges – temporal ordering of neighborhood exposure relative to outcome, and better exposure measurements – would need to be addressed to advance the science in this area. For temporal ordering of exposures, I was fortunate to work with faculty members who had done research using the renowned Harvard Cohort studies. These studies, some of which have collected nearly forty years of data, have generated insights into how long-term diet and physical activity can influence chronic disease risk. Participants provide a “snapshot” of their health, lifestyles, and medications every two years, allowing researchers to clearly define the period between exposure and disease onset. For studying cancer incidence, this is essential, because the underlying changes that influence risk are believed to occur many years before the disease is eventually detected.

To address the second limitation in earlier work – lack of a clearly defined neighborhood exposure – I began reading about the growing body of public health research on the many benefits of natural green vegetation in neighborhood environments. It was remarkable – benefits included lower rates of depression (10, 11), higher levels of physical activity (12), lower cardiovascular disease mortality (13), lower all-cause mortality (14-16) – the list went on and on (17- 20). Neighborhood greenness represented a specific aspect of the neighborhood

environment that could be further interrogated. I was fortunate to have faculty advisors who encouraged me to take a risk in exploring these questions, and pushed me to justify my analyses using conceptual framework grounded in theory (4, 5). The collective wisdom shared by my advisors, whose expertise spans cancer epidemiology, causal inference, and environmental epidemiology, ensured the methodological and scientific rigor of the studies contained in this thesis.

Prostate cancer was chosen as the cancer of focus for several reasons. Prostate cancer is the most commonly diagnosed non-cutaneous cancer among men in the United States of America, and accounts for roughly one out of five cancer-related deaths each year (21). Few modifiable risk factors have been identified for prostate cancer, though obesity and physical activity have been linked with more aggressive forms of disease (22, 23). Geographic differences in prostate cancer incidence and mortality have been identified previously, but the literature on environmental risk factors for prostate cancer is quite limited (24, 25). We felt studies of neighborhood greenness and prostate cancer could therefore help us better understand these geographic patterns of risk, and how neighborhood context might influence lifestyles, thereby influencing risk and mortality (26).

This thesis is organized as follows. In chapter 1, we estimated the association between neighborhood greenness and lethal prostate cancer risk over 28 years of follow-up in a cohort of male health professionals. We applied mediation methods to specifically evaluate whether vigorous physical activity could explain this association. In chapter 2, given that many men with prostate cancer die of causes other than prostate cancer, and given strong, replicated associations between neighborhood greenness and mortality, we evaluated the association between neighborhood greenness and mortality in a population-based registry in the United States. Use of

a population-based dataset allowed us to evaluate whether these relationships between neighborhood greenness and mortality existed for Black and White men, and whether a hypothetical intervention to fix neighborhood greenness for all men with prostate cancer to a level experienced by the wealthiest White men with prostate cancer could reduce disparities. In chapter 3, to further explore possible mechanisms, we conducted a cross-sectional multi-country study in sub-Saharan Africa to evaluate the association between neighborhood greenness, cancer metabolic risk factors and prevalence of non-communicable diseases.

When taken together, these results provide support for the hypothesis that men living in greener neighborhoods experience lower rates of lethal prostate cancer, and men with prostate cancer living in greener neighborhoods experience lower mortality. Mechanisms could include metabolic risk factors, but are likely to vary by race, urbanicity, and geography. Further studies using more precise measurements of nature contact and behavioral risk factors are needed to confirm these findings.

REFERENCES

1. Fee E. Disease and Discovery: A History of the Johns Hopkins School of Hygiene and Public Health, 1916-1939. Baltimore, Maryland: Johns Hopkins University Press, 2016. Chapter 2: Competition for the First School of Hygiene and Public Health. pp. 26-56.
2. Tomasetti C, Li L, Vogelstein B. Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention. *Science*. 2017;355(6331):1330-1334.
3. Song M, Vogelstein B, Giovannucci EL, Willett WC, Tomasetti C. Cancer prevention: Molecular and epidemiologic consensus. *Science*. 2018;361(6409):1317-1318.
4. Krieger N. Epidemiology and the People's Health: Theory and Context. Oxford: Oxford University Press. 2011. Chapter 7: Ecosocial Theory of Disease Distribution
Embodying Societal & Ecologic Context.
5. Lynch SM & Rebbeck TR. Bridging the gap between biologic, individual, and macroenvironmental factors in cancer: a multilevel approach. *Cancer Epidemiol Biomarkers Prev*. 2013;22(4):485-96.
6. Gomez SL, Shariff-Marco S, DeRouen M, Keegan TH, Yen IH, Mujahid M et al. The impact of neighborhood social and built environment factors across the cancer continuum: Current research, methodological considerations, and future directions. *Cancer*. 2015;121(14):2314-30.
7. Bind M-A. Causal Modeling in Environmental Health. *Annu Rev Public Health*. 2019;40:23-43.
8. Kaufman JS. Commentary: Causal Inference for Social Exposures. *Annu Rev Public Health*. 2019;40:7-21.

9. Chaix B, Meline J, Duncan S, Merrien C, Karusisi N, Perchoux C et al. GPS tracking in neighborhood and health studies: A step forward for environmental exposure assessment, a step backward for causal inference? *Health Place*. 2013;46-51.
10. Banay RF, James P, Hart JE, Kubzansky LD, Spiegelman D, Okereke OI et al. Greenness and Depression Incidence among Older Women. *Environ Health Perspect*. 2019;127(2):27001. Doi: 10.1298/EHP1229.
11. Association of Urban Green Space With Mental Health and General Health Among Adults in Australia. Astell-Burt T, Feng X. *JAMA Netw Open*. 2019;2(7):e198209. Doi: 10.1001/jamabetworkopen.2019.8209.
12. James P, Hart JE, Hipp JA, Mitchell JA, Kerr J, Hurvitz PM et al. GPS-Based Exposure to Greenness and Walkability and Accelerometry-Based Physical Activity. *Cancer Epidemiol Biomarkers Prev*. 2017;26(4):525-532.
13. Wilker EH, Wu CD, McNeely E, Mostofsky E, Spengler J, Wellenius GA et al. *Environ Res*. 2014;133:42-8.
14. James P, Hart JE, Banay RF, Laden F. Exposure to Greenness and Mortality in a Nationwide Prospective Cohort Study of Women. *Environ Health Perspect*. 2016;124(9):1344-52.
15. Crouse DL, Pinault L, Balram A, Hystad P, Peters PA, Chen H et al. Urban greenness and mortality in Canada's largest cities: a national cohort study. *Lancet Planet Health*. 2017;1(7):e289-e297.
16. Vienneau D, de Hoogh K, Faeh D, Kaufmann M, Wunderli JM, Rösli M et al. More than clean air and tranquility: Residential green is independently associated with decreasing mortality. *Environ Int*. 2017;108:176-184.

17. James P, Banay RF, Hart JE, Laden F. A Review of the Health Benefits of Greenness. *Curr Epidemiol Rep.* 2015;2(2):131-142.
18. Fong KC, Hart JE, James P. A Review of Epidemiologic Studies on Greenness and Health: Updated Literature Through 2017. *Curr Environ Health Rep.* 2018;5(1):77-87.
19. Twohig-Bennett C & Jones A. The health benefits of the great outdoors: A systematic review and meta-analysis of greenspace exposure and health outcomes. *Environ Res.* 2018;166:628-637.
20. Kondo MC, Fluehr JM, McKeon T, Branas CC. Urban Green Space and Its Impact on Human Health. *Int J Environ Res Public Health.* 2018;15(3): pii. E445.
21. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin.* 2019;69(1):7-34.
22. Pernar CH, Ebot EM, Wilson KM, Mucci LA. The Epidemiology of Prostate Cancer. *Cold Spring Harb Perspect Med.* 2018;8(12): pii: a030361.
23. Hammarsten J, Damber JE, Hagsheno MA, Mellström D, Peeker R. A stage-dependent link between metabolic syndrome components and incident prostate cancer. *Nat Rev Urol.* 2018;15(5):321-333.
24. Shimizu H, Ross RK, Bernstein L, Yatani R, Henderson BE, Mack TM. Cancers of the prostate and breast among Japanese and white immigrants in Los Angeles County. *Br J Cancer.* 1991;63(6):963-6.
25. Occupation and prostate cancer. Parent ME, Siemiatycki J. *Epidemiol Rev.* 2011;23(1):138-43.

26. Demoury C, Thierry B, Richard H, Sigler B, Kestens Y, Parent ME. Residential greenness and risk of prostate cancer: A case-control study in Montreal, Canada. *Environ Int.* 2017;98:129-136.

Chapter 1: The Association between Neighborhood Greenness and Incidence of Lethal Prostate Cancer: A Prospective Cohort Study

Hari S. Iyer¹, Peter James², Linda Valeri³, Jaime E. Hart^{4,5}, Claire H. Pernar¹, Lorelei A. Mucci^{1,4}, Michelle D. Holmes^{1,4}, Francine Laden^{1,4}, Timothy R. Rebbeck^{1,6}

Affiliations:

¹Harvard T. H. Chan School of Public Health, Department of Epidemiology, Boston, MA 02115; United States

²Harvard Medical School and Harvard Pilgrim Health Care Institute, Department of Population Medicine, Boston, MA 02215; United States

³Columbia University Mailman School of Public Health, Department of Biostatistics, New York, NY 10032; United States

⁴Brigham and Women's Hospital, Department of Medicine, Channing Division of Network Medicine, Boston, MA 02115; United States

⁵Harvard T. H. Chan School of Public Health, Department of Environmental Health, Boston, MA 02115; United States

⁶Dana Farber Cancer Institute, Department of Medical Oncology, Boston, MA 02115; United States

ABSTRACT

Background

Growing evidence suggests that neighborhood contextual environment could influence risk factors, and therefore incidence, of lethal prostate cancer. We studied the association between neighborhood greenness and lethal prostate cancer incidence, and assessed mediation by vigorous physical activity.

Methods

47,958 participants were followed in the Health Professionals Follow-up Study from 1986 to 2014. Neighborhood greenness exposure was estimated using normalized difference vegetation index (NDVI) with 1 kilometer resolution, assigned to home or work addresses at start of follow-up. Hazard ratios (aHR) and 95% confidence intervals (CI) were estimated using sequentially adjusted Cox models with individual and contextual prostate cancer risk factors as covariates. Analyses were compared among those whose addresses did not change over follow-up, and stratified by population density and address type.

Results

We observed 898 cases over 1,054,743 person-years. An interquartile range increase in NDVI was associated with 5% lower rate of lethal prostate cancer (aHR: 0.95, 95% CI: 0.88, 1.03), with stronger associations in non-movers (aHR: 0.92, 95% CI: 0.85, 1.01). Inverse associations were observed among men in high (aHR: 0.90, 95% CI: 0.82, 0.99) but not low (aHR: 1.11, 95% CI: 0.95, 1.29, $P_{\text{het}}=0.086$) population density areas, and those reporting from work (aHR: 0.87, 95% CI: 0.75, 1.01) but not home (aHR: 1.04, 95% CI: 0.91, 1.17, $P_{\text{het}}=0.10$) addresses. There was no evidence of mediation by vigorous physical activity.

Conclusion

We report inverse associations between neighborhood greenness and lethal prostate cancer when restricting to non-movers, and in high population density areas. Replication could confirm findings and clarify mechanisms.

INTRODUCTION

Prostate cancer is the most common non-cutaneous malignancy among men in the United States of America (US), with an estimated 174,650 new cases and 31,620 deaths in 2019 (1).

Prostate cancer is considered to be a heterogeneous disease, contrasting indolent, screen-detected cancer with advanced or lethal prostate cancer defined by clinical stage and grade (2, 3). Most risk factors for total prostate cancer (age, family history, African American race, height, genetic risk loci) are not modifiable. However, modifiable risk factors, including smoking, obesity and physical activity, have been identified for lethal prostate cancer (4-7). Focusing purely on individual-level risk factors ignores the broader societal and environmental context in which the individual is embedded (8). Therefore, studying contextual environmental risk factors could help develop a multi-level model of lethal prostate cancer risk (9).

Natural vegetation in a given area (referred to hereafter as “neighborhood greenness”) is increasingly considered to be a health promoting contextual environmental factor (10-14). Large observational studies have reported beneficial associations between greenness and health, including all-cause mortality, depression, physical activity, and obesity (15-19). Neighborhood greenness exposure could offer psychological benefits that increase adherence to healthy lifestyles, or spaces to exercise which increase physical activity (20-22). In addition, neighborhood greenness is associated with stronger community cohesion and greater social capital, which are associated with increased use of preventive health care services (23-25). Together, these pathways could reduce risk of lethal prostate cancer (26).

We studied the association between baseline neighborhood greenness and lethal prostate cancer incidence in a nation-wide prospective cohort of male health professionals in the US. We hypothesized that neighborhood greenness would be associated with lower rates of lethal prostate

cancer, and that this protective association would be mediated in part through higher levels of vigorous physical activity among participants in greener neighborhoods (2, 27, 28). Since prior studies had focused on urban areas using residential neighborhood greenness as the primary exposure (26), we further sought to evaluate whether associations varied by population density or exposure at home compared to work.

METHODS

Study population and design

We used data from the Health Professionals' Follow-up Study (HPFS), an ongoing prospective cohort study based at the Harvard T. H. Chan School of Public Health. Since 1986, 51,529 participating male health professionals across the US have completed biennial questionnaires that record information about lifestyle and health-related factors, as well as diagnosis of new illnesses. Cohort participants could choose to mail their questionnaire to a home or work address over follow-up. Geocoded addresses were available from questionnaire mailing records from 1988 to 2012. In 1988, participants indicated if the address was their home, work, or other address. Upon receipt of a new diagnosis, study personnel conduct a detailed review of medical and pathological information for validation purposes. The questionnaire response rate is 90%, with mortality follow-up over 98% (29). Participants with prior history of prostate cancer or non-melanoma skin cancer (n=2084), missing a geocoded address (n=1447) or date of birth (n=36), or died before returning their first questionnaire (n=4) were excluded, resulting in a study population of 47,958. The study protocol was approved by the Institutional Review Boards of the Brigham and Women's Hospital and Harvard T. H. Chan School of Public Health, and those of participating registries as required.

Lethal prostate cancer assessment

Incident prostate cancer diagnoses were ascertained from biennial questionnaires. Study personnel and clinical staff reviewed medical records and pathology reports to confirm reported diagnosis. Lethal prostate cancer was defined by presence of distant metastasis (stage M1), or indication that prostate cancer was the primary cause of death for the study participant, over follow-up. Study staff were notified of cohort deaths from family members, as well as linkages with the National Death Index (29).

Exposure to neighborhood greenness

Exposure to neighborhood greenness was estimated by linking satellite data on greenness to geocoded participant addresses from the 1988 questionnaire, allowing us to compare greenness exposure measurements at home and work. We used the normalized difference vegetation index (NDVI), calculated by taking a ratio of the difference of near infrared and visible light divided by the sum of near infrared and visible light (30). Longitudinal NDVI data were obtained from images produced by the Advanced Very High Resolution Radiometer satellite of the National Oceanic and Atmospheric Administration. Images were taken every 16 days at 1000m resolution and began in 1989, earlier than other sources. The NDVI scale ranges from -1 to 1, with 1 representing maximal vegetation, values close to 0 representing barren areas of rock, sand or snow, and values approaching -1 indicating bodies of water (31).

The HPFS follows a biennial questionnaire cycle and cohort participants reside across the US, reflecting a broad range of regional and seasonal variation in neighborhood greenness. Since

prostate cancer has a long natural history, we modeled associations between neighborhood greenness at the start of follow-up and lethal prostate cancer. We took an average of measurements of NDVI corresponding to different seasons (January, April, July, and September) to account for seasonal changes in and geographic differences in duration of greenness. Seasonal average NDVI measurements from 1989, the earliest year that NDVI data were available, were assigned to participant's geocoded address within a 1000m buffer.

We modeled baseline neighborhood greenness as our primary exposure rather than cumulative updated average because we felt that earlier exposure to neighborhood greenness, rather than duration and intensity of exposure up to diagnosis, would be more likely to occur during the etiologic window for lethal prostate cancer. As a secondary exposure, we estimated cumulative updated average NDVI, incorporating four seasonal images per year over follow-up (Supplementary Appendix).

Longitudinal measures of physical activity

Physical activity was reported by participants on biennial questionnaires. Participants were asked questions about the average time spent each week engaging in different types of physical activity (walking or hiking outdoors, jogging, running, bicycling, lap swimming, tennis, squash or racquetball, and calisthenics or rowing). In subsequent questionnaire cycles, additional activities were included: heavy outdoor work (from 1988), weightlifting (1990), moderate outdoor work (2004), lower intensity exercise and other aerobic exercise (2010). Additional activities included flights of stairs traversed daily, and usual walking pace. Each activity was assigned a metabolic equivalent of task (MET) (32). Non-vigorous activities were classified as

those with MET<6, while vigorous activities were classified as MET >= 6. Total physical activity was reported in MET-hours per week, calculated by summing the product of MET-hours and average hours per week for all physical activity reported by participants. Validation studies comparing MET-hours per week in questionnaires to weekly diaries found generally high correlations (33).

Statistical analysis

Participant follow-up began with return of the first questionnaire (1986) until diagnosis of lethal prostate cancer, death from another cause, or administrative censoring on January 1, 2014, whichever came first. We used Cox proportional hazards models with study follow-up as the primary time scale to estimate hazard ratios and 95% confidence intervals for the association between rate of lethal prostate cancer and NDVI. We modeled NDVI as quintiles and estimated p-values for linear trend using the median value for each NDVI quintile. We also estimated the change in rate of lethal prostate cancer associated with a linear interquartile range (IQR) unit increase in continuous NDVI (0.11 units). We tested for non-linearity of continuous NDVI using splines and proportional hazards by fitting a time by greenness interaction. In addition, to more precisely examine long-term exposure to neighborhood greenness, we repeated the main analysis restricting to participants who did not move during follow-up.

To assess the impact of covariate adjustment on effect estimates, we fit sequentially adjusted models (Model 1: age (continuous), calendar time at 2-year questionnaire cycle (continuous) included as covariates in the baseline hazard; Model 2: All covariates included in model 1, plus race (categorical: White, African-American, Other), diabetes mellitus, BMI at age

21 (kilogram/meter²; <20, 20 to <22.5, 22.5 to <25, ≥25), height (inches, <66, 66 to <68, 68 to <70, 70 to <72, ≥72), smoking (never smokers, current and/or quit smoking ≤ 10 years ago, quit > 10 years ago), family history of prostate cancer, PSA testing over follow-up using two variables: ever had PSA screening prior to diagnosis (lagged to reflect screened, rather than diagnostic PSA test) and intensity of PSA screening prior to diagnosis (defined as having reported having PSA screening in over half of prior visits since 1994), census tract median income (USD, continuous), census tract median home value (USD, continuous); Model 3: All covariates in model 2, plus vigorous physical activity, non-vigorous physical activity (quintiles), and current BMI (kilogram/meter², <21, 21 to <23, 23 to <25, 25 to <27.5, 27.5 to <30, ≥30). Baseline measures of all lifestyle covariates described above were used in our primary analysis. Vigorous physical activity was modeled as a five-level variable, with the lowest level corresponding to 0 METS of vigorous physical activity, and the remaining levels modeled as quartiles of non-zero vigorous METs (2). Model 2 corresponds to a confounding adjusted model, and model 3 corresponds to the controlled direct effects model specified in our mediation analysis (Supplementary Appendix).

Given that participant data covered a broad geographic area, and that exposure source varied between participants, we evaluated multiplicative effect modification of the association between continuous NDVI and lethal prostate cancer by population density (≥1,000 people per square mile compared to <1000 people per square mile) and address type (home compared to work) using likelihood ratio tests. For home and work addresses, we rescaled the quintiles based on the NDVI distribution for participants at each location. We further evaluated multiplicative effect modification by census region (US Census defined North, South, East, West) and PSA screening history and intensity.

Further details regarding mediation analysis as well as sensitivity analysis for cumulative updated average exposure and unmeasured confounding using e-values are provided in the Supplementary Appendix.

RESULTS

After exclusions, 47,958 participants (93%) remained in our analytic sample, giving rise to 898 cases of incident lethal prostate cancer accrued over 1,054,743 person-years of follow-up. Age-adjusted characteristics of the study population are described in Table 1.1 across quintiles of NDVI. Most participants were white (95%) with an average age of 64.4 years over follow-up. Participants in the highest quintile of NDVI reported higher non-vigorous physical activity (NDVI Q5: 17.6 v. Q1: 15.6 MET-hours/week) and lower vigorous physical activity (NDVI Q5: 8.4 v. Q1: 9.7 MET hours/week) compared to participants in the lowest quintile. These patterns held in adjusted models (Table S1.1). Average census tract population density (NDVI Q5: 1,720 v. Q1: 8,870 people per mi²) decreased with increasing quintiles of NDVI, while median income increased (NDVI Q5: \$58,870 v. Q1: \$52,270). Maps of participant locations (Figure S1.1) and NDVI in July 1989 (Figure S1.2) display the geographic spread of exposure locations.

Table 1.1. Age-standardized Characteristics by Quintile of Normalized Difference Vegetation Index (NDVI) Among Men in the Health Professionals Follow-up Study From 1986 to 2014^{a,b}

Characteristics	Total	Baseline Normalized Difference Vegetation Index (NDVI)				
		Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Participants, no.	47,958	9,504	9,562	9,688	9,718	9,486
Age ^{c,d} , years	64.4 (11.2)	64.9 (11.4)	64.6 (11.3)	64.4 (11.2)	64.2 (11.2)	64.0 (11.3)
Baseline NDVI ^d	0.28 (0.09)	0.14 (0.05)	0.23 (0.02)	0.28 (0.01)	0.33 (0.02)	0.41 (0.04)
NDVI ^d (cumulative updated average)	0.31 (0.09)	0.19 (0.07)	0.27 (0.05)	0.31 (0.05)	0.35 (0.04)	0.41 (0.05)
Vigorous Activity ^d , MET-hours/week	8.9 (17.9)	9.7 (18.7)	9.1 (18.2)	8.7 (17.0)	8.5 (17.8)	8.4 (17.6)
Non-vigorous Activity ^d , MET-hours/week	16.7 (22.1)	15.6 (20.9)	16.8 (22.1)	16.4 (21.8)	17.0 (22.3)	17.6 (23.1)
Total activity ^d , MET-hours/week	28.4 (30.6)	28.2 (30.8)	28.9 (30.9)	27.9 (29.7)	28.4 (30.5)	28.9 (31.1)
Height ^d , inches	70.2 (2.8)	70.1 (2.9)	70.2 (2.8)	70.2 (2.8)	70.2 (2.7)	70.2 (2.8)
BMI at age 21 ^d , kilogram/meter ²	23.1 (3.0)	23.1 (2.9)	23.1 (2.9)	23.1 (3.0)	23.1 (3.0)	23.0 (2.9)
Current BMI ^d , kilogram/meter ²	26.0 (3.8)	25.9 (3.8)	26.0 (3.8)	26.1 (3.8)	26.1 (3.7)	26.0 (3.8)
Race						
White, %	95	94	96	97	96	97
African American, %	1	1	1	1	1	1
Asian, %	2	3	1	1	1	1
Other, %	2	2	2	1	2	1
Smoking status						
Non-smoker, %	57	59	58	56	57	56
Past, quit >10 years ago, %	29	27	28	29	29	30
Current & past, quit ≤10 years ago, %	14	14	14	15	14	14
Diabetes, %	6	6	6	6	6	5
Family history of CaP, %	12	12	12	12	11	12
Prostate Specific Antigen (PSA) screening history						
Had PSA test prior to diagnosis, %	35	34	35	36	36	36
PSA test on at least half of all questionnaires, 1994-2012, %	33	31	32	33	33	34

Table 1.1. Age-standardized Characteristics by Quintile of Normalized Difference Vegetation Index (NDVI) Among Men in the Health Professionals Follow-up Study From 1986 to 2014^{a,b} (continued)

Characteristics	Total	Baseline Normalized Difference Vegetation Index (NDVI)				
		Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Census Region						
Northeast, %	22	16	13	20	27	34
Midwest, %	26	20	38	38	27	8
South, %	29	20	21	25	33	45
West, %	23	44	28	17	13	12
Population density ^d , 1,000 people/mi ²	4.0 (9.5)	8.9 (18.4)	3.9 (6.4)	3.0 (4.1)	2.5 (3.5)	1.7 (3.8)
Census tract median income ^d , 1,000 USD	54.3 (28.4)	52.3 (30.4)	52.3 (26.3)	53.0 (26.1)	54.9 (27.3)	58.9 (31.3)
Census tract median home value ^d , 1000 USD	162.9 (145.3)	183.6 (173.0)	152.6 (135.7)	146.6 (128.8)	152.1 (128.1)	179.8 (152.0)
Address Type (1988)						
Home, %	33	24	30	33	37	42
Work, %	41	49	44	43	38	32
Other, %	1	2	1	1	1	2
Not reported, %	24	26	25	23	24	23
Moved during follow-up, %	12	13	12	12	12	11

Abbreviations: BMI, body mass index; MET, metabolic equivalent of task; NDVI, normalized difference vegetation index; USD, United States Dollar.

