



# HHS Public Access

Author manuscript

*J Urol.* Author manuscript; available in PMC 2021 July 13.

Published in final edited form as:

*J Urol.* 2019 December ; 202(6): 1209–1216. doi:10.1097/JU.0000000000000416.

## The Relative Importance of Race Compared to Health Care and Social Factors in Predicting Prostate Cancer Mortality: A Random Forest Approach

**Heidi A. Hanson**<sup>\*,†</sup>,

Department of Surgery and Population Sciences, University of Utah, Salt Lake City, Utah

**Christopher Martin**<sup>†</sup>,

Division of Urology, Huntsman Cancer Institute, University of Utah, Salt Lake City, Utah

**Brock O'Neil**,

Department of Surgery and Population Sciences, University of Utah, Salt Lake City, Utah

**Claire L. Leiser**,

Department of Population Sciences, University of Utah, Salt Lake City, Utah

**Erik N. Mayer**,

Division of Urology, Huntsman Cancer Institute, University of Utah, Salt Lake City, Utah

**Ken R. Smith**,

Department of Population Sciences, University of Utah, Salt Lake City, Utah; Department of Family and Consumer Studies, University of Utah, Salt Lake City, Utah

**William T. Lowrance**

Division of Urology, Huntsman Cancer Institute, University of Utah, Salt Lake City, Utah

### Abstract

**Purpose:** There is a differential in prostate cancer mortality between black and white men.

Advances in precision medicine have shifted the research focus toward underlying genetic differences. However, nonbiological factors may have a large role in these observed disparities. Therefore, we sought to measure the relative importance of race compared to health care and social factors on prostate cancer specific mortality.

**Materials and Methods:** Using the SEER (Surveillance, Epidemiology, and End Results) database we identified 514,878 men diagnosed with prostate cancer at age 40 years or greater between 2004 and 2012. We also selected a subset of black and white men matched by age, stage and birth year. We stratified patients by age 40 to 54, 55 to 69 and 70 years or older and disease

\*Correspondence: Departments of Surgery and Population Science, University of Utah, 2000 Circle of Hope, Room 1501, Salt Lake City, Utah 84112 (Heidi.hanson@hci.utah.edu).

<sup>†</sup>Equal study contribution.

The corresponding author certifies that, when applicable, a statement(s) has been included in the manuscript documenting institutional review board, ethics committee or ethical review board study approval; principles of Helsinki Declaration were followed in lieu of formal ethics committee approval; institutional animal care and use committee approval; all human subjects provided written informed consent with guarantees of confidentiality; IRB approved protocol number; animal approved project number.

No direct or indirect commercial, personal, academic, political, religious or ethical incentive is associated with publishing this article.

stage, resulting in 18 groups. By applying random forest methods with variable importance measures we analyzed 15 variables and interactions across 4 categories of factors (tumor characteristics, race, and health care and social factors) and the relative importance for prostate cancer specific mortality.

**Results:** Tumor characteristics at diagnosis were the most important factors for prostate cancer mortality. Across all groups race was less than 5% as important as tumor characteristics and only more important than health care and social factors in 2 of the 18 groups. Although race had a significant impact, health care and social factors known to be associated with racial disparities had greater or similarly important effects across all ages and stages.

**Conclusions:** Eradicating disparities in prostate cancer survival will require a multipronged approach, including advances in precision medicine. Disparities will persist unless health care access and social equality are achieved among all populations.

### Keywords

prostatic neoplasms; SEER program; race factors; mortality; health services accessibility

---

BLACK men have a 1.6-fold higher incidence of PCa and more than twice the rate of PCa mortality compared to white men in the United States.<sup>1</sup> The PCa death rate has declined since the 1990s, a phenomenon commonly attributed to improved diagnostic and treatment regimens.<sup>2,3</sup> However, the black/white PCa mortality differential increased during this time, in contrast to the reduced differential of overall cancer mortality.<sup>4</sup> Genetic differences in PCa susceptibility are likely not the sole contributors to disparities but the disparities are likely due to a combination of genetic and environmental factors, including underlying social and health care factors.

Genetic differences in PCa susceptibility and tumor behavior are associated with increased PCa mortality in black men.<sup>5,6</sup> Additionally, black men have higher PSA values with equivalently staged cancer.<sup>7</sup> Patients with low risk disease are at increased risk for clinical upstaging, progression to treatment on active surveillance,<sup>8</sup> adverse pathology findings after prostatectomy<sup>9</sup> and mortality,<sup>10</sup> suggesting that identifying genetic factors leading to differential mortality would potentially improve treatment and survival but only in patients with access to care.

