



Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2016 January ; 25(1): 83–89. doi:10.1158/1055-9965.EPI-15-0456.

The role of stage at diagnosis in colorectal cancer black-white survival disparities: a counterfactual causal inference approach

Linda Valeri^{1,2}, Jarvis T. Chen³, Xabier Garcia-Albeniz⁴, Nancy Krieger³, Tyler J. VanderWeele^{4,5}, and Brent A. Coull⁵

¹Psychiatric Biostatistics Laboratory, McLean Hospital, Belmont MA USA

²Harvard Medical School, Boston MA USA

³Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston MA USA

⁴Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston MA USA

⁵Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston MA USA

Abstract

Background—To date, a counterfactual framework has not been used to study determinants of social inequalities in cancer. Considering the case of colorectal cancer (CRC), for which racial/ethnic differences in stage at diagnosis and survival are well-documented, we quantify the extent to which black versus white survival disparities would be reduced had disparities in stage at diagnosis been eliminated in a large patient population.

Methods—We obtained data on CRC patients (diagnosed between 1992–2005 and followed until 2010) from US-SEER cancer registries. We employed a counterfactual approach to estimate the mean survival time up to the 60th month since diagnosis for black CRC patients had black-white disparities in stage at diagnosis been eliminated.

Results—Black patients survive approximately 4.0 (CI=4.6,3.2) months less than White patients within five years since diagnosis. Had disparities in stage at diagnosis been eliminated, survival disparities decrease to 2.6 (CI=3.4,1.7) months, a ~35% reduction. For patients diagnosed after the age of 65 disparities would be halved, while reduction of ~30% is estimated for younger patients. Survival disparities would be reduced by ~44% for women and ~26% for men.

Conclusions—Employing a counterfactual approach and allowing for heterogeneities in black-white disparities across patients' characteristics, we give robust evidence that elimination of disparities in stage at diagnosis contributes to a substantial reduction in survival disparities in CRC.

Impact—We provide the first evidence in the SEER population that elimination of inequities in stage at diagnosis might lead to larger reductions in survival disparities among elderly and women.

Keywords

Colorectal cancer; black-white disparities; survival; stage at diagnosis; counterfactual framework

INTRODUCTION

Use of a counterfactual causal inference framework is recognized as a valuable contribution to quantifying the causal effects of potential interventions [1]. To our knowledge, however, this framework has not been applied to analysis of the contribution of stage at diagnosis to social inequalities in cancer outcomes. Here, we adopt a counterfactual causal inference framework to investigate determinants of black-white disparities as proposed by VanderWeele and Robinson [2], framed by an understanding that inequitable race relations, not “race” per se, are the cause of racial/ethnic health inequities, that is, unjust, unfair, and preventable social inequalities in health [3–8].

The case we have chosen is colorectal cancer (CRC), as black-white disparities in both stage at diagnosis and survival are well documented. Differences in CRC mortality rates between black and white patients started manifesting in 1980 [9] and the black-white mortality rate ratio peaked for women in 2005 at 1.5 and for men in 2006 at 1.6 [10]. Since CRC is the third leading cause of cancer death in both men and women [11], to reduce the uneven burden of cancer in the United States it is crucial to investigate rigorously why these disparities exist.

The determinants of CRC mortality disparities are multifactorial, with black-white disparities often attributed to differences in socioeconomic and behavioral factors, as well as in quality and access to prevention practices and to care. Since the 1970’s blacks have lagged behind whites in access to both screening and improved treatment options [12]. Therefore, many have suggested that the large differences in stage at diagnosis and poorer prognosis for distant-stage disease may be driving mortality disparities [12].

Mixed evidence is found on black-white disparities in tumor characteristics [13,14] and differences in tumor location are not considered to be main drivers of mortality disparities [14].

Previous studies reported that differences in stage at diagnosis accounted for up to 60% of the observed mortality disparities [15–17]. These studies, however, were conducted in smaller patient populations and using regression approaches valid only under strong assumptions, which are violated in this context. In particular, a popular approach to estimate the extent to which racial/ethnic disparities in survival is explained by intermediate factors is to compare estimates of the effect of interest from a regression model with and without adjustment for the hypothesized factor. As is now well documented in the epidemiological and biostatistics literature, this approach is biased if there are interactions between race/ethnicity and stage at diagnosis [18].

In this paper, we apply a counterfactual framework, using data from the National Cancer Institute’s (NCI) Surveillance, Epidemiology, and End Results (SEER) Program, to estimate

the extent to which black-white survival disparities among CRC patients would be reduced had disparities in the intermediate factor, stage at diagnosis, been eliminated. We estimate a clinically meaningful, model-free measure of survival to quantify black-white disparities and the impact of this hypothetical intervention across baseline patients' characteristics, providing important insights to patients, clinicians, and policy makers. In particular, we allow for interactions in race and stage in their impact on survival and, recognizing that people embody diverse forms of inequality simultaneously, we investigate differences in survival disparities according to age and gender.