^aValues are standardized to the age distribution of the study population

^bValues of polytomous variables may not sum to 100% due to rounding

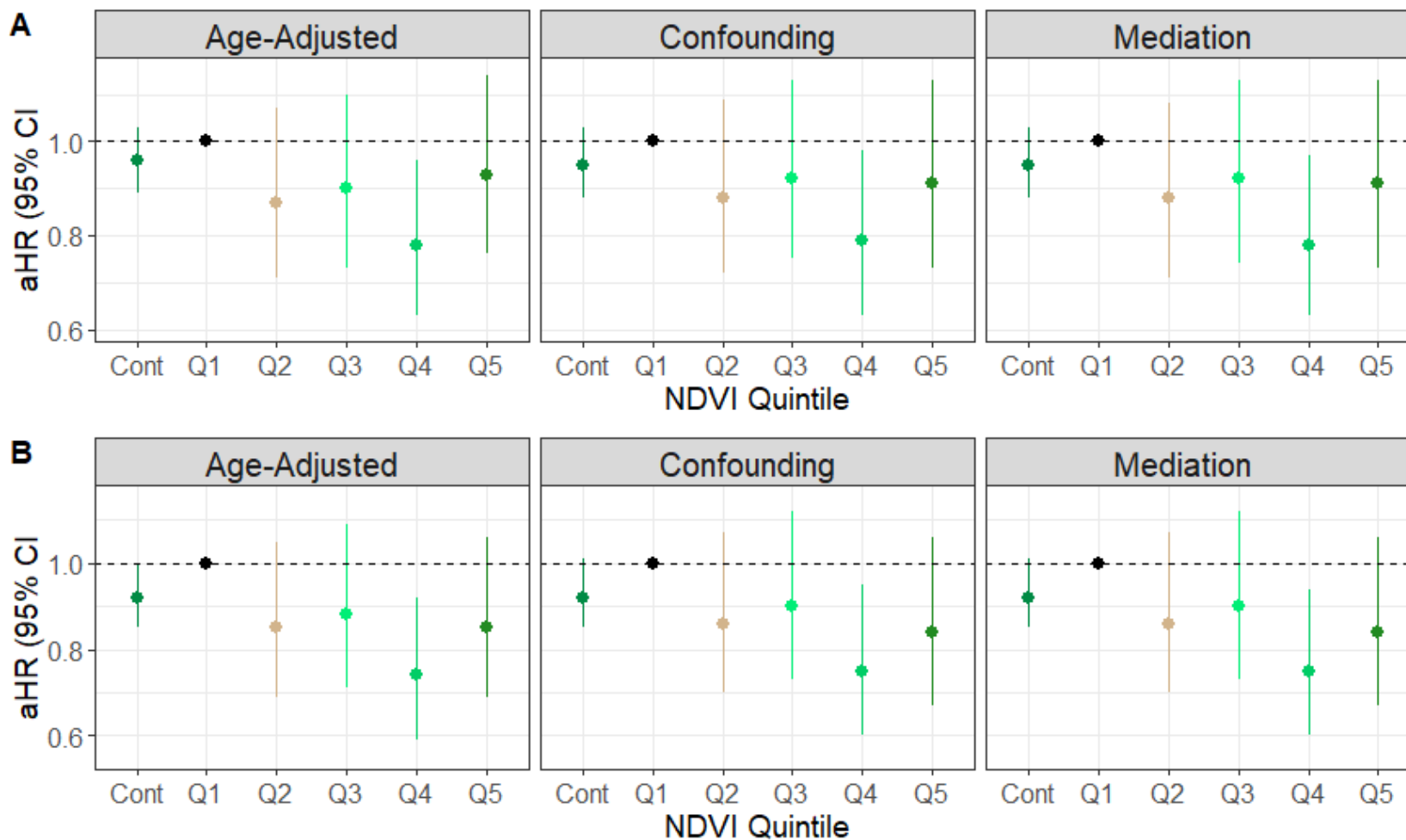
^cNot age-adjusted

^dValues are expressed as mean (standard deviation)

In our analysis of the full cohort (Figure 1.1, Table S1.2), increasing quintiles of baseline NDVI were not significantly associated with lower rates of lethal prostate cancer compared to the lowest quintile (Q1) in age and calendar year- and confounding-adjusted models. Only the estimate for Q4 was statistically significant at the 0.05 level (aHR: 0.78, 95% CI: 0.63, 0.96, $P_{\text{trend}} = 0.25$). Inverse associations were stronger among the 42,492 (89%) participants who did not change addresses during follow-up (813 cases over 930,033 person-years) (Figure 1.1, Table S1.2). Among non-movers, we observed an 8% lower rate of lethal prostate cancer associated with an IQR increase in NDVI (aHR: 0.92, 95% CI: 0.85, 1.01), with weak evidence of lower rates of lethal prostate cancer associated with increasing NDVI quintiles ($P_{\text{trend}} = 0.068$). Results from models further adjusting for vigorous physical activity and body mass index were similar to those from confounding models in the total and restricted populations (Figure 1.1, Table S1.2).

Table 1.2 presents results from models evaluating the association between NDVI and incidence of lethal prostate cancer within levels of population density, and address type. Stronger inverse associations were observed in high (>1000 people/mi²) compared to low population density neighborhoods (<1000 people/mi²) though the p-value for heterogeneity did not reach statistical significance ($P_{\text{het}} = 0.086$). In high population density areas, an IQR increase in NDVI was associated with a 10% lower rate of lethal prostate cancer (aHR 0.90, 95% CI: 0.82, 0.99), while in low population density areas, the direction of this association was reversed (aHR: 1.11, 95% CI: 0.95, 1.29).

Figure 1.1. Hazard Ratios and Confidence Intervals for the Association Between Baseline Normalized Difference Vegetation Index (NDVI) and Lethal Prostate Cancer Incidence in the Health Professionals Follow-up Study, United States, 1986-2014



Sequentially adjusted for age in months and calendar time as strata (**Age-Adjusted**); race, diabetes mellitus, height, family history of prostate cancer, BMI at age 21, smoking status in 1986, 1990 census tract median income (USD), and 1990 census tract median home value (USD), history of prostate-specific antigen testing, and intensity of prostate-specific antigen testing (**Confounding**); vigorous physical activity (categorical), non-vigorous physical activity (quintiles), and current BMI (categorical) (**Mediation**). Panel (A) Total population (N=47,958); (B) Participants who did not move over follow-up (N=42,492). Abbreviations: aHR=adjusted hazard ratio, Cont=continuous, Q=quintile.

Table 1.2. Hazard Ratios for the Association Between Baseline Normalized Difference Vegetation Index (NDVI)^a and Lethal Prostate Cancer Incidence in the Health Professionals Follow-up Study, United States, 1986-2014, Stratified by Population Density (high: ≥ 1000 , low: < 1000 people/mi²) and Address Type (Work, Home)

Model	Baseline Normalized Difference Vegetation Index (NDVI)														P_{trend}	$P_{\text{het}}^{\text{e}}$
	Continuous ^b		Quintile 1		Quintile 2		Quintile 3		Quintile 4		Quintile 5					
	aHR	95% CI	aHR	95% CI	aHR	95% CI	aHR	95% CI	aHR	95% CI	aHR	95% CI				
Total Population																
Population Density^c																
High (N=34,229)	0.90	0.82, 0.99	1.00	Referent	0.89	0.71, 1.12	0.94	0.74, 1.18	0.76	0.59, 0.97	0.76	0.57, 1.01	0.022	0.086		
Low (N=13,729)	1.11	0.95, 1.29	1.00	Referent	0.65	0.37, 1.13	0.87	0.52, 1.45	0.92	0.56, 1.51	1.08	0.69, 1.70	0.15			
Address Type^d																
Work (N=18,742)	0.87	0.75, 1.01	1.00	Referent	0.80	0.57, 1.12	0.53	0.35, 0.79	0.73	0.50, 1.07	0.66	0.45, 0.98	0.027	0.10		
Home (N=16,732)	1.04	0.91, 1.17	1.00	Referent	1.14	0.83, 1.57	1.07	0.77, 1.48	0.91	0.65, 1.29	1.19	0.85, 1.66	0.66			
Non-movers																
Population Density^c																
High (N=30,259)	0.88	0.80, 0.97	1.00	Referent	0.88	0.70, 1.11	0.93	0.73, 1.19	0.68	0.52, 0.89	0.73	0.54, 0.99	0.0072	0.15		
Low (N=12,233)	1.07	0.91, 1.26	1.00	Referent	0.59	0.32, 1.07	0.83	0.47, 1.44	1.04	0.61, 1.75	0.97	0.60, 1.59	0.26			
Address Type^d																
Work (N=16,967)	0.85	0.73, 0.99	1.00	Referent	0.80	0.56, 1.14	0.51	0.34, 0.78	0.69	0.46, 1.02	0.63	0.42, 0.95	0.014	0.15		
Home (N=14,466)	0.99	0.87, 1.14	1.00	Referent	1.11	0.79, 1.56	1.12	0.79, 1.59	0.86	0.60, 1.25	1.07	0.75, 1.53	0.89			

Abbreviations: aHR, adjusted hazard ratio; BMI, body mass index; CI, confidence interval, IQR, interquartile range; NDVI, normalized difference vegetation index; USD, United States Dollars.

^aAdjusted for age in months and calendar time as strata, race (categorical), diabetes mellitus (yes or no), height (categorical), family history of prostate cancer (yes or no), BMI at age 21 (categorical), smoking status in 1986 (categorical), 1990 census tract median income (USD), 1990 census tract median home value (USD), history of prostate-specific antigen testing, and intensity of prostate-specific antigen testing.

^bEstimate corresponds to an IQR increase in continuous NDVI of 0.11 units.

^cModels additionally adjusted for address type (categorical: home, work, missing category).

^dRestricted to only participants who reported home or work address. Quintiles reflect within group distributions for home and work. Models additionally adjusted for population density (binary: ≥ 1000 , < 1000 people/mi²).

^eLikelihood ratio test with one degree of freedom for interaction between continuous NDVI and stratification variable.

When stratifying by address type, we observed stronger inverse associations among participants for whom NDVI was assessed at work ($P_{\text{trend}} = 0.027$) compared to home though evidence for effect modification was weak ($P_{\text{het}} = 0.10$). There was a 13% lower rate of lethal prostate cancer associated with an IQR increase in NDVI (aHR: 0.87, 95% CI: 0.75, 1.01) among men for whom NDVI was assessed at work, compared to a 4% increased rate among those with residential NDVI (aHR: 1.04, 95% CI: 0.91, 1.17). Linear associations for high population density and among those with work addresses were strengthened when restricting to non-movers (Table 1.2). Further examination of effect modification by additional factors (PSA screening intensity, prior history of PSA screening, or geographic region) did not reveal any differences (Table 1.3).

Table 1.3. Hazards Ratios for the Association Between Baseline Normalized Difference Vegetation Index (NDVI)^{a,b} and Lethal Prostate Cancer Incidence in the Health Professionals Follow-up Study, United States, 1986-2014, Stratified by Address, Prostate-specific Antigen Screening, and Census Region

Subgroups	Cases/Person-years	Incidence Rate per		95% CI	P_{het}^c
		100K Person-	aHR		
		Years			
Had PSA test prior to diagnosis					0.42
Yes	257/376,050	68	1.00	0.87, 1.16	
No	641/678,702	94	0.93	0.85, 1.03	
PSA screening intensity					0.12
≥50% of questionnaires	226/342,450	66	1.06	0.91, 1.23	
<50% of questionnaires	672/712,302	94	0.92	0.84, 1.01	
Region					0.99
Northeast	226/232,433	97	0.96	0.83, 1.10	
Midwest	213/278,845	76	0.96	0.78, 1.19	
South	235/304,930	77	0.98	0.84, 1.13	
West	224/238,535	94	0.94	0.81, 1.10	

Abbreviations: aHR, adjusted hazard ratio; BMI, body mass index, CI, confidence interval, IQR, interquartile range; NDVI, normalized difference vegetation index; Prostate-specific antigen (PSA), USD, United States Dollar.

^aEstimate corresponds to an IQR increase in continuous NDVI of 0.11 units.

^bAll models are adjusted for age in months and calendar time as strata, race (categorical), diabetes mellitus (yes or no), height (categorical), family history of prostate cancer (yes or no), BMI at age 21 (categorical), smoking status in 1986 (categorical), 1990 census tract median income (USD), 1990 census tract median home value (USD), population density (binary: ≥1000, <1000 people/mi²), history of PSA testing, and intensity of PSA testing.

^cHeterogeneity test performed using 1-degree of freedom likelihood ratio test for interaction between continuous NDVI and address type, Had PSA test prior to diagnosis, PSA screening intensity; 3-degree of freedom test for region.

In sensitivity analyses using cumulative updated average NDVI, inverse associations were weaker and failed to reveal evidence of effect modification by population density (Table S1.3). Strongest e-values for point estimates and confidence intervals were observed for NDVI Q4 and Q5 compared to Q1 among men who did not move with addresses in high population density neighborhoods, suggesting these estimates are less likely to be completely explained by unmeasured confounding (Table S1.4).

DISCUSSION

We observed an inverse association between baseline neighborhood greenness and lethal prostate cancer, though this finding was restricted to those in high population density areas. Contrary to expectation, we found that neighborhood greenness was associated with lower levels of vigorous physical activity in this population of health professionals. We did not observe evidence of a mediating role of vigorous physical activity. Restricting to men who remained at the same address over follow-up strengthened the inverse association between neighborhood greenness and lethal prostate cancer incidence, suggesting that mechanisms are related to environmental context.

Few studies have assessed the association between neighborhood greenness and prostate cancer (26, 34). Our findings corroborate results from a population-based case control study conducted by Demoury and colleagues in Montreal, the second-largest city in Canada (26). In an urban population, they reported effect estimates of similar magnitude to ours, though they used maximal annual residential NDVI at diagnosis, and 10 years prior to diagnosis. They also found no evidence of physical activity as a mediating pathway. Since we used different exposure and outcome measures, our studies are not directly comparable, but both are consistent with a

hypothesis that green spaces and contextual environment could play a role in prostate cancer risk.

There is limited evidence for direct effects of exposure to neighborhood greenness and carcinogenesis. However, physiologic changes that arise from spending time in green environments could serve as a mechanism. Interventional studies conducted in Japan comparing visits to urban areas with forests observed higher parasympathetic activation, lower cardiometabolic response, and lower natural killer cell activity following forest visits (35, 36). Cross-sectional studies in the US reported inverse associations between neighborhood greenness and allostatic load, a composite index derived from biomarkers to capture physiologic adaptation to stress (37). One of these inflammatory biomarkers, interleukin-8, could drive cancer progression by decoupling tumor growth from androgen hormone regulation (38, 39). Further studies are needed to clarify biological mechanisms.

The magnitude and direction of the association between neighborhood greenness and incidence rate of lethal prostate cancer varied by levels of high and low population density, though we lacked power to detect statistically significant effect modification. Since neighborhood greenness varies spatially, these different relationships could be related to different geographic patterns of care seeking and treatment for lethal prostate cancer. Geographic patterns of prostate cancer care have been observed in the US; for example, rural prostate cancer patients are less likely to receive radiotherapy and surgery compared to urban patients (40, 41). In rural areas, benefits of greenness could be offset by increased lethal prostate cancer mortality resulting from absence of these treatment modalities. Environmental factors could also explain this effect heterogeneity. Ultraviolet light exposure, which has been linked with reduced rates of prostate cancer in prospective studies, could be influenced by neighborhood greenness and vary by

population density (42, 43). There is some evidence of increased risk of aggressive prostate cancer from pesticide use among farmers (44). Prospective studies exploring relationships between neighborhood greenness and other environmental factors could clarify these urban and rural differences.

Interpretation of our results warrants consideration of our study limitations. Unmeasured confounding is a major threat to validity. Using e-values, we quantified the magnitude of bias needed to change our inference and found that moderate bias conditional on covariates would be required. Our prospective design allowed us to control for major individual clinical, lifestyle, and socioeconomic contextual factors, making it unlikely for an unmeasured covariate to exhibit associations with neighborhood greenness and lethal prostate cancer as extreme as those presented in our sensitivity analysis. Address type was missing for 24% of participants, and for the remainder, only available for either home or work. We consider this to be an issue of measurement error, in which we have randomly sampled greenness exposure for some participants at home and others at work within strata of confounding variables. This non-differential measurement error means that our reported associations are weaker than what one would expect to see with perfect exposure assessment. Finally, results obtained from this select population of predominantly white health professionals may not extend to other populations. However, restriction based on socioeconomic status and race strengthens internal validity of our study.

In a 28-year prospective study of 47,958 health professionals, we observed an inverse association between neighborhood greenness and rate of lethal prostate cancer in high population density areas. These findings suggest that health benefits of neighborhood greenness could include reduced incidence of lethal prostate cancer. Future studies should apply more precise

measurements of exposure to greenness, clarify mechanisms, and assess transportability of these findings.

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2019. *CA Cancer J Clin.* 2019;69(1):7-34.
2. Giovannucci E, Liu Y, Platz, EA, et al. Risk factors for prostate cancer incidence and progression in the health professionals follow-up study. *Int J Cancer.* 2007;121(7):1571-1578.
3. Hammarsten J, Damber J-E, Haghsheno MA, et al. A stage-dependent link between metabolic syndrome components and incident prostate cancer. *Nat Rev Urol.* 2018;15(5):321-333.
4. Pernar CH, Ebot EM, Wilson KM, et al. The Epidemiology of Prostate Cancer. *Cold Spring Harb Perspect Med.* 2018;8(12): pii: a030361.
5. Brookman-May SD, Campi R, Henríquez, JDS, et al. Latest Evidence on the Impact of Smoking, Sports, and Sexual Activity as Modifiable Lifestyle Risk Factors for Prostate Cancer Incidence, Recurrence, and Progression: A Systematic Review of the Literature by the European Association of Urology Section of Oncological Urology (ESOU). *Eur Urol Focus.* 2018;pii:S2405-4569(18)30069-5.
6. Wekesa A, Harrison M, Watson RW. Physical activity and its mechanistic effects on prostate cancer. *Prostate Cancer Prostatic Dis.* 2015;18(3):197-207
7. Allott EH, Masko EM, Freedland SJ. Obesity and prostate cancer: Weighing the Evidence. *Eur Urol.* 2013;63(5):800-809
8. Krieger N. Epidemiology and the web of causation: Has anyone seen the spider? *Soc Sci Med.* 1994;39(7):887-903.

9. Lynch S, Rebbeck TR. Bridging the gap between biologic, individual, and macroenvironmental factors in cancer: A multilevel approach. *Cancer Epidemiol. Biomarkers Prev.* 2013;22(4):485-495.
10. James P, Banay RF, Hart JE, et al. A Review of the Health Benefits of Greenness. *Curr Epidemiol Rep.* 2015;2(2):131-142.
11. Markevych I, Schoierer J, Hartig T, et al. Exploring pathways linking greenspace to health: Theoretical and methodological guidance. *Environ. Res.* 2017;158:301-317.
12. Frumkin H, Bratman GN, Breslow SJ, et al. Nature Contact and Human Health: A Research Agenda. *Environ. Health Perspect.* 2017;125(7):075001.
13. Fong KC, Hart JE, & James P. A Review of Epidemiologic Studies on Greenness and Health: Updated Literature Through 2017. *Curr Environ Health Rep.* 2018;5(1):77-87.
14. Twohig-Bennett C, Jones A. The health benefits of the great outdoors: A systematic review and meta-analysis of greenspace exposure and health outcomes. *Environ. Res.* 2018;166:628-637.
15. James P, Hart JE, Banay RF, et al. Exposure to Greenness and Mortality in a Nationwide Prospective Cohort Study of Women. *Environ. Health Perspect.* 2016;124(9):1344-1352.
16. South EC, Hohl BC, Kondo MC, et al. Effect of Greening Vacant Land on Mental Health of Community-Dwelling Adults: A Cluster Randomized Trial. *JAMA Netw Open.* 2018;1(3):e180298.
17. James P, Hart JE, Aaron H, et al. GPS-Based Exposure to Greenness and Walkability and Accelerometry-Based Physical Activity. *Cancer Epidemiol. Biomarkers Prev.* 2017;26(4):525-532.

18. Villeneuve PJ, Jerrett M, Su JG, et al. Association of residential greenness with obesity and physical activity in a US cohort of women. *Environ. Res.* 2018;160:372-384.
19. Persson A, Pyko A, Lind T, et al. Urban residential greenness and adiposity: A cohort study in Stockholm County. *Environ Int.* 2018;121(Pt 1):832-841.
20. Azevedo Da Silva M, Singh-Manoux A, Brunner EJ, et al. Bidirectional association between physical activity and symptoms of anxiety and depression: The Whitehall II study. *Eur J Epidemiol* 2012;27(7):537-546.
21. Win S, Parakh K, Eze-Nliam CM, et al. Depressive symptoms, physical inactivity and risk of cardiovascular mortality in older adults: the Cardiovascular Health Study. *Heart.* 2011;97(6):500-505.
22. Hunter RF, Christian H, Veitch J, et al. The impact of interventions to promote physical activity in urban green space: a systematic review and recommendations for future research. *Soc Sci Med.* 2015;124:246-256.
23. Sugiyama T, Leslie E, Giles-Corti B, et al. Associations of neighbourhood greenness with physical and mental health: do walking, social coherence and local social interaction explain the relationships? *J Epidemiol Community Health.* 2008;62(5):e9.
24. de Vries S, van Dillen SM, Groenewegen PP, et al. Streetscape greenery and health: stress, social cohesion and physical activity as mediators. *Soc Sci Med.* 2013;94:26-33.
25. Kim ES, Kawachi I. Perceived Neighborhood Social Cohesion and Preventive Healthcare Use. *Am J Prev Med.* 2017;53(2):e35-e40.
26. Demoury C, Thierry B, Richard H, et al. Residential greenness and risk of prostate cancer: A case-control study in Montreal, Canada. *Environ Int.* 2017;98:129-136.

27. Giovannucci EL, Liu Y, Leitzmann MF, et al. A prospective study of physical activity and incident and fatal prostate cancer. *Arch. Intern. Med.* 2005;165(9):1005-1010.
28. Pernar CH, Ebot EM, Pettersson A, et al. A Prospective Study of the Association between Physical Activity and Risk of Prostate Cancer Defined by Clinical Features and TMPRSS2:ERG. *Eur Urol.* 2019;76(1):33-40.
29. Rich-Edwards J, Corsano KA, Stampfer MJ. Test of the National Death Index and Equifax Nationwide Death Search. *Am. J. Epidemiol.* 1994;140(11):1016-1019.
30. Kriegler FJ, Malila WA, Nalepka, RF, et al. (1969). Preprocessing transformations and their effects on multispectral recognition. In: *Proceedings of the Sixth International Symposium on Remote Sensing of Environment*. Ann Arbor, MI: Institute of Science and Technology of the University of Michigan; 1969:97-131.
31. Weier J, Herring D. Measuring Vegetation (NDVI & EVI). <http://earthobservatory.nasa.gov/Features/MeasuringVegetation/>. Published August 30, 2000. Accessed May 28, 2019.
32. Ainsworth BE, Haskell WL, Herrmann SD, et al. 2011 Compendium of Physical Activities: A Second Update of Codes and MET Values. *Med Sci Sports Exerc.* 2011;43(8):1575-81.
33. Chasan-Taber S, Rimm EB, Stampfer MJ, et al. Reproducibility and validity of a self-administered physical activity questionnaire for male health professionals. *Epidemiology.* 1996;7(1):81-86.
34. Datzmann T, Markevych I, Trautmann F, Heinrich J, Schmitt J, Tesch F. Outdoor air pollution, green space, and cancer incidence in Saxony: a semi-individual cohort study. *BMC Public Health.* 18(1):715.

35. Li Q, Morimoto K, Kobayashi M, et al. Visiting a forest, but not a city, increases human natural killer activity and expression of anti-cancer proteins. *Int J Immunopathol Pharmacol.* 2008;21(1):117-127.
36. Park BJ, Tsunetsugu Y, Kasetani T, et al. The physiologic effects of Shinrin-yoku (taking in the forest atmosphere or forest bathing): evidence from field experiments in 24 forests across Japan. *Environ Health Prev Med.* 2010;15(1):18-26.
37. Egorov AI, Griffin SM, Converse RR, et al. Vegetated land cover near residence is associated with reduced allostatic load and improved biomarkers of neuroendocrine, metabolic and immune functions. *Environ Res.* 2017;158:508-521.
38. Waugh DJ, Wilson C. The interleukin-8 pathway in cancer. *Clin Cancer Res.* 2008;14(21):6735-6741.
39. Michaud DS, Daugherty SE, Berndt SI, et al. Genetic polymorphisms of interleukin-1B (IL-1B), IL-6, IL-8, and IL-10 and risk of prostate cancer. *Cancer Res.* 2006;66(8):4525-4530.
40. Muralidhar V, Rose BS, Chen YW, et al. Association Between Travel Distance and Choice of Treatment for Prostate Cancer: Does Geography Reduce Patient Choice? *Int. J. Radiat. Oncol. Biol. Phys.* 2016;96(2):313-317.
41. Cary C, Odisho AY, Cooperberg MR. Variation in prostate cancer treatment associated with population density of the county of residence. *Prostate Cancer Prostatic Dis.* 2016;19(2):174-179.
42. Gilbert R, Metcalfe C, Oliver SE, et al. Life course sun exposure and risk of prostate cancer: Population-based nested case-control study and meta-analysis. *Int J. Cancer.* 2009;125(6):1414-1423.

43. Lin S-W, Wheeler DC, Park Y, et al. Prospective study of ultraviolet radiation exposure and risk of cancer in the United States. *Int J. Cancer*. 2012;131(6):E1015-E1023.
44. Koutros S, Beane Freeman LE, Lubin JH, et al. Risk of Total and Aggressive Prostate Cancer and Pesticide Use in the Agricultural Health Study. *Am. J. Epidemiol*. 2013;177(1):59-74.
45. Jemal A, Ward E, Wu X, et al. Geographic patterns of prostate cancer mortality and variations in access to medical care in the United States. *Cancer Epidemiol. Biomarkers Prev*. 2005;14(3):590-595.
46. Krieger N. Ecosocial Theory of Disease Distribution: Embodying Societal & Ecological Context. In: Krieger N, ed. *Epidemiology and the people's health : theory and context*. New York, NY: Oxford University Press; 2011:202-236.
47. VanderWeele T, ed. *Explanation in causal inference : Methods for mediation and interaction*. New York, NY: Oxford University Press; 2015.
48. Valeri L, VanderWeele TJ. Mediation Analysis Allowing for Exposure-Mediator Interactions and Causal Interpretation: Theoretical Assumptions and Implementation With SAS and SPSS Macros. *Psychol Methods*. 2013;18(2):137-150.
49. Robins JM & Greenland S. Identifiability and Exchangeability for Direct and Indirect Effects. *Epidemiology*. 1992;3(2):143-155.
50. VanderWeele T & Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann. Intern. Med*. 2017;167(4):268-274

Supplementary Appendix

Covariate Selection

We selected individual confounding variables *a priori* based on previous studies that assessed the association between greenness and cancer, greenness and physical activity, and physical activity and cancer. Studies of physical activity and cancer have been conducted previously in HPFS, and adjusted models included prostate cancer risk factors that could be associated with neighborhood greenness, including age, BMI at age 21, height, family history of prostate cancer, smoking, diabetes mellitus, and race (2, 27). Since prostate cancer screening using prostate specific antigen (PSA) testing is associated with diagnosis of lethal prostate cancer and urbanicity (45) and thus could be associated with neighborhood greenness, we considered it to be a confounder.

Greenness is a contextual covariate and so adjustment must be made for other contextual environmental factors that may also be associated with greenness and lethal prostate cancer. Our selection of contextual covariates was guided by Krieger's ecosocial theory of epidemiology (46). We assumed that participants living in different geographic areas would be exposed to different social contexts, urban environments, and healthcare access that would impact their risk of prostate cancer. Contextual socioeconomic status was estimated using data from the 1990 US decennial census at tract level. We used median income and median home value to capture income and wealth of study participants' neighborhoods.

Use of green spaces varies between urban and rural areas (11), so we examined effect modification by population density at census tract level, using a cutpoint of 1000 people/mi² to group people into high compared to low population density areas. We further examined effect

modification by greenness exposure at home compared to work address among those participants for whom address type was documented in 1988 (N=35,474).

Mediation Analysis

We used causal mediation analysis to evaluate the importance of vigorous physical activity as a mediating pathway. Briefly, causal mediation analysis differs from traditional mediation analysis by specifying counterfactual targets that correspond to a decomposition of the total effect into a “direct” effect (effect of exposure independent of a mediator) and “indirect” effect (effect of exposure due to mediator) (47, 48). In addition, a controlled direct effect can be estimated under slightly weaker assumptions of (1) no exposure-outcome confounding and (2) no mediator-outcome confounding (48, 49).

To evaluate the importance of vigorous physical activity as a mediating pathway, we fit multinomial logistic regression models at baseline with our mediator (categories of vigorous physical activity) as the dependent variable and our exposure (NDVI, continuous and using quintiles) as our independent variable to determine the strength of the exposure-mediator association in our analysis (48). We fit a multiplicative interaction term between continuous NDVI and population density to evaluate possible differences in the association between NDVI and vigorous physical activity in high compared to low population density areas. To estimate the unbiased effect of NDVI on vigorous physical activity, we adjusted for the covariates described above, as well as additional confounders of the hypothesized exposure-mediator effect (1986 measures of non-vigorous physical activity, current BMI).

