Published in final edited form as:

Cancer. 2020 January 01; 126(9): 1949–1957. doi:10.1002/cncr.32734.

Utilization of Empiric Methods to Inform Prostate Cancer Health Disparities: Comparison of Neighborhood-Wide Association Study (NWS) “Hits” in Black and White Men

Shannon M. Lynch, PhD, MPH¹, Kristen Sorice, BA¹, Erin Tagai, PhD¹, Elizabeth A. Handorf, PhD²

¹Cancer Prevention and Control, Fox Chase Cancer Center, Philadelphia, PA

²Population Studies Facility, Fox Chase Cancer Center, Philadelphia, PA

Abstract

Background: Black men are more likely to die of prostate cancer (PCa) than White men. Factors from genetics to neighborhood environment contribute to these disparities. However, unlike genetics, agnostic investigations that identify candidate variables from large-scale data, and that allow for empiric investigations into differential associations between neighborhood and PCa by race/ethnicity, are not well-explored. Thus, we build on our previously-developed, empiric neighborhood-wide association study (NWS) in White men and conduct NWS in Black men to determine if findings differ by race.

Methods: Pennsylvania Cancer Registry data were linked to U.S. Census data. For the NWS in non-Hispanic Black men, we evaluated the association between 14,663 neighborhood census variables and advanced PCa ($n=911$ high stage/grade cases; $n=8632$ low stage/grade cases), adjusting for age, diagnosis year, spatial correlation, and multiple-testing. Odds ratios (OR) and 95% credible intervals (CI) are reported. Replication of NWS findings across Black/White races was assessed using Bayesian mixed-effects models.

Results: Five variables related to housing ($n=3$), education ($n=1$), and employment/transportation ($n=1$) were significantly associated with advanced PCa in Black men, compared to 17 socioeconomic variables (mostly related to poverty/income) in White men. The top hit in Black men related to crowding in renter-occupied housing ($OR=1.10$; $CI=1.001-1.12$). Nine of 22 NWS hits (4/5 hits in Black men) replicated across race/ethnic groups.

Conclusions: Different neighborhood variables, or “candidates,” were identified across race-specific NWSs. These findings and empiric approaches warrant additional study and may inform

Corresponding Author: Shannon M. Lynch, Ph.D., MPH, 333 Cottman Avenue, Philadelphia, PA 19111, Tel: 215-728-5377, Shannon.Lynch@fccc.edu.

Authors’ contributions: SML designed, analyzed, interpreted and wrote the manuscript; EH analyzed, interpreted, and lead the methods/results; KS extracted/geocoded participant data and edited the manuscript; ET provided consultation on study interpretation/manuscript writing. All authors read and approved the final manuscript.

Data sharing and data accessibility: The data from this study are available from the Pennsylvania State Cancer Registry upon request and the publically-available U.S. Census Data.

Conflicts of interest: The authors declare that they have no conflicts of interest.

PCa racial disparities, particularly future gene-environment studies aimed at identifying patients/communities at-risk for advanced PCa.

Keywords

NWAS; Prostate Cancer; Neighborhood; Disparities; Multilevel

Introduction

Despite high survival, prostate cancer (PCa) remains the second leading cause of cancer death in U.S. men.¹ Few risk factors for PCa have been identified, with only race/ethnicity, older age, PCa family history, and some inherited genetic conditions (e.g., Lynch syndrome, BRCA1/2 mutations) considered standard risk factors.¹ Compared to other cancers, health disparities play a major role in PCa incidence and mortality, particularly as Black men are more than twice as likely to be diagnosed with, and die of, PCa than White men.¹ Despite improvement in PCa treatment, this disparity has persisted for three decades and is the largest of any cancer.¹

A number of multilevel conceptual frameworks in cancer suggest that biology, race/ethnicity, and socioeconomic background (e.g., education, income), as well as factors outside an individual—including the macroenvironment or neighborhood in which a person lives—are related to PCa health disparities.² Recent studies have centered on the role of genes and biology, including the identification of genetic susceptibility markers from family-based studies,³ *a priori* selected candidate-gene association studies,⁴ and empiric genome-wide association studies (GWAS),^{3,5} overshadowing the important role of neighborhood in PCa disparities.

Social environment, defined as the neighborhood in which one lives, is often characterized by U.S. Census variables related to socioeconomic status (SES) that describe economic (e.g., employment, income); physical (e.g., housing/transportation structure); and social (e.g., poverty, education) characteristics. Census tracts (smaller geographic boundaries than counties) are often used to define neighborhood boundaries.^{6–8} Previous studies have found that low SES/more deprived neighborhoods are associated with greater risk of late-stage/high-grade PCa^{6,9} in both Black and White men, and less aggressive treatment.¹⁰ Associations between neighborhood SES and PCa remain, even after race/ethnicity adjustments.^{6,7,11–14} Thus, these findings suggest that neighborhood is independently associated with PCa, perhaps under a chronic stress mechanism, whereby residents from disadvantaged neighborhoods experience constant emotional distress that over time affects cancer initiation and progression.^{15–17} Further, neighborhood variables are often correlated with access to care, including cancer screening utilization,¹⁸ suggesting their relevance as markers of PCa disparities.