PCa racial disparities due to exogenous factors, such as health care access, are well documented and may fully explain the differences.<sup>11-14</sup> Compared to white men, black men have higher stage disease at diagnosis,<sup>15</sup> particularly in younger age groups.<sup>16</sup> It is necessary to understand the impact and complex interplay of disease specific characteristics, race, health care access and quality, and other social factors to identify the factors determining PCa mortality and improve outcomes in black men.

Identifying the most important factors for predicting PCa mortality requires methodological approaches which simultaneously consider multiple variables and their interactions. Classic regression methods use reductionist approaches and investigate risk factors in isolation or specify a priori interactions to test for differences in the effect of predictors. Random forest regression, which is a supervised machine learning method, incorporates all multiway

interactions and provides measures of variable importance, quantifying the importance of each variable for predicting survival classes.<sup>17</sup>

Using cohorts from the NCI (National Cancer Institute) SEER Program, we took advantage of random forest regression to identify strong predictors of PCa mortality. We hypothesized that tumor characteristics (Gleason score and PSA) would strongly predict PCa mortality among men with all stages of disease. We also hypothesized that race would have a measurable impact on mortality, although a smaller impact than tumor characteristics, and it would be similarly as impactful as health care factors (access and quality) and social factors (median family income, population density and social vulnerability).

## METHODS

Data for this study were obtained from the 18 registries comprising the SEER database.<sup>18</sup> This study includes all 514,878 males more than 40 years old at diagnosis who were diagnosed with PCa from 2004 to 2012. County level health care data were obtained from the HHS (Department of Health and Human Services) AHRF, the CDC (Centers for Disease Control and Prevention) and the AHRQ (Agency for Healthcare Research and Quality). County level measurements were based on the county of residence of an individual at diagnosis.

### Sample Selection

Figure 1 shows the sample selection tree. We excluded individuals who were diagnosed at death or autopsy. We required individuals to be linked to the AHRF and have a known stage at diagnosis and a link to county level SES measures.

### Measures

SEER data were obtained for individual level measures. Time was measured as months from diagnosis to PCa death and observations were censored at death from another cause or the end of followup. Race/ethnicity was categorized as black, nonHispanic white, Hispanic white, other race or unknown race/ethnicity. Hispanic ethnicity of nonwhite races was grouped with the corresponding race due to small cell sizes. The supplementary material (<https://www.jurology.com>) provides a detailed description of stage, grade and PSA measures.

Measures of county level access to health care were drawn from the AHRF database, including the number of physicians, radiation oncologists and urologists, and the number of chemotherapy treatment centers. The AHRQ PQI was used to measure the quality of ambulatory care. The supplementary material (<https://www.jurology.com>) shows a detailed description and a map of the spatial distribution.

Social factors which may affect PCa mortality were derived from multiple sources. The rural/urban designation was obtained from the AHRF. The county level median family income from the 2000 United States Census was included. We considered the proportion of the Medicare population in each county which was also eligible for Medicaid enrollment (dual enrollees).<sup>19</sup> The CDC Social Vulnerability Index, a measure of resilience of

communities confronted by external stresses on health (natural or human caused disasters, disease outbreaks, etc) was also included.<sup>20</sup>

### Statistical Methods

Random forests are an extension of CART (Classification and Regression Tree) modeling. In this method  $n$  trees are grown using bootstrapped samples from the learning sample.<sup>17,21</sup> As these trees grow, each branch point (node) separates groups based on the factor which best explains the variability in mortality in the groups. For example, if a Gleason score greater than 8 explains more variance in mortality in the group compared to PSA greater than 20 ng/ml, the next node separates the single group into 2 groups based on Gleason score regardless of PSA and other factors. It identifies the most relevant variables when multicollinearity is present and reduces the variables of interest to those with the most explanatory value.<sup>22</sup>

Trees are a random 70% subset of the original data set. Therefore, the remaining 30% of the data which were not selected (eg OOBs) could be used to calculate variable importance. OOB data create permutation accuracy VIMPs by predicting class membership in the OOB sample and then randomly permuting the values and calculating the decrease in predictive accuracy with permuted variables. The average difference in accuracy of the OOB vs the permuted OOB observations over all trees is the VIMP with a VIMP close to zero implying that the variable has no predictive power. The increasing value of the VIMP implies increasing importance for predicting the outcome.