Next, we estimated the controlled direct effect of greenness on lethal prostate cancer, fixing vigorous physical activity across levels of NDVI (48, 49). Valid estimation of controlled

direct effects requires the assumption of no unmeasured mediator-outcome confounding along with the assumption of no exposure-outcome confounding, so we further adjusted for non-vigorous physical activity and current BMI as confounders of the effect of vigorous physical activity on lethal prostate cancer. Finally, to ensure correct model specification, we additionally tested for exposure-mediator interaction by fitting multiplicative interaction terms between NDVI (continuous) and quintiles of vigorous physical activity.

Cumulative Updated Average NDVI Exposure

As a sensitivity analysis, we evaluated the association between NDVI and rate of lethal prostate cancer using cumulative updated average, rather than baseline, exposure. This exposure metric was calculated by updating NDVI exposure at each follow-up point (every two years) with the average across four seasonal images per year (January, April, July, September). We analyzed the data by fitting time-varying Cox proportional hazards models, sequentially adjusting for the same variable sets as described in the primary analysis. However, for regression models using cumulative updated average NDVI, we included time-varying covariates (NDVI, smoking, vigorous physical activity, non-vigorous physical activity, current BMI, every two years; census socioeconomic status measures every 10 years, PSA screening prior to diagnosis, and PSA screening intensity). Since population density patterns were most strongly pronounced, we provided stratified estimates by high and low population density.

Sensitivity Analysis for Unmeasured Confounding

To determine robustness of our analyses to assumptions regarding unmeasured confounding, we calculated e-values corresponding to fully adjusted hazard ratios and

confidence intervals for baseline NDVI and lethal prostate cancer in the full population and among non-movers (50). The e-value reflects the minimum strength of the ratio effect measure describing 1) the association between an unmeasured confounder and outcome and 2) confounder and exposure, conditional on covariates, required to attenuate a reported association to the null. For confidence intervals, the e-value can be interpreted as the minimum bias required to shift the closer bound such that it includes the null. Stated another way, e-values provide bounds on potential bias arising from failure to adjust for unmeasured confounding. Larger e-values (farther from 1) provide stronger evidence that reported estimates are unlikely to be explained by unmeasured confounding; smaller e-values (closer to 1) provide weaker evidence

Figure S1.1. Health Professionals Follow-up Study geocoded locations (Home, Work, Other) at baseline (1988)

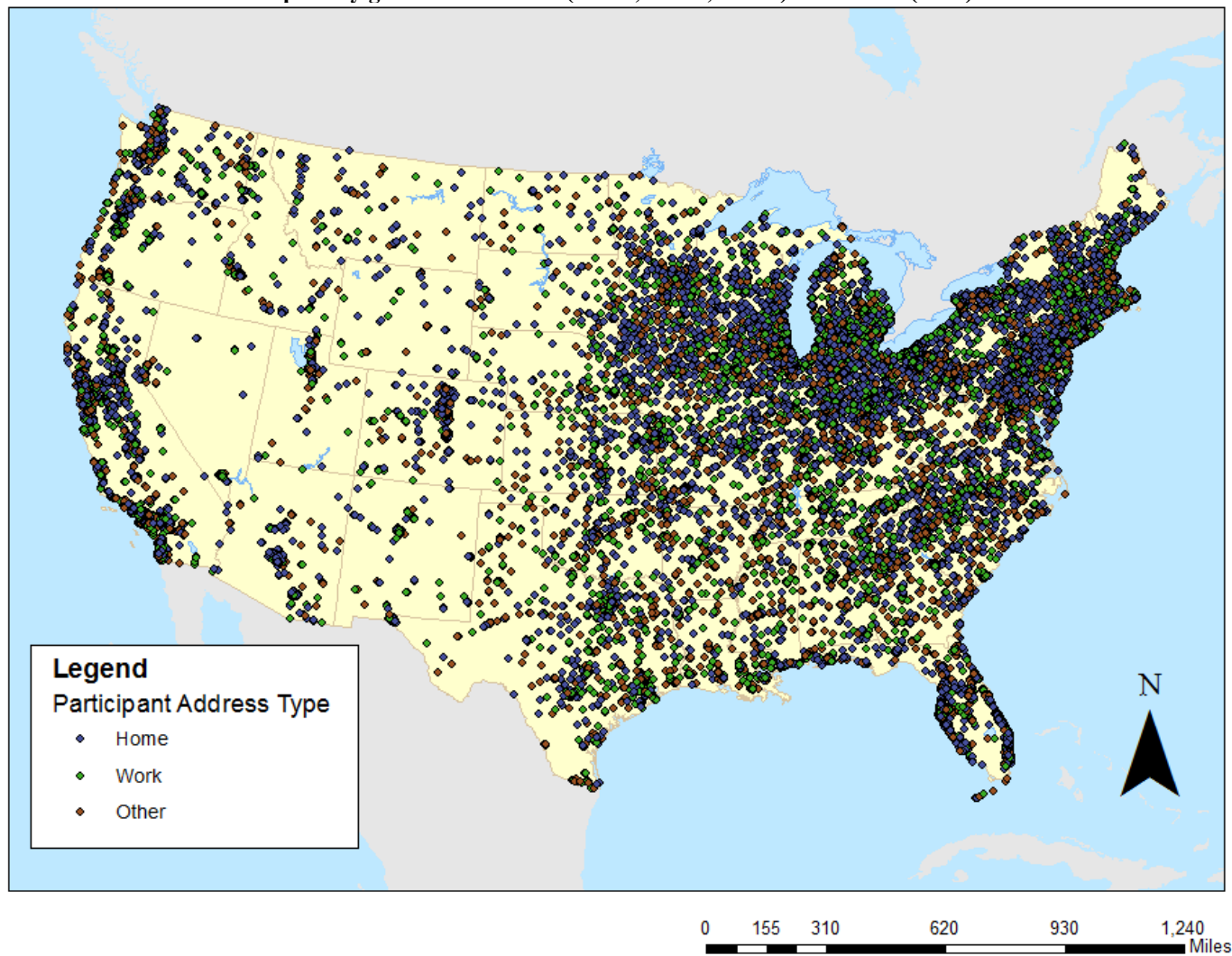


Figure S1.2. NDVI 1989 values at Health Professionals Follow-up Study geocoded locations (Home, Work, Other)

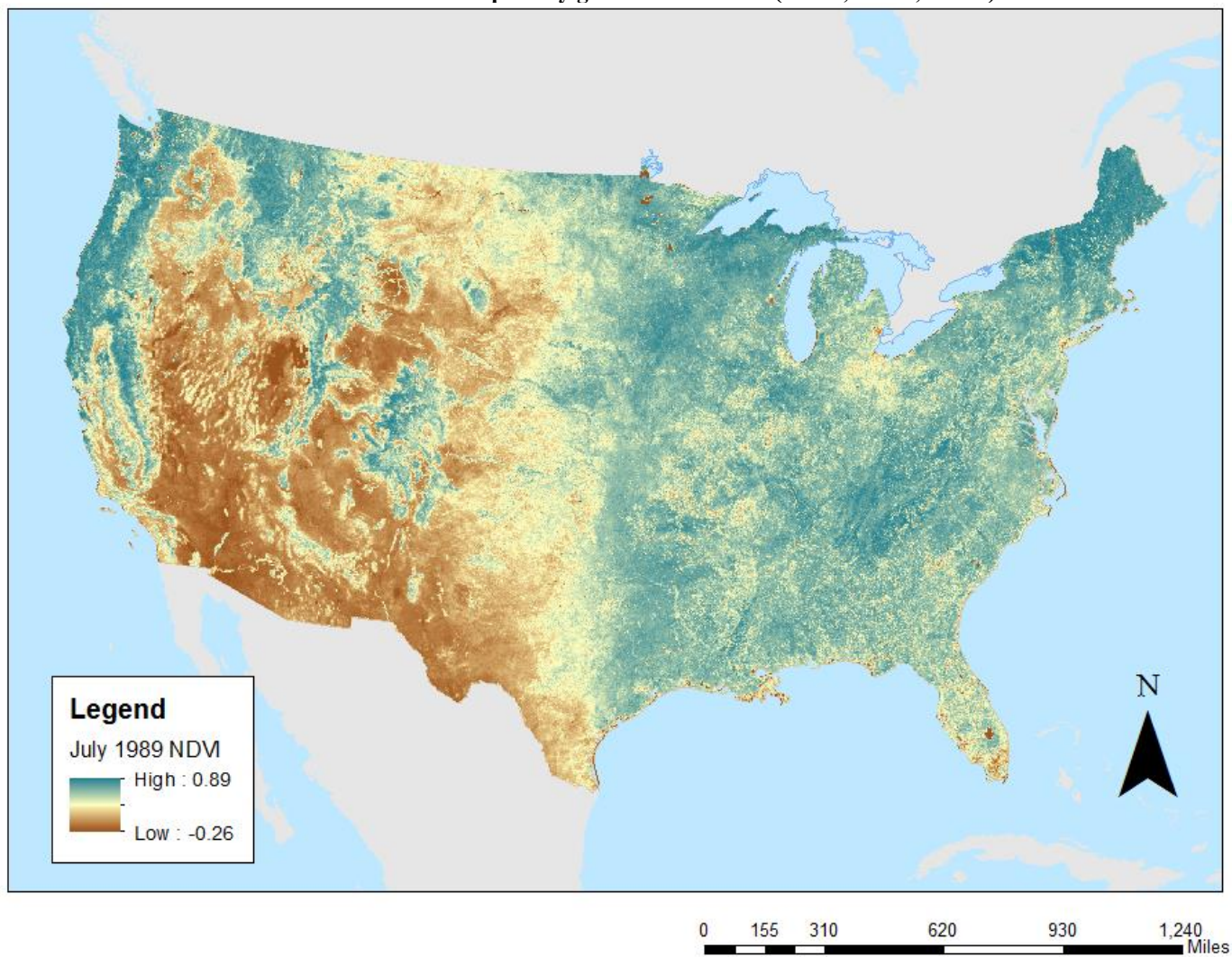


Table S1.1. Odds Ratios for the Association Between Normalized Difference Vegetation Index (NDVI) and Quintiles of Vigorous Physical Activity (PA)^a, Health Professionals Follow-up Study, 1986 (N=47,958)

Model	Baseline Normalized Difference Vegetation Index (NDVI)												
	Continuous ^b		Quintile 1		Quintile 2		Quintile 3		Quintile 4		Quintile 5		<i>P</i> _{trend}
	aOR	95% CI	aOR	95% CI	aOR	95% CI	aOR	95% CI	aOR	95% CI	aOR	95% CI	
Vigorous Physical Activity Level^c													
L1	1.00	Referent	1.00	Referent	1.00	Referent	1.00	Referent	1.00	Referent	1.00	Referent	
L2	0.94	0.91, 0.98	1.00	Referent	0.99	0.90, 1.08	0.90	0.82, 0.98	0.91	0.83, 0.99	0.89	0.81, 0.97	0.0023
L3	0.91	0.88, 0.94	1.00	Referent	0.92	0.85, 1.01	0.89	0.81, 0.97	0.85	0.78, 0.93	0.81	0.74, 0.89	<.0001
L4	0.91	0.88, 0.94	1.00	Referent	0.93	0.85, 1.01	0.85	0.78, 0.93	0.83	0.76, 0.90	0.79	0.72, 0.87	<.0001
L5	0.89	0.85, 0.92	1.00	Referent	0.95	0.86, 1.06	0.83	0.75, 0.93	0.81	0.72, 0.90	0.75	0.67, 0.84	<.0001

Abbreviations: aOR, adjusted odds ratio; BMI, body mass index; CI, confidence interval; IQR, interquartile range; L, level; NDVI, normalized difference vegetation index, USD, United States Dollars

^aAdjusted for age in months and calendar time as baseline strata, race (categorical), diabetes mellitus (yes or no), height (categorical), family history of prostate cancer (yes or no), BMI at age 21 (categorical), smoking status in 1986 (categorical), 1990 census tract median income (USD), and 1990 census tract median home value (USD), population density (binary: ≥ 1000 , < 1000 people/mi²), non-vigorous physical activity in 1986 (quintiles), current BMI in 1986 (categorical)

^bEstimate corresponds to an IQR increase in continuous NDVI of 0.11 units.

^cInterpretation of model: In a multinomial logistic regression model, odds ratios are calculated for multiple binary outcomes, in which nominal categories are compared to the referent (vigorous physical activity quintile 1). Results are provided for outcomes defined using vigorous physical activity levels (L) 2-5, where Level 1 (referent) corresponds to 0 MET-hours/week, and Levels 2-5 correspond to quartiles among participants who reported any vigorous physical activity.

Table S1.2 Hazard Ratios for the Association Between Baseline Normalized Difference Vegetation Index (NDVI) and Lethal Prostate Cancer Incidence in the Health Professionals Follow-up Study, United States, 1986-2014

Model	Baseline Normalized Difference Vegetation Index (NDVI)													<i>P</i> _{trend}
	Continuous ^d		Quintile 1		Quintile 2		Quintile 3		Quintile 4		Quintile 5			
	aHR	95% CI	aHR	95% CI	aHR	95% CI	aHR	95% CI	aHR	95% CI	aHR	95% CI		
Total population (N=47,958)														
Age and Calendar-time ^a	0.96	0.89, 1.03	1.00	Referent	0.87	0.71, 1.07	0.90	0.73, 1.10	0.78	0.63, 0.96	0.93	0.76, 1.14	0.25	
Confounding ^b	0.95	0.88, 1.03	1.00	Referent	0.88	0.72, 1.09	0.92	0.75, 1.13	0.79	0.63, 0.98	0.91	0.73, 1.13	0.23	
Controlled Direct Effects ^c	0.95	0.88, 1.03	1.00	Referent	0.88	0.71, 1.08	0.92	0.74, 1.13	0.78	0.63, 0.97	0.91	0.73, 1.13	0.21	
Men who did not move during follow-up (N=42,492)														
Age and Calendar-time ^a	0.92	0.85, 1.00	1.00	Referent	0.85	0.69, 1.05	0.88	0.71, 1.09	0.74	0.59, 0.92	0.85	0.69, 1.06	0.057	
Confounding ^b	0.92	0.85, 1.01	1.00	Referent	0.86	0.70, 1.07	0.90	0.73, 1.12	0.75	0.60, 0.95	0.84	0.67, 1.06	0.068	
Controlled Direct Effects ^c	0.92	0.85, 1.01	1.00	Referent	0.86	0.70, 1.07	0.90	0.73, 1.12	0.75	0.60, 0.94	0.84	0.67, 1.06	0.067	

Abbreviations: aHR, adjusted hazard ratio; BMI, body mass index; CI, confidence interval, IQR, interquartile range; NDVI, normalized difference vegetation index; USD, United States Dollars.

^aAdjusted for age in months and calendar time as strata.

^bAdjusted for age in months and calendar time as strata, race (categorical), diabetes mellitus (yes or no), height (categorical), family history of prostate cancer (yes or no), BMI at age 21 (categorical), smoking status in 1986 (categorical), 1990 census tract median income (USD), 1990 census tract median home value (USD), population density, history of prostate-specific antigen testing, and intensity of prostate-specific antigen testing.

^cAdjusted for age in months and calendar time as strata, race (categorical), diabetes mellitus (yes or no), height (categorical), family history of prostate cancer (yes or no), BMI at age 21 (categorical), smoking status in 1986 (categorical), 1990 census tract median income (USD), 1990 census tract median home value (USD), population density, history of prostate-specific antigen testing, intensity of prostate-specific antigen testing, vigorous physical activity (categorical), non-vigorous physical activity (quintiles), and current BMI (categorical).

^dEstimate corresponds to an IQR increase in continuous NDVI of 0.11 units.

Table S1.3. Hazard ratios for the Association Between Cumulative Updated Average Normalized Difference Vegetation Index (NDVI) and Lethal Prostate Cancer Incidence in the Health Professionals Follow-up Study, United States, 1986-2014, Stratified by Population Density

Model	Cumulative Updated Average Normalized Difference Vegetation Index (NDVI)												
	Continuous ^c		Quintile 1		Quintile 2		Quintile 3		Quintile 4		Quintile 5		<i>P</i> _{trend}
	aHR	95% CI	aHR	95% CI	aHR	95% CI	aHR	95% CI	aHR	95% CI	aHR	95% CI	
Total population^d (N=47,958)													
Age and Calendar-time ^a	0.96	0.89, 1.04	1.00	Referent	0.85	0.69, 1.04	0.84	0.68, 1.03	0.86	0.70, 1.05	0.87	0.71, 1.06	0.19
Confounding ^b	0.96	0.88, 1.05	1.00	Referent	0.85	0.69, 1.05	0.85	0.69, 1.05	0.87	0.71, 1.08	0.86	0.69, 1.07	0.21
Controlled Direct Effects ^c	0.96	0.88, 1.05	1.00	Referent	0.85	0.70, 1.06	0.85	0.69, 1.05	0.88	0.71, 1.08	0.86	0.69, 1.07	0.21
High (≥1000 people/mi²) population density (N=34,229)													
Age and Calendar-time ^a	0.93	0.84, 1.03	1.00	Referent	0.84	0.67, 1.05	0.80	0.63, 1.01	0.79	0.63, 1.00	0.81	0.62, 1.05	0.048
Confounding ^b	0.93	0.84, 1.04	1.00	Referent	0.84	0.67, 1.05	0.80	0.63, 1.02	0.80	0.63, 1.02	0.81	0.61, 1.06	0.065
Controlled Direct Effects ^c	0.93	0.84, 1.04	1.00	Referent	0.84	0.67, 1.06	0.80	0.63, 1.02	0.80	0.63, 1.03	0.80	0.61, 1.05	0.063
Low (<1000 people/mi²) population density (N=13,729)													
Age and Calendar-time ^a	1.04	0.88, 1.22	1.00	Referent	0.85	0.50, 1.44	1.06	0.64, 1.75	1.11	0.68, 1.80	1.02	0.65, 1.61	0.67
Confounding ^b	1.05	0.89, 1.24	1.00	Referent	0.91	0.53, 1.56	1.17	0.70, 1.95	1.22	0.74, 1.99	1.10	0.69, 1.75	0.51
Controlled Direct Effects ^c	1.06	0.89, 1.25	1.00	Referent	0.90	0.52, 1.55	1.15	0.69, 1.92	1.23	0.75, 2.03	1.11	0.70, 1.77	0.45

Abbreviations: aHR, adjusted hazard ratio; BMI, body mass index; CI, confidence interval, IQR, interquartile range; NDVI, normalized difference vegetation index; USD, United States Dollars.

^aAdjusted for age in months and calendar time as baseline strata

^bAdjusted for age in months and calendar time as baseline strata, race (categorical), diabetes mellitus (yes or no), height (categorical), family history of prostate cancer (yes or no), BMI at age 21 (categorical), history of prostate-specific antigen testing, smoking (categorical), intensity of prostate-specific antigen testing (categorical), census tract median income (USD), and census tract median home value (USD)

^cAdjusted for age in months and calendar time as baseline strata, race (categorical), diabetes mellitus (yes or no), height (categorical), family history of prostate cancer (yes or no), BMI at age 21 (categorical), history of prostate-specific antigen testing, smoking (categorical), intensity of prostate-specific antigen testing (categorical), census tract median income (USD), and census tract median home value (USD), vigorous physical activity (categorical), non-vigorous physical activity (quintiles), and current BMI (categorical)

^dFor models fit in total population, models 2 and 3 additionally adjusted for population density (binary: ≥1000, <1000 people/mi²)

^eEstimate corresponds to an IQR increase in continuous NDVI of 0.11 units.

Table S1.4. E-values for Hazard Ratios for the Association Between Baseline Normalized Difference Vegetation Index (NDVI) and Lethal Prostate Cancer^a, Total Population and Men who did not Move During Follow-up, Health Professionals Follow-up Study, United States, 1986-2014, stratified by population density

E-Value	Baseline Normalized Difference Vegetation Index (NDVI)											
	Continuous ^b		Quintile 1		Quintile 2		Quintile 3		Quintile 4		Quintile 5	
	Estimate	CI	Estimate	CI	Estimate	CI	Estimate	CI	Estimate	CI	Estimate	CI
Total population												
Full cohort ^c (N=47,958)	1.29	1.00	Referent		1.53	1.00	1.39	1.00	1.85	1.16	1.36	1.00
High (≥ 1000 people/mi ²) population density (N=34,229)	1.46	1.11	Referent		1.50	1.00	1.32	1.00	1.96	1.21	1.96	1.00
Low (<1000 people/mi ²) population density (N=13,729)	1.46	1.00	Referent		2.45	1.00	1.56	1.00	1.39	1.00	1.37	1.00
Participants who did not change addresses												
Full subcohort ^c (N=42,492)	1.39	1.00	Referent		1.60	1.00	1.45	1.00	2.00	1.29	1.67	1.00
High (≥ 1000 people/mi ²) population density (N=30,259)	1.53	1.21	Referent		1.53	1.00	1.36	1.00	2.30	1.50	2.08	1.11
Low (<1000 people/mi ²) population density (N=12,233)	1.34	1.00	Referent		2.78	1.00	1.70	1.00	1.24	1.00	1.21	1.00

Abbreviations: BMI, body mass index; CI, confidence interval, IQR, interquartile range; NDVI, normalized difference vegetation index; USD, United States Dollars.

^aAdjusted for age in months and calendar time as baseline strata, race (categorical), diabetes mellitus (yes or no), height (categorical), family history of prostate cancer (yes or no), BMI at age 21 (categorical), smoking status in 1986 (categorical), 1990 census tract median income (USD), 1990 census tract median home value (USD), history of prostate-specific antigen testing, and intensity of prostate-specific antigen testing

^bEstimate corresponds to an IQR increase in continuous NDVI of 0.11 units.

^cFor models fit in full cohort and subcohorts, additionally adjusted for population density (binary: ≥ 1000 , <1000 people/mi²)

Chapter 2: The Contribution of Neighborhood Greenness to Mortality among Men with Prostate Cancer: A Registry-based Cohort Study of Black and White Men

Hari S. Iyer¹, Linda Valeri², Peter James³, Jarvis T. Chen⁴, Jaime E. Hart^{5,6}, Francine Laden^{1,5}, Michelle D. Holmes^{1,5}, Timothy R. Rebbeck^{1,7}

Affiliations:

¹Harvard T. H. Chan School of Public Health, Department of Epidemiology, Boston, MA 02115; United States

²Columbia University Mailman School of Public Health, Department of Biostatistics, New York, NY 10032; United States

³Harvard Medical School and Harvard Pilgrim Health Care Institute, Department of Population Medicine, Boston, MA 02215; United States

⁴Harvard T. H. Chan School of Public Health, Department of Social and Behavioral Sciences, Boston, MA 02115; United States

⁵Brigham and Women's Hospital, Department of Medicine, Channing Division of Network Medicine, Boston, MA 02115; United States

⁶Harvard T. H. Chan School of Public Health, Department of Environmental Health, Boston, MA 02115; United States

⁷Dana Farber Cancer Institute, Department of Medical Oncology, Boston, MA 02115; United States

ABSTRACT

Background

Black men with prostate cancer (CaP) experience excess mortality compared to White men. We studied the association between neighborhood greenness, a health promoting contextual factor, and mortality among men with CaP.

Methods

We identified Pennsylvania Cancer Registry cases diagnosed between January 2000 and December 2015. Participants were followed until death or January 1, 2018, whichever occurred first. Exposure was characterized using the Normalized Difference Vegetation Index (NDVI). NDVI and indices of area-level socioeconomic status, geographic healthcare access, and segregation were assigned to participants' residential address. We estimated hazard ratios using adjusted Cox models. To determine whether increasing neighborhood greenness could reduce racial disparities, we compared standardized 10-year mortality Black-White risk differences under a hypothetical intervention fixing NDVI to the 75th percentile of NDVI experienced by White men.

Results

We observed 29,978 deaths over 916,590 person-years. Comparing men in highest to lowest NDVI quintile, all-cause (aHR: 0.88, 95% CI: 0.84, 0.92, $P_{\text{trend}} < 0.0001$), prostate-specific (aHR: 0.88, 95% CI: 0.80, 0.99, $P_{\text{trend}} = 0.0021$), and cardiovascular-specific (aHR: 0.82, 95% CI: 0.74, 0.90, $P_{\text{trend}} < 0.0001$) mortality were lower. Inverse associations between continuous NDVI (0.14 units) and cardiovascular-specific mortality were observed in White (aHR: 0.90, 95% CI: 0.86, 0.93) but not Black men (aHR: 0.97, 95% CI: 0.89, 1.06; $P_{\text{het}} = 0.067$). Hypothetical interventions

to increase NDVI led to non-significant reductions in all-cause (-5.3%) and prostate-specific (-23.2%), but not cardiovascular-specific mortality disparities (+50.5%).

Conclusion

Neighborhood greenness was associated with lower mortality among men with CaP, but findings suggest increasing neighborhood greenness would have limited impact on disparities.

INTRODUCTION

Cancer of the prostate (CaP) is the most commonly diagnosed cancer and the second leading cause of cancer-related death among men in the United States of America (US), accounting for 1 out of every 10 cancer deaths in men (1). In the US, Black men experience more than double the mortality from CaP compared with White men (2). While racial gaps in access to CaP care have narrowed over time, disparities in mortality rates among men with CaP have persisted for as long as reliable registry data have been available (3-5). Causes of racial disparities in mortality among men with CaP are multifactorial, requiring a multilevel framework that considers genetic and lifestyle risk factors along with historical and social context (6, 7). Recent advances in epidemiologic methods have provided investigators with analytic tools to quantify the impact of social and environmental policy changes on racial disparities (8-11). While most research to date on cancer disparities has focused on biological and social factors, few studies have investigated the mediating role of the built and contextual environment on racial disparities in cancer (12).

A growing literature describes numerous health benefits of neighborhood greenness, defined as the extent of green, natural vegetation within a given area. More comprehensive than “green space”, the term neighborhood “greenness” includes all vegetation in a given area, regardless of type, (e.g. parks, forests, gardens, and street trees). Neighborhood greenness is hypothesized to confer health benefits through promotion of healthy lifestyles and social cohesion, and reduction of harmful environmental exposures and bio-psychosocial stressors (13-16). Cohort studies have reported inverse associations between neighborhood greenness and several diseases, including all-cause mortality, cardiovascular disease (CVD) and depression (17-20).

We studied the association between neighborhood greenness and mortality in a cohort of Black and White men with CaP in Pennsylvania. Since earlier studies reported stronger associations between neighborhood greenness and specific causes of death (17, 18, 21, 22), we assessed the magnitude of the association between neighborhood greenness and all-cause mortality, prostate-specific mortality, and CVD mortality. We further evaluated whether the mortality disparity between Black and White men with CaP could be mediated by neighborhood greenness.

METHODS

Study design and participants

We used data from the population-based Pennsylvania Cancer Registry. We included 145,399 Black and White men with CaP diagnosed from 2000 to 2015 and followed them until death, 10 years post-diagnosis, or 1 January 2018, whichever came first. Participant addresses were geocoded using ArcGIS software version 10.2. We excluded cases who were diagnosed with *in situ* cancers (n=69), missing address at diagnosis (n=85), or missing stage or grade (n=16,677). A total of 128,568 (88%) men with CaP were included in the study. The Institutional Review Board of Harvard T. H. Chan School of Public Health approved this study protocol. Since existing data sources were used, no written consent was required for participation in the study.