MATERIALS AND METHODS

Study population

We obtained data from NCI's SEER Program, which provides information on tumor site, stage, and individual demographic information for cancers occurring in approximately 28% of the US population. Patients from the Atlanta, Connecticut, Detroit, Iowa, New Mexico, San Francisco-Oakland, Seattle-Puget Sound, Utah, San Jose-Monterey, Los Angeles and Rural Georgia registries were included in the study. The California registries were combined and the Rural Georgia and Atlanta registries were combined. The study population consisted of non-Hispanic white and non-Hispanic black cancer patients diagnosed between 1992–2005 and followed until 2010. We selected patients diagnosed above the age of 18, with microscopically confirmed adenocarcinoma histology and CRC as first diagnosis. We included patients diagnosed at stage I–IV according to the American Joint Committee on Cancer (AJCC) staging criteria [19]. Patients from the Hawaii registry were excluded from the analysis due to the low percentage of non-Hispanic blacks. Finally, we excluded patients for whom the reporting source was nursing home/hospice, autopsy or death certificate. Supplementary Figure S1 and Supplementary Table S1 display the flowchart of eligibility criteria for the present study and the codes we used to extract information about tumor characteristics from SEER data. Census and American Community Survey (ACS) data were merged to SEER data to obtain area-based measures of socioeconomic position [20–21].

Stage at diagnosis was categorized as Stage I–IV, according to the AJCC staging criteria. For the purpose of the present analyses, patients with in situ cancers and unstaged cancers were excluded from the primary analyses. Approximately 6% of the sample had unstaged cancer at diagnosis (n=9,192).

Survival time in months from date of diagnosis was calculated under the assumption that, if last known vital status was alive, then the patients were presumed alive and the database study cut-off date was used as date of last contact. Survival time was censored for 27% of black and 30% of white population. Individuals who died within 5 years since diagnosis had survival time information reliably recorded, since registries conduct active patients' follow-up for at least 5 years.

Age at diagnosis, categorized as age<50, 50–65, and >65, year at diagnosis (1992–1995, 1996–2000, 2001–2005), gender (men versus women), grade of tumor differentiation (I–IV), and registry were considered as potential confounders of the stage-survival association. Because individual-level SEP data are not available in the SEER registry, median household

income at county level, an aggregate measure of SEP, was linearly interpolated from 1990 and 2000 Census, and 2009 American Community Survey data and categorized in quintiles.

Statistical analysis

The distribution of baseline characteristics between the two groups was assessed for differences using chi-square statistics. For non-Hispanic whites and non-Hispanic blacks within each confounder stratum (denoted by x), we estimated non-parametrically the restricted mean survival time (RMST) up to 60 months of follow up. This quantity is estimated as the area under the Kaplan-Meier curve from diagnosis to 60 months [22]. We denote the difference of RMST for the blacks versus whites as the *disparity* in survival.

$$Disparity(x) = RMST(black, x) - RMST(white, x) \quad (\text{eq.1})$$

The overall disparity measure was then obtained via random effect meta-analysis [23–24] across all confounder strata (x) to provide more precise inferences by borrowing strength while allowing for heterogeneous underlying disparity measures in each stratum.

We estimated the *residual disparity* between black and white individuals that would be observed had we intervened on stage at diagnosis. The causal contrast can be formally defined as the difference of the RMST in black versus white patients had we intervened to shift the stage at diagnosis distribution among blacks to match that of the whites, the most advantaged patients [2]. Effectively, the residual disparity is calculated as the difference in the RMST for blacks versus whites once the stage-specific RMST's for black patients are multiplied by the probability of being diagnosed at that stage in the reference population (white patients), adjusting for confounders of the stage-survival relationship. The quantity is estimated using the following standardization formula:

$$Residual\ Disparity(x) = \sum_{stage} \{RMST(black, stage, x) - RMST(white, stage, x)\} p(stage|white, x) \quad (\text{eq. 2})$$

To compute the residual disparity after the hypothetical intervention on stage, we estimated non-parametrically the area under the Kaplan Meier curve, truncated at month 60, for white and black individuals in each *stage* and confounder (x) strata; the probability of being diagnosed at a certain stage in the white population was also empirically estimated. Equation 2 correctly estimates the effect of a shift in the distribution of stage at diagnosis in the black population towards that of the white population under the assumption that all confounders of the stage-survival relationship are adjusted for in the analysis [2]. The overall *residual disparity* measure was again obtained via random effect meta-analysis across all covariate strata.