Models were stratified by patient age at diagnosis (40 to 54, 55 to 69 and 70 years or greater) and stage at diagnosis (I/II, III and IV). This allowed for examination of differences in stage, thus netting out racial differences in PCa survival due to stage. Analyses were run in R using the randomForestSRC and survival packages (<https://cran.r-project.org/web/packages/>). We performed our analysis in a subset of the sample in which all black individuals were matched to a nonHispanic white counterpart by birth year, and stage and age at diagnosis. This allowed us to compare the VIMP without biasing it by the unequal distribution of race in the total population and directly compare nonHispanic white men to black men.

Each model included 15 variables. For descriptive purposes the variables were categorized broadly as tumor characteristics, social factors, health care factors and race. To measure relative importance we summed the VIMPs across categories and compared them to the sum of tumor characteristic VIMPs to standardize across models stratified by age and stage. All variables were categorical to avoid a bias in VIMPs due to the inclusion of categorical and continuous variables in a single model. We repeated these analyses in the total patient cohort to measure the impact of race more generally.

## RESULTS

The table lists statistics on all variables. A higher proportion of men of Hispanic, black and other races was diagnosed with stage IV disease relative to nonHispanic white men. Men of other races were more likely to be diagnosed with Gleason 7+ disease and PSA greater than

40 ng/ml than nonHispanic white men. A higher proportion of Hispanic and black men lived in areas in the bottom 25th percentile of median family income and the black racial group had the highest percent of the population living in the bottom 10th percentile. Black and Hispanic men were more likely to live in socially vulnerable areas. There were minor racial differences in the number of physicians by county but stark differences in the PQI with nearly twice the black population living in low PQI areas relative to nonHispanic white populations. PCa was diagnosed in 55,493 men 40 to 54 years old, in 282,358 who were 55 to 69 years old and in 174,213 who were 70 years old or older. There was more racial heterogeneity in younger age groups relative to men 70 years old or older.

Tumor characteristics (Gleason score and PSA) were the most important predictors of survival across all ages and stages. The supplementary material (<https://www.jurology.com>) shows the range of importance of each factor. Figure 2 shows VIMPs of social factors, health care factors and race relative to the VIMP of tumor characteristics by age and stage (1/2, 3 and 4). Figure 2, A includes results of the matched subset analysis of black and nonHispanic white men. Figure 2, B includes results in the broader population.

Across analyses and groups race was never greater than 1/20th as important as tumor characteristics. In fact, in many subgroups, including younger patients and patients with stage 4 disease, race was not a measurably important factor for mortality. Race was only more important than health care and social factors in 2 of the 18 groups, including stage 3, and ages 55 to 69 and 70 years or greater on matched subset analysis. Health care factors were comparatively important predictors of PCa mortality across all groups. Social factors were important predictors of mortality in every group except the stage 3, age 40 to 54-year age group on the matched subset analysis.

The VIMP pattern differed in the overall cohort compared to the matched subset. In the matched subset social factors were more important in men with lower disease stages who were 55 to 69 years old and in general factors other than tumor characteristics were less important as stage increased. Comparatively in the full cohort health care and social factors were less important relative to tumor characteristics at lower disease stages and in general these factors increased in importance as stage increased, particularly in the younger age groups.

## DISCUSSION

Tumor characteristics were the most important factors for predicting PCa mortality at all ages and stages. All other factors combined ranged from a quarter to a tenth as important when directly compared to tumor characteristics. While race was a measurably important predictor of PCa mortality, health care and social factors (which are factors associated with racial disparities) were more important in all but 2 of the 18 groups in our analysis.

The relationship between race and PCa outcomes is complicated and it is difficult to disentangle the social and biological factors leading to disparities. For example, recent research shows that black men with Gleason 6 disease are twice as likely to die of PCa than nonblack men and the investigators suggested that there were biological differences.<sup>10</sup>

However, results from the SEARCH (Shared Equal Access Regional Cancer Hospital) database are conflicting, showing that lower SES black men with equal access to care have decreased PCa mortality.<sup>23</sup> Dess et al also found no racial difference in PCa mortality when access to care and standardized treatment were equal.<sup>24</sup>

Genetic heterogeneity and its relationship to PCa outcomes is an important question. However, it should not be couched in the framework of racial disparities. Identifying genetic factors contributing to differences in tumor aggressiveness should be blinded to race, which is a social construct. Rather, genetic differences in outcomes should be studied across subgroups of the population in a biological framework. Known genetic differences which lead to differences in outcomes, such as transforming growth factor  $\beta$ ,<sup>5</sup> should be screened for across all populations. The interaction between biology and social aspects at multiple levels is important but care should be taken not to conflate the 2 constructs. Personalized approaches to health care based on genetics can only equalize outcomes if all populations have the same access to care.