Mortality assessment

CaP diagnoses were staged according to the 2000 Surveillance, Epidemiology, and End Results (SEER) summary staging guidelines (23). Race was extracted from facility medical records and included in data provided by state health providers to the Pennsylvania Cancer

Registry. Each year, the Pennsylvania Cancer Registry conducts a Death Clearance in which reportable cause of death information from Pennsylvania Death Certificates is linked with Pennsylvania Cancer Registry files. If deaths occur out of state, linkage is done through data exchanges. Causes of death were categorized based on ICD-09 and ICD-10 codes. For CaP-specific mortality, we included deaths coded as 185 (ICD-09) and C61 (ICD-10). For CVD mortality, we included deaths coded as 390-459 and I00-I99.

Exposure assessment

To estimate residential exposure to neighborhood greenness for CaP cases at time of diagnosis, we used the normalized difference vegetation index (NDVI), a satellite-derived spatial measure of neighborhood greenness (24). NDVI values range from -1 to 1 and quantify the amount of infrared light absorbed vs reflected by plant life. NDVI values approaching 1 correspond to lush forests, close to 0 reflect barren areas, and below 1 indicate bodies of water. In this study, NDVI values were restricted to 0 and above to focus specifically on associations related to green vegetation. Moderate Resolution Imaging Spectroradiometer (MODIS) data capturing NDVI at a 250 meter resolution were obtained using Google Earth Engine. We used Google Earth Engine's cloud cover algorithm to extract the least cloudy image in January, April, July, and September for every year from 2000 to 2015, representing seasonal variation in neighborhood greenness. Exposure was modeled using NDVI averaged over seasonal measures during calendar year of diagnosis (baseline), as well as cumulative updated average NDVI measurements over each participant's entire follow-up period as a sensitivity analysis. Participants were assigned the NDVI value for the 250 meter square pixel containing their residential address. We obtained area-level socioeconomic data at census Block Group level in 2000 from the National Historical Geographic Information System Database (25) and spatially

joined these data to cohort participant addresses using the R statistical package. When Block Group data were not available (n=101), we used census tract level data.

Statistical analysis

We estimated adjusted hazard ratios (aHR) and 95% confidence intervals for the association between NDVI and each of the mortality outcomes (all-cause, prostate, CVD) using multiple Cox proportional hazards models in SAS. NDVI was modeled using quintiles and as a continuous exposure using restricted cubic splines with 3 knots to test for non-linearity. When no evidence of non-linearity was observed, we modeled continuous NDVI using a linear term scaled in units of interquartile range for the study population (0.14 units). We estimated p-values for linear trend in categorical models using the median for each quintile. Since NDVI and CaP rates vary by urbanicity, we stratified our analyses by population density (≥ 1000 people/mi² vs. < 1000 people/mi²). This threshold was chosen to differentiate more rural settings from suburban and urban settings.

We sequentially adjusted for the variables that could be interpreted as confounders, or as mediators (advanced stage, marital status). In model 1, we stratified by age in 10-year categories and diagnosis year (categorical), and adjusted for race (Black vs. White); census Block Group median income (continuous: US\$); median home value (continuous: US\$); percent poverty (continuous); percent adults aged 25 and older with less than high school education (continuous); the joint race and income Index of Concentration at Extremes (ICE), a measure of inequality based on income- or race-based privilege in a given geographic area (quintiles) (26); four indicator variables for receipt of care at a currently NCI-designated cancer center (The University of Pennsylvania, University of Pittsburgh Medical Center, Fox Chase Cancer Center, or Thomas Jefferson University Hospital); population density (continuous), distance between

each participant's geocoded address to the closest cancer center using road network distances (continuous, minutes, calculated using ArcMap 10.2). In model 2, we additionally adjusted for stage (categorical: localized, regional, distant) and grade (categorical: I – IV). In model 3, we further adjusted for marital status, using logistic regression models and Monte Carlo imputation with 10 repetitions to impute missing marital status (n=46,519) conditional on the covariates used in model 1. We assessed whether primary associations varied by race (binary: Black vs White), stage (binary: Localized vs Regional/Distant), and population density (≥ 1000 people/mi², < 1000 people/mi²). Tests for effect modification were performed by fitting interaction terms between these modifiers and NDVI (continuous and as quintiles).

In order to evaluate the role of neighborhood greenness as a potential mediator of racial disparities in cause-specific mortality among men with CaP, we estimated racial disparities among men with CaP following hypothetical interventions that fix NDVI for all participants to a specific value using previously described statistical methods (9-11, 27, 28). This approach assumes no unmeasured confounding of race and cause-specific mortality, no unmeasured confounding between neighborhood greenness and cause-specific mortality, and correct model specification. Technical details are provided in the Supplementary Appendix.

First, we fit the outcome model described above (Cox model 1) for each mortality outcome, omitting NDVI. Resulting model parameters were used to estimate 10-year mortality among Black and White men, standardized to covariates described above. The difference in standardized 10-year mortality for Black and White men with CaP was defined as the racial disparity. Since most Black men with CaP in our study lived in high population density areas, we repeated this procedure separately among men living in high and low population density areas. Next, we estimated the racial disparity that would remain following hypothetical interventions to

fix NDVI to target values for all study participants. We again estimated expected racial disparities between Black and White men using our outcome model, with two additional parameters (continuous NDVI and an NDVI-race interaction). Bootstrapping with 500 repetitions was used to estimate 95% confidence intervals.

Three levels of NDVI were chosen to reflect a range of plausible values that could result from a policy change: (1) the observed racial disparity with no change in NDVI, (2) the 25th percentile of NDVI among Black men with CaP, and (3) the 75th percentile of NDVI among White men with CaP. We then estimated the proportion of racial disparity that could be eliminated by implementing policy change (3) (8, 29). Details regarding sensitivity analyses for competing risks, and estimation of bounds for bias due to unmeasured confounding using E-values (30) are provided in the Supplementary Appendix.

RESULTS

After exclusions, we observed 29,978 deaths over 916,590 person-years of post-diagnosis follow-up. Study population characteristics are presented in Table 2.1 overall and by NDVI in the year before diagnosis. Median age at diagnosis was 66 and did not vary by quintile of NDVI. Black men made up 11% of the study population and were less likely than Whites to reside in neighborhoods in the highest quintile of NDVI (NDVI Q1: 33% v. Q5: 3%). Most participants were diagnosed with localized disease (85%). Participants in greener neighborhoods (Q5) had lower population density, higher census Block Group income and median home value than participants in less green neighborhoods (Q1). Study participants were concentrated in the Southeast and Western parts of Pennsylvania, corresponding to the Pittsburgh and Philadelphia

metropolitan areas where NDVI was relatively lower than in other regions of the state (Figures 2.1 and 2.2). Cardiovascular disease was leading cause of death (n=7,677), followed by CaP (n=6,515).

Table 2.1. Descriptive Characteristics of Pennsylvania Cancer Registry Cohort Stratified by Quintile of Baseline Normalized Difference Vegetation Index (NDVI) Quintile, from 2000 to 2015

	Quintile of Normalized Difference Vegetation Index (NDVI)					Total
	Q1	Q2	Q3	Q4	Q5	
Total Population	25,708	25,723	25,719	25,717	25,701	128,568
	0.36	0.50	0.56	0.62	0.67	0.56
NDVI baseline^a	(0.30, 0.41)	(0.48, 0.52)	(0.55, 0.58)	(0.60, 0.63)	(0.65, 0.68)	(0.48, 0.63)
Age at Diagnosis, years^a	66 (59, 73)	67 (60, 74)	67 (60, 74)	66 (60, 73)	66 (60, 73)	66 (60, 73)
Black Reported Race [n (%)]	8,599 (33)	2,512 (10)	1,371 (5)	1,028 (4)	654 (3)	14,164 (11)
Married [n (%)]						
Yes	9,868 (38)	11,887 (46)	13,142 (51)	13,956 (54)	14,515 (56)	63,368 (49)
No ^b	5,910 (23)	3,752 (15)	3,342 (13)	3,001 (12)	2,676 (10)	18,681 (15)
(Missing)	9,930 (39)	10,084 (39)	9,235 (36)	8,760 (34)	8,510 (33)	46,519 (36)
Stage at Diagnosis [n (%)]						
Localized	21,665 (84)	22,007 (86)	22,060 (86)	22,097 (86)	21,990 (86)	109,819 (85)
Regional	2,721 (11)	2,679 (10)	2,672 (10)	2,804 (11)	2,902 (11)	13,778 (11)
Distant	1,322 (5)	1,037 (4)	987 (4)	816 (3)	809 (3)	4,971 (4)
Grade [n (%)]						
I	1,329 (5)	1,387 (5)	1,244 (5)	1,348 (5)	991 (4)	6,299 (5)
II	13,227 (51)	13,505 (53)	13,489 (52)	13,425 (52)	13,307 (52)	66,953 (52)
III	10,993 (43)	10,698 (42)	10,863 (42)	10,821 (42)	11,284 (44)	54,659 (43)
IV	159 (1)	133 (1)	123 (0)	123 (0)	119 (0)	657 (1)
Year of Diagnosis [n (%)]						
2000-2004	8,791 (34)	9,082 (35)	8,309 (32)	7,884 (31)	7,985 (31)	42,051 (33)
2005-2009	8,059 (31)	7,406 (29)	8,884 (35)	9,842 (38)	10,014 (39)	44,205 (34)
2010-2015	8,858 (34)	9,235 (36)	8,526 (33)	7,991 (31)	7,702 (30)	42,312 (33)
Contextual Factors						
Population Density (100 people/mi ²) ^a	24.5 (9.9, 59.3)	6.1 (1.2, 14.1)	2.4 (0.5, 8.6)	1.9 (0.5, 6.3)	1.3 (0.5, 4.3)	4.0 (0.8, 12.3)
<i>Census Block Group</i>						
Census Block Income (1000 US\$) ^a	31.1 (23.6, 39.4)	40.6 (32.2, 51.0)	43.7 (35.5, 55.0)	47.6 (38.0, 60.6)	52.9 (40.4, 72.0)	42.453 (33.2, 55.6)
Census Block % Poverty ^a	0.11 (0.05, 0.21)	0.05 (0.02, 0.09)	0.04 (0.02, 0.07)	0.03 (0.01, 0.06)	0.03 (0.01, 0.05)	0.04 (0.02, 0.09)

Table 2.1. Descriptive Characteristics of Pennsylvania Cancer Registry Cohort Stratified by Quintile of Baseline Normalized Difference Vegetation Index (NDVI) Quintile, from 2000 to 2015 (continued)

	Quintile of Normalized Difference Vegetation Index (NDVI)					Total
	Q1	Q2	Q3	Q4	Q5	
% over 25 with less than high school education ^a	0.16 (0.11, 0.21)	0.11 (0.07, 0.15)	0.1 (0.06, 0.14)	0.09 (0.06, 0.13)	0.08 (0.04, 0.13)	0.11 (0.06, 0.15)
Median home value (1000 US\$) ^a	63.2 (42.3, 87.3)	94.7 (72.6, 123.6)	104.6 (80.3, 135.8)	116.3 (88.5, 148.6)	134.1 (96.3, 188.0)	99.5 (72.4, 139.0)
Index Concentration at the Extremes						
Income ^a	-0.09 (-0.14, -0.03)	-0.02 (-0.07, 0.04)	0.0 (-0.05, 0.06)	0.02 (-0.04, 0.09)	0.05 (-0.02, 0.13)	-0.01 (-0.07, 0.06)
Joint Race/Income ^a	0 (-0.18, 0.02)	0.02 (0, 0.06)	0.03 (0.01, 0.07)	0.04 (0.02, 0.1)	0.07 (0.03, 0.17)	0.03 (0, 0.08)
Race ^a	0.74 (-0.31, 0.94)	0.95 (0.85, 0.98)	0.96 (0.9, 0.98)	0.96 (0.92, 0.98)	0.96 (0.92, 0.98)	0.95 (0.86, 0.98)
Quintiles Joint Race-Income ICE [n (%)]						
Q1	12,886 (50)	4,216 (16)	2,250 (9)	1,243 (5)	573 (2)	21,168 (16)
Q2	6,290 (24)	7,302 (28)	6,742 (26)	5,632 (22)	4,295 (17)	30,261 (24)
Q3	3,256 (13)	5,663 (22)	6,192 (24)	5,818 (23)	4,791 (19)	25,720 (20)
Q4	2,152 (8)	5,072 (20)	5,969 (23)	6,695 (26)	5,835 (23)	25,723 (20)
Q5	1,124 (4)	3,470 (13)	4,566 (18)	6,329 (25)	10,207 (40)	25,696 (20)
Geographic Access to Oncology Services [n (%)]						
<i>Accredited Cancer Center where Patient was Diagnosed</i>						
University of Pennsylvania	1,131 (4)	275 (1)	253 (1)	250 (1)	537 (2)	2,446 (2)
University of Pittsburgh Medical Center	1,579 (6)	2,533 (10)	2,818 (11)	2,857 (11)	2,514 (10)	12,301 (10)
Fox Chase	829 (3)	487 (2)	443 (2)	535 (2)	798 (3)	3,092 (2)
Jefferson Health	834 (3)	159 (1)	142 (1)	126 (0)	212 (1)	1,473 (1)
Network Distance to Closest Facility (Minutes) ^a	3.12 (1.97, 5.01)	5.73 (3.22, 11.39)	7.69 (4.34, 15.61)	8.9 (4.89, 17.32)	9.15 (5.38, 16.35)	6.45 (3.45, 13.65)
Mortality and Follow-up [n (%)]						
Follow-up (months) ^a	89.4 (49.1, 120.0)	91.6 (52.2, 120.0)	95.7 (54.0, 120.0)	101.2 (56.2, 120.0)	106.3 (63.6, 120.0)	96.4 (54.9, 120.0)
All Deaths over 10 years	6,971 (27)	6,150 (24)	6,029 (23)	5,624 (22)	5,204 (20)	29,978 (23)
Of Deaths, Prostate Cancer	1,633 (23)	1,305 (21)	1,257 (21)	1,159 (21)	1,161 (22)	6,515 (22)
Of Deaths, Cardiovascular Disease	1,831 (26)	1,647 (27)	1,505 (25)	1,428 (25)	1,266 (24)	7,677 (26)

^aMedian [Interquartile Range]

^bSingle, Divorced, Widowed

Figure 2.1. Participant Residential Address Locations in Pennsylvania Cancer Registry Prostate Cancer Cohort Study from 2000 to 2015

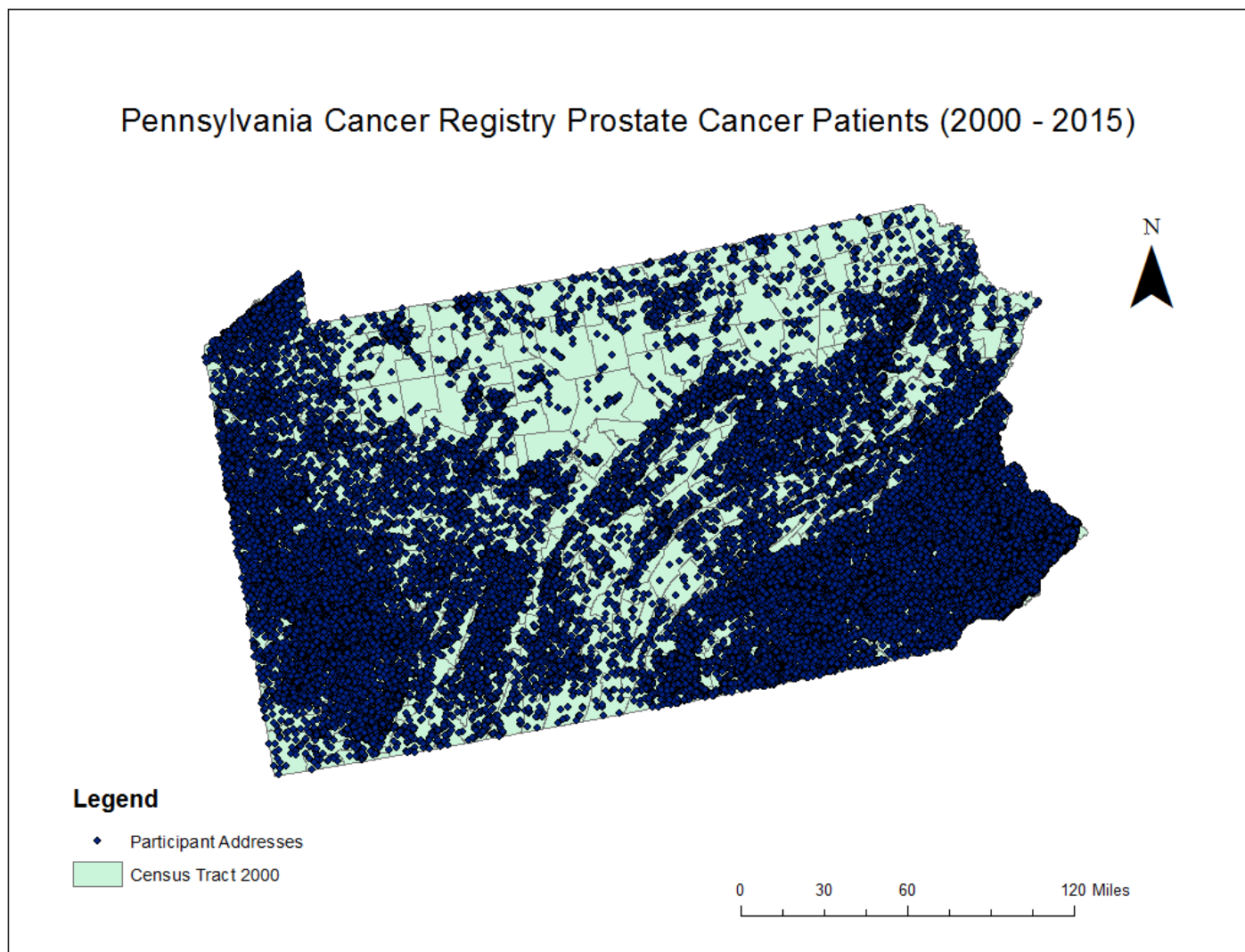
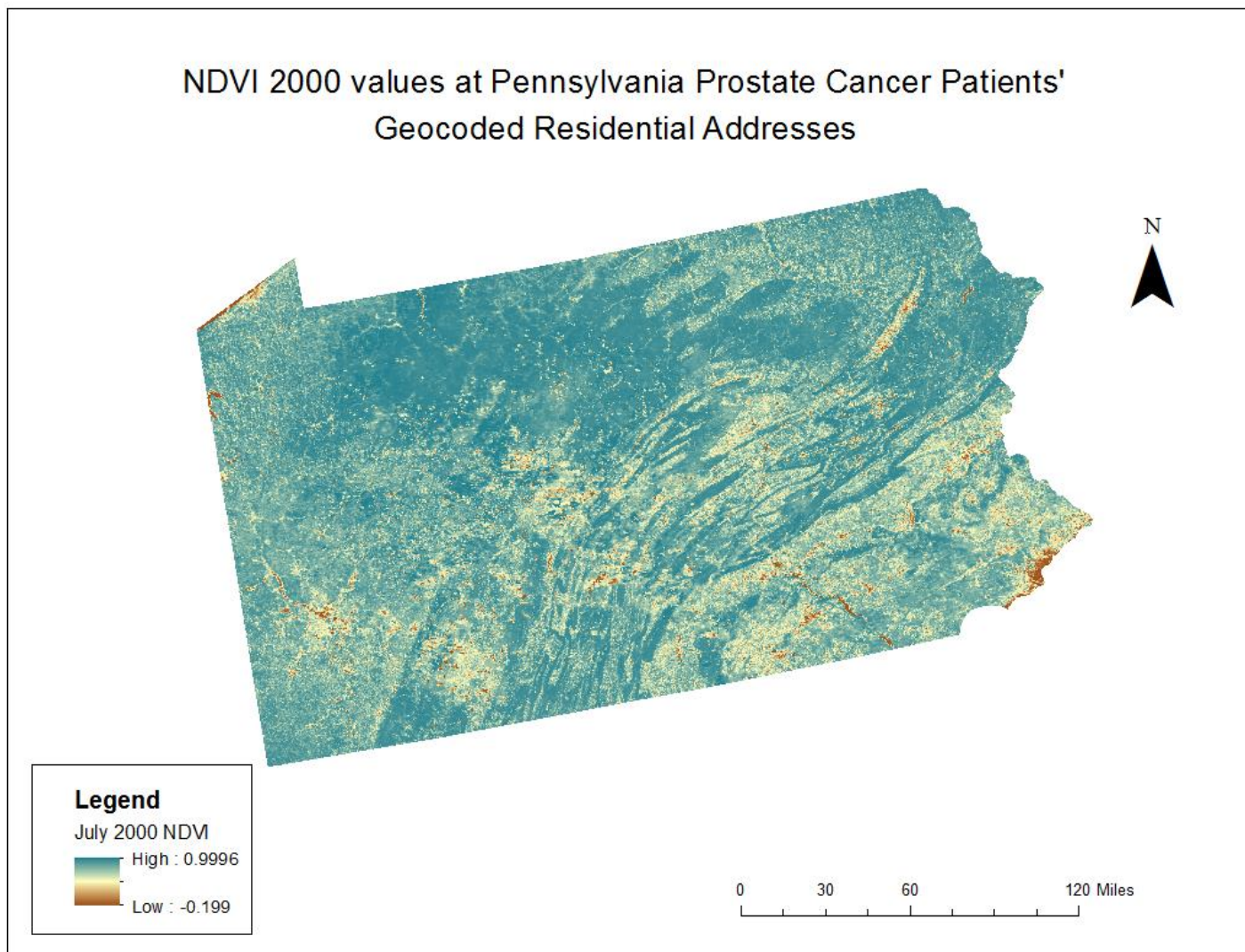


Figure 2.2. Normalized Difference Vegetation Index (NDVI) July 2000 values at Pennsylvania Cancer Registry participants' residential address locations from 2000 to 2015



In adjusted analysis, we observed statistically significant inverse associations between NDVI in the year of diagnosis and rates of mortality using quintiles and continuous exposure parameterizations (Table 2.2). Tests for splines were not significant, so we assumed linear dose response between continuous NDVI and mortality. When considering confounding factors (Model 1), there was a 12% lower rate of all-cause mortality comparing participants with NDVI Quintile 5 to 1 (Q5 to 1) (aHR: 0.88, 95% CI: 0.84, 0.92, $P_{trend}<0.0001$). This association was similar for prostate-specific mortality (aHR: 0.88, 95% CI: 0.80, 0.98, $P_{trend}=0.0021$). For CVD mortality, there was an 18% lower rate comparing NDVI Q5 to 1 (aHR: 0.82, 95% CI: 0.74, 0.90, $P_{trend}<.0001$).

Table 2.2. Cox Proportional Hazards Models for Association between Normalized Difference Vegetation Index (NDVI) at Diagnosis and Cause-specific Mortality among Pennsylvania Prostate Cancer Patients Diagnosed between 2000 and 2015

		Full	+ Stage/Grade	+ Marital Status ^a
	Cases/Person-Years	aHR (95% CI)	aHR (95% CI)	aHR (95% CI)
<i>All-cause mortality</i>				
	29,978/916,590			
Linear (per 0.14 units)		0.94 (0.92, 0.96)	0.95 (0.93, 0.97)	0.96 (0.94, 0.98)
Q1	6,972/175,795	Ref	Ref	Ref
Q2	6,149/179,406	0.94 (0.90, 0.98)	0.95 (0.92, 0.99)	0.96 (0.93, 1.00)
Q3	6,026/182,719	0.94 (0.90, 0.98)	0.95 (0.91, 0.99)	0.96 (0.93, 1.01)
Q4	5,624/185,936	0.89 (0.85, 0.93)	0.92 (0.88, 0.96)	0.94 (0.90, 0.98)
Q5	5,207/192,733	0.88 (0.84, 0.92)	0.89 (0.85, 0.94)	0.92 (0.87, 0.96)
<i>P_{trend}</i>		<.0001	<.0001	0.0004
<i>Prostate-specific Mortality</i>				
	6,515/916,590			
Linear (per 0.14 units)		0.93 (0.89, 0.98)	0.95 (0.91, 1.00)	0.96 (0.92, 1.01)
Q1	1,633/175,795	Ref	Ref	Ref
Q2	1,305/179,406	0.89 (0.82, 0.97)	0.93 (0.86, 1.01)	0.94 (0.87, 1.02)
Q3	1,256/182,719	0.88 (0.81, 0.96)	0.92 (0.84, 1.01)	0.94 (0.86, 1.02)
Q4	1,157/185,936	0.83 (0.76, 0.92)	0.92 (0.84, 1.01)	0.94 (0.85, 1.03)
Q5	1,164/192,733	0.88 (0.80, 0.98)	0.93 (0.84, 1.03)	0.96 (0.86, 1.06)
<i>P_{trend}</i>		0.0021	0.10	0.28
<i>Cardiovascular Mortality</i>				
	7,677/916,590			
Linear (per 0.14 units)		0.90 (0.87, 0.94)	0.90 (0.87, 0.94)	0.91 (0.88, 0.95)
Q1	1,832/175,795	Ref	Ref	Ref
Q2	1,646/179,406	0.95 (0.88, 1.03)	0.95 (0.88, 1.03)	0.96 (0.89, 1.04)
Q3	1,505/182,719	0.88 (0.81, 0.96)	0.88 (0.81, 0.96)	0.90 (0.83, 0.98)
Q4	1,428/185,936	0.86 (0.78, 0.93)	0.86 (0.79, 0.94)	0.88 (0.81, 0.96)
Q5	1,266/192,733	0.82 (0.74, 0.90)	0.82 (0.74, 0.90)	0.84 (0.76, 0.92)
<i>P_{trend}</i>		<.0001	<.0001	<.0001

Models adjusted for age (deciles), diagnosis year, race, census block group socioeconomic status (% poverty, median income, median home value, % 25 and older with less than high school education, joint race-income index concentration at extremes (quintiles), site at diagnosis (University of Pennsylvania Medical Center, University of Pittsburgh Medical Center, Fox Chase, Jefferson Health), network distance to closest cancer facility (minutes), population density

^aMissing values for marital status were obtained using multiple imputation

The associations with prostate-specific mortality were attenuated in models additionally adjusting for stage and grade (aHR NDVI Q5 to 1: 0.93, 95% CI: 0.84, 1.03) and then marital status (aHR NDVI Q5 to 1: 0.96, 95% CI: 0.86, 1.06). Adjusting for stage and grade did not result in major changes in inference with respect to all-cause mortality or CVD mortality. However, adjusting for marital status resulted in modest attenuation of the association with all-cause mortality (aHR NDVI Q5 to 1: 0.92, 95% CI: 0.87, 0.96), but not CVD mortality.

In stratified analyses, we found no evidence of effect modification by race, stage or population density with respect to all-cause mortality (Table 2.3). With prostate-specific mortality, the inverse association between an IQR increase in continuous NDVI and prostate-specific mortality was stronger among participants with localized (aHR: 0.92, 95% CI: 0.87, 0.97) compared to distant CaP (aHR: 0.98, 95% CI: 0.93, 1.03, $P_{het}=0.032$). In addition, the inverse association was stronger among participants in high (aHR: 0.88, 95% CI: 0.83, 0.93) compared to low (aHR: 0.96, 95% CI: 0.91, 1.01) population density areas ($P_{het}=0.028$). There was no association between continuous NDVI and CVD mortality among Black men with CaP (aHR: 0.97, 95% CI: 0.89, 1.06), but there was an inverse association among White men with CaP (aHR: 0.90, 95% CI: 0.86, 0.93, $P_{het}=0.067$), suggesting increasing levels of NDVI could increase disparities by preferentially benefiting White but not Black men with CaP.