Finally, percent disparity reduction was estimated as:

$$\%Disparity\ Reduction = \frac{disparity - residual\ disparity}{disparity} \quad (\text{eq.3})$$

Data were analyzed using the R software package [25].

Sensitivity Analyses

We explored the robustness of our results to selection bias introduced by missing data on stage at diagnosis. Information on stage at diagnosis was missing for ~6% of the sample. Although the number of unstaged patients was not large, we observed an important difference in the 5-year survival probability between individuals with complete information and individuals with missing this information (50% vs 35%). The survival probability and Kaplan-Meier curves for unstaged patients in the present study lied between those calculated for stage III and stage IV patients. Therefore, we assessed the sensitivity of our results to the extreme assumption that all unstaged individuals had stage IV at diagnosis.

RESULTS

Our analytic sample consists of 166,722 colorectal cancer patients (148,010 non-Hispanic whites and 18,712 non-Hispanic blacks) with no missing age at diagnosis. As shown in Table 1, black and white individuals display significant differences in many of the baseline characteristics. White individuals display a survival advantage (log rank test p-value<0.0001) and are more likely than black patients to survive past 5 years from diagnosis (observed 5-year survival from death due to all causes: 49% vs 40%, p<0.001). Blacks are more likely to be diagnosed at stage IV than whites (24% vs 17%; p<0.001).

In the SEER population we find that RMST by stage at diagnosis for whites differs substantially from that estimated for black patients. For white patients diagnosed at stage I, the area under the Kaplan-Meier curve is 51.7 months: that is, one expects a white CRC patient diagnosed at stage I to be alive for 51.7 months of the 60 months of follow-up. In contrast, the RMST for blacks is 48.5 months. For patients diagnosed at stage II, we observe RMST of 47.9 for whites vs. 45.5 months for blacks; for stage III patients, 42.6 months for whites vs. 41.8 months for blacks; and for stage IV patients, 18.3 months for whites vs. 15.5 months for blacks. In this population we find the survival differences to be larger among patients diagnosed at stage I (Table 2). Race-stage interaction was significant in an accelerated failure time model, indicating that disparities were significantly larger for patients diagnosed at stage I (LRT p-value<0.001, Supplementary Table S2). Figure 1 summarizes the estimates of the observed RMST for black and white patients, and the counterfactual RMST for black patients under the scenario of a shift of stage at diagnosis distribution. The estimates are reported by age at diagnosis. Individuals diagnosed before the age of 65 survive longer than older patients. Comparing the histograms on the first row (RMST's for blacks) with the histograms on the second row (RMST's for whites) we see that whites experience longer survival. Comparing the histograms on the second row (RMST's for whites) with the histograms on the third row (RMST's for blacks after the hypothetical intervention) we observe a shift of the RMST's for the black population towards that of the white; however, the disparity is not eliminated.

Figure 2 displays the density of the estimated *disparity* (black line) and *residual disparity* (gray line) across covariate strata, separating the three age groups. The distribution of the disparity measures displays high variability. This variability reflects both the heterogeneity in the underlying disparities across covariate strata as well as the uncertainty in the estimates themselves. Greater disparities are noted for patients diagnosed before the age of 65. Had we

intervened eliminating disparities in stage at diagnosis, the residual disparity gets closer to zero but still important differences are observed.

By combining the estimates employing random effect meta-analysis, a significant *disparity* in overall survival is observed for which the difference in RMST truncated at year 5 for black versus white is -4.0 months (CI= $-4.6, -3.3$, Table 3).

The *residual disparity* in CRC survival had we intervened on the stage at diagnosis distribution is -2.6 months (CI= $-3.4, -1.7$, Table 3), indicating that elimination of disparities in stage at diagnosis may lead to a reduction of overall survival disparities in the first 5 years after diagnosis of 35% (CI=11%, 55%, Table 3). Our non-parametric analyses account for heterogeneities across the several factors. Black-white disparities are slightly larger among men and the impact of the hypothetical elimination of stage at diagnosis disparities appears stronger among women (44% vs 26%). Estimates of the disparity by age groups reveal that patients diagnosed before the age of 65 experience significant greater disparities in survival with respect to the elderly population ($disparity_{age<50} = -5.9$, $disparity_{age\ 50-65} = -5.1$ vs $disparity_{age>65} = -2.4$) and that equalizing stage at diagnosis might have more impact in the patient population diagnosed after age 65 (percent disparity reduction: 50%) with respect to the younger age groups (percent disparity reduction= $\sim 30\%$).

Sensitivity analysis

Our primary analyses appear robust to selection bias due to stage at diagnosis being not missing at random. Assuming that all unstaged individuals had stage IV at diagnosis leads to an estimate of the percent disparity reduction of 34%, similar to the primary analyses (See Supplementary Table S3).