While there appears to be an association between neighborhood SES and PCa, whether such associations with advanced PCa could change by race/ethnicity is still unclear. With current “candidate-based” variable selection approaches in neighborhood and cancer studies, this is challenging to investigate. *Within* a single neighborhood and cancer study, the same neighborhood SES factors or domains (i.e., poverty, education) are often selected *a priori* for

evaluation within each racial/ethnic group.^{13,14,19} More specifically, 5–100 U.S. Census variables are typically hand-selected as candidates, and various analytic approaches are used²⁰ to identify relevant SES domains, reduce the number of variables to study (by removing overlapping/highly correlated variables), and/or develop composite neighborhood SES/deprivation indices that create single summary scores of a neighborhood's overall employment, education, income, etc.^{6,7,20} These approaches do not take into account the full range of available U.S. census variables (>14,000) and lack consistency because one study might define neighborhood SES in terms of percentage of female heads of households with children; another in terms of poverty based on household income.²¹ Inconsistencies in variable selection affect comparability and generalizability across studies, and can make etiologic inferences²² and investigations into race/ethnic differences difficult. Further, these studies are often conducted in majority White populations where variable selection could be heavily weighted by White population associations. This is important because similar to genetic studies where there are racial/ethnic differences in allele frequencies that impact genetic associations with disease,^{3,4} individuals often self-select into neighborhoods based on their race/ethnicity and/or income level,⁷ thus frequencies of commonly-used neighborhood SES measures likely also differ by race/ethnicity in a way that can impact neighborhood associations with cancer outcomes.

In genetic studies, to account for population stratification,²³ or identify differences by race/ethnicity, empiric analysis or agnostic investigations, such as genome-wide association studies (GWAS), that comprehensively look for associations between *all available* genetic data and disease without a prior hypothesis, are often conducted in separate race/ethnic groups, and coupled with replication studies to test whether significant findings in one race/ethnic group are also significant in other groups.^{3–5} Empiric analyses, like GWAS studies, have proven to be hypothesis-generating for the study of PCa disparities, identifying potentially important genetic differences in Black and White men that may inform the identification of high-risk populations for advanced PCa.²⁴ Similarly, in neighborhood studies, we previously designed a novel neighborhood-wide association study (NWAS),²⁵ which is a multi-phase, successively more stringent, empiric variable selection method derived from GWAS,⁵ that agnostically identified 17 neighborhood variables (out of >14,000 U.S. census variables) significantly associated with advanced PCa in White men.²⁵ To investigate whether neighborhood associations with advanced PCa differ by race, in this study, we conduct an NWAS in Black men and compare our results to NWAS results in White men. The primary goal of our study is to empirically identify neighborhood measures or “candidates” associated with advanced PCa in Black men and to evaluate whether neighborhood measures and associations with advanced PCa differ by race, an understudied question in PCa disparities. Given few individual-level factors, beyond race/ethnicity, exist to identify at-risk populations for advanced PCa, our empiric assessment of neighborhood could serve to be hypothesis-generating by identifying new neighborhood variables which may be of interest when studying Black men, White men, or both. These variables and race-specific differences could then potentially serve as candidates to inform the identification of high-risk populations for advanced PCa in future studies.

Methods

Study Population.

All primary, incident PCa cases diagnosed in Pennsylvania (PA) between 1995 and 2005 were requested from the PA State Department of Health Cancer Registry. Incident cases were identified according to ICD-0–3 site and morphology coding. We ascertained patient-level variables related to PCa tumor stage/grade, age at diagnosis, diagnosis year, and race (Black/White). There were 80,575 White cases and 9,785 non-Hispanic Black PCa cases. We excluded cases with a P.O. Box address (n=112), and those missing tumor grade/stage (n=36,603), age (n=10), or year of diagnosis (n=6). A total of 9,543 non-Hispanic Black men were included in the analysis. This study utilized an existing data source and was approved by the Institutional Review Board of Fox Chase Cancer Center.

Neighborhood Variables.

Residential addresses of PCa patients were geocoded at the census tract level and assigned a Federal Information Processing Standard (FIPS) code²⁶ using ArcGIS software. Using this geocode, PCa cases were linked to the neighborhood variable values of the census tract in which they live using SAS statistical software 12.1.