Table 2.3. Cox Proportional Hazards Models for Association between Normalized Difference Vegetation Index (NDVI) at Diagnosis and Cause-specific Mortality among Pennsylvania Prostate Cancer Patients Diagnosed between 2000 and 2015, Stratified by Race, Stage and Population Density

	All-cause Mortality		Prostate-specific Mortality		Cardiovascular Mortality	
	Cases/Person-Years	aHR (95% CI)	Cases/Person-Years	aHR (95% CI)	Cases/Person-Years	aHR (95% CI)
<i>Race^a</i>						
White	26,472/819,457		5,601/819,457		6,796/819,457	
Linear (per 0.14 units)		0.94 (0.92, 0.96)		0.94 (0.90, 0.98)		0.90 (0.86, 0.93)
Q1	4,694/118,068	Ref	1,018/118,068	Ref	1,271/118,068	Ref
Q2	5,545/161,961	0.94 (0.90, 0.98)	1,166/161,961	0.90 (0.83, 0.99)	1,486/161,961	0.94 (0.87, 1.02)
Q3	5,721/172,952	0.93 (0.89, 0.98)	1,172/172,952	0.88 (0.80, 0.96)	1,436/172,952	0.88 (0.80, 0.95)
Q4	5,411/178,657	0.89 (0.85, 0.93)	1,107/178,657	0.84 (0.76, 0.92)	1,361/178,657	0.84 (0.76, 0.91)
Q5	5,101/187,820	0.88 (0.84, 0.92)	1,138/187,820	0.89 (0.80, 0.99)	1,242/187,820	0.81 (0.74, 0.90)
<i>P_{trend}</i>		<.0001		0.0046		<.0001
Black	3,506/97,132		914/97,132		881/97,132	
Linear (per 0.14 units)		0.94 (0.89, 0.98)		0.90 (0.82, 0.98)		0.97 (0.89, 1.06)
Q1	2,278/57,728	Ref	615/57,728	Ref	561/57,728	Ref
Q2	604/17,445	0.94 (0.86, 1.04)	139/17,445	0.82 (0.68, 0.99)	160/17,445	1.01 (0.84, 1.21)
Q3	305/9,767	0.96 (0.84, 1.08)	84/9,767	0.99 (0.78, 1.26)	69/9,767	0.88 (0.68, 1.14)
Q4	213/7,279	0.95 (0.82, 1.10)	50/7,279	0.83 (0.62, 1.12)	67/7,279	1.23 (0.94, 1.60)
Q5	106/4,913	0.75 (0.62, 0.92)	26/4,913	0.69 (0.46, 1.03)	24/4,913	0.70 (0.46, 1.06)
<i>P_{trend}</i>		0.022		0.048		0.66
<i>P_{het}</i> (quintiles)		0.47		0.41		0.076
<i>P_{het}</i> (linear)		0.84		0.33		0.067

Models adjusted for age (deciles), diagnosis year, race, census block group socioeconomic status (% poverty, median income, median home value, % 25 and older with less than high school education, joint race-income index concentration at extremes (quintiles), site at diagnosis (University of Pennsylvania Medical Center, University of Pittsburgh Medical Center, Fox Chase, Jefferson Health), network distance to closest cancer facility (minutes), population density

^aModel includes Race and NDVI interaction (1 degree of freedom)

^bModel includes Stage and NDVI interaction (1 degree of freedom)

^cModel includes Population Density (High: ≥ 1000 people/mi², Low: < 1000 people/mi²) and NDVI interaction (1 degree of freedom)

Table 2.3. Cox Proportional Hazards Models for Association between Normalized Difference Vegetation Index (NDVI) at Diagnosis and Cause-specific Mortality among Pennsylvania Prostate Cancer Patients Diagnosed between 2000 and 2015, Stratified by Race, Stage and Population Density (continued)

	All-cause Mortality		Prostate-specific Mortality		Cardiovascular Mortality	
	Cases/Person-Years	aHR (95% CI)	Cases/Person-Years	aHR (95% CI)	Cases/Person-Years	aHR (95% CI)
<i>Stage^b</i>						
<i>Localized</i>	23,930/804,433		2,926/804,433		6,947/804,433	
Linear (per 0.14 units)		0.94 (0.92, 0.96)		0.92 (0.87, 0.97)		0.90 (0.87, 0.94)
Q1	5,484/153,397	Ref	731/153,397	Ref	1,654/153,397	Ref
Q2	4,940/157,537	0.95 (0.91, 0.99)	605/157,537	0.91 (0.81, 1.02)	1,481/157,537	0.94 (0.87, 1.01)
Q3	4,849/161,217	0.95 (0.90, 0.99)	566/161,217	0.88 (0.78, 0.99)	1,366/161,217	0.88 (0.81, 0.96)
Q4	4,548/163,499	0.91 (0.87, 0.95)	531/163,499	0.85 (0.75, 0.96)	1,293/163,499	0.85 (0.77, 0.93)
Q5	4,109/168,784	0.88 (0.83, 0.92)	493/168,784	0.84 (0.74, 0.96)	1,153/168,784	0.82 (0.74, 0.90)
<i>P_{trend}</i>		<.0001		0.0038		<.0001
<i>Regional/Distant</i>	6,048/112,156		3,589/112,156		730/112,156	
Linear (per 0.14 units)		0.95 (0.92, 0.98)		0.98 (0.93, 1.03)		0.91 (0.83, 0.99)
Q1	1,488/22,398	Ref	902/22,398	Ref	178/22,398	Ref
Q2	1,209/21,870	0.93 (0.86, 1.01)	700/21,870	0.94 (0.84, 1.04)	165/21,870	1.07 (0.86, 1.32)
Q3	1,177/21,503	0.94 (0.87, 1.02)	690/21,503	0.95 (0.85, 1.06)	139/21,503	0.93 (0.74, 1.16)
Q4	1,076/22,437	0.89 (0.82, 0.96)	626/22,437	0.90 (0.80, 1.01)	135/22,437	0.94 (0.74, 1.18)
Q5	1,098/23,949	0.94 (0.86, 1.02)	671/23,949	0.98 (0.87, 1.11)	113/23,949	0.83 (0.65, 1.06)
<i>P_{trend}</i>		0.024		0.42		0.12
<i>P_{het}</i> (quintiles)		0.34		0.33		0.79
<i>P_{het}</i> (linear)		0.55		0.032		0.89

Models adjusted for age (deciles), diagnosis year, race, census block group socioeconomic status (% poverty, median income, median home value, % 25 and older with less than high school education, joint race-income index concentration at extremes (quintiles), site at diagnosis (University of Pennsylvania Medical Center, University of Pittsburgh Medical Center, Fox Chase, Jefferson Health), network distance to closest cancer facility (minutes), population density

^aModel includes Race and NDVI interaction (1 degree of freedom)

^bModel includes Stage and NDVI interaction (1 degree of freedom)

^cModel includes Population Density (High: ≥ 1000 people/mi², Low: < 1000 people/mi²) and NDVI interaction (1 degree of freedom)

Table 2.3. Cox Proportional Hazards Models for Association between Normalized Difference Vegetation Index (NDVI) at Diagnosis and Cause-specific Mortality among Pennsylvania Prostate Cancer Patients Diagnosed between 2000 and 2015, Stratified by Race, Stage and Population Density (continued)

	All-cause Mortality		Prostate-specific Mortality		Cardiovascular Mortality	
	Cases/Person-Years	aHR (95% CI)	Cases/Person-Years	aHR (95% CI)	Cases/Person-Years	aHR (95% CI)
<i>Population Density^c</i>						
<i>Low</i>	20,020/644,185		4,306/644,185		5,082/644,185	
Linear (per 0.14 units)		0.94 (0.91, 0.96)		0.96 (0.91, 1.01)		0.89 (0.85, 0.94)
Q1	1,787/44,248	Ref	392/44,248	Ref	488/44,248	Ref
Q2	3,841/114,433	0.95 (0.89, 1.00)	823/114,433	0.92 (0.82, 1.04)	1,018/114,433	0.95 (0.85, 1.06)
Q3	4,662/143,322	0.94 (0.89, 1.00)	989/143,322	0.91 (0.81, 1.03)	1,156/143,322	0.88 (0.79, 0.98)
Q4	4,834/160,757	0.90 (0.85, 0.96)	998/160,757	0.86 (0.76, 0.97)	1,227/160,757	0.86 (0.77, 0.97)
Q5	4,896/181,426	0.89 (0.83, 0.94)	1,104/181,426	0.91 (0.80, 1.04)	1,193/181,426	0.82 (0.73, 0.92)
<i>P_{trend}</i>		<.0001		0.090		0.0002
<i>High</i>	9,958/272,404		2,209/272,404		2,595/272,404	
Linear (per 0.14 units)		0.94 (0.92, 0.97)		0.88 (0.83, 0.93)		0.93 (0.88, 0.98)
Q1	5,185/131,548	Ref	1,241/131,548	Ref	1,344/131,548	Ref
Q2	2,308/64,973	0.93 (0.88, 0.98)	482/64,973	0.85 (0.76, 0.95)	628/64,973	0.96 (0.87, 1.06)
Q3	1,364/39,397	0.93 (0.88, 0.99)	267/39,397	0.83 (0.72, 0.95)	349/39,397	0.90 (0.79, 1.02)
Q4	790/25,179	0.86 (0.79, 0.93)	159/25,179	0.78 (0.66, 0.93)	201/25,179	0.82 (0.70, 0.95)
Q5	311/11,307	0.85 (0.76, 0.96)	60/11,307	0.74 (0.56, 0.96)	73/11,307	0.77 (0.60, 0.98)
<i>P_{trend}</i>		<.0001		<.0001		0.0028
<i>P_{het}</i> (quintiles)		0.87		0.58		0.91
<i>P_{het}</i> (linear)		0.84		0.028		0.24

Models adjusted for age (deciles), diagnosis year, race, census block group socioeconomic status (% poverty, median income, median home value, % 25 and older with less than high school education, joint race-income index concentration at extremes (quintiles), site at diagnosis (University of Pennsylvania Medical Center, University of Pittsburgh Medical Center, Fox Chase, Jefferson Health), network distance to closest cancer facility (minutes), population density

^aModel includes Race and NDVI interaction (1 degree of freedom)

^bModel includes Stage and NDVI interaction (1 degree of freedom)

^cModel includes Population Density (High: ≥ 1000 people/mi², Low: < 1000 people/mi²) and NDVI interaction (1 degree of freedom)

Racial disparities in 10-year mortality without adjustment for NDVI were greatest for all-cause mortality, resulting in 29.3 excess deaths per 1,000 (95% CI: 22.1, 36.5) among Black men with CaP, and least for CVD mortality (11.5, 95% CI: 6.4, 16.7 excess deaths per 1,000). Disparities were greater in low (All-cause: 33.9, 95% CI: 20.9, 47.8; Prostate: 22.1, 95% CI: 13.0, 31.2; CVD: 16.8 (95% CI: 7.3, 26.3 per 1,000) compared to high population density areas (All-cause: 25.1, 95% CI: 15.2, 35.0; Prostate: 15.1, 95% CI: 8.1, 22.1; CVD: 8.5, 95% CI: 1.6, 15.3 per 1,000). There were no statistically significant differences in racial disparities among men with CaP following hypothetical interventions fixing residential NDVI to the 25th percentile (Black), observed values of NDVI, or the 75th percentile (White) (Table 2.4). Fixing NDVI to the 75th percentile (White) resulted in the lowest cause-specific mortality, and fixing NDVI to the 25th percentile (Black) resulted in the highest cause-specific mortality in all scenarios except for CVD mortality among Black men in low population density areas (Figure 2.3).

Table 2.4. Cause-specific 10-year Mortality Risks^a, Disparities^b, and 95% Confidence Intervals Under Three Levels of Normalized Difference Vegetation Index (NDVI) at Diagnosis among Black and White Men with Prostate Cancer in Pennsylvania, 2000 to 2015

10-year mortality risk ^a	All-cause Mortality			Prostate-specific Mortality			Cardiovascular-specific Mortality		
	Black/100	White/100	Disparity/1,000	Black/100	White/100	Disparity/1,000	Black/100	White/100	Disparity/1,000
Total Population									
No NDVI intervention	21.1 (20.4, 21.8)	18.2 (18.0, 18.3)	29.3 (22.1, 36.5)	6.4 (5.8, 6.9)	4.6 (4.5, 4.8)	17.3 (11.6, 22.9)	6.5 (6.0, 7.0)	5.4 (5.2, 5.5)	11.5 (6.4, 16.7)
NDVI threshold									
25 th percentile (Black)	22.8 (21.7, 23.9)	19.6 (19.1, 20.1)	31.8 (21.7, 41.9)	7.2 (6.3, 8.1)	5.1 (4.7, 5.4)	21.3 (13.2, 29.4)	7.1 (6.4, 7.9)	6.3 (5.9, 6.6)	8.8 (1.8, 15.8)
Observed	21.1 (20.2, 21.9)	18.2 (18.0, 18.3)	29.0 (20.1, 37.8)	6.2 (5.6, 6.8)	4.6 (4.5, 4.7)	15.6 (9.0, 22.2)	6.9 (6.2, 7.5)	5.4 (5.2, 5.5)	14.9 (8.4, 21.5)
75 th percentile (White)	20.3 (19.1, 21.5)	17.6 (17.3, 17.8)	27.8 (15.7, 39.9)	5.8 (4.9, 6.6)	4.4 (4.3, 4.6)	13.3 (4.5, 22.0)	6.7 (5.9, 7.6)	5.0 (4.8, 5.2)	17.4 (8.4, 26.3)
Proportion Eliminated^c			-0.053			-0.232			+0.505
Urban (≥1000 people/mi²)									
No NDVI intervention	22.1 (21.3, 22.9)	19.6 (19.2, 19.9)	25.1 (15.2, 35.0)	6.5 (5.9, 7.1)	5.0 (4.7, 5.3)	15.1 (8.1, 22.1)	6.8 (6.2, 7.4)	5.9 (5.7, 6.2)	8.5 (1.6, 15.3)
NDVI threshold									
25 th percentile (Black)	23.0 (21.9, 24.1)	20.5 (19.8, 21.2)	25.6 (13.0, 38.2)	7.1 (6.2, 7.9)	5.5 (5.0, 5.9)	15.8 (6.5, 25.1)	7.1 (6.3, 7.9)	6.5 (6.0, 7.0)	6.3 (-2.4, 15.1)
Observed	22.0 (21.1, 22.9)	19.6 (19.2, 20.0)	24.0 (13.6, 34.5)	6.5 (5.8, 7.1)	5.0 (4.7, 5.3)	14.6 (7.4, 21.8)	6.9 (6.2, 7.5)	6.0 (5.7, 6.2)	9.1 (1.8, 16.5)
75 th percentile (White)	21.2 (19.8, 22.6)	18.9 (18.4, 19.5)	22.8 (7.8, 37.8)	6.0 (5.1, 6.9)	4.6 (4.3, 5.0)	13.6 (3.7, 23.5)	6.7 (5.7, 7.7)	5.6 (5.2, 5.9)	11.1 (0.7, 21.6)
Proportion Eliminated			-0.092			-0.098			+0.314
Rural (<1000 people/mi²)									
No NDVI intervention	21.0 (19.7, 22.3)	17.6 (17.4, 17.8)	33.9 (20.9, 47.8)	6.7 (5.8, 7.6)	4.5 (4.4, 4.6)	22.1 (13.0, 31.2)	6.8 (5.9, 7.7)	5.1 (5.0, 5.3)	16.8 (7.3, 26.3)
NDVI threshold									
25 th percentile (Black)	21.9 (20.4, 23.4)	18.3 (17.9, 18.7)	36.0 (20.6, 51.4)	7.4 (6.3, 8.5)	4.6 (4.4, 4.9)	27.7 (16.8, 38.6)	6.9 (5.8, 8.0)	5.5 (5.3, 5.8)	14.3 (3.0, 25.5)
Observed	20.7 (19.4, 22.0)	17.6 (17.4, 17.8)	31.2 (17.8, 44.5)	6.3 (5.4, 7.2)	4.5 (4.4, 4.6)	17.9 (8.4, 27.4)	7.0 (6.0, 7.9)	5.1 (5.0, 5.3)	18.6 (9.0, 28.2)
75 th percentile (White)	20.0 (18.3, 21.8)	17.2 (16.9, 17.4)	28.5 (11.1, 46.0)	5.7 (4.5, 6.8)	4.4 (4.2, 4.6)	12.4 (0.7, 24.1)	7.0 (5.8, 8.2)	4.9 (4.7, 5.1)	21.2 (8.7, 33.7)
Proportion Eliminated^c			-0.163			-0.439			+0.265

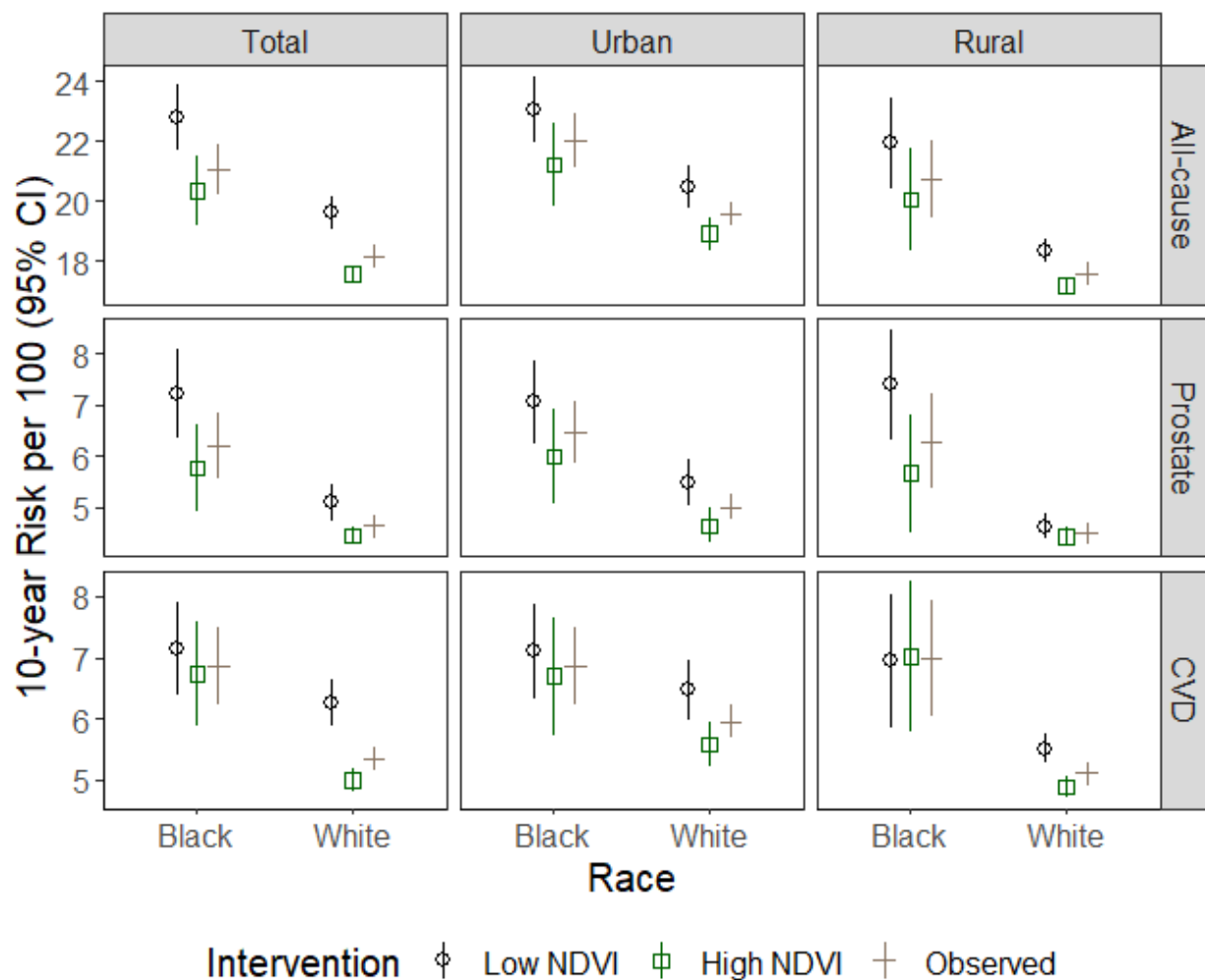
Table 2.4. Cause-specific 10-year Mortality Risks^a, Disparities^b, and 95% Confidence Intervals Under Three Levels of Normalized Difference Vegetation Index (NDVI) at Diagnosis among Black and White Men with Prostate Cancer in Pennsylvania, 2000 to 2015 (continued)

^a10-year risks among Black and White men estimated using Cox models with continuous NDVI (0.14 unit increase), adjusted for age (deciles), diagnosis year, race, census block group socioeconomic status (% poverty, median income, median home value, % 25 and older with less than high school education, joint race-income index concentration at extremes (quintiles), site at diagnosis (University of Pennsylvania Medical Center, University of Pittsburgh Medical Center, Fox Chase, Jefferson Health), network distance to closest cancer facility (minutes), population density (total only), and interaction term between NDVI and race, standardized to distribution of confounders in total Urban and Rural population

^bDisparity denotes racial disparity, estimated by taking difference in 10-year mortality risks in Black and White men standardized to covariates. Confidence intervals were estimated using bootstrapping with 500 repetitions.

^cProportion disparity eliminated by increasing NDVI estimated by taking the difference of racial disparity under No NDVI intervention and 75th percentile of NDVI (White), divided by disparity under No NDVI intervention. (-) indicates disparity would be reduced, (+) indicates disparity would be increased.

Figure 2.3. Standardized 10-year Cause-specific (All-Cause, Cardiovascular-, Prostate-specific) Mortality Risk and 95% Confidence Intervals Under Three Scenarios^a of Normalized Difference Vegetation Index (NDVI) at Diagnosis among Black and White Men with Prostate Cancer in Pennsylvania)



^aScenario 1: Observed NDVI, Scenario 2: Low NDVI (25th percentile of Black men with CaP), Scenario 3: High NDVI (75th percentile of White men with CaP), CVD = cardiovascular disease specific-mortality, (Urban: ≥ 1000 people/mi², Rural: < 1000 people/mi²)

Estimated proportions of racial disparity in mortality that would be eliminated by fixing neighborhood greenness to the 75th percentile of NDVI (White) were modest for all-cause (5.3%) and prostate-specific (23.2%) mortality. However, for CVD mortality, we estimated a relative 50.5% increase in the racial disparity following this hypothetical intervention (Table 2.4). These findings are consistent with results from our race-stratified models, in which NDVI was associated with lower CVD mortality among White but not Black men with CaP. Stratification by population density preserved these patterns, though estimated proportions of racial disparities eliminated for all-cause and prostate mortality were greater in low compared to high population density areas (Table 2.4).

In sensitivity analysis for competing risks, results for stratified associations between NDVI and prostate- and CVD-specific mortality resulted in slightly weaker estimates compared to primary results and no change to inference so we did not use competing risk models for our main analysis (Table S2.1). E-values summarizing bounds of bias due to unmeasured confounding for our primary effect estimates are provided in Table S2.2. Effect sizes for the association between socioeconomic status, a likely confounding variable, and mortality among men with CaP from previous registry-based studies range from 1.14 to 1.52 (31). E-values for all-cause mortality and prostate-specific mortality lie within this range, meaning that if an unmeasured factor exhibited patterns of association with NDVI and mortality similar to that of socioeconomic status, adjusting for that factor could explain away these results. However, this unmeasured factor would need to be sufficiently correlated with NDVI and mortality even after adjusting for the demographic, socioeconomic, and geographic access variables already included in our analysis. Our strongest e-values corresponding to the hazard ratio for CVD-mortality comparing men in Q5 to Q1 are 1.75 for point estimate, and 1.47 for confidence interval,

suggesting that these results are unlikely to be explained by unmeasured confounding bias.

Associations between cumulative updated average NDVI and mortality exhibited non-linear dose response, with increased all-cause mortality and prostate-specific mortality in lowest and highest quintiles of NDVI. (Table S2.3).

DISCUSSION

In this cohort of Black and White men with CaP, we observed inverse associations between NDVI and lower all-cause, prostate-, and CVD-specific mortality after adjusting for demographics, neighborhood socioeconomic context, and geographic health care access. Our results suggest that increasing levels of neighborhood greenness could result in modest, non-significant decreases in racial disparities in all-cause and prostate-specific mortality. However, we estimated increases in racial disparities in CVD mortality among men with CaP following hypothetical interventions to increase neighborhood greenness. In our sensitivity analysis using cumulative updated average, we observed different dose-response patterns compared to analyses using NDVI at time of diagnosis. Increased all-cause and prostate specific mortality observed with increasing cumulative updated average NDVI could be attributable to reverse causation, resulting from tree planting and greening interventions such as Philadelphia's "Green Works" program, implemented from 2009 to present. These interventions were targeted precisely at those urban areas which were most deprived and experienced worse outcomes during the study period (32).

While few studies have reported associations between neighborhood greenness and mortality among men with CaP, our findings are consistent with results from earlier prospective

population-based and occupational cohort studies in the US, Canada, and Europe, which have also reported inverse associations between neighborhood greenness and all-cause mortality (17, 19, 21, 22). Most men in our study were diagnosed with localized CaP. Ten-year survival is relatively high among these men, and deaths from prostate cancer are few relative to deaths from other causes like CVD (33, 34). This suggests mortality risks for these men could be similar to the general population. Cohort studies in Canada, Europe and the US have also reported inverse associations between neighborhood greenness and cardiovascular mortality (18, 19, 21, 22, 35). Though we did not have data to evaluate lifestyle risk factors, prior research shows that physical activity is associated with lower mortality risk (36-38), and obesity is associated with higher risk (39, 40) among men with CaP. In our study, CVD-specific mortality was the leading cause of death among men with CaP, so inverse associations between neighborhood greenness and mortality reported here could be due to reduced CVD-specific mortality, possibly through pathways related to physical activity and obesity (13, 14, 16).

The second question we sought to answer was whether increasing neighborhood greenness could reduce racial disparities in mortality among men with CaP. No differences in the association between NDVI and either all-cause or prostate-specific mortality comparing Black to White men with CaP were observed. However, for CVD mortality, we observed an inverse association with NDVI in White but not Black men. Wide confidence intervals for the cause-specific racial disparities from our simulation-based approach limited our ability to statistically evaluate differences in disparities under hypothetical interventions to fix NDVI to different thresholds. However, estimates of the proportion of disparity eliminated suggest that increasing neighborhood greenness could lead to modest reductions in disparities in all-cause mortality. Estimated reductions in racial disparities for CaP mortality were offset by increases in disparities

for CVD mortality. Cohort studies in the general population from Canada and Europe looking at all-cause and CVD mortality have also reported stronger inverse associations among high income or privileged racial groups (21, 22). Better understanding of how contextual environment and CaP outcomes vary by race in different US and global settings will be essential to informing policy interventions.

While we lacked data to explain racial differences in the association between NDVI and CVD mortality, literature on differing patterns of park use between Black and White men and women offers some guidance. Parks are a major contributor to urban neighborhood greenness. The ways in which Black and White men experience neighborhood greenness could be different, which in turn could have consequences for potential health benefits of exposure to high levels of greenness. Parks in predominantly Black neighborhoods may be used less frequently due to fewer resources for security and maintenance (41). Black men in the US may use parks differently because they were historically excluded from public parks through segregation (42, 43). Residents' perception of higher crime rates, lower levels of walkability, and lack of upkeep could make parks less welcoming for physical activity and socializing, particularly for older community members (44-46). Surveys of park users in the US have found that Black community members often cite greater obstacles to using parks compared to White users, including feeling unwelcome, inconvenient schedules, and financial barriers (42, 47). These findings suggest that merely introducing neighborhood green spaces in these communities, without any attempts to ensure that the space fits the needs and social mores of that community, could fail to produce any health benefits.