Our results support research showing that efforts to equalize access to care will have a great impact on reducing racial disparities in PCa outcomes. Further research is needed to identify relative contributions of factors across multiple levels, including the patient, the provider and the health care system. The same care should be given to all individuals. Urological care deserts should be defined by multiple factors, not solely distance to care, to target intervention efforts in resource depleted areas.

Leveraging random forests enabled us to explore the complex associations between multiple variables and measure the importance of each variable in predicting PCa mortality. This highly predictive modeling approach considers all possible combinations of variable interactions and addresses multicollinearity issues. This data driven, reverse engineering approach allowed us to disentangle the role of individual and neighborhood factors correlating with race/ethnicity and to identify the strength of that contribution to PCa outcomes. These models may have better predictive power than traditional regression methods and they can identify complex hidden relationships, especially for questions with multicollinearity between measures for which it is difficult to determine the importance of effects.<sup>25</sup>

There are challenges in measuring the impact of race on PCa mortality. It is important to note that the methods used in this analysis were designed to optimize prediction and not assess causation. The OOB error rate, which ranged from 0.15 to 0.25 for the different age groups and disease stages, suggests that the factors which we considered explain large amounts of variation in PCa mortality in this cohort. This suggests that while race is often an important predictor of PCa mortality, risk is also tied up in social forces which influence mortality. The variation between the matched subset analysis and the cohort overall suggests that the pattern of racial disparities among other groups may differ from those found in black men. Whatever these specific patterns may be, clearly the methods used to reduce racial disparities in PCa outcomes must involve the elimination of racial disparities for social and health care factors.

There are important considerations when interpreting our analysis. While it is likely that fundamental differences exist in regard to PCa incidence and disease behavior, it is difficult for clinicians to incorporate racial factors when clinically treating patients with PCa, particularly when considering confounding factors such as access to and quality of care and the paucity of clinical research devoted to black patients. SEER, which is the only comprehensive, population based cancer database in the United States, represents a unique opportunity to study ethnic and racial differences in PCa mortality.

However, certain limitations must be considered when interpreting the results of observational studies and SEER data in particular. Race and ethnicity are self-reported variables and, thus, subject to sociological and historical factors shaping self-identification into categories.<sup>26</sup> County level income and access to care measurements do not directly measure individual access but provide information about the interplay between area and individual level characteristics. Future studies should incorporate individual level measures of access to care and SES.

Our analyses do not suggest that race is unimportant with regard to PCa mortality. Our cohort consisted of patients who were already diagnosed with PCa and our analysis did not consider differences in PCa incidence or variation in stage at diagnosis between races, each of which is a known racial disparity.<sup>1,15,16</sup> However, in patients with equivalent disease specific factors it appears that health care and social factors associated with race are as important, if not more important, when considering PCa specific mortality.

## CONCLUSIONS

Attempting to reduce racial disparities through personalized, medicine driven approaches without addressing racial disparities in social and health care factors may close the gap for economically advantaged patients while reinforcing disparities for disadvantaged persons. Although race had a significant impact in many groups, health care and social factors had greater or similarly important effects across most disease stages and age groups. Reducing disparities in social and health care factors may provide large gains in improving PCa mortality and minimizing racial disparities.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

Supported by the Genitourinary Malignancies Disease-Oriented Team, Huntsman Cancer Institute Cancer Center Support Grant P30CA042014 and NIH (National Institutes of Health) Grant 1K12HD085852-01.