DISCUSSION

This is the first study, to our knowledge, to employ a counterfactual causal inference approach to quantifying the role of stage in diagnosis in black-white cancer survival disparities. The method relies on the assumption that we adjust for factors that affect survival and stage at diagnosis and that differences in stage at diagnosis are mostly the result of differential access to preventive care and may be mitigated by fostering equal access to screening and follow-up practices.

Using this approach, we assessed in the SEER population of CRC patients how much black-white survival disparities would be reduced if disparities in stage at diagnosis were eliminated. We find that black patients within five years since diagnosis survive approximately 4.0 months less than whites. The elimination of disparities in stage at diagnosis would contribute to a reduction of about 35% of this difference among CRC patients.

Our patient population, diagnosed between 1992 and 2005, experienced important disparities in access to screening [26,27]. However, in the 2010 CDC Behavioral Risk Factor Surveillance System survey, blacks and whites reported more similar screening rates

(66.3% for whites and 65% for blacks) [28]. Universal insurance coverage through Medicare and the recent elimination of out-of-pocket costs for screening likely contributed to these changes. It remains to be seen whether reductions in disparities in screening will result in equalization of the distribution of stage at diagnosis and reductions in black-white disparities as anticipated by our analysis. Our data suggest that, to the extent that similar screening rates lead to the elimination of disparities in stage at diagnosis, survival disparities could substantially reduce in the near future. Importantly, we find that such improvements might be particularly apparent for subpopulations (e.g. women and the elderly). The reasons for these heterogeneities should be further researched. We hypothesize that differentials in comorbidities might explain the differences in the impact of stage at diagnosis.

Our study carries important implications for future work. We show the importance of accounting for heterogeneities across gender and age to assess the role of stage at diagnosis in explaining colorectal cancer survival disparities. This might be relevant for the study of determinants of black-white disparities in other cancer outcomes as well. Further, we show that there are still determinants of the black-white disparities that are unexplained. Therefore, other factors may play a larger role in explaining such disparities among younger groups and in males. In particular, socioeconomic characteristics are important drivers of disparities in colorectal cancer survival [29] and their interplay need to be further investigated.

A popular approach to estimate the extent to which racial/ethnic disparities are explained by intermediate factors is to compare estimates of the effect of interest from regression models with and without adjustment for the hypothesized factor. Previous studies employing this approach yielded mixed results. The NCI-sponsored black/white Cancer Survival Study (B/WCSS) [15] found that stage differences accounted for 60% of the black-white survival disparities. Another study, based on data from Pennsylvania State Tumor Registry [16], found that differences in tumor stage, tumor site and grade combined explained one third of the excess mortality among blacks. That study however, did not investigate the unique contribution of stage at diagnosis. As is now well documented in the epidemiological and biostatistics literature, this approach is biased if there are interactions between race and stage at diagnosis [18]. In our study we confirm that such interaction is present. We further find evidence of race-age, stage-age, and stage-gender interactions (Supplementary Table S2).

Along with observational studies, micro-simulation studies can inform interventions to eliminate black-white disparities in cancer outcomes. Recent micro-simulation studies, evaluating the role of screening and stage-specific relative CRC survival in a population aged 50 years and older, estimated that screening and stage-specific survival explained a little more than 50% of the disparity in CRC mortality between blacks and whites [30]. Although a powerful tool that can reflect the natural history of a complex disease, micro-simulation comes with some limitations. Large-scale micro-simulation models require many parameters that may not be available from existing data sources. In addition, the degree of model detail does not go hand in hand with overall prediction power and potential for bias in complex models is often overlooked.

Our approach has several advantages. First, we formalized the quantities of interest accounting for the causal nature of the question. Adopting the counterfactual framework was key for clarifying the objects of inference and the identification conditions. Moreover, informed by an ecosocial understanding of how people jointly embody diverse types of inequality, we considered different CRC patients' subgroups according to gender and age. We estimated differences in the white and black patients using restricted mean survival times, a clinically meaningful measure of disparity, employing a non-parametric approach adjusting for potential confounders. Moreover, we report disparities on the difference scale, which has clearer public health interpretation. Finally, we evaluated the sensitivity of our analyses to selection bias due to missing stage at diagnosis.

There are limitations to this analysis. Our estimates could be biased due to unmeasured residual confounding. Comorbidities were not measured in the SEER cancer registry and these factors affect stage at diagnosis as well as cancer survival. We expect this bias to lead to an over-estimation of the impact of the hypothetical intervention. Further, as survival is our outcome, lead-time bias might affect our estimates. In the study we have mitigated the impact of this bias by excluding patients diagnosed with in-situ cancers. The direction of such bias in our study would again be an over estimation. Micro-simulations would be valuable complement to our analyses to assess the sensitivity of the results to this bias. Finally, stage at diagnosis measure is a crude representation of actual cancer stage; this might lead to an underestimation of the role of stage at diagnosis.