Results from our study should be interpreted in light of its limitations. This study was conducted in the state of Pennsylvania, which has a unique history, geography, and racial

composition. Thus, our results may not be generalizable to dissimilar populations. NDVI is a popular objective measure of neighborhood greenness, but does not capture quality or accessibility of green spaces, which may be necessary to inform appropriate interventions. It is possible that our exposure assessment approach may have introduced measurement error because participants may spend considerable time outside of the home in green areas, may have moved during follow up, and spent significant amounts of time away from their residence, though this measurement error is likely to be non-differential after accounting for other important confounding factors. We did not have measurements of screening, diet, and lifestyle factors at diagnosis, which could influence neighborhood selection and mortality, leading to unmeasured confounding (a threat to all observational studies). Our sensitivity analysis using E-values suggests that unmeasured variables would be unlikely to completely explain our effect estimates, particularly for CVD mortality. If lifestyle factors lie on the causal path between neighborhood greenness and mortality, we would not adjust even if those data were available. Strengths of our study include a cohort design with long follow-up, a large, racially diverse population, adjustment for major sociodemographic, clinical, and contextual environmental confounders, and analysis of the contribution of environment to racial disparities.

In conclusion, we report an inverse association between neighborhood greenness and rate of all-cause, prostate- and CVD-specific mortality among men with CaP in Pennsylvania. While interventions to set thresholds of neighborhood greenness could have limited impact on reducing racial disparities, increases in greenness were associated with reduced all-cause and prostate-specific mortality rates among both Black and White men with CaP. Enhanced understanding of differences in how Black and White men interact with green spaces could inform targeted nature-based interventions to allow all men with CaP to experience those benefits.

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2019. *CA Cancer J Clin.* 2019;69(1):7-34, PMID: 30620402, <https://doi.org/10.3322/caac.21551>
2. Hsing AW, Tsao L, Devesa SS. International trends and patterns of prostate cancer incidence and mortality. *Int J Cancer.* 2000;85(1): 60-7, PMID: 10585584, [https://doi.org/10.1002/\(sici\)1097-0215\(20000101\)85:1<60::aid-ijc11>3.0.co;2-b](https://doi.org/10.1002/(sici)1097-0215(20000101)85:1<60::aid-ijc11>3.0.co;2-b)
3. Kelly SP, Rosenberg PS, Anderson WF, Andreotti G, Younes N, Cleary SD, Cook MB. Trends in the Incidence of Fatal Prostate Cancer in the United States by Race. *Eur Urol.* 2017;71(2): 195-201, PMID: 27476048, <https://doi.org/10.1016/j.eururo.2016.05.011>
4. Chornokur G, Dalton K, Borysova ME, Kumar NB. Disparities at presentation, diagnosis, treatment, and survival in African American Men, affected by Prostate Cancer. *Prostate.* 2011;71(9):985-97, PMID: 21541975, <https://doi.org/10.1002/pros.21314>
5. Rebbeck TR. Prostate Cancer Disparities by Race and Ethnicity: From Nucleotide to Neighborhood. *Cold Spring Harbor Perspect Med.* 2018;8(9): pii: a030387, PMID: 29229666, <https://doi.org/10.1101/cshperspect.a030387>
6. Krieger N. Epidemiology and the web of causation: Has anyone seen the spider? *Soc Sci Med.* 1994;39(7): 887-903, PMID: 7992123, [https://doi.org/10.1016/0277-9536\(94\)90202-x](https://doi.org/10.1016/0277-9536(94)90202-x)
7. Lynch SM, Rebbeck TR. (2013). Bridging the gap between biologic, individual, and macroenvironmental factors in cancer: A multilevel approach. *Cancer Epidemiol Biomarkers Prev.* 2013;22(4): 485-95, PMID: 23462925 <https://doi.org/10.1158/1055-9965.EPI-13-0010>

8. VanderWeele T. ed. *Explanation in causal inference : Methods for mediation and interaction*. New York: Oxford University Press; 2015.
9. Valeri L, Chen JT, Garcia-Albeniz X, Krieger N, Vanderweele TJ, Coull BA. The Role of Stage at Diagnosis in Colorectal Cancer Black-White Survival Disparities: A Counterfactual Causal Inference Approach. *Cancer Epidemiol Biomarkers Prev*. 2016;25(1): 83-9, PMID: 26503034, <https://doi.org/10.1158/1055-9965.EPI-15-0456>
10. VanderWeele TJ, Robinson WR. On the Causal Interpretation of Race in Regressions Adjusting for Confounding and Mediating Variables. *Epidemiology*. 2014;25(4): 473-484, PMID: 24887159, <https://doi.org/10.1097/EDE.000000000000105>
11. Jackson JW, VanderWeele TJ. Decomposition analysis to identify intervention targets for reducing disparities. *Epidemiology*. 2018;29(6): 825-835, PMID: 30063540, <https://doi.org/10.1097/EDE.0000000000000901>
12. Gomez SL, Shariff-Marco S, DeRouen M, Keegan TH, Yen IH, Mujahid M et al. The impact of neighborhood social and built environment factors across the cancer continuum: Current research, methodological considerations, and future directions. *Cancer* 2015;121(14): 2314-30, PMID:25847484, <https://doi.org/10.1002/cncr.29345>
13. James P, Banay RF, Hart JE, Laden F. A Review of the Health Benefits of Greenness. *Current Epidemiol Rep*. 2015;2(2): 131-142, PMID:26185745 <https://doi.org/10.1007/s40471-015-0043-7>
14. Markevych I, Schoierer J, Hartig T, Chudnovsky A, Hystad P, Dzhambov AM et al. Exploring pathways linking greenspace to health: Theoretical and methodological

- guidance. *Environ Res.* 2017;158: 301-317, PMID:28672128,
<https://doi.org/10.1016/j.envres.2017.06.028>
15. Frumkin H, Bratman GN, Breslow SJ, Cochran B, Kahn PH Jr, Lawler JJ et al. Nature Contact and Human Health: A Research Agenda. *Environ Health Perspect.* 2017;125(7): 075001, PMID:28796634, <https://doi.org/10.1289/EHP1663>
16. Fong KC, Hart JE, James P. A Review of Epidemiologic Studies on Greenness and Health: Updated Literature Through 2017. *Curr Environ Health Rep.* 2018;5(1): 77-87, PMID:29392643, <https://doi.org/10.1007/s40572-018-0179-y>
17. James P, Hart JE, Banay RF, Laden F. Exposure to Greenness and Mortality in a Nationwide Prospective Cohort Study of Women. *Environ Health Perspect.* 2016;124(9): 1344-52: PMID:27074702, <https://doi.org/10.1289/ehp.1510363>
18. Wilker EH, Wu CD, McNeely E, Mostofsky E, Spengler J, Wellenius GA et al. Green space and mortality following ischemic stroke. *Environ Res.* 2014;133: 42-8. PMID:24906067, <https://doi.org/10.1016/j.envres.2014.05.005>
19. Twohig-Bennett C, Jones A. The health benefits of the great outdoors: A systematic review and meta-analysis of greenspace exposure and health outcomes. *Environ Res* 2018;166: 628-637. PMID:29982151, <https://doi.org/10.1016/j.envres.2018.06.030>
20. South EC, Hohl BC, Kondo MC, MacDonald JM, Branas CC. Effect of Greening Vacant Land on Mental Health of Community-Dwelling Adults: A Cluster Randomized Trial. *JAMA Netw Open.* 2018;1(3): e180298. PMID:30646029
<https://doi.org/10.1001/jamanetworkopen.2018.0298>
21. Vienneau D, de Hoogh K, Faeh D, Kaufmann M, Wunderli JM, & Rösli M. More than clean air and tranquillity: Residential green is independently associated with decreasing

- mortality. *Environ Int.* 2017;108: 176-184. PMID:28863390,
<https://doi.org/10.1016/j.envint.2017.08.012>
22. Crouse DL, Pinault L, Balram A, Hystad P, Peters PA, Chen H et al. Urban greenness and mortality in Canada's largest cities: A national cohort study. *Lancet Planet Health.* 2017;1(7): e289-e297. PMID:29851627, [https://doi.org/10.1016/S2542-5196\(17\)30118-3](https://doi.org/10.1016/S2542-5196(17)30118-3)
23. Young JL Jr, Roffers SD, Ries LAG, Fritz AG, Hurlbut AA. eds. *SEER Summary Staging Manual – 2000: Codes and Coding Instructions*, Bethesda: National Cancer Institute, NIH Pub. No. 01-4969; 2001.
24. Kriegler FJ, Malila WA, Nalepka, RF, Richardson W. (1969). Preprocessing transformations and their effects on multispectral recognition. In: *Proceedings of the Sixth International Symposium on Remote Sensing of Environment*. Ann Arbor, MI: Institute of Science and Technology of the University of Michigan; 1969: 97–131.
25. Manson S, Schroeder J, Van Riper D, Ruggles S. IPUMS National Historical Geographic Information System: Version 13.0 [Database]. Minneapolis: University of Minnesota; 2018. Accessed at [<https://data2.nhgis.org/main>] on 18 July 2019.
26. Krieger N, Waterman PD, Spasojevic J, Li W, Maduro G, Van Wye G. Public health monitoring of privilege and deprivation with the index of concentration at the extremes. *Am J Public Health.* 2016;106(2): 256-63. PMID:26691119,
<https://doi.org/10.2105/AJPH.2015.302955>
27. Naimi AI, Schnitzer ME, Moodie EM, Bodnar LM. Mediation Analysis for Health Disparities Research. *Am J Epidemiol.* 2016;184(4):315-24. PMID:27489089,
<https://doi.org/10.1093/aje/kwv329>

28. Hernán MA, Robins JM. Causal Inference. Boca Raton: Chapman & Hall/CRC, forthcoming. Available at: <https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/>
29. Valeri L, VanderWeele TJ. Mediation Analysis Allowing for Exposure–Mediator Interactions and Causal Interpretation: Theoretical Assumptions and Implementation With SAS and SPSS Macros. *Psychol Methods*. 2013;18(2):137-150. PMID:23379553, <https://doi.org/10.1037/a0031034>
30. VanderWeele T & Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann. Intern. Med.* 2017;167(4): 268-274. PMID:28693043, <https://doi.org/10.7326/M16-2607>
31. Klein J, von dem Knesebeck O. Socioeconomic inequalities in prostate cancer survival: A review of the evidence and explanatory factors. *Soc Sci Med*. 2015;142:9-18.
32. Greenworks Philadelphia. City of Philadelphia, Mayor’s Office of Sustainability. Available at: <https://www.phila.gov/media/20160419140515/2009-greenworks-vision.pdf>. Accessed on 18 July 2019.
33. Albertsen PC, Hanley JA, Gleason DF, Barry MJ. Competing risk analysis of men aged 55 to 84 years at diagnosis managed conservatively for clinically localized prostate cancer. *JAMA*. 1998;280(11):975-80. PMID:9749479, <https://doi.org/10.1001/jama.280.11.975>
34. Lu-Yao GL, Albertsen PC, Moore DF, Lin Y, DiPaola RS, Yao SL. Fifteen-year Outcomes Following Conservative Management Among Men Aged 65 Years or Older with Localized Prostate Cancer. *Eur Urol*. 2015;68(5):805-11. PMID:25800944, <https://doi.org/10.1016/j.eururo.2015.03.021>

35. Yitshak-Sade M, James P, Kloog I, Hart JE, Schwartz JD, Laden F et al. Neighborhood Greenness Attenuates the Adverse Effect of PM_{2.5} on Cardiovascular Mortality in Neighborhoods of Lower Socioeconomic Status. *Int J Environ Res Public Health*. 2019;16(5): pii: e814. PMID:30845676, <https://doi.org/10.3390/ijerph16050814>
36. Wekesa A, Harrison M, Watson RW. Physical activity and its mechanistic effects on prostate cancer. *Prostate Cancer Prostatic Dis*. 2015;18(3):197-207. PMID:25800589, <https://doi.org/10.1038/pcan.2015.9>
37. Kenfield SA, Stampfer MJ, Giovannucci E, Chan JM. Physical activity and survival after prostate cancer diagnosis in the health professionals follow-up study. *J Clin Oncol*. 2011;29(6):726-32. PMID:21205749, <https://doi.org/10.1200/JCO.2010.31.5226>
38. Wang Y, Jacobs EJ, Gapstur SM, Maliniak ML, Gansler T, McCullough ML et al. Recreational Physical Activity in Relation to Prostate Cancer-specific Mortality Among Men with Nonmetastatic Prostate Cancer. *Eur Urol*. 2017;72(6):931-939. PMID:28711382, <https://doi.org/10.1016/j.eururo.2017.06.037>
39. Allott EH, Masko EM, Freedland SJ. Obesity and prostate cancer: weighing the evidence. *Eur Urol*. 2013;63(5):800-9. PMID:23219374, DOI:[10.1016/j.eururo.2012.11.013](https://doi.org/10.1016/j.eururo.2012.11.013)
40. Zhang X, Zhou G, Sun B, Zhao G, Liu D, Sun J et al. Impact of obesity upon prostate cancer-associated mortality: A meta-analysis of 17 cohort studies. *Oncol Lett*. 2015;9(3):1307-1312. PMID:25663903, <https://doi.org/10.3892/ol.2014.2841>
41. Bruton CM, Floyd M. Disparities in Built and Natural Features of Urban Parks: Comparisons by Neighborhood Level Race/Ethnicity and Income. *J Urban Health*. 2014;91(5): 894-907. PMID:25078037, <https://doi.org/10.1007/s11524-014-9893-4>

42. Byrne J, Wolch J. Nature, race, and parks: Past research and future directions for geographic research. *Prog Hum Geogr.* 2009;33(6), 743-765.
43. Wolch JR, Byrne J, Newell JP. Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough'. *Landsc Urban Plan.* 2014;125: 234-244.
44. Cohen DA, McKenzie TL, Sehgal A, Williamson S, Golinelli D, Lurie N. Contribution of public parks to physical activity. *Am J Public Health.* 2007;97(3), 509-14, PMID:17267728, <https://doi.org/10.2105/AJPH.2005.072447>
45. Kaczynski AT, Wilhelm Stanis SA, Hastmann TJ, Besenyi GM. Variations in observed park physical activity intensity level by gender, race, and age: individual and joint effects. *J Phy Act Health.* 2011;8 Suppl 2: S151-60. PMID:28829710, <https://doi.org/10.1123/jpah.8.s2.s151>
46. Gobster PH. Managing Urban Parks for a Racially and Ethnically Diverse Clientele. *Leis Sci,* 2002;24(2): 143-159.
47. Das KV, FanY, French SA. Park-use Behavior and Perceptions by Race, Hispanic Origin, and Immigrant Status in Minneapolis, MN: Implications on Park Strategies for Addressing Health Disparities. *J Immigr. Minor Health.* 2017;19(2): 318-327. PMID:26724925, <https://doi.org/10.1007/s10903-015-0339-1>.
48. VanderWeele TJ. Policy-relevant proportions for direct effects. *Epidemiology.* 2013;24(1): 175-176.
49. Breslow NE. Discussion of the paper by D. R. Cox J. *R Statist Soc B.* 1972;34: 216-217.
50. Hernán MA, Hernandez-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004;15(5): 615-25

Supplementary Appendix

This appendix provides additional details regarding procedures used for the following analyses conducted in our study:

1. Methods to estimate the impact of hypothetical interventions to increase neighborhood greenness on racial disparities in mortality among men with prostate cancer
2. Sensitivity analyses for competing risks
3. Sensitivity analyses to estimate minimum bounds of bias due unmeasured confounding

1. Methods to Evaluate Impact of Policy to Increase Neighborhood Greenness on Racial Disparities in Mortality

We used epidemiologic methods for causal mediation analysis to evaluate whether increasing neighborhood greenness could reduce racial disparities in mortality among men with prostate cancer. Briefly, causal mediation analysis is an analytic approach that allows decomposition of a total effect into a portion that is attributable only to exposure in the absence of a mediator (defined as the *natural direct effect*), and a portion that is attributable to the exposure's effect on the outcome which flows through the mediator (*natural indirect effect*) (8, 29). In this framework, an effect is defined as the difference in potential outcomes that would be observed in the presence or absence of an exposure.

Causal mediation analysis also allows estimation of the *controlled direct effect*, or the effect of an exposure on an outcome that would occur if the mediator were, possibly contrary to fact, fixed to a specific value (8, 29). Estimating the controlled direct effect is particularly useful

when attempting to understand if an intervention that occurs downstream of a particular exposure can reduce the effect of that exposure (48). Estimation of the controlled direct effect requires fewer assumptions than natural direct and indirect effects, namely that (1) the covariate set accounts for confounding of the exposure and outcome (which is implicitly assumed in any observational study), and (2) confounding of the mediator and outcome (8, 29, 48).

Epidemiologists have applied this causal mediation analysis framework to study how policy changes could impact racial and socioeconomic disparities (9-11, 27). In these settings, the type of disparity (for example, the difference in mortality comparing Black and White men) is treated as the “exposure”, and the policy change is treated as the “mediator” (10). Assuming that we have sufficient covariate data to control for confounding of race and our outcome, as well as our policy and our outcome, we can estimate controlled direct effects at different levels of our policy. These controlled direct effects can be interpreted as the disparity that would remain following implementation of the policy under study. Concretely, in our study, controlled direct effects correspond to the residual racial disparities in cause-specific mortality following implementation of hypothetical interventions or policies to fix Normalized Difference Vegetation Index (NDVI) for all men with prostate cancer to specific values.

Since the key focus of this analysis was understanding the extent to which racial disparities could be reduced by increasing neighborhood greenness, we estimated the proportion of racial disparity that would be eliminated by fixing NDVI to the 75th percentile experienced by White men with prostate cancer using the equation below (48):

Equation 1.

Proportion Eliminated =

$$\frac{\text{Racial Disparity-Controlled Direct Effect (NDVI= NDVI at 75th percentile for White Men)}}{\text{Racial Disparity}}$$

We modeled counterfactual 10-year mortality risks for Black and White CaP cases under different levels of NDVI using Cox proportional hazards models. For estimation of the overall racial disparity in mortality, we fit Cox model 1 described in the methods, omitting the NDVI variable. For estimation of controlled direct effects, we fit Cox models using the covariates in model 1, and additionally including a term for continuous NDVI and a term for the interaction between NDVI and race (8, 29). We calculated the survival proportion at each failure time using the Breslow method (49). We then standardized these race-specific counterfactual 10-year mortality risks to the total study population using Robins' g-formula, a simulation-based procedure. We made three copies of the dataset, one which contained the observed data, one in which the race of all participants was set to White, and one in which the race of all participants was set to Black. Since estimates preserve the distribution of covariates defined above, model estimates correspond to 10-year cause-specific mortality in each racial group, standardized to confounders in the total population (28).

We then computed 10-year mortality risk differences using the standardized 10-year mortality for Black men and White men with CaP, and obtained 95% confidence intervals using 500 bootstrapped samples (28). We repeated this approach separately for each mortality outcome in the total population. We then repeated the analysis separately for high (≥ 1000 people/mi²) and low (< 1000 people/mi²) population density areas because most Black study participants lived in high population density areas. In addition, since NDVI varied between high and low population density areas, thresholds for each policy were set separately by level of population density.

2. Competing Risks

When assessing prostate- and CVD-specific mortality, to assess sensitivity to competing risks, we calculated inverse probability of censoring weights for all other causes of death using multiple logistic regression (50). Briefly, the weights simulate experience of a “pseudopopulation” in which death from other causes cannot occur, and so models associations of interest in the absence of those competing risks of death. We compared estimates from these weighted models to our primary analysis to determine whether competing risks would lead to changes in our inference. Since we did not find evidence that competing risks resulted in major changes in inference, we did not apply this approach to our primary analyses.

3. Unmeasured Confounding

We computed E-values to estimate bounds of bias due to unmeasured confounding for each hazard ratio point estimate and confidence interval reported in Table 2 (30). The E-value quantifies the minimum strength of the relative association between an unmeasured confounder and either exposure or outcome, conditional on covariates, required to attenuate the observed point estimate to the null value of 1. For confidence intervals, the E-value represents the minimum association needed to shift the confidence interval limit closest to 1 to contain that null value. Larger E-values suggest stronger bias due to unmeasured confounding would be needed to explain the reported association. Smaller E-values provide weaker evidence against unmeasured confounding bias as an explanation for reported findings.

We computed E-values for point estimates and confidence intervals for each of the adjusted hazard ratios estimated for the association between NDVI and cause-specific mortality, under the different sets of confounders in Table 2 of our main manuscript.

Table S2.1. Sensitivity Analysis for Competing Risks: Cox Proportional Hazards Model Estimates of Association between Normalized Difference Vegetation Index (NDVI) at Diagnosis and Prostate-specific Mortality and Cardiovascular Mortality using Inverse Probability of Censoring Weights

Prostate-specific Mortality			Cardiovascular-specific Mortality		
	Cases/Person-Years ^a	aHR (95% CI)		Cases/Person-Years ^a	aHR (95% CI)
<i>Population Density</i>			<i>Race</i>		
<i>Low</i>	6,742/700,116		<i>White</i>	10,472/881,646	
Linear (per 0.14 units)		0.94 (0.89, 1.00)	Linear (per 0.14 units)		0.90 (0.86, 0.94)
Q1	615/49,334	Ref	Q1	2,063/129,441	Ref
Q2	1,259/124,658	0.98 (0.85, 1.11)	Q2	2,287/174,725	0.94 (0.86, 1.02)
Q3	1,518/156,210	0.95 (0.83, 1.08)	Q3	2,223/185,991	0.89 (0.81, 0.97)
Q4	1,495/174,455	0.88 (0.77, 1.01)	Q4	2,053/191,172	0.84 (0.76, 0.92)
Q5	1,587/195,459	0.93 (0.81, 1.06)	Q5	1,849/200,318	0.82 (0.75, 0.91)
<i>P_{trend}</i>		0.10	<i>P_{trend}</i>		<.0001
<i>High</i>	3,446/301,159		<i>Black</i>	1,289/104,019	
Linear (per 0.14 units)		0.89 (0.83, 0.94)	Linear (per 0.14 units)		0.98 (0.89, 1.08)
Q1	1,951/146,463	Ref	Q1	827/62,223	Ref
Q2	738/71,551	0.90 (0.80, 1.02)	Q2	239/18,538	1.06 (0.87, 1.29)
Q3	433/43,202	0.90 (0.77, 1.04)	Q3	101/10,297	0.96 (0.72, 1.28)
Q4	238/27,660	0.78 (0.64, 0.94)	Q4	91/7,722	1.21 (0.91, 1.61)
Q5	88/12,284	0.75 (0.57, 0.99)	Q5	32/5,239	0.69 (0.43, 1.1)
<i>P_{trend}</i>		0.0028	<i>P_{trend}</i>		0.9446
<i>P_{het}</i> (quintiles)		0.61	<i>P_{het}</i> (quintiles)		0.094
<i>P_{het}</i> (linear)		0.11	<i>P_{het}</i> (linear)		0.064

^aEstimates from pseudopopulation generated using inverse probability of censoring weights. This model simulates the experience of a population that can only exit the cohort through death from the specified cause.

Table S2.2. E-values for Robustness to Unmeasured Confounding for Cox Model Hazard Ratio Estimates of the Association between Normalized Difference Vegetation Index (NDVI) at Diagnosis and Rate of Cause-specific Mortality among men with prostate cancer

	Full	+ Stage/Grade	+ Marital Status ^a
	aHR (Lower Bound)	aHR (Lower Bound)	aHR (Lower Bound)
<i>All-cause mortality</i>			
Linear (per 0.14 units)	1.26 (1.20)	1.24 (1.18)	1.21 (1.15)
Q1	Ref	Ref	Ref
Q2	1.26 (1.14)	1.23 (1.09)	1.19 (1.00)
Q3	1.27 (1.15)	1.23 (1.09)	1.19 (1.00)
Q4	1.38 (1.27)	1.32 (1.20)	1.26 (1.13)
Q5	1.41 (1.31)	1.37 (1.26)	1.32 (1.20)
<i>Prostate-specific Mortality</i>			
Linear (per 0.14 units)	1.35 (1.19)	1.28 (1.06)	1.23 (1.00)
Q1	Ref	Ref	Ref
Q2	1.49 (1.21)	1.35 (1.00)	1.31 (1.00)
Q3	1.53 (1.24)	1.39 (1.00)	1.34 (1.00)
Q4	1.69 (1.41)	1.40 (1.00)	1.34 (1.00)
Q5	1.52 (1.19)	1.36 (1.00)	1.27 (1.00)
<i>Cardiovascular Mortality</i>			
Linear (per 0.14 units)	1.45 (1.32)	1.45 (1.32)	1.41 (1.27)
Q1	Ref	Ref	Ref
Q2	1.29 (1.00)	1.29 (1.00)	1.24 (1.00)
Q3	1.52 (1.26)	1.52 (1.26)	1.47 (1.19)
Q4	1.61 (1.35)	1.61 (1.34)	1.53 (1.25)
Q5	1.75 (1.47)	1.75 (1.47)	1.67 (1.39)

Models adjusted for age (deciles), diagnosis year, race, census block group socioeconomic status (% poverty, median income, median home value, % 25 and older with less than high school education, joint race-income index concentration at extremes (quintiles)), site at diagnosis (University of Pennsylvania Medical Center, University of Pittsburgh Medical Center, Fox Chase, Jefferson Health), network distance to closest cancer facility (minutes), population density

^aMarital Status estimated using multiple imputation

Table S2.3. Cox Proportional Hazards Models for Association between Cumulative Updated Average NDVI over follow-up and Cause-specific Mortality among Pennsylvania Prostate Cancer Patients Diagnosed between 2000 and 2015

		Full	+ Stage/Grade	+ Marital Status ^a
	Cases/Person-Years	aHR (95% CI)	aHR (95% CI)	aHR (95% CI)
<i>All-cause mortality</i>	29,978/916,590			
Linear (per 0.14 units)		1.03 (1.01, 1.05)	1.02 (1.00, 1.05)	1.04 (1.02, 1.06)
Q1	6,966/176,384	Ref	Ref	Ref
Q2	6,431/179,028	0.97 (0.93, 1.01)	0.97 (0.93, 1.01)	0.98 (0.94, 1.02)
Q3	5,606/187,427	0.87 (0.83, 0.91)	0.89 (0.85, 0.93)	0.91 (0.87, 0.95)
Q4	5,253/190,346	0.86 (0.82, 0.90)	0.88 (0.84, 0.92)	0.89 (0.85, 0.94)
Q5	5,722/183,404	1.11 (1.06, 1.17)	1.11 (1.06, 1.16)	1.13 (1.08, 1.19)
<i>P_{trend}</i>		0.007	0.033	0.002
<i>Prostate-specific Mortality</i>	6,515/916,590			
Linear (per 0.14 units)		1.04 (1.00, 1.09)	1.03 (0.99, 1.07)	1.04 (1.00, 1.08)
Q1	1,612/176,384	Ref	Ref	Ref
Q2	1,394/179,028	0.98 (0.90, 1.06)	0.97 (0.89, 1.05)	0.97 (0.90, 1.06)
Q3	1,140/187,427	0.83 (0.76, 0.91)	0.90 (0.82, 0.98)	0.91 (0.83, 1.00)
Q4	1,019/190,346	0.78 (0.71, 0.86)	0.81 (0.74, 0.90)	0.83 (0.75, 0.92)
Q5	1,350/183,404	1.22 (1.10, 1.35)	1.19 (1.07, 1.31)	1.21 (1.09, 1.34)
<i>P_{trend}</i>		0.021	0.24	0.0948
<i>Cardiovascular Mortality</i>	7,677/916,590			
Linear (per 0.14 units)		0.97 (0.94, 1.01)	0.97 (0.94, 1.01)	0.99 (0.95, 1.02)
Q1	1,854/176,384	Ref	Ref	Ref
Q2	1,692/179,028	0.93 (0.86, 1.00)	0.92 (0.86, 1.00)	0.94 (0.87, 1.01)
Q3	1,400/187,427	0.79 (0.72, 0.86)	0.79 (0.72, 0.86)	0.81 (0.74, 0.88)
Q4	1,381/190,346	0.83 (0.76, 0.90)	0.83 (0.76, 0.90)	0.85 (0.78, 0.93)
Q5	1,350/183,404	0.96 (0.87, 1.05)	0.96 (0.87, 1.05)	0.98 (0.89, 1.08)
<i>P_{trend}</i>		0.33	0.30	0.61

Models adjusted for age (deciles), diagnosis year, race, census block group socioeconomic status (% poverty, median income, median home value, % 25 and older with less than high school education, joint race-income index concentration at extremes (quintiles)), site at diagnosis (University of Pennsylvania Medical Center, University of Pittsburgh Medical Center, Fox Chase, Jefferson Health), network distance to closest cancer facility (minutes), population density

^aMissing values for marital status were obtained using multiple imputation

Chapter 3: Neighborhood Greenness and Burden of Non-communicable Diseases in Sub-Saharan Africa: A Multi-country Cross-sectional Study

Hari S. Iyer^{1§}, Peter James², Linda Valeri³, Francis Bajunirwe⁴, Joan Nankya-Mutyoba⁵, Marina Njelekela^{6,7,8}, Faraja Chiwanga⁷, Vikash Sewram⁹, IkeOluwapo Ajayi¹⁰, Clement Adebamowo^{11,12}, Shona Dalal¹, Todd G. Reid^{13,14}, Timothy R. Rebbeck^{1,15}, Hans Olov-Adami^{1,16,17}, Michelle D. Holmes^{1,18}

Affiliations:

¹Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, United States of America

²Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, United States of America

³Department of Biostatistics, Columbia University Mailman School of Public Health, New York, United States of America

⁴Department of Community Health, Mbarara University of Science and Technology, Mbarara, Uganda

⁵Department of Epidemiology & Biostatistics, School of Public Health, Makerere University College of Health Sciences, Kampala, Uganda

⁶Department of Physiology, Muhimbili University of Health and Allied Sciences, Dar es Salaam, Tanzania

⁷Department of Internal Medicine, Muhimbili National Hospital, Dar es Salaam, Tanzania

⁸Deloitte Consulting Limited, Dar es Salaam, Tanzania

⁹African Cancer Institute, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa

¹⁰Department of Epidemiology and Medical Statistics, Faculty of Public Health, College of Medicine, University of Ibadan, Ibadan, Nigeria

¹¹Institute of Human Virology, Abuja, Nigeria

¹²Greenebaum Comprehensive Cancer Center and Institute of Human Virology, University of Maryland School of Medicine, Baltimore, United States of America

¹³Media Lab and Connection Science, Massachusetts Institute of Technology, Boston, United States of America

¹⁴Department of Social and Behavioral Sciences, Harvard T. H. Chan School of Public Health, Boston, United States of America

¹⁵Department of Medical Oncology, Dana Farber Cancer Institute, Boston, United States of America

¹⁶Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

¹⁷Clinical Effectiveness Research Group, Institute of Health and Society, University of Oslo, Oslo, Norway

¹⁸Channing Division of Network Medicine, Brigham & Women's Hospital, Harvard Medical School, Boston, United States of America.