For empiric NWAS assessments in Black men (see statistical analysis), 14,663 census tract variables from the 2000 U.S. Census Summary File 1 (SF1) and Summary File 3 (SF3) were downloaded from Social Explorer (<http://www.socialexplorer.com>) and linked to patient-level data from the PA State Cancer Registry. The nomenclature for these variables follow the format Filename (SF1_, SF3_, PCT_SF1_, PCT_SF3_) followed by the unique variable code from the SF/SF3 codebooks (e.g., SF3_P03_P089002, which corresponds to percentage of individuals living below poverty). For replication assessments across Black and White men, significant variables from the NWAS in Black men were tested in White men, and the 17 significant (“top hit”) variables from a previously published NWAS²⁵ in White men were tested in Black men (Supporting File 1), as described in the statistical analysis below.

Outcome Definition.

We defined our outcome, PCa aggressiveness, with a variable that defined high stage cases as patients with high tumor stage (stage 3) and high tumor grade (grade 7), and low stage cases as patients with <stage 3 or <Gleason 7 PCa.²⁵ This outcome is relevant as most men die with, and not of PCa; making it important to study cases who are likely to be diagnosed with advanced PCa, where prognosis is poorer.²⁵ Further, this was the outcome used in the previous NWAS in White men,²⁵ which will allow for comparisons between White and Black men.

Statistical Analysis.

This cross-sectional analysis utilizing existing data sources was divided into two parts: (1) an empiric NWAS analysis in Black men; (2) replication assessment comparing empiric NWAS findings in Black and White men using a candidate, cross-over approach.

NWAS Analysis: As previously reported, NWAS is an empiric, data reduction method derived from approaches and concepts in GWAS.²⁵ Briefly, we utilized a successively more stringent, multi-phase approach, where we ran a regression model for each neighborhood variables within each phase, carrying forward the most statistically significant variables from each phase after adjustment for multiple comparisons. Neighborhood variables were Z-score transformed in order to compare odds ratios from many regressions.²⁵ In Phase 1, we account for neighborhood clustering effects via Generalized Estimating Equations (GEE) with a logistic regression model to estimate odds ratio (OR) and 95% confidence interval.²⁷ In Phase 2, we account for spatial effects, which assume that nearby neighborhoods have similar characteristics. This is an important consideration because spatial variance can affect disease associations, a consideration lacking in prior studies.²² More specifically, we specified a Bayesian hierarchical logistic regression model²⁸ using Integrated Nested Laplace Approximations (INLA)²⁹ that calculated ORs and adjusted 95% credible intervals (CI). In Phase 1 and 2, all models were adjusted for continuous age and year of diagnosis. To account for multiple testing, Bonferroni-corrected thresholds corresponding to p-values less than 0.05 were considered statistically significant in both phases. In Phase 3, we address the high degree of correlation among remaining census variables⁵ using principal components analysis²⁰ implemented in R statistical software. The most significant variables (“top hits”) in the NWAS in Black men were selected by identifying variables within each principal component (variables with the highest loading on the given component) with the tightest credible interval. All phases used R statistical software.

Replication Assessment: We first compared findings from the 2 empiric NWAS analyses in White and Black men to determine if similar components or domains, as well as variables, replicated across both groups. Association of the NWAS hits found in the Black population versus those found in the White men were summarized using Pearson Correlation. Similar to replication approaches in genetic association studies that use GWAS findings as *a priori* findings, or “candidates” that can then be tested in different race/ethnic groups,⁴ we formally tested whether NWAS hits from one race replicated in the other race using regression models stratified by race and adjusted for age and year of diagnosis. Specifically, each hit was tested in the other population using GEE models (similar to NWAS Phase 1), and those which were significant (after adjustment for multiple testing) were confirmed using more stringent Bayesian Hierarchical models^{28,29} (similar to NWAS Phase 2). Variables that remained significant across each race-specific NWAS, or GEE and Bayesian models stratified by race, were considered to be relevant in the overall study population. NWAS variables that were only significant in race-stratified NWAS were considered to have race-specific associations. We also evaluated race and NWAS hit interactions, and results were the same as in race-stratified models (data not shown).

Results

In White men, there were 6,416 high-stage cases and 70,670 low-stage cases; in Black men, 911 high-stage cases and 8,632 low-stage cases. The average age of the total study population was 68.5 (standard deviation [sd]: 9.2). The average age of White men was 68.8 (sd=9.2) compared to 66.3 (sd=9.3) in Black men. The average ages of White and Black

men with aggressive PCa were 69.8 (sd=10.4) and 67.1 (sd=10.1), respectively. Of the reported 3,135 census tracts in PA in 2000, 3,080 (98%) census tracts are represented in our study sample. Aggressive PCa cases for both Black and White men were clustered mainly in urban areas, namely Pittsburgh and Philadelphia (data not shown²⁵).