## Abbreviations and Acronyms

<b>AHRF</b>	Area Health Resource File
<b>OOB</b>	out of bag observation
<b>PCa</b>	prostate cancer



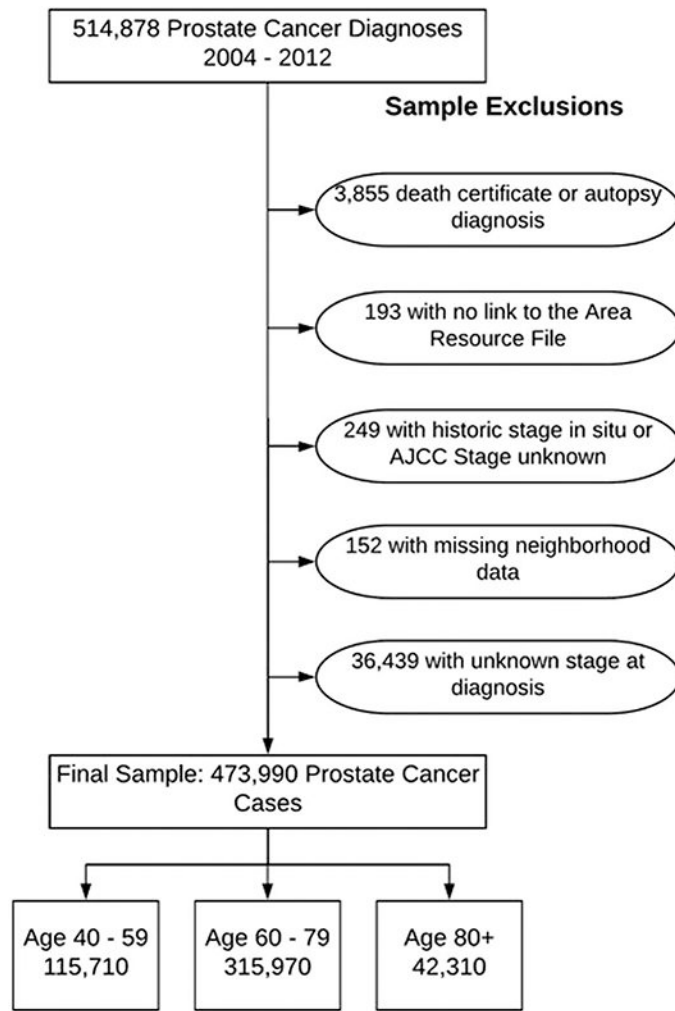
<b>PQI</b>	Physician Quality Index
<b>PSA</b>	prostate specific antigen
<b>SEER</b>	Surveillance, Epidemiology, and End Results
<b>SES</b>	socioeconomic status
<b>VIMP</b>	variable importance measure