Other factors, such as socioeconomic position and quality of treatment have also been reported to contribute to explain both differences in CRC incidence and survival among racial/ethnic groups [31–32], but their doing so does not mean that interventions to modify stage at diagnosis are irrelevant. We note that a recent comprehensive screening program deployed in the state of Delaware resulted in the elimination of stage at diagnosis disparities and in the reduction of survival disparities [33]. Finally, in our study we evaluate the role of stage at diagnosis, a more direct intervention concerns screening uptake and follow-up rates. Linked SEER-Medicare data would allow evaluating potential screening regimes in the elderly population. A recent report from the CDC [28] found that colorectal cancer screening rates among whites were substantially higher than those among Hispanics, Asian/Pacific Islanders, and American Indian/Alaska Natives (AI/AN). Important disparities in cancer survival have been observed comparing white patients with Hispanic and Asian patients. Therefore, future research will consider other racial/ethnic groups as well.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank the senior editor, associate editor, and referees for helpful comments on the paper, and Dr. Lorenzo Trippa for helpful discussions.

Funding: All authors were supported by the National Institute of Health grant 5P01CA134294. T. J. VanderWeele was supported by the National Institute of Health grant ES017876. J. T. Chen, N. Krieger, and B. A. Coull were supported by the National Institute of Health grant P50CA148596.

REFERENCES

1. Rubin D. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol.* 1974; 66:688–701.
2. VanderWeele TJ, Robinson WR. On the causal interpretation of race in regressions adjusting for confounding and mediating variables. *Epidemiology.* 2014; 25:473–484. [PubMed: 24887159]
3. Krieger N. Defining and investigating social disparities in cancer: critical issues. *Cancer Causes Control.* 2005; 16:5–14. [PubMed: 15750853]
4. Krieger, N. Discrimination and health inequities. In: Berkman, LF.; Kawachi, I.; Glymour, M., editors. *Social Epidemiology.* 2nd ed.. New York, NY: Oxford University Press; 2014. p. 63-125.
5. National Cancer Institute. [accessed: April 6, 2015] Health Disparities. Available at: <http://cancer-control.cancer.gov/research-emphasis/health-disparities.html>
6. Whitehead M. The concepts and principles of equity and health. *Int JHealth Serv.* 1992; 22:429–445. [PubMed: 1644507]
7. Braveman P. Health disparities and health equity: Concepts and measurement. *Annu Rev Public Health.* 2006; 27:167–194. [PubMed: 16533114]
8. Blank, RM.; Babady, R.; Citro, CF., editors. National Research Council. *Measuring racial discrimination.* Washington, D.C.: National Academies Press; 2004.
9. Irby K, Anderson WF, Henson DE, Devesa SS. Emerging and widening colorectal carcinoma disparities between Blacks and Whites in the United States (1975–2002). *Cancer Epidemiology Biomarkers & Prevention.* 2006; 15(4):792–797.
10. Siegel R, DeSantis C, Jemal H. Colorectal cancer statistics, 2014. *CA: Cancer J Clin.* 2014; 64:104–117. [PubMed: 24639052]
11. American Cancer Society. *Cancer facts and figures for African Americans 2011–2012.* Atlanta, GA: American Cancer Society; 2011.
12. Robbins AS, Siegel R, Jemal J. Racial disparities in stage-specific colorectal cancer mortality rates from 1985 to 2008. *J Clin Oncol.* 2012; 30:401–405. [PubMed: 22184373]
13. Nelson RL, Doller T, Freels S, Persky V. The relation of age, race and gender to the subsite location of colorectal carcinoma. *Cancer.* 1997; 80(2):193–197. [PubMed: 9217029]
14. Laiyemo AO, Doubeni C, Pinsky PF, Doria-Rose VP, Bresalier R, Lamerato LE, Berg CD. Race and colorectal cancer disparities: health-care utilization vs different cancer susceptibilities. *Journal of the National Cancer Institute.* 2010; 102(8):538–546. [PubMed: 20357245]
15. Mayberry RM, Coates RJ, Hill HA, Click LA, Chen VW, Austin DF, Edwards BK. Determinants of black/white differences in colon cancer survival. *Journal of the National Cancer Institute.* 1995; 87(22):1686–1693. [PubMed: 7473817]
16. Marcella S, Miller JE. Racial differences in colorectal cancer mortality: The importance of stage and socioeconomic status. *J Clin Epidemiol.* 2001; 54:353–366.
17. Doubeni CA, Field TS, Buist DS, Korner EJ, Bigelow C, Lamerato L, Wagner EH. Racial differences in tumor stage and survival for colorectal cancer in an insured population. *Cancer.* 2007; 109(3):612–620. [PubMed: 17186529]
18. VanderWeele TJ. Causal mediation analysis with survival data. *Epidemiology.* 2011; 22:582–585. [PubMed: 21642779]
19. Byrd, DR.; Compton, CC.; Fritz, AG.; Greene, FL.; Trotti, AIII. *AJCC cancer staging manual.* Vol. 649. New York: Springer; 2010.
20. U.S. Census Bureau. [accessed 2014 Dec 2.] Census 1990 and 2000; using American FactFinder. <http://factfinder2.census.gov>
21. U.S. Census Bureau. [accessed 2014 Dec 2] 2010 American Community Survey; using American FactFinder. <http://factfinder2.census.gov>
22. Uno H, Claggett B, Tian L, Inoue E, Gallo P, Miyata T, Wei LJ. Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. *Journal of Clinical Oncology.* 2014; 32(22):2380–2385. [PubMed: 24982461]
23. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clin Trials.* 1986; 7:177–188. 527. [PubMed: 3802833]