NWAS in Black Men Analysis:

Figure 1 summarizes Phases 1–3 of the NWAS in Black men. In Phase 1, from 14,663 variables, we identified 99 neighborhood variables significantly associated with PCa aggressiveness (Supporting File 2). After Phase 2, twenty variables were still significant (Supporting File 3). In Phase 3, five uncorrelated principal components were identified from these 20 neighborhood variables (Supporting File 4), explaining 90% of the variance: Renter-Occupied Crowded Housing (Component 1–7 variables loaded), Housing Facilities (Component 2–6 variables), Owner-Occupied Crowded Housing (Component 3;3 variables), Education (Component 4;1 variable) and Employment/Transportation (Component 5;3 variables). The top 5 significant variables (i.e. tightest credible intervals) selected from each of the 5 principal components are presented in Figure 1 and Table 1; three of these variables (from Components 1–3) related to housing characteristics. The most significant variable in the analysis loaded on Component 1, *% renter-occupied housing units with householder age 45–54, with 1.01–1.5 occupants per room (OR=1.11, CI=1.01–1.20)*.

Replication Assessment:

Different top hits were identified across racial groups in empiric NWAS analyses, i.e., there was no overlap in the significant variables identified in Black and White men (Supporting File 1). However, two components or domains related to housing and employment/transportation were identified in both NWAS analyses. For example, the top hit in the NWAS in White men (*% workers 16 years taking trolley or street car public transportation to work [OR=1.05, CI=1.001–1.09]*), but in Black men a different variable related to employment/transportation was identified: *% workers 16 with an aggregate travel time to work less than 30 minutes by public transportation (OR=1.13, CI=1.01–1.27)*. In contrast, other domains in the NWAS in White men related to poverty, immigration, and social support were not empirically identified as significant in the NWAS in Black men. Further, education, which was not a significant domain in the NWAS in White men, was found to be associated with aggressive PCa in Black men (*% females 18–24 with a graduate/professional degree [OR=1.11, CI=1.01–1.22]*).

Given different variables and unique domains were identified across the White and Black NWAS analyses, correlation analyses and replication studies were subsequently conducted among the 17 top NWAS hits in White men and 5 top NWAS hits in Black men to determine which were relevant in the full study population. Correlations between top hits were small to moderate, with the largest correlations for the Black hits (out of their correlations with all 17 White hits) ranging from 0.22–0.70 (Figure 2; Supporting File 1). As shown in Table 2, 5 of the 17 top hits in White men (representing social support, income, and immigration domains) validated in the Black population, while 4 of the 5 top hits for Black men (representing housing and employment/transportation domains) validated in the White population. These 9 NWAS variables in Table 2 also tended to be the same variables with

larger correlations in the full study population (Figure 2), but their effect sizes were attenuated (i.e., ORs for the NWAS hits in White men that replicated in Black men [Table 2] were smaller than any of the original top hits from the NWAS in Black men [Table 1]). The top hit in the NWAS in Black men, which related to crowded housing (*% renter-occupied housing units with householder age 45–54, with 1.01–1.5 occupants per room*), was replicated in White men; however, the top hit in White men, which related to employment/transportation, did not replicate in Black men, despite employment/transportation being identified as a significant domain in NWAS Black only analysis.

Discussion

In this study, we used a data-driven, systematic approach to comprehensively assess differences in neighborhood associations with advanced PCa in White and Black men.²² We applied our novel NWAS approach²⁵ to empirically and comprehensively assess the association of 14,663 neighborhood variables with PCa aggressiveness in Black men, and we identified 5 top hits associated with PCa aggressiveness. Three of the 5 top hits related to housing, and the remaining two hits related to education and employment/transportation. Although this cross-sectional study cannot be used to infer causality, only association, our findings are consistent with other association studies in literature. Studies show that housing factors, including mortgage discrimination,³⁰ are associated with disparities in the Black community, given decades of unequal housing policies have led to poor living conditions (i.e., crowding) that can negatively impact health.^{31–33} In most prior PCa and neighborhood studies, neighborhood SES is commonly measured with composite deprivation indices and single, *a priori* selected U.S. Census variables related to education, income, poverty, housing, transportation, and employment.^{6,9} In contrast to previous studies, the NWAS in Black men identified more complex variables indicative of potential interactions among these common domains (i.e., *% workers 16 with an aggregate travel time to work less than 30 minutes by public transportation* vs *%employed workers 16*). Employment and transportation variables are often used as surrogates for access to care and it's possible that the addition of the public transportation information in the NWAS hit could be an indicator of location in urban versus rural settings.³⁴ Education also was a significant domain in the NWAS in Black men, but the variable identified for this domain (*% females 18–24 with a graduate/professional degree*) was more specific, rare (average frequency of less than 1% in the study population), but highly correlated ($r=0.81$) with one commonly used variable to represent education (*% with less than a high school education*). This variable (*%females 18–24 with a graduate/professional degree*) further demonstrates that variables selected from NWAS may serve as “candidates” or markers for domains, such as education, but to truly understand if neighborhood variables have etiologic significance, additional studies with different designs would be needed (e.g. going into neighborhoods and surveying communities with specific characteristics). Interestingly, poverty and income, arguably the two most commonly-used variables to represent neighborhood SES³⁵ were not significantly associated with aggressive PCa in Black men in empiric analysis. This could suggest these factors may play less of a role in PCa disparities in Black men compared to White men, but additional replication studies would be needed.