## REFERENCES

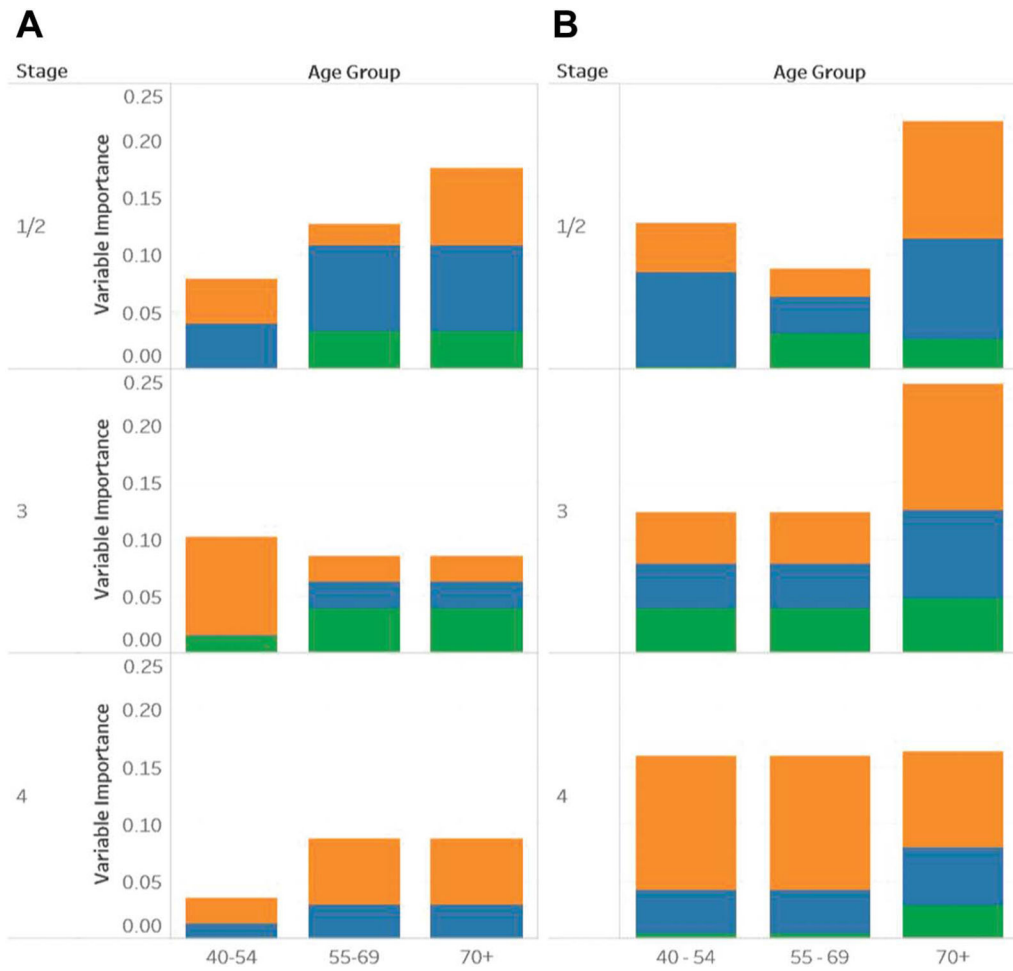
1. Siegel RL, Miller KD and Jemal A: Cancer statistics, 2018. *CA Cancer J Clin* 2018; 68: 7. [PubMed: 29313949]
2. Hankey BF, Feuer EJ, Clegg LX et al.: Cancer surveillance series: interpreting trends in prostate cancer—part I: evidence of the effects of screening in recent prostate cancer incidence, mortality, and survival rates. *J Natl Cancer Inst* 1999; 91: 1017. [PubMed: 10379964]
3. Chu KC, Tarone RE and Freeman HP: Trends in prostate cancer mortality among black men and white men in the United States. *Cancer* 2003; 97: 1507. [PubMed: 12627516]
4. DeSantis CE, Siegel RL, Sauer AG et al.: Cancer statistics for African Americans, 2016: progress and opportunities in reducing racial disparities. *CA Cancer J Clin* 2016; 66: 290. [PubMed: 26910411]
5. Elliott B, Zackery DL, Eaton VA et al.: Ethnic differences in TGFbeta-signaling pathway may contribute to prostate cancer health disparity. *Carcinogenesis* 2018; 39: 546. [PubMed: 29474521]
6. Lachance J, Berens AJ, Hansen MEB et al.: Genetic hitchhiking and population bottlenecks contribute to prostate cancer disparities in men of African descent. *Cancer Res* 2018; 78: 2432. [PubMed: 29438991]
7. Vijayakumar S, Winter K, Sause W et al.: Prostate-specific antigen levels are higher in African-American than in white patients in a multicenter registration study: results of RTOG 94-12. *Int J Radiat Oncol Biol Phys* 1998; 40: 17. [PubMed: 9422553]
8. Abern MR, Bassett MR, Tsivian M et al.: Race is associated with discontinuation of active surveillance of low-risk prostate cancer: results from the Duke Prostate Center. *Prostate Cancer Prostatic Dis* 2013; 16: 85. [PubMed: 23069729]
9. Sundi D, Ross AE, Humphreys EB et al.: African American men with very low-risk prostate cancer exhibit adverse oncologic outcomes after radical prostatectomy: should active surveillance still be an option for them? *J Clin Oncol* 2013; 31: 2991. [PubMed: 23775960]
10. Mahal BA, Berman RA, Taplin ME et al.: Prostate cancer-specific mortality across Gleason scores in black vs nonblack men. *JAMA* 2018; 320: 2479. [PubMed: 30561471]
11. McClelland S, Page BR, Jaboin JJ et al.: The pervasive crisis of diminishing radiation therapy access for vulnerable populations in the United States, part 1: African-American patients. *Adv Radiat Oncol* 2017; 2: 523. [PubMed: 29204518]
12. Schwartz K, Powell IJ and Underwood W III et al.: Interplay of race, socioeconomic status, and treatment on survival of patients with prostate cancer. *Urology* 2009; 74: 1296. [PubMed: 19962532]
13. Bernard B, Muralidhar V, Chen YH et al.: Impact of ethnicity on the outcome of men with metastatic, hormone-sensitive prostate cancer. *Cancer* 2017; 123: 1536. [PubMed: 28055108]
14. Bach PB, Schrag D, Brawley OW et al.: Survival of blacks and whites after a cancer diagnosis. *JAMA* 2002; 287: 2106. [PubMed: 11966385]
15. Powell IJ, Bock CH, Ruterbusch JJ et al.: Evidence supports a faster growth rate and/or earlier transformation to clinically significant prostate cancer in black than in white American men, and influences racial progression and mortality disparity. *J Urol* 2010; 183: 1792. [PubMed: 20299055]
16. He T and Mullins CD: Age-related racial disparities in prostate cancer patients: a systematic review. *Ethn Health* 2017; 22: 184. [PubMed: 27706949]
17. Breiman L: Random forests. *Mach Learn* 2001; 45: 5.