ABSTRACT

Background: Population growth, demographic transitions and urbanization in sub-Saharan Africa (SSA) will increase non-communicable disease (NCD) burden. Studies in North America, Europe and Asia suggest that increasing neighborhood greenness, or natural green vegetation, reduces NCDs. Proposed mechanisms include reduced inflammation and stress, causing improved cardiometabolic profiles. However, few studies have evaluated these associations in SSA.

Objectives: We studied the association between neighborhood greenness and NCDs in a multi-country cross-sectional study.

Methods: In 2011, study participants from Uganda, South Africa, and Tanzania provided data on NCD risk factors and socioeconomic status through surveys and biospecimen collection. Work or residential locations were geocoded using Google Maps and verified by site-based teams. Neighborhood greenness exposure was estimated using satellite-derived Normalized Difference Vegetation Index (NDVI) with 250m resolution. Outcomes based on study measurements and self-reported prior history of disease or treatment included: body mass index, diabetes, hypertension, cholesterol, and allostatic load, a composite outcome representing physiologic adaptation to stress. Odds ratios (OR) were modeled using sequentially adjusted multiple linear, logistic, and multinomial regression models. We used multiple imputation with 10 samples for missing data, and assigned Bonferroni correction to p-values from fully adjusted models.

Results: Among 1178 participants, in adjusted models, a 0.11 unit NDVI increase was associated with lower BMI (β : -1.01, 95% CI: -1.35, -0.67), and lower odds of overweight/obesity (aOR: 0.73, 95% CI: 0.62, 0.85), diabetes (aOR: 0.77, 95% CI: 0.62, 0.96)),

and having ≥ 3 allostatic load components compared to none (aOR: 0.66, 95% CI: 0.52, 0.85).

Except for diabetes, these remained statistically significant after Bonferroni correction. We observed no association between NDVI and hypertension or cholesterol.

Discussion: Our findings are consistent with health benefits of neighborhood greenness reported in other countries, suggesting greening strategies could be considered as part of broader public health interventions for NCDs.

INTRODUCTION

Sub-Saharan Africa's (SSA) population is projected to increase from 1.05 billion in 2018 to 2.2 billion in 2050, resulting in a more urban, older population (1, 2). Non-communicable disease (NCD) burden is expected to continue rising alongside these demographic changes (3-6). The Global Burden of Disease Study estimated a 67% increase in disability-adjusted life-years attributable to NCDs in SSA from 1990 to 2017, with cancer (79.5%) and diabetes (83.1%) burden presenting the largest relative increases (7). National health ministries in SSA are developing strategies to control NCDs by expanding medical services and promoting lifestyle changes (8-13). Increased investments in health care delivery are necessary to limit the spread of NCDs in SSA countries (14). However, experience from other parts of the world suggests that in order to be effective, public health strategies for NCD control must consider socioeconomic and contextual factors, such as income and education, catastrophic health expenditures, availability of spaces for active transport and healthy diets, and incentives for promoting healthy lifestyles (15). Analytic models that incorporate relationships between risk factors and underlying social and environmental context are essential to accurately inform NCD prevention policy (16, 17).

Neighborhood greenness, or the natural green vegetation in a given neighborhood area, is increasingly understood as an important health promoting contextual environmental factor. African scientists have called for more parks and street trees in cities, listing benefits including reductions of adverse health impacts of temperature and air pollution, as well as space for active transport and exercise (18-21). Cohort studies in western countries have reported multiple health benefits of green spaces, including promotion of physical activity, lower risk of obesity and diabetes, improved mental health, enhanced social capital, and reduced all-cause and cardiovascular mortality (22-24). Researchers have estimated effects of nature visits on

physiologic stress response, and reported reduced inflammation, enhanced glucose and cardiometabolic profile, and reduced sympathetic nervous system activation (25, 26). Cross-sectional studies in the United States reported associations between neighborhood greenness and allostatic load, an index made up of biomarkers capturing physiologic response to cumulative biologic stress (27), and sympathetic stress response (28, 29). While studies in SSA have assessed the association between neighborhood greenness and mental health (30), little evidence exists regarding relationships with other NCDs.

We conducted a cross-sectional study using data from a multi-country sample of urban and rural community members in SSA to evaluate the association between neighborhood greenness and cardiometabolic NCDs. We hypothesized that neighborhood greenness would be associated with lower NCD prevalence, lower prevalence of lifestyle risk factors, and lower cardiometabolic allostatic load.

METHODS

Study setting and population

We used participant data from the Africa/Harvard Partnership for Cohort Research and Training (PaCT), launched in 2011 as a collaboration between the Harvard T. H. Chan School of Public Health (USA), Makerere University (Uganda), Mbarara University (Uganda), Muhimbili University (Tanzania), University of Ibadan (Nigeria), and Stellenbosch University (South Africa). For this study, we included participants from sites with available location data (Uganda, South Africa, Tanzania). In Uganda, participants were recruited from villages in peri-urban (Kampala) and rural (Mbarara) settings. In Tanzania and South Africa, occupational samples of teachers were recruited from cities (Dar Es Salaam, Tanzania and Cape Town, South Africa).

Standardized, culturally-adapted questionnaires were administered to capture information on demographics, socioeconomic status, diet and lifestyle factors, medical history, and care seeking behaviors. Anthropometric data and blood measurements were taken by trained study staff. Participants provided names of the village where they lived (Uganda sites) or school where they worked (Tanzania, South Africa sites). Full details regarding study procedures are described elsewhere (31).

One of the investigators (HSI) collected geocodes for locations by entering names of villages and schools into Google Maps and extracting coordinates at the center of the village or school. Site investigators (FB, JM, MN, VS) verified the accuracy of these coordinates and resolved discrepancies (10/81, 12%). Of 1215 participants, 3% (n=37) were excluded due to missing geocodes resulting in a final sample of 1178.

Exposure

We estimated exposure to neighborhood greenness at participant locations using the Normalized Difference Vegetation Index (NDVI), a satellite-derived measure of natural green vegetation (32). NDVI is calculated by taking the difference of near infrared light (reflected by leaves) and visible red light (absorbed), and dividing by the sum of these measures. Values range from -1 to +1, with values below 0 reflecting bodies of water. NDVI was obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) with resolution of 250 meters. We extracted the NDVI value for the area that included the participant's school (Tanzania, South Africa) or village (Uganda) geocodes, and collected four seasonal (January, April, July, September) images each year from 2010 to 2011. Google Earth Engine was used to extract images and select the least cloudy image in each month (33). In order to capture recent exposure and account for seasonal differences, we averaged over the eight seasonal measures at each

location. Sensitivity analyses using zonal statistics to assign mean NDVI to village locations defined by polygons did not change conclusions, so we used a point-based approach for consistency across villages and schools.

Outcomes

We studied the association between neighborhood greenness and multiple NCD endpoints including specific diseases (diabetes, hypertension), lifestyle risk factors (obesity), and cardiometabolic profile (allostatic load, total cholesterol). Whenever possible, we relied on measurements taken by trained study staff. However, because access to NCD services varied across sites and prior disease status could affect measurements taken during the study, we also included information about prior history of diagnosis and treatment to define outcomes.

Diabetes was reported on questionnaires by participants and, when available, estimated based on blood glucose measures taken during the study. Participants were classified as diabetic if they reported history of diabetes (62/139, 45%), or if they reported a fasting glucose level ≥ 7 mmol/l (77/139, 55%) (34). Hypertension was reported on questionnaires, and, when available, confirmed by blood pressure measurements. We used WHO classifications for hypertension as follows: Hypertensive (SBP: 140 mm Hg and/or DBP: 90 mm Hg based on blood pressure measurement (297/448, 66%), self-reported being on regular anti-hypertensive therapy (94/448, 21%), or self-reported history of hypertension (57/448, 13%)) (35). We modeled obesity using the body mass index (BMI) as a continuous scale, and using WHO classification (under-weight: BMI < 18.5 kg/m²; normal weight: BMI in 18.5-24.9 kg/m²; overweight: BMI in 25-29.9 kg/m²; obese: 30 kg/m² and over). We also dichotomized BMI categories into overweight/obese vs normal/under-weight. Total cholesterol was ascertained by trained nurses who visited schools to collect blood samples at pre-specified times, while in villages, trained staff conducted blood

draws in the community. Cholesterol was modeled as a continuous variable, and dichotomized, defining high cholesterol as any value above 6.22 mmol/L based on clinical guidelines used in 2011 (36).

We developed a minimal allostatic load score to assess the association between NDVI and chronic stress in this population focusing on metabolic measures available in our data, adapting the full allostatic load measure developed by Seeman et al. which incorporated ten components reflecting multiple homeostatic regulatory systems (37). Our score was comprised of the following subset of four variables: systolic blood pressure, diastolic blood pressure, natural log of blood glucose, and body mass index (as a proxy for waist-to-hip ratio). We used this abridged score because it was consistent with cardiovascular contributions to allostatic load used elsewhere, and because other individual components were missing for entire sites (27, 37). To calculate the allostatic load score, we first created quartiles for each of the components, and then assigned a score of 1 if the component was in the highest quartile and 0 otherwise (37). Allostatic load was then calculated using the sum of component scores, ranging from 0 to 4, with four being the least healthy. Individuals with three or more components were grouped together due to few participants with four allostatic load components (n=17), yielding a four-level categorical variable (0, 1, 2 or ≥ 3 components).

Ethics

Written informed consent from each participant was provided through a signed form along with mailed completed questionnaire (South Africa, Tanzania) or during enrollment interviews with study staff (Uganda). This study was approved by the Harvard T. H. Chan School of Public Health Institutional Review Board; the Health Research Ethics Committee of the Faculty of Health Sciences, Stellenbosch University; Makerere University School of Public

Health Higher Degrees Research and Ethics Committee; National Institute for Medical Research, Tanzania; Mbarara University of Science and Technology Research Ethics Committee; and the Uganda National Council of Science and Technology.

Statistical Analysis

Since 369 participants (31%) were missing values for at least one outcome or covariate, we used multiple imputation with chained equations to fit models for missingness, assuming data were missing at random (Rubin 1987). To multiply impute missing continuous covariates (age, BMI, systolic blood pressure, diastolic blood pressure, blood glucose) we used linear regression, and for missing categorical and binary covariates (private health care seeking, piped water source, overweight/obesity, allostatic load category) we used the discriminant function. We used the following covariates as predictors of missingness: age, sex, marital status, educational attainment, cooking fuel type, NDVI, smoking status, and site. For models estimating the association between NDVI and cholesterol, we excluded the peri-urban Uganda site because no cholesterol measurements were taken. Using these models, we imputed values for covariates with missing data to generate 10 datasets.

Next, we fit multiple linear regression models for continuous outcomes (BMI, cholesterol) and logistic regression models for binary outcomes (obese/overweight, diabetes, hypertension, high cholesterol). Models for categorical allostatic load were fit using multiple multinomial logistic regression. NDVI was modeled as a continuous exposure, and effect estimates correspond to a 0.11 unit increase in NDVI (0.58 standard deviation) to facilitate comparisons with previous studies. We tested for linearity between NDVI and each outcome using quadratic terms. Models were sequentially adjusted for (1) age at interview (continuous), sex; (2) educational attainment (primary school or lower vs other), fuel source (electric or gas vs

other), marital status (married/cohabiting vs other), smoking status (current smoker or not), site (Uganda/rural, Uganda/peri-urban, Tanzania/urban, South Africa/urban), care sought in private sector (binary). Covariates were selected based on previous studies by PaCT investigators (34, 35, 38, 39) and literature on common causes of NDVI and NCDs (22-24). Since these models were fit in imputed datasets, standard errors were estimated using Rubin's Rules (40). We provided results from a complete case analysis as a sensitivity analysis. As an additional sensitivity analysis, we restricted the analysis to urban sites only (excluding the rural Uganda site).

In addition, we explored the role of BMI as a mediator of the association between NDVI and NCD outcomes. Following methods described by Valeri and VanderWeele (41), we fit regression models adjusting for the full set of covariates above, and tested for interaction between NDVI and BMI at the $\alpha=0.05$ level. If interactions were present, we presented estimates for the NDVI-outcome association at three levels of BMI (20, 25, 30). For binary outcomes, we replaced the logistic model with a log-linear model with Poisson distribution to model prevalence ratios which can be directly compared in models with and without the mediator, a property not shared by odds ratios (41).

Since we studied the association between NDVI and multiple outcomes, we used VanderWeele's Outcome-wide Epidemiology approach (42). This approach facilitates comparisons between multiple exposure-outcome associations simultaneously. We report E-values as a sensitivity analysis to quantify the minimum unmeasured confounding bias required to explain away our findings (43). The E-value for a point estimate corresponds to the minimum bias, conditional on covariates, necessary to attenuate the point estimate to the null value of 1. Confidence interval E-value quantifies the bias required to shift the confidence interval to

include 1. To limit type 1 error, we applied a Bonferroni correction to all significance tests for association between NDVI and NCD outcomes in fully adjusted outcome models performed (7 tests in total population, 5 tests in urban sites only, for a total of 12 tests). This correction resulted in an alpha type 1 error cutoff of 0.0042. All analyses were done using SAS version 9.4.

RESULTS

After exclusions, 1178 participants were included in the study, with 275 (23%) from Peri-urban Uganda, 200 (17%) from rural Uganda, 477 (40%) from South Africa, and 226 (19%) from Tanzania. The NDVI distribution varied by site, with Tanzania and South Africa exhibiting lower levels of NDVI than the Uganda sites (Table 3.1). Higher levels of NDVI were associated with lower median age (NDVI Q5: 36, NDVI Q1: 46.7), lower proportion reporting female gender (NDVI Q5: 51%, NDVI Q1: 81%), and lower educational attainment (NDVI Q5: 82% with primary school or lower, NDVI Q1: 3%). We observed high reported prevalence of overweight (31%) and obesity (33%) in the total study population. Prevalence of diabetes was 12%, and prevalence of hypertension was 38%. In our analysis of NDVI and cholesterol excluding the peri-urban Uganda site, the prevalence of high cholesterol was 14%.

Table 3.1. Descriptive Characteristics of Study Participants from four Sites in Three sub-Saharan African Countries

No.	Normalized Difference Vegetation Index (NDVI)					
	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Total
	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
	233	229	245	240	231	1178
Age^a	46.7 [40.7, 52.3]	44.2 [37.1, 51.4]	43.4 [34.29, 51.14]	33.2 [26, 47.42]	36.1 [27.2, 42.94]	41.9 [32.6, 50.32]
Female	189 (81.1)	178 (77.7)	160 (65.3)	125 (52.1)	117 (50.7)	769 (65.3)
BMI^a	30.1 [26.8, 34.8]	29.9 [25.7, 33.2]	27.6 [23.7, 32.3]	24.1 [21.7, 28.6]	24.2 [21.8, 27.3]	27.1 [23.3, 31.4]
MET-Hours/week^a	38 [10.7, 65]	52 [24, 81.7]	45.2 [16, 76]	56 [22, 100]	60 [52, 72]	52 [24, 76]
Ever Smoke	27 (11.6)	20 (8.7)	26 (10.6)	28 (11.7)	28 (12.1)	129 (11)
Systolic Blood Pressure (mm/Hg)^a	133.8 [124, 145.5]	130 [120, 145]	124.3 [116, 137]	123.5 [114, 133.5]	118.8 [109.5, 125]	125 [115.5, 138.5]
Diastolic Blood Pressure (mm/Hg)^a	79.3 [73, 87.5]	80 [74.5, 90]	80 [70, 87]	75.5 [68.5, 82.5]	77.3 [70.5, 82.5]	79 [70.5, 85.5]
Blood glucose^a	5.2 [4.6, 5.8]	4.8 [4.4, 5.5]	5.2 [4.5, 5.8]	5.1 [4.5, 5.7]	5.7 [5.1, 6.4]	5.2 [4.6, 5.8]
NDVI (two-year seasonal average)^a	0.23 [0.23, 0.24]	0.31 [0.29, 0.33]	0.40 [0.37, 0.41]	0.61 [0.54, 0.67]	0.73 [0.72, 0.75]	0.40 [0.29, 0.67]
Site						
Kampala, Uganda (peri-urban)	.	.	89 (36.3)	169 (70.4)	17 (7.4)	275 (23.3)
Mbarara, Uganda (rural)	200 (86.6)	200 (17)
Cape Town, South Africa (urban)	195 (83.7)	119 (52)	93 (38)	70 (29.2)	.	477 (40.5)
Dar Es Salaam, Tanzania (urban)	38 (16.3)	110 (48)	63 (25.7)	1 (0.4)	14 (6.1)	226 (19.2)
Primary School Education or Lower	6 (2.6)	14 (6.1)	38 (15.5)	63 (26.3)	190 (82.3)	311 (26.4)
Currently Married/Cohabiting	160 (68.7)	159 (69.4)	163 (66.5)	162 (67.5)	190 (82.3)	834 (70.8)
Fuel (not electricity, natural gas, biogas)	39 (16.7)	111 (48.5)	149 (60.8)	165 (68.8)	231 (100)	695 (59)
Piped drinking water	193 (82.8)	136 (59.4)	122 (49.8)	136 (56.7)	5 (2.2)	592 (50.3)
Seek health care at private facility						
Yes	181 (77.7)	195 (85.2)	143 (58.4)	117 (48.8)	77 (33.3)	713 (60.5)
No	52 (22.3)	34 (14.9)	98 (40)	122 (50.8)	153 (66.2)	459 (39.0)
(Missing)	.	.	4 (1.6)	1 (0.4)	1 (0.4)	6 (0.5)
Have you ever been told by a doctor or other health worker that you have DM?						
Yes	24 (10.3)	23 (10.0)	21 (8.6)	6 (2.5)	2 (0.9)	76 (6.5)
No	162 (69.5)	171 (74.7)	198 (80.8)	225 (93.8)	228 (98.7)	984 (83.5)
(Missing)	47 (20.2)	35 (15.3)	26 (10.6)	9 (3.8)	1 (0.4)	118 (10)

Table 3.1. Descriptive Characteristics of Study Participants from four Sites in Three sub-Saharan African Countries (continued)

	Normalized Difference Vegetation Index (NDVI)					Total N (%)
	Quintile 1 N (%)	Quintile 2 N (%)	Quintile 3 N (%)	Quintile 4 N (%)	Quintile 5 N (%)	
Family member with diabetes						
Yes	105 (45.1)	87 (38.0)	85 (34.7)	63 (26.3)	28 (12.1)	368 (31.2)
No	111 (47.6)	131 (57.2)	152 (62.0)	174 (72.5)	202 (87.5)	770 (65.4)
(Missing)	17 (7.3)	11 (4.8)	8 (3.3)	3 (1.3)	1 (0.4)	40 (3.4)
Have you ever been told by a doctor or other health worker that you have HPT?						
Yes	75 (32.2)	78 (34.1)	57 (23.3)	23 (9.6)	14 (6.1)	247 (21.0)
No	145 (62.2)	140 (61.1)	174 (71.0)	209 (87.1)	217 (93.9)	885 (75.1)
(Missing)	13 (5.6)	11 (4.8)	14 (5.7)	8 (3.3)	.	46 (3.9)
Family member with hypertension						
Yes	137 (58.8)	118 (51.5)	122 (49.8)	118 (49.2)	66 (28.6)	561 (47.6)
No	84 (36.1)	99 (43.2)	116 (47.4)	120 (50.0)	164 (71.0)	583 (49.5)
(Missing)	12 (5.2)	12 (5.2)	7 (2.9)	2 (0.8)	1 (0.4)	34 (2.9)

^aMedian [Interquartile Range]

Results of our outcome-wide regression analysis using multiple imputation are presented in Table 3.2. In fully adjusted models, a 0.11 unit increase in NDVI was associated with a 1.01 unit decrease in BMI (b: -1.01, 95% CI: -1.35, -0.67). Similarly, in fully adjusted models, we observed 27% lower odds of overweight/obesity associated with a 0.11 unit increase in NDVI (aOR: 0.73, 95% CI: 0.62, 0.85) in the total study population. Both were statistically significant after applying the Bonferroni correction. In fully adjusted models, a 0.11 unit increase in NDVI was associated with a 23% lower odds of prevalence of diabetes in the total study population (aOR: 0.77, 95% CI: 0.62, 0.96), but this association was not statistically significant after correcting for multiple testing. We did not observe statistically significant associations between NDVI and prevalence of hypertension, or between NDVI and cholesterol (continuous or dichotomous outcome) after adjustment for covariates. Results from fully adjusted multinomial regression models for allostatic load showed that a 0.11 unit increase NDVI was associated with 34% lower odds of having 3 or more allostatic load components (worst) compared to none (best) (aOR: 0.66, 95% CI: 0.52, 0.85), and 33% lower odds of having 2 allostatic load components compared to none (aOR: 0.67, 95% CI: 0.52, 0.85). This association was statistically significant after correction for multiple testing. Magnitude and direction of reported associations were similar when restricting to urban populations only.

Table 3.2. Cross-sectional associations between Normalized Difference Vegetation Index (0.11 unit increase) and components of allostatic load using Multiple Imputation^a among Participants from four Sites in Three sub-Saharan African Countries

Outcome	Total Population (N=1178)				Urban Population (N=978)			
	β	OR	(95% CI)	<i>p</i>	β	OR	(95% CI)	<i>p</i>
BMI (continuous)								
Unadjusted	-1.65		(-1.84, -1.45)	<.0001	-1.87		(-2.17, -1.58)	<.0001
Age + Sex	-1.38		(-1.58, -1.17)	<.0001	-1.52		(-1.84, -1.21)	<.0001
Full Adjustment ^b	-1.01		(-1.35, -0.67)	<.0001	-0.97		(-1.37, -0.57)	<.0001
Overweight/Obese								
Unadjusted		0.57	(0.53, 0.62)	<.0001		0.52	(0.46, 0.58)	<.0001
Age + Sex		0.64	(0.58, 0.70)	<.0001		0.60	(0.53, 0.68)	<.0001
Full Adjustment ^b		0.73	(0.62, 0.85)	<.0001		0.73	(0.63, 0.85)	<.0001
Diabetes								
Unadjusted		0.87	(0.79, 0.97)	0.0153		0.74	(0.63, 0.87)	0.0003
Age + Sex		0.93	(0.83, 1.05)	0.2381		0.76	(0.64, 0.91)	0.0032
Full Adjustment ^b		0.77	(0.62, 0.96)	0.0193		0.77	(0.62, 0.95)	0.0167
Hypertension								
Unadjusted		0.68	(0.63, 0.74)	<.0001		0.75	(0.68, 0.82)	<.0001
Age + Sex		0.76	(0.70, 0.82)	<.0001		0.84	(0.75, 0.94)	0.0017
Full Adjustment ^b		0.92	(0.80, 1.05)	0.1996		0.91	(0.79, 1.04)	0.1531
Cholesterol (continuous)^c								
Unadjusted	-0.24		(-0.30, -0.18)	<.0001				
Age + Sex	-0.17		(-0.23, -0.11)	<.0001				
Full Adjustment ^b	0.04		(-0.07, 0.14)	0.4875				
Cholesterol (binary)^c								
Unadjusted		0.99	(0.90, 1.11)	0.9252				
Age + Sex		1.07	(0.95, 1.20)	0.2626				
Full Adjustment ^b		1.15	(0.93, 1.42)	0.2093				

Table 3.2. Cross-sectional associations between Normalized Difference Vegetation Index (0.11 unit increase) and components of allostatic load using Multiple Imputation^a among Participants from four Sites in Three sub-Saharan African Countries (continued)

Outcome	Total Population (N=1178)				Urban Population (N=978)			
	β	OR	(95% CI)	<i>p</i>	β	OR	(95% CI)	<i>p</i>
Allostatic Load								
Unadjusted								
0 (Low)								
1	0.86	(0.79, 0.94)	0.0016		0.80	(0.71, 0.91)	0.0005	
2	0.66	(0.58, 0.75)	<.0001		0.60	(0.50, 0.70)	<.0001	
≥ 3 (High)	0.61	(0.52, 0.70)	<.0001		0.58	(0.48, 0.69)	<.0001	
Age + Sex								
0 (Low)		Ref						
1	0.94	(0.86, 1.04)	0.1376		0.90	(0.78, 1.03)	0.1376	
2	0.74	(0.65, 0.84)	<.0001		0.66	(0.55, 0.79)	<.0001	
≥ 3 (High)	0.69	(0.59, 0.81)	<.0001		0.66	(0.54, 0.8)	<.0001	
Full Adjustment ^b								
0 (Low)		Ref						
1	0.87	(0.74, 1.02)	0.1702		0.88	(0.74, 1.05)	0.1702	
2	0.67	(0.52, 0.85)	0.0004		0.65	(0.51, 0.82)	0.0004	
≥ 3 (High)	0.66	(0.52, 0.85)	0.0014		0.67	(0.52, 0.85)	0.0014	

^aImputation models were fit using multiple chained equations to impute missing data for body mass index and the following covariates: age, sex, educational attainment, fuel source, marital status, NDVI, smoking status, site.