In the replication assessments, when comparing NWAS findings across White vs Black men, important domains were identified and replicated in both studies related to housing and employment/transportation. Similarly, 9 of 22 NWAS hits were considered validated for the full study population, covering domains representing social support, income, and immigration. However, domains related to poverty, employment, and physical environment from the NWAS in White men were not replicated (Supporting File 1). Lack of representation of relevant domains, as well as lack of validation of specific neighborhood variables (particularly from the White NWAS in the Black population), suggest that although there are a few variables (9/22) which apply to the full study population, other neighborhood variables, and related domains, could be missed if race-specific associations are not considered. This has implications for future studies of PCa disparities. However, we were able to identify a set of variables that were relevant in the total study population utilizing both empiric (NWAS) and “candidate” association study approaches. This validated group of NWAS hits may be informative for intervention studies, particularly improving the identification of high risk populations. These nine NWAS variables generally covered the majority of significant domains identified in race-specific analyses, and were often moderately correlated with non-validated NWAS hits from the same domain. Replication of neighborhood findings is becoming increasingly important as a number of professional organizations are now supporting efforts towards recommending a standardized set of neighborhood measures to include in health disparity research and medical record data.³⁶ However, given the differences in identified neighborhood measures by race in this study, creating a variable set that is generalizable across different outcomes, geographies, and race/ethnic groups for use in other U.S. research studies would require additional investigations.

There are a number of limitations to note in this study. The lack of replication across both NWAS hits and related domains could indicate that pathways to PCa are different by race, but racial differences could also be the result of the NWAS data structure itself. U.S. Census variables are generally highly collinear, and by design, the NWAS method only selects one variable from each group of highly interrelated variables. Additionally, the frequency or percentages of census variables may vary, not only by race (Supporting File 1), but also by geographic location and the geographic boundary selected as the unit of study (e.g., county vs. census tract).^{37,38} Further, due to the smaller sample size in the Black population, we may have lacked statistical power to validate the 17 hits from the white NWAS.

Additionally, the use of census tracts as neighborhood boundaries could lead to the modifiable areal unit problem. Thus, given the complexity and variability of this dataset, subsequent attempts to replicate the same variables across independent study populations may be challenging, as these specific variables may not be generalizable elsewhere. For this reason and because single association studies do not allow for etiologic inferences, incorporating other systematic, empiric approaches that can account for complex correlation and that can allow for standardized replication efforts and comparisons within and across datasets, such as machine learning methods, may be warranted.³⁹ Additionally, we utilized publically available datasets where systematic reporting bias exists²⁶ and where relevant individual-level SES data are lacking. Further, this analysis was based on an analysis of cases only, which can be subject to selection bias,^{40–42} but was useful to allow for comparisons between the previously published NWAS in White men and the new NWAS in

Black men presented here. Additionally, future NWAS should be conducted in study populations that can account for individual-level SES factors, which are known to be correlated with neighborhood circumstances.⁷

Despite limitations, this is the first study to systematically, empirically, and comprehensively evaluate and compare the role of neighborhood-level factors in PCa across racial groups. Our NWAS has led to a list of potentially interesting candidate variables that could be useful for future investigations that could impact cancer prevention and control.³⁹ For instance, NWAS findings could potentially be incorporated in patient and population risk assessment studies, or gene-environment studies, aimed at more precisely identifying and targeting both high-risk patients and geographic areas with unfavorable neighborhood characteristics.³⁹ While most PCas are slow-growing and can be safely monitored with active surveillance, a percentage of men will be diagnosed with aggressive prostate cancer. NWAS findings may help clinicians identify patient populations at elevated risk to inform shared decision-making for PCa screening, as well as management of diagnoses.^{43,44} Further, NWAS findings may serve as additional “markers”, along with disease rates, to help researchers identify neighborhoods or communities in need of cancer prevention efforts.³⁹ Additionally, our NWAS *method* could serve as a useful variable selection/reduction method for large-scale, multilevel, gene-environment studies interested in studying PCa racial disparities. In this era of Big Data and Precision Medicine,⁴⁵ access to and interest in the utilization of neighborhood data will only continue to grow. Thus, moving forward, systematic approaches, like NWAS, will likely be useful to evaluate the generalizability of significant neighborhood findings across race/ethnic groups and the potential utility of neighborhood data for studying PCa disparities.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding support: This research was supported by grants from the American Cancer Society IRG-92-027-20 and MRSG-CPHPS-130319 to SML and the NIH/NCI Cancer Center Support Grant P30CA00692.