18. Mérette C, King MC and Ott J: Heterogeneity analysis of breast cancer families by using age at onset as a covariate. *Am J Hum Genet* 1992; 50: 515. [PubMed: 1539592]
19. Kaiser Family Foundation: Kaiser Commission on Medicaid and the Uninsured and Urban Institute estimates based on data from FY 2008 MSIS, 2012. Available at <https://kaiserfamilyfoundation.files.wordpress.com/2013/04/medicaid-its-role-today-and-under-the-affordable-care-act-jama-medicaid-082212.pptx>. Accessed June 12, 2019.
20. Flanagan BE, Gregory EW, Hallisey EJ et al.: A social vulnerability index for disaster management. *J Homel Secur Emerg Manag* 2011; 8: 1.
21. Ishwaran H, Kogalur UB, Blackstone EH et al.: Random survival forests. *Ann Appl Stat* 2008; 2: 841.
22. Archer KJ and Kimes RV: Empirical characterization of random forest variable importance measures. *Comput Stat Data Anal* 2008; 52: 2249.
23. Everist MM, Howard LE, Aronson WJ et al.: Socioeconomic status, race, and long-term outcomes after radical prostatectomy in an equal access health system: results from the SEARCH database. *Urol Oncol* 2018; 37: 289.e17.
24. Dess RT, Hartman HE, Mahal BA et al.: Association of black race with prostate cancer-specific and other-cause mortality. *JAMA Oncol* 2019; 5: 975. [PubMed: 31120534]
25. Chen X and Ishwaran H: Random forests for genomic data analysis. *Genomics* 2012; 99: 323. [PubMed: 22546560]
26. Bryc K, Durand EY, Macpherson JM et al.: The genetic ancestry of African Americans, Latinos, and European Americans across the United States. *Am J Hum Genet* 2015; 96: 37. [PubMed: 25529636]



**Figure 1.** Sample selection criteria. *AJCC*, American Joint Committee on Cancer.



**Figure 2.** Summary of relative importance of VIMPs compared to tumor characteristic VIMPs. Mortality factors include health care factors (orange bars), social factors (blue bars) and race/ethnicity (green bars). *A*, balanced subset of black and white men with prostate cancer. *B*, full cohort.

## Full sample descriptive statistics

	No. NonHispanic White (%)	No. Hispanic (%)	No. Black (%)	No. Other (%)	No. Unknown (%)
Overall	353,324	43,113	77,440	25,703	12,484
<i>Individual level variables</i>					
Stage:					
I/II	300,178 (84.96)	35,561 (82.48)	66,158 (85.43)	21,158 (82.32)	11,834 (94.79)
III	28,397 (8.04)	3,429 (7.95)	4,756 (6.14)	2,253 (8.77)	339 (2.72)
IV	24,749 (7)	4,123 (9.56)	6,526 (8.43)	2,292 (8.92)	311 (2.49)
PCa mortality:					
Alive or dead of other cause	334,721 (94.73)	40,592 (94.15)	72,526 (93.65)	24,392 (94.9)	12,274 (98.32)
PCa death	18,603 (5.27)	2,521 (5.85)	4,914 (6.35)	1,311 (5.1)	210 (1.68)
<i>Tumor characteristics</i>					
Gleason category:					
6 or Less	155,508 (44.01)	19,041 (44.17)	30,736 (39.69)	9,825 (38.23)	6,401 (51.27)
7 (3, 4)	91,971 (26.03)	10,287 (23.86)	21,671 (27.98)	6,147 (23.92)	2,773 (22.21)
7 (4, 3)	37,701 (10.67)	4,416 (10.24)	8,913 (11.51)	3,066 (11.93)	1,182 (9.47)
8	28,539 (8.08)	3,768 (8.74)	7,258 (9.37)	2,881 (11.21)	970 (7.77)
9 + 10	25,035 (7.09)	3,253 (7.55)	5,402 (6.98)	2,559 (9.96)	572 (4.58)
998, 999	14,570 (4.12)	2,348 (5.45)	3,460 (4.47)	1,225 (4.77)	586 (4.69)
PSA category (ng/ml):					
Less than 40	234,614 (66.4)	25,836 (59.93)	45,682 (58.99)	15,248 (59.32)	6,693 (53.61)
40–99	43,509 (12.31)	6,619 (15.35)	11,647 (15.04)	4,482 (17.44)	1,463 (11.72)
100–199	32,601 (9.23)	5,818 (13.49)	11,629 (15.02)	3,633 (14.13)	935 (7.49)
200	42,600 (12.06)	4,840 (11.23)	8,482 (10.95)	2,340 (9.1)	3,393 (27.18)
<i>Social factors</i>					
Statistical area:					
Rural	21,560 (6.1)	397 (0.92)	2,702 (3.49)	109 (0.42)	179 (1.43)
Metropolitan	303,961 (86.03)	41,278 (95.74)	70,341 (90.83)	24,283 (94.48)	11,972 (95.9)
Metropolitan	27,803 (7.87)	1,438 (3.34)	4,397 (5.68)	1,311 (5.1)	333 (2.67)
Median family income:					
Bottom 10%	15,221 (4.31)	1,682 (3.9)	5,635 (7.28)	280 (1.09)	165 (1.32)