^bMultiple regression models were fit for categorical (multinomial logistic), binary (logistic) or continuous (linear) outcome variables, using the following covariates: age, sex, educational attainment, fuel source, marital status, smoking status, site, seeking health care in private sector.

^cSince cholesterol was missing for one site (peri-urban Uganda), these models were fit only on remaining sites (Tanzania, South Africa, and Rural Uganda) with a sample size of 903.

Bold: $p < 0.05$ after Bonferroni correction applied to final fully adjusted outcome models (the p -value cutoff for Bonferroni correction = $0.05/12$ outcomes = 0.0042).

Table 3.3 presents results from our exploratory mediation analysis. Further adjustment for BMI attenuated the association between NDVI and diabetes prevalence towards the null (aPR: 0.87, 95% CI: 0.71, 1.06). There was a significant interaction between NDVI and BMI ($P_{het} = 0.019$), so we estimated prevalence ratios at different levels of BMI. Results suggest that a joint hypothetical intervention to increase NDVI while fixing BMI of all participants to 20 would be associated with a 13% reduced prevalence of hypertension (aPR: 0.87, 95% CI: 0.77, 0.98). However, fixing BMI at higher levels would attenuate the association between NDVI and hypertension prevalence.

Table 3.3. Cross-sectional associations between Normalized Difference Vegetation Index (0.11 unit increase) and Non-Communicable Disease Prevalence Using Log-Linear Models using Multiple Imputation^a to Evaluate Mediation by Body Mass Index among Participants from four Sites in Three sub-Saharan African Countries

Outcome	Total Population (N=1178)				Urban Population (N=978)			
	β	PR	(95% CI)	<i>p</i>	β	PR	(95% CI)	<i>p</i>
Diabetes								
Full Adjustment		0.80	(0.65, 0.98)	0.030	0.80	(0.65, 0.97)		0.026
Full + BMI ^b		0.87	(0.71, 1.06)	0.17	0.86	(0.71, 1.06)		0.16
Hypertension								
Full Adjustment		0.95	(0.86, 1.05)	0.32	0.95	(0.86, 1.04)		0.27
Full + BMI ^b		0.98	(0.89, 1.09)	0.75	0.98	(0.88, 1.08)		0.64
BMI: 20 ^c		0.87	(0.77, 0.98)	0.025	0.87	(0.77, 0.99)		0.029
BMI: 25 ^c		0.93	(0.85, 1.01)	0.10	0.93	(0.85, 1.01)		0.088
BMI: 30 ^c		0.99	(0.92, 1.07)	0.83	0.99	(0.92, 1.06)		0.71
Cholesterol (continuous)								
Full Adjustment	0.04		(-0.07, 0.14)	0.49				
Full + BMI ^b	0.02		(-0.08, 0.13)	0.66				
Cholesterol (binary)								
Full Adjustment		1.11	(0.92, 1.34)	0.28				
Full + BMI ^b		1.09	(0.90, 1.32)	0.37				

Abbreviations: Normalized Difference Vegetation Index (NDVI), Body Mass Index (BMI), PR (prevalence ratio)

^aImputation models were fit using multiple chained equations to impute missing data for BMI and the following covariates: age, sex, educational attainment, fuel source, marital status, NDVI, smoking status, site.

^bMultiple log-linear regression models were fit for the outcome, using the following covariates: age, sex, educational attainment, fuel source, marital status, smoking status, site, seeking health care in private sector, BMI

^cMultiple log-linear regression models were fit for prevalence of hypertension, using the following covariates: age, sex, educational attainment, fuel source, marital status, smoking status, site, seeking health care in private sector, interaction between BMI and NDVI ($P_{het} = 0.019$ for Total Population, $P_{het} = 0.042$ for Urban Population).

In complete case analysis, associations between NDVI and NCD outcomes remained largely unchanged (Table S3.1), but there were a few differences. A 0.11 unit increase in NDVI was associated with 20% lower odds of hypertension (aOR: 0.80, 95% CI: 0.68, 0.93), remaining statistically significant after multiple testing in urban sites only. Associations between NDVI and allostatic load in complete case analysis were of similar magnitude but less precise, and did not meet statistical significance after applying Bonferroni correction.

E-values for point estimates and 95% confidence intervals are reported in Table S3.2. Estimates for BMI, overweight/obesity, and allostatic load would require moderate amounts of unmeasured confounding bias to attenuate point estimates to the null or shift confidence intervals to contain 1 (point estimates range: 1.59-1.76, confidence intervals range: 1.39-1.44). An omitted variable would have to be associated with a 59%-76% increased relative risk or prevalence comparing exposed to unexposed, conditional on covariates, to attenuate point estimates to the null. E-values for estimates and confidence intervals did not change when restricting only to urban sites.

DISCUSSION

In this multi-country sample of urban and rural study participants in SSA, we observed statistically significant inverse associations between NDVI and a number of NCDs and risk factors. Participants in locations with higher NDVI had lower BMI and lower allostatic load based on cardiovascular components. Though we observed an inverse association between NDVI and prevalence of diabetes, this finding was not statistically significant after accounting for multiple testing. The strength of associations between NDVI and NCDs were attenuated

following adjustment for BMI, suggesting that BMI could play an explanatory role, though this cannot be confirmed in our cross-sectional study. Restricting analyses to urban sites alone did not result in major changes to our inference.

Our results are largely in agreement from studies of neighborhood greenness, BMI and diabetes conducted in other parts of the world. Following results from a large cross-sectional study in Australia (44), several authors have reported an association between NDVI and lower odds of diabetes. A recent meta-analysis estimated 28% lower odds of type II diabetes associated with high compared to low levels of greenness (aOR: 0.72, 95% CI: 0.61, 0.85) (24). In addition, others have previously reported inverse associations between NDVI and BMI (22, 45). Studies in Europe and China have reported inverse associations between NDVI and glucose levels in children and adults, consistent with our findings (46-48).

We did not observe statistically significant associations between NDVI and hypertension or cholesterol. While studies from Western countries largely find that increased levels of neighborhood greenness are associated with lower cardiovascular disease, evidence of the relationship between neighborhood greenness, blood pressure and hypertension is mixed (24, 49). A large study in Korea reported lower levels of lipidemia associated with increasing proximity to parks, while a study among 10-15 year old children in Lithuania using NDVI as greenness exposure found no association (50, 51). Another cross-sectional study in China reported associations between neighborhood greenness and lower hypertension prevalence, but did not control for income (52). While we did not observe an association between NDVI and total cholesterol, others have reported inverse associations between NDVI and high-density lipoprotein cholesterol (30). Future studies employing prospective designs will be needed to clarify the relationship between NDVI, hypertension, and cholesterol in SSA.

We were unable to specifically evaluate the role of physical activity as a mediator of the estimated effect between neighborhood greenness and NCDs in this study due to differences in how types and frequency of physical activity were assessed at each site. However, rural populations in SSA experience much higher levels of physical activity from occupation and travel compared to urban populations (53, 54). Thus, the fact that associations between NDVI and NCDs in our study did not change when restricting to urban populations only suggests that differences in unmeasured physical activity between sites cannot completely explain our findings. Cross-sectional studies in Europe and New Zealand attempting to evaluate the role of physical activity as a mediator of the effect of neighborhood greenness on cardiovascular disease and mortality generally have found little explanatory power, possibly due to difficulties in accurately measuring physical activity (55, 56). A Chinese study reported strong inverse associations between neighborhood greenness and physical activity, but did not control for income, an important confounder (52). Longitudinal designs incorporating precise physical activity measurements could improve assessment of this pathway.

Limitations of our study include its cross-sectional design which hampers causal inference. Our study further relied on discrete, spatially diffuse locations, assuming that these locations (villages and schools) would be sufficient to capture relevant exposure to neighborhood greenness. If exposure to neighborhood greenness occurs outside these locations, we could introduce non-differential exposure measurement error. For diabetes and hypertension, outcomes were assessed in part using self-report, which could have led to outcome misclassification. Higher NCD detection in urban compared to rural areas is possible, and since urban areas had lower NDVI, failure to account for this relationship could have resulted in residual bias. Control for factors associated with urbanicity, NDVI and NCDs, and our sensitivity analysis restricting to

urban sites should mitigate this bias. While NDVI is an objective measure of neighborhood greenness, it does not capture specific types of vegetation that could drive the associations, which would be needed to develop interventions. Study participants represent a unique mix of urban vs rural, as well as occupational- vs population-based samples, which limits the overall generalizability of these findings because neighborhood greenness exposure, behavioral risk factors, and NCD risk could vary in other African populations. Future studies should explore more detailed residential histories and assess movement patterns at the individual- rather than area-level.

CONCLUSION

We found that higher levels of neighborhood greenness are associated with lower BMI and lower cardiometabolic allostatic load in a multi-country study of SSA participants, supporting evidence from other countries. Studies with prospective follow-up and more detailed measurement of contextual environmental exposure are needed to strengthen evidence for these associations. Given growing interest in green infrastructure and urban planning for public health in SSA, confirmation could help inform evidence-based urban public health interventions to control NCD burden in SSA.

REFERENCES

1. UN Economic & Social Affairs. World Urbanization Prospects: the 2018 Revision. Available at: <https://esa.un.org/unpd/wup/Publications/Files/WUP2018-KeyFacts.pdf>. (accessed June 26, 2019).
2. Population Reference Bureau. 2018 World Population Data Sheet. Available at: https://www.prb.org/wp-content/uploads/2018/08/2018_WPDS.pdf. (accessed June 26, 2019).
3. Dalal S, Beunza JJ, Volmink J, Adebamowo C, Bajunirwe F, Njelekela M et al. Non-communicable diseases in sub-Saharan Africa: what we know now. *Int J Epidemiol*. 2011;40(4):885-901.
4. NCD Risk Factor Collaboration (NCD-RisC) – Africa Working Group. Trends in obesity and diabetes across Africa from 1980 to 2014: an analysis of pooled population-based studies. *Int J Epidemiol*. 2017;46(5):1421-1432.
5. Jemal A, Bray F, Forman D, O'Brien M, Ferlay J, Center M et al. Cancer burden in Africa and opportunities for prevention. *Cancer*. 2012;118(18):4372-84.
6. Mensah GA, Sampson UK, Roth GA, Forouzanfar MH, Naghavi M, Murray CL et al. Mortality from cardiovascular diseases in sub-Saharan Africa, 1990-2013: a systematic analysis of data from the Global Burden of Disease Study 2013. *Cardiovasc J Afr*. 2015;26(2 H3Africa Suppl): S6-S10.
7. Gouda HN, Charlson F, Sorsdahl K, Ahmadzada S, Ferrari AJ, Kerskine H et al. Burden of non-communicable diseases in sub-Saharan Africa, 1990-2017: results from the Global Burden of Disease Study 2017. *Lancet Glob Health*. 2019;7(10):e1375-e1387.

8. Farmer P, Basilio M, Kerry V, et al. Global Health Priorities for the Early Twenty-First Century. In: Farmer PE, Kim JY, Basilio M, Kleinman A, eds. Reimagining Global Health. Berkeley and Los Angeles, California: *University of California Press*, 2013;320-329.
9. Tapela NM, Tshisimogo G, Shatera BP, et al. Integrating noncommunicable disease services into primary care, Botswana. *Bull World Health Organ*. 2019;97(2):142-153.
10. Binagwaho A, Muhimpundu MA, Bukhman G. 80 under 40 by 2020: an equity agenda for NCDs and injuries. *Lancet*. 2014;383(9911):3-4.
11. Kwan GF, Bongani MM, Mocumbi AO, et al. Endemic Cardiovascular Diseases of the Poorest Billion. *Circulation*. 2016;133(24):2561-75.
12. WHO. Time to deliver: report of the WHO Independent High-level Commission on Noncommunicable Diseases. Geneva: World Health Organization; 2018. Licence: CC BY-NC-SA 3.0 IGO.
13. Stefan DC, Elzawawy AM, Khaled HM, Ntaganda F, Asiimwe A, Addai BW et al. Developing cancer control plans in Africa: examples from five countries. *Lancet Oncol*. 2013;14(4):e189-95.
14. Alwan A, Evans D, Edejer TT, et al. Scaling up action against noncommunicable diseases: How much will it cost? Geneva. *World Health Organization*. 2011.
15. Nugent R, Bertram MY, Jan S et al. Investing in non-communicable disease prevention and management to advance the Sustainable Development Goals. *Lancet*. 2018;391(10134):2029-2035.

16. Krieger N. ed. *Epidemiology and the People's health: Theory and Context*. New York: Oxford University Press; 2011. Chapter 7: Ecosocial Theory of Disease Distribution: Embodying Societal and Ecologic Context. pp.202-235
17. Frumkin H, Haines A. Global Environmental Change and Noncommunicable Disease Risks. *Annu. Rev. Public Health*. 2019;40:261-82.
18. Ezeh A. Addressing planetary health challenges in Africa. *Public Health Reviews*. 2016;37:27.
19. Abernethy K, Maisels F, White LTJ. Environmental Issues in Central Africa. *Annu. Rev. Environ. Resour.* 2016;41:1-33.
20. Cilliers S, Cilliers J, Lubbe R, Siebert S. Ecosystem services of urban green spaces in African countries-perspectives and challenges. *Urban Ecosyst.* 2013;16:681-702.
21. Toit MJ, Cilliers SS, Dallimer M, et al. Urban green infrastructure and ecosystem services in sub-Saharan Africa. *Landscape and Urban Planning*. 2018;180:249-261.
22. James P, Banay RF, Hart JE, et al. A Review of the Health Benefits of Greenness. *Curr Epidemiol Rep.* 2015;2(2):131-142.
23. Fong KC, Hart JE, & James P. A Review of Epidemiologic Studies on Greenness and Health: Updated Literature Through 2017. *Curr Environ Health Rep.* 2018;5(1):77-87.
24. Twohig-Bennett C, Jones A. The health benefits of the great outdoors: A systematic review and meta-analysis of greenspace exposure and health outcomes. *Environ. Res.* 2018;166:628-637.
25. Li Q, Morimoto K, Kobayashi M, et al. Visiting a forest, but not a city, increases human natural killer activity and expression of anti-cancer proteins. *Int J Immunopathol Pharmacol.* 2008;21(1):117-127.

26. Park BJ, Tsunetsugu Y, Kasetani T, et al. The physiologic effects of Shinrin-yoku (taking in the forest atmosphere or forest bathing): evidence from field experiments in 24 forests across Japan. *Environ Health Prev Med.* 2010;15(1):18-26.
27. McEwen BS. Stress, adaptation and disease. Allostasis and allostatic load. *Ann N Y Acad Sci.* 1998;840:33-44.
28. Egorov AI, Griffin SM, Converse RR, et al. Vegetated land cover near residence is associated with reduced allostatic load and improved biomarkers of neuroendocrine, metabolic and immune functions. *Environ Res.* 2017;158:508-521.
29. Yeager R, Riggs DW, DeJarnett N, et al. Association Between Residential Greenness and Cardiovascular Disease Risk. *J Am Heart Assoc.* 2018;7(24):e009117.
30. Tomita A, Vandormael AM, Cuadros D, et al. Green environment and incident depression in South Africa: a geospatial analysis and mental health implications in a resource-limited setting. *Lancet Planet Health.* 2017;1(4):e152-e162.
31. Dalal S, Holmes MD, Laurence C, et al. Feasibility of a large cohort study in sub-Saharan Africa assessed through a four-country study. *Glob Health Action.* 2015;8:27422.
32. Kriegler FJ, Malila WA, Nalepka, RF, et al. (1969). Preprocessing transformations and their effects on multispectral recognition. In: *Proceedings of the Sixth International Symposium on Remote Sensing of Environment.* Ann Arbor, MI: Institute of Science and Technology of the University of Michigan; 1969:97-131.
33. Mutanga O and Kumar L. Google Earth Engine Applications. *Remote Sens.* 2019;11:591.
34. Chiwanga FS, Njelekela MA, Diamond MB, et al. Urban and rural prevalence of diabetes and pre-diabetes and risk factors associated with diabetes in Tanzania and Uganda. *Glob Health Action.* 2016;9:31440.

35. Guwatudde D, Nankya-Mutyoba J, Kalyesubula R, et al. The burden of hypertension in sub-Saharan Africa: a four-country cross sectional study. *BMC Public Health*. 2015;15:1211.
36. Craig SR, Amin RV, Russell DW et al. Blood cholesterol screening influence of fasting state on cholesterol results and management decisions. *J Gen Intern Med* 2000;15(6):395-399.
37. Seeman TE, Singer BH, Rowe JW, et al. Price of adaptation—allostatic load and its health consequences. MacArthur studies of successful aging. *Arch Intern Med*. 1997;157(19):2259-68.
38. Ajayi IO, Adebamowo C, Adami HO, et al. Urban-rural and geographic differences in overweight and obesity in four sub-Saharan African adult population: a multi-country cross-sectional study. *BMC Public Health*. 2016;16(1):1126.
39. Laurence EC, Volmink J, Esterhuizen TM, et al. Risk of cardiovascular disease among teachers in Cape Town: Findings of the South African PaCT pilot study. *S Afr Med J*. 2016;106(10):996-1001.
40. Rubin DB. Multiple Imputation for Nonresponse in Surveys. *Wiley*. 1987.
41. Valeri L, VanderWeele TJ. Mediation Analysis Allowing for Exposure-Mediator Interactions and Causal Interpretation: Theoretical Assumptions and Implementation With SAS and SPSS Macros. *Psychol Methods*. 2013;18(2):137-150.
42. VanderWeele TJ. Outcome-wide Epidemiology. *Epidemiology*. 2017;28(3):399-402.
43. VanderWeele T, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann. Intern. Med*. 2017;167(4):268-274.

44. Astell-Burt T, Feng X, Kolt GS. Is Neighborhood Green Space Associated With a Lower Risk of Type 2 Diabetes? Evidence from 267,072 Australians. *Diabetes Care*. 2014;37:197-201.
45. Villeneuve P, Jerrett M, Su JG, et al. Association of residential greenness with obesity and physical activity in a US cohort of women. *Environ Res*. 2018;160:372-384.
46. Thiering E, Markevych I, Bruske I, et al. Associations of Residential Long-Term Air Pollution Exposures and Satellite-Derived Greenness with Insulin Resistance in German Adolescents. *Environ Health Perspect*. 2016;124(8):1291-8.
47. Tamosiunas A, Grazuleviciene R, Luksiene D, et al. Accessibility and use of urban green spaces, and cardiovascular health: findings from a Kaunas cohort study. *Environ Health*. 2014;13(1):20.
48. Yang B, Markevych I, Heinrich J, et al. Associations of greenness with diabetes mellitus and glucose-homeostasis markers: The 33 Communities Chinese Health Study. *International Journal of Hygiene and Environmental Health*. 2019;222:283-290.
49. Brown SC, Lombard J, Wang K, et al. Neighborhood Greenness and Chronic Health Conditions in Medicare Beneficiaries. *Am J Prev Med*. 2016;51(1):78-89.
50. Kim HJ, Min JY, Kim HJ, Min KB. Parks and Green Areas Are Associated with Decreased Risk for Hyperlipidemia. *Int J Environ Res Public Health*. 2016;13(12). Pii: E1205.
51. Markevych I, Standl M, Sugiri D, et al. Residential greenness and blood lipids in children: A longitudinal analysis in GINIplus and LISApplus. *Environ Res*. 2016;151:168-173.

52. Jia X, Yu Y, Xia W, et al. Cardiovascular diseases in middle aged and older adults in China: the joint effects and mediation of different types of physical exercise and neighborhood greenness and walkability. *Environ Res.* 2018;167:175-183.
53. Guwatudde D, Kirunda BE, Wesonga R et al. Physical Activity Levels Among Adults in Uganda: Findings From a Countrywide Cross-Sectional Survey. *J Phys Act Health.* 2016;13(9):938-45.
54. Mashili FL, Kagaruki GB, Mbatia J et al. Physical Activity and Associated Socioeconomic Determinants in Rural and Urban Tanzania: Results from the 2012 WHO-STEPS Survey. *International Journal of Population Research.* 2018;4965193;
55. Lachowycz K, Jones AP. Does walking explain associations between access to greenspace and lower mortality? *Soc Si Med.* 2014;107:9-17.
56. Richardson EA, Pearce J, Mitchell R, Kingham S. Role of physical activity in the relationship between urban green space and health. *Public Health.* 2013;127(4):318-24.

Supplementary Appendix

Table S3.1. Cross-sectional associations between Normalized Difference Vegetation Index (0.11 unit increase) and components of allostatic load among Participants from four Sites in Three sub-Saharan African Countries (complete case)

Outcome	Total Population (N=984)				Urban Population (N=822)			
	β	OR	(95% CI)	<i>p</i>	β	OR	(95% CI)	<i>p</i>
BMI (continuous)								
Unadjusted	-1.65		(-1.86, -1.45)	<.0001	-1.89		(-2.18, -1.59)	<.0001
Age + Sex	-1.39		(-1.61, -1.17)	<.0001	-1.53		(-1.85, -1.22)	<.0001
Full Adjustment ^a	-0.98		(-1.34, -0.61)	<.0001	-0.97		(-1.36, -0.58)	<.0001
Overweight/Obese								
Unadjusted		0.57	(0.52, 0.62)	<.0001		0.51	(0.45, 0.58)	<.0001
Age + Sex		0.64	(0.58, 0.70)	<.0001		0.60	(0.53, 0.69)	<.0001
Full Adjustment ^a		0.73	(0.62, 0.86)	0.0001		0.72	(0.61, 0.85)	0.0001
Diabetes								
Unadjusted		0.87	(0.77, 0.98)	0.0219		0.72	(0.6, 0.87)	0.0004
Age + Sex		0.93	(0.82, 1.06)	0.3005		0.75	(0.62, 0.91)	0.0033
Full Adjustment ^a		0.76	(0.6, 0.96)	0.0226		0.75	(0.59, 0.95)	0.0157
Hypertension								
Unadjusted		0.66	(0.60, 0.72)	<.0001		0.69	(0.62, 0.77)	<.0001
Age + Sex		0.73	(0.67, 0.80)	<.0001		0.78	(0.69, 0.88)	<.0001
Full Adjustment ^a		0.80	(0.68, 0.93)	0.0043		0.79	(0.67, 0.92)	0.0026
Cholesterol (continuous)^b								
Unadjusted	-0.26		(-0.32, -0.2)	<.0001				
Age + Sex	-0.18		(-0.25, -0.12)	<.0001				
Full Adjustment ^a	0.03		(-0.09, 0.15)	0.6461				
Cholesterol (binary)^b								
Unadjusted		0.99	(0.89, 1.11)	0.8804				
Age + Sex		1.06	(0.94, 1.21)	0.3341				
Full Adjustment ^a		1.11	(0.88, 1.39)	0.3698				

Table S3.1. Cross-sectional associations between Normalized Difference Vegetation Index (0.11 unit increase) and components of allostatic load among Participants from four Sites in Three sub-Saharan African Countries (complete case) (continued)

Outcome	Total Population (N=984)				Urban Population (N=822)			
	β	OR	(95% CI)	<i>p</i>	β	OR	(95% CI)	<i>p</i>
Allostatic Load^c								
Unadjusted								
0 (Low)								
1		0.87	(0.79, 0.96)	0.0064	0.80	(0.70, 0.92)	0.0011	
2		0.69	(0.61, 0.79)	<.0001	0.65	(0.54, 0.77)	<.0001	
≥ 3 (High)		0.64	(0.55, 0.75)	<.0001	0.62	(0.51, 0.76)	<.0001	
Age + Sex								
0 (Low)		Ref						
1		0.96	(0.86, 1.07)	0.4557	0.91	(0.78, 1.05)	0.205	
2		0.79	(0.69, 0.91)	0.0012	0.74	(0.61, 0.90)	0.0022	
≥ 3 (High)		0.75	(0.64, 0.89)	0.0007	0.72	(0.58, 0.90)	0.0033	
Full Adjustment ^a								
0 (Low)		Ref						
1		0.87	(0.72, 1.05)	0.1532	0.87	(0.72, 1.05)	0.1377	
2		0.72	(0.56, 0.92)	0.0091	0.70	(0.55, 0.90)	0.006	
≥ 3 (High)		0.74	(0.56, 0.98)	0.0353	0.72	(0.55, 0.96)	0.0234	

^aMultiple regression models were fit for categorical (multinomial logistic), binary (logistic) or continuous (linear) outcome variables, using the following covariates: age, sex, educational attainment, fuel source, marital status, smoking status, site, seeking health care in private sector.

^bSince cholesterol was missing for one site (peri-urban Uganda), these models were fit only on remaining sites (Tanzania, South Africa, and Rural Uganda) with a sample size of 903.

^cSince glucose was missing for several participants, these models were only fit among those with non-missing data (total: N=809; urban: N=662).

Bold: $p < 0.05$ after Bonferroni correction applied to final fully adjusted outcome models (the p -value cutoff for Bonferroni correction = $0.05/12$ outcomes = 0.0042).

Table S3.2. Sensitivity Analysis for Unmeasured Confounding using E-values for Associations between Normalized Difference Vegetation Index (0.11 unit increase) and allostatic load components^a

Allostatic Load Component	Total Population		Urban Population Only	
	E-value for Point Estimate	E-value for Confidence Interval	E-value for Point Estimate	E-value for Confidence Interval
BMI (continuous)	1.59	1.44	1.56	1.39
Overweight/Obese	1.62	1.39	1.61	1.38
Diabetes	1.53	1.17	1.54	1.18
Hypertension	1.26	1.00	1.28	1.00
Cholesterol (continuous)	1.17	1.00	-	-
Cholesterol (binary)	1.35	1.00	-	-
Allostatic Load				
0 (Low)	Ref	-	Ref	-
1	1.36	1.00	1.33	1.00
2	1.75	1.39	1.79	1.45
≥ 3 (High)	1.75	1.39	1.75	1.38

^aAll models were fit using multiple imputation with chained equations to impute body mass index and the following covariates: age, sex, educational attainment, fuel source, marital status, NDVI, smoking status, site. Outcome models were fit adjusting for the following covariates: age, sex, educational attainment, fuel source, marital status, smoking status, site, seeking health care in private sector.