References:

1. American Cancer Society. Cancer Facts and Figures. Atlanta: American Cancer Society;2009.
2. Lynch SM, Rebbeck TR. Bridging the Gap between Biologic, Individual, and Macroenvironmental Factors in Cancer: A Multilevel Approach. *Cancer Epidemiology Biomarkers & Prevention*. 2013;22(4):485–495.
3. Rebbeck TR. Prostate Cancer Genetics: Variation by Race, Ethnicity, and Geography. *Semin Radiat Oncol*. 2017;27(1):3–10. [PubMed: 27986209]
4. Chang B-L, Spangler E, Gallagher S, et al. Validation of genome-wide prostate cancer associations in men of African descent. *Cancer Epidemiol Biomarkers Prev*. 2011;20(1):23–32. [PubMed: 21071540]
5. Eeles RA, Kote-Jarai Z, Al Olama AA, et al. Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. *Nat Genet*. 2009;41(10):1116–1121. [PubMed: 19767753]

6. Ziegler-Johnson C, Tierney, Rebbeck, Rundle A. Prostate Cancer Severity Associations with Neighborhood Deprivation. *Prostate Cancer*. 2011;2011.
7. Diez Roux AV, Mair C. Neighborhoods and health. *Annals of the New York Academy of Sciences*. 2010;1186(1):125–145. [PubMed: 20201871]
8. Krieger N, Chen JT, Waterman PD, Soobader M-J, Subramanian SV, Carson R. Geocoding and Monitoring of US Socioeconomic Inequalities in Mortality and Cancer Incidence: Does the Choice of Area-based Measure and Geographic Level Matter?: The Public Health Disparities Geocoding Project. *American Journal of Epidemiology*. 2002;156(5):471–482. [PubMed: 12196317]
9. Carpenter W, Howard D, Taylor Y, Ross L, Wobker S, Godley P. Racial differences in PSA screening interval and stage at diagnosis. *Cancer Causes and Control*. 2010;21(7):1071–1080. [PubMed: 20333462]
10. Lyratzopoulos G, Barbiere JM, Greenberg DC, Wright KA, Neal DE. Population based time trends and socioeconomic variation in use of radiotherapy and radical surgery for prostate cancer in a UK region: continuous survey. *BMJ*. 2010;340.
11. Ellis L, Canchola AJ, Spiegel D, Ladabaum U, Haile R, Gomez SL. Racial and Ethnic Disparities in Cancer Survival: The Contribution of Tumor, Sociodemographic, Institutional, and Neighborhood Characteristics. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2018;36(1):25–33. [PubMed: 29035642]
12. DeRouen MC, Schupp CW, Koo J, et al. Impact of individual and neighborhood factors on disparities in prostate cancer survival. *Cancer epidemiology*. 2018;53:1–11. [PubMed: 29328959]
13. DeRouen MC, Schupp CW, Yang J, et al. Impact of individual and neighborhood factors on socioeconomic disparities in localized and advanced prostate cancer risk. *Cancer Causes Control*. 2018;29(10):951–966. [PubMed: 30136012]
14. Rebbeck TR, Weber AL, Walker AH, et al. Context-dependent effects of genome-wide association study genotypes and macroenvironment on time to biochemical (prostate specific antigen) failure after prostatectomy. *Cancer Epidemiol Biomarkers Prev*. 2010;19(9):2115–2123. [PubMed: 20826827]
15. Ellison GL, Coker AL, Hebert JR, Sanderson SM, Royal CD, Weinrich SP. Psychosocial stress and prostate cancer: a theoretical model. *Ethnicity & disease*. 2001;11(3):484–495. [PubMed: 11572415]
16. Geronimus AT, Hicken M, Keene D, Bound J. “Weathering” and Age Patterns of Allostatic Load Scores Among Blacks and Whites in the United States. *Am J Public Health*. 2006;96:826–833. [PubMed: 16380565]
17. Protective McEwen B. and Damaging effects of stress Mediators. *N Engl J Med*. 1998;338:171–179. [PubMed: 9428819]
18. Shenoy D, Packianathan S, Chen AM, Vijayakumar S. Do African-American men need separate prostate cancer screening guidelines? *BMC Urology*. 2016;16:19. [PubMed: 27165293]
19. Gomez SL, Glaser SL, McClure LA, et al. The California Neighborhoods Data System: a new resource for examining the impact of neighborhood characteristics on cancer incidence and outcomes in populations. *Cancer Causes Control*. 2011;22(4):631–647. [PubMed: 21318584]
20. Messer L, . Laraia B, Kaufman J, Eyster J, Holzman C, Culhane J, et al. . The development of a standard neighborhood deprivation index. . *Journal of Urban Health*. 2006;83(6):1041–1062. [PubMed: 17031568]
21. Gomez SL, Shariff-Marco S, DeRouen 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):2314–2330. [PubMed: 25847484]
22. Sampson RJ, Morenoff JD, Gannon-Rowley T. *Ann Rev Sociol* 2002;28:443–478. “Assessing Neighborhood Effects”: Social Processes and New Directions in Research. *Ann Rev Sociol*. 2002;28:443–478.
23. Wang Y, Localio R, Rebbeck TR. Evaluating Bias due to Population Stratification in Epidemiologic Studies of Gene-Gene or Gene-Environment Interactions. *Cancer Epidemiology Biomarkers & Prevention*. 2006;15(1):124–132.
24. Wang X, Oldani MJ, Zhao X, Huang X, Qian D. A review of cancer risk prediction models with genetic variants. *Cancer Inform*. 2014;13(Suppl 2):19–28.