24. Gasparrini A, Armstrong B. Multivariate meta-analysis: a method to summarize non-linear associations. *Stat Med*. 2011; 30:204–206.
25. R Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing; 2013.
26. Malley AS, Forrest CB, Feng S, Mandelbalatt J. Disparities despite coverage. *Arch Intern Med*. 2005; 165:2129–2135. [PubMed: 16217003]
27. White A, Vernon SW, Franzini L, Du XL. Racial and ethnic disparities in colorectal cancer screening persisted despite expansion of Medicare's screening reimbursement. *Cancer Epidemiol Biomarkers Prev*. 2011; 20:811–817. [PubMed: 21546366]
28. Joseph DA, King JB, Miller JW, Richardson LC. CDC. Prevalence of colorectal cancer screening among adults—Behavioral Risk Factor Surveillance System, U.S., 2010. *MMWR Morb Mortal Wkly Rep*. 2012; 61:51–56. [PubMed: 22695464]
29. Siegel R, Ward E, Brawley O, Jemal A. The impact of eliminating socioeconomic and racial disparities on premature cancer deaths. *Cancer J Clin*. 2011; 61:212–236.
30. Lansdorp-Vogelaar I, Kuntz KM, Knudsen AB, van Ballegooijen M, Zauber AG, Jemal A. Contribution of screening and survival differences to racial disparities in colorectal cancer rates. *Cancer Epidemiology Biomarkers & Prevention*. 2012; 21(5):728–736.
31. Haas JS, Brawarsky P, Iyer A, Fitzmaurice GM, Neville BA, Earle C. Association of area sociodemographic characteristics and capacity for treatment with disparities in colorectal cancer care and mortality. *Cancer*. 2011; 117(18):4267–4276. [PubMed: 21413000]
32. Le H, Ziogas A, Lipkin SM, Zell JA. Effects of socioeconomic status and treatment disparities in colorectal cancer survival. *Cancer Epidemiol Biomarkers Prev*. 2008; 17:1950–1962. [PubMed: 18708384]
33. Grubbs SS, Polite BN, Carney J, Bowser W, Rogers J, Katurakes N, Paskett ED. Eliminating racial disparities in colorectal cancer in the real world: it took a village. *Journal of Clinical Oncology*. 2013; 31(16):1928–1930. [PubMed: 23589553]