	No. NonHispanic White (%)	No. Hispanic (%)	No. Black (%)	No. Other (%)	No. Unknown (%)
10%–25%	97,305 (27.54)	19,939 (46.25)	25,017 (32.31)	6,879 (26.76)	3,666 (29.37)
Top 75%	240,798 (68.15)	21,492 (49.85)	46,788 (60.42)	18,544 (72.15)	8,653 (69.31)
Dual enrollment:					
Bottom 90%	315,512 (89.3)	41,195 (95.55)	72,566 (93.71)	23,409 (91.07)	11,621 (93.09)
Top 10%	37,812 (10.7)	1,918 (4.45)	4,874 (6.29)	2,294 (8.93)	863 (6.91)
Social vulnerability:					
Low (less than 0.75)	254,635 (72.07)	18,445 (42.78)	39,282 (50.73)	17,898 (69.63)	7,307 (58.53)
High (0.75 or greater)	98,689 (27.93)	24,668 (57.22)	38,158 (49.27)	7,805 (30.37)	5,177 (41.47)
<i>Health care access</i>					
No. urologists/100,000 people:					
0	44,259 (12.53)	1,721 (3.99)	7,565 (9.77)	380 (1.48)	459 (3.68)
1+	309,065 (87.47)	41,392 (96.01)	69,875 (90.23)	25,323 (98.52)	12,025 (96.32)
No. chemotherapy treatment centers:					
1+	301,753 (85.4)	39,613 (91.88)	67,240 (86.83)	24,661 (95.95)	11,832 (94.78)
0	51,571 (14.6)	3,500 (8.12)	10,200 (13.17)	1,042 (4.05)	652 (5.22)
No. physicians:					
Bottom 25%	97,601 (27.62)	7,958 (18.46)	19,587 (25.29)	2,031 (7.9)	2,025 (16.22)
Middle 50%	168,124 (47.58)	28,887 (67)	32,344 (41.77)	15,384 (59.85)	7,738 (61.98)
Top 25%	87,599 (24.79)	6,268 (14.54)	25,509 (32.94)	8,288 (32.25)	2,721 (21.8)
No. radiation oncologists:					
Bottom 25%	98,199 (27.79)	8,350 (19.37)	17,975 (23.21)	2,916 (11.34)	2,260 (18.1)
Middle 50%	160,290 (45.37)	28,196 (65.4)	35,140 (45.38)	17,296 (67.29)	7,656 (61.33)
Top 25%	94,835 (26.84)	6,567 (15.23)	24,325 (31.41)	5,491 (21.36)	2,568 (20.57)
Physician Quality Index (county level):					
Top 75%	309,101 (87.48)	40,331 (93.55)	57,501 (74.25)	24,843 (96.65)	11,385 (91.2)
Bottom 25%	44,223 (12.52)	2,782 (6.45)	19,939 (25.75)	860 (3.35)	1,099 (8.8)
African-American top 75%	254,968 (72.16)	26,932 (62.47)	43,509 (56.18)	17,809 (69.29)	8,263 (66.19)
African-American bottom 25%	98,356 (27.84)	16,181 (37.53)	33,931 (43.82)	7,894 (30.71)	4,221 (33.81)

For individual level variables, tumor characteristics, social factors and health care access chi-square test  $p < 0.001$ .