25. Lynch SM, Mitra N, Ross M, et al. A Neighborhood-Wide Association Study (NWS): Example of prostate cancer aggressiveness. *PLoS One*. 2017;12(3):e0174548-e0174548.
26. Bureau USC. 2000 Census Technical Documentation for SF 3. United States Department of Commerce. 2007.
27. Hubbard AE, Ahern J, Fleischer NL, et al. To GEE or Not to GEE: Comparing Population Average and Mixed Models for Estimating the Associations Between Neighborhood Risk Factors and Health. *Epidemiology*. 2010;21(4):467–474 410.1097/EDE.1090b1013e3181caeb1090. [PubMed: 20220526]
28. Pearson TA, Manolio TA. How to interpret a genome-wide association study. *JAMA*. 2008;299(11):1335–1344. [PubMed: 18349094]
29. Ru H, Martino S. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *JR Statist Soc B*. 2008;71(2):319–392.
30. Zhou Y, Bermanian A, Beyer KMM. Housing Discrimination, Residential Racial Segregation, and Colorectal Cancer Survival in Southeastern Wisconsin. *Cancer Epidemiology Biomarkers & Prevention*. 2017;26(4):561–568.
31. Beyer KMM, Laud PW, Zhou Y, Nattinger AB. Housing discrimination and racial cancer disparities among the 100 largest US metropolitan areas. *Cancer*. 2019;125(21):3818–3827. [PubMed: 31287559]
32. White K, Lipsitz G. Using Fair Housing to Achieve Health Equity In. *Stanford Social Innovation Review*. California: Stanford University Press; 201.
33. Williams DR, Mohammed SA. Discrimination and racial disparities in health: evidence and needed research. *Journal of behavioral medicine*. 2009;32(1):20–47. [PubMed: 19030981]
34. Guidry JJ, Aday LA, Zhang D, Winn RJ Transportation as a barrier to cancer treatment. *Cancer Pract*. 1997;5(6):361–366. [PubMed: 9397704]
35. McIntire RK, Keith SW, Boamah M, et al. A Prostate Cancer Composite Score to Identify High Burden Neighborhoods. *Preventive medicine*. 2018;112:47–53. [PubMed: 29625131]
36. Polite BN, Adams-Campbell LL, Brawley OW, et al. Charting the Future of Cancer Health Disparities Research: A Position Statement from the American Association for Cancer Research, the American Cancer Society, the American Society of Clinical Oncology, and the National Cancer Institute. *Cancer Research*. 2017;77(17):4548–4555. [PubMed: 28739629]
37. Baade PD, Yu XQ, Smith DP, Dunn J, Chambers SK. Geographic disparities in prostate cancer outcomes--review of international patterns. *Asian Pacific journal of cancer prevention : APJCP*. 2015;16(3):1259–1275. [PubMed: 25735366]
38. Dasgupta P, Baade PD, Aitken JF, Ralph N, Chambers SK, Dunn J. Geographical Variations in Prostate Cancer Outcomes: A Systematic Review of International Evidence. *Frontiers in oncology*. 2019;9:238. [PubMed: 31024842]
39. Lynch SM. Towards Systematic Methods in an Era of Big Data: A Neighborhood Wide Association Study In: Springer, ed. *Geospatial Approaches to Energy Balance and Breast Cancer*. Springer; 2019:99–115.
40. Walsh MC, Trentham-Dietz A, Gangnon RE, Nieto FJ, Newcomb PA, Palta M. Selection bias in population-based cancer case-control studies due to incomplete sampling frame coverage. *Cancer Epidemiol Biomarkers Prev*. 2012;21(6):881–886. [PubMed: 22490319]
41. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15(5):615–625. [PubMed: 15308962]
42. Smith LH, VanderWeele TJ. Bounding Bias Due to Selection. *Epidemiology*. 2019;30(4):509–516. [PubMed: 31033690]
43. Sanda MG, Cadeddu JA, Kirkby E, et al. Clinically Localized Prostate Cancer: AUA/ASTRO/SUO Guideline. Part I: Risk Stratification, Shared Decision Making, and Care Options. *The Journal of urology*. 2018;199(3):683–690. [PubMed: 29203269]
44. Mohler JL, Antonarakis ES, Armstrong AJ, et al. Prostate Cancer, Version 2.2019, NCCN Clinical Practice Guidelines in Oncology. *Journal of the National Comprehensive Cancer Network : JNCCN*. 2019;17(5):479–505. [PubMed: 31085757]
45. Collins FS, Varmus H. A New Initiative on Precision Medicine. *New England Journal of Medicine*. 2015;372(9):793–795. [PubMed: 25635347]