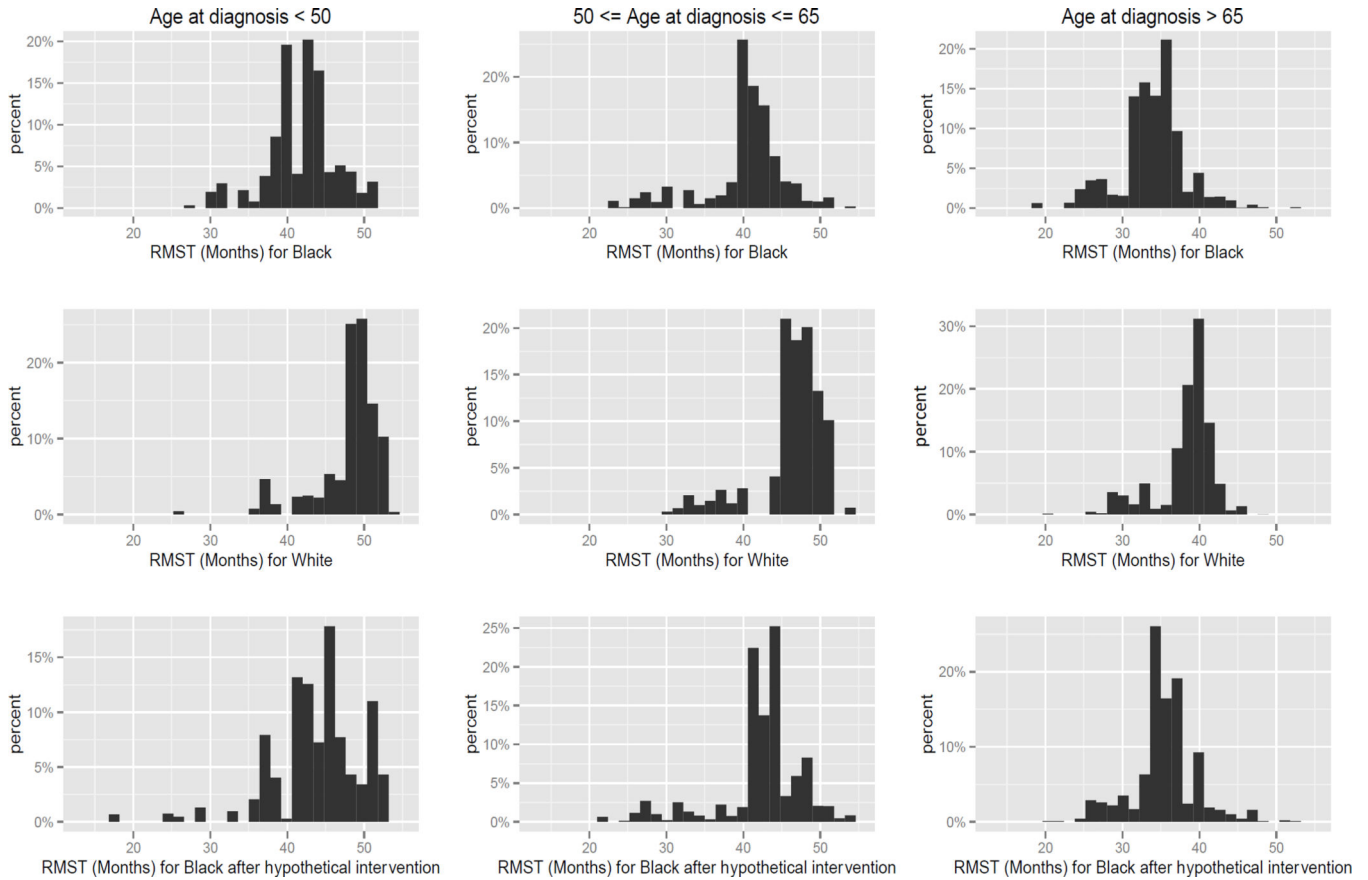


Figure 1. Histograms of restricted mean survival time truncated at month 60 by age at diagnosis for white, black, and black after the hypothetical intervention. The intervention consists on the shift of the stage distribution of the black population to match that of the white population, resulting in the elimination of the disparities in stage at diagnosis. The restricted mean survival time is estimated in strata of confounding factors (age at diagnosis, grade of tumor differentiation, county median income, date at diagnosis, gender, and registry). The y-axis represents the percent of individuals for which a certain RMST is estimated and the x-axis represents the RMST in months.