Significant Conclusions:

A new empiric method, a neighborhood-wide association study (NWAS) derived from genetics and applied to social environmental data, demonstrates differential associations between neighborhood circumstances and aggressive prostate cancer by race. These findings suggest that, like genetic risk factors, social environmental associations with cancer may differ by race, thus informing future gene-environment studies of prostate cancer disparities aimed at identifying patients and communities at risk for advanced prostate cancer.

Figure 1.
Summary of Neighborhood-Wide Association Study (NWAS) Findings in Black Men in Pennsylvania (911 high stage cases; 8632 low stage cases)

Figure 2.
Correlation Matrix Comparing Neighborhood-wide Association Study (NWAS) Top Hits in Black and White Men

Table 1.

Neighborhood-Wide Association Study (Nwas) Top 5 Hits in Black Men

			Phase 2 Associations Reported		Phase 3
			OR [*]	CI ^b	Component
Census Variable (code)	Mean (sd) ^a	Range			
% Renter-occupied housing unit, householder age 45–54, with 1.01–1.5 occupants per room (PCT_SF3_H021045)	5.99 (4.46)	0–100	1.109	1.01–1.20	1 – Renter-Occupied Crowded Housing
Imputation of kitchen facilities (PCT_SF3_H111002)	6.6 (4.74)	0–100	1.123	1.02–1.23	2 – Housing Facilities
% Owner-occupied housing unit with 3 bedrooms (PCT_SF3_H042006)	34.85 (16.88)	0–100	0.871	0.77–0.99	3 – Owner-Occupied Crowded Housing
% Females 18–24 with a graduate/professional degree (PCT_SF3_PCT025051)	0.05 (0.18)	0–5.56	1.112	1.01–1.22	4 – Education
% Workers 16 who do not work at home, aggregate travel time to work less than 30 minutes by public transportation (PCT_SF3_P033003)	3.25 (3.05)	0–43.64	1.133	1.01–1.27	5 – Employment/Transportation

^aStandard deviation (sd);^bConfidence or Credible Interval (CI);^{*}Odds ratios (OR) reported are based on centered/standardized values of neighborhood variables

Table 2.

Replication Assessment: Validation of NWAS Top Hits Across Race/Ethnic Groups

White hits validating in Black Men				
Census variable (code)	Component/ Domain	OR^a	95% LCL^b	95% UCL
%Male householder living alone (PCT_SF3_H019093)	Social Support	1.089	1.020	1.161
%Renter occupied housing unit built 1939 or earlier with householder aged 15-24 years (PCT_SF3_HCT005083)	Housing(Rent-Occupied)	1.087	1.023	1.150
%Household income \$60K-74,999 (PCT_SF3_P052012)	Income	0.919	0.853	0.988
%Foreign born naturalized citizen at or above poverty level (PCT_SF3_PCT051020)	Immigration	0.922	0.853	0.995
%Male householder >65 living alone in nonfamily household (PCT_SF1_P030012)	Social Support	1.093	1.023	1.165
Black hits validating in White Men				
Census variable (code)	Component/ Domain	OR	95% LCL	95% UCL
%Renter-occupied housing unit, householder age 45–54, with 1.01–1.5 occupants per room (PCT_SF3_H021045)	Crowded Housing (Rent-occupied)	1.041	1.015	1.066
Imputation of kitchen facilities (PCT_SF3_H111002)	Housing facilities	1.038	1.014	1.062
%Workers 16 who do not work at home, aggregate travel time to work less than 30 minutes by public transportation (PCT_SF3_P033003)	Employment/ Transportation	1.032	1.005	1.060
%Owner-occupied housing unit with 3 bedrooms (PCT_SF3_H042006)	Housing(owner-occupied)	0.953	0.928	0.979

^aOdds Ratio^bLower/Upper Credible Interval (LCL/UCL)