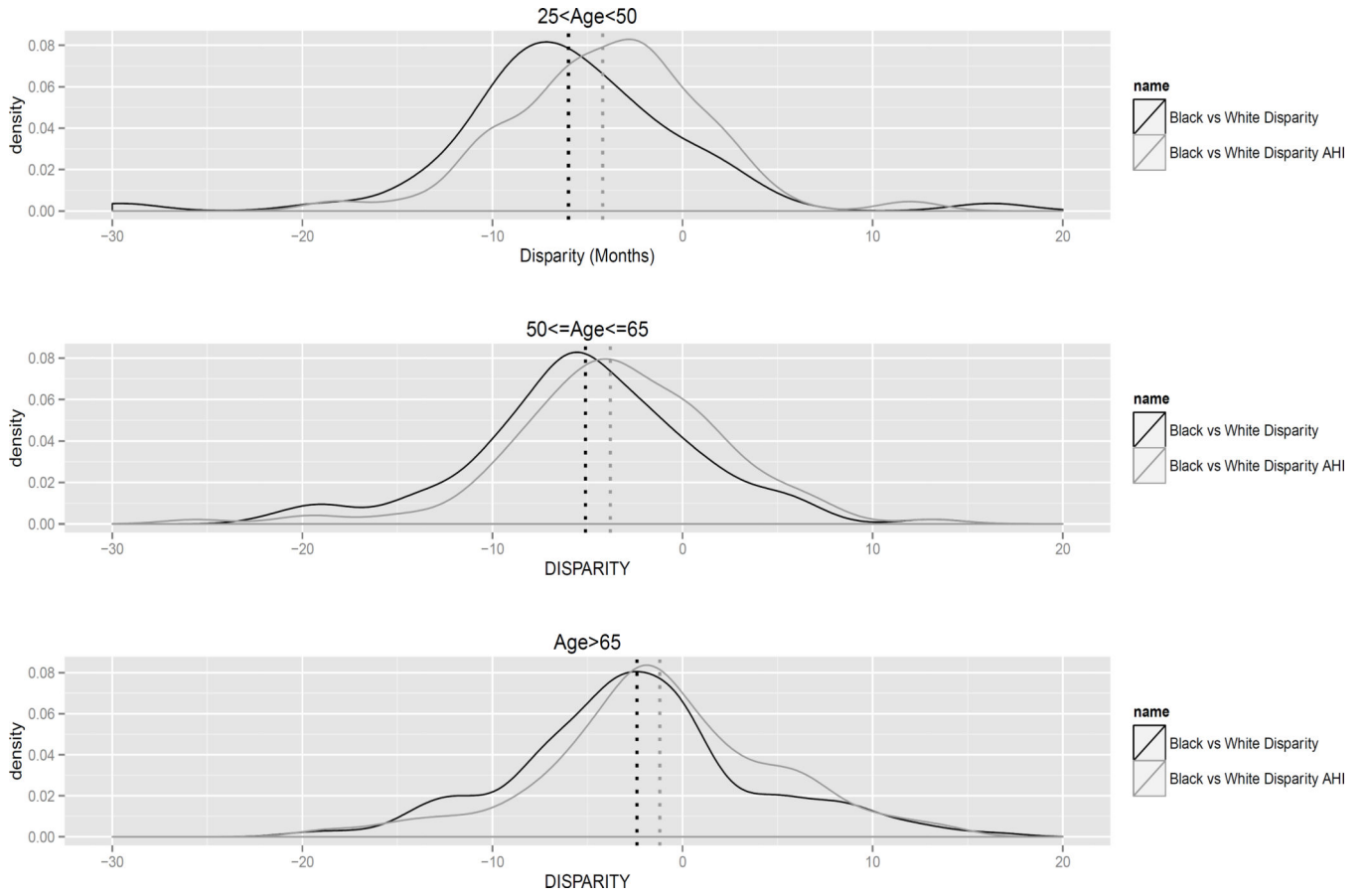


Figure 2. Density of difference in restricted mean survival time by age at diagnosis for white versus black (*disparity*, black) and white versus black under the hypothetical intervention (*residual disparity*, gray) across confounders' strata. The intervention consists on the shift of the stage distribution of the black population to match that of the white population. Dashed lines represent the estimated disparity and residual disparity using random effect meta-analysis for each age group.

Table 1

Baseline Characteristics of Colorectal Cancer patients (n (%)): US SEER registries, 1992–2005.

	black (n=18,712)	white (n=148,010)	P-value
Stage			
I	3,131 (17%)	31,998 (22%)	<0.001
II	5,079 (27%)	46,195 (31%)	<0.001
III	4,724 (25%)	36,631 (25%)	0.1242
IV	4,442 (24%)	25,298 (17%)	<0.001
Unstaged	1,296 (7%)	7,888 (5%)	<0.001
5-year survival	7,659 (40%)	72,076 (49%)	<0.001
Grade			
I	1,393 (7.4%)	11,555 (8%)	0.0834
II	12,403 (66%)	93,578 (63%)	<0.001
III	2,852 (15.2%)	29,163 (20%)	<0.001
IV	84 (0.4%)	1,078 (0.1%)	<0.001
Unknown	1,956 (10.4%)	12,363 (8.9%)	<0.001
Age at diagnosis			
<50	2,396(13%)	10,623 (7%)	<0.001
50–65	5,921 (32%)	33,246 (23%)	<0.001
>65	10,395 (55%)	104,141 (70%)	<0.001
Male	8,962 (48%)	74,431 (50%)	<0.001
Date at diagnosis			
1992–1995	4,838 (26%)	42,783 (29%)	<0.001
1996–2000	6,541 (35%)	54,606 (37%)	<0.001
2001–2005	7,333 (39%)	50,621 (34%)	<0.001
Registry			
San Francisco-Oakland Los Angeles, San Jose-Monterey	7,300 (39%)	45,574 (31%)	<0.001
Connecticut	1,371 (7.3%)	22,009 (15%)	<0.001
Atlanta and Rural Georgia	3,588 (19%)	7,663 (5.2%)	<0.001
Iowa	220 (1.2%)	23,444 (16%)	<0.001
Detroit	5,485 (29%)	19,047 (13%)	<0.001
New Mexico	115 (0.6%)	5,278 (3.5%)	<0.001
Utah	36 (0.2%)	6,456 (4.3%)	<0.001
Seattle-Puget Sound	597 (3.2%)	18,539 (12%)	<0.001
Median County Income (in 2010 \$)			
\$21,116–\$50,254	3,303 (21%)	25,928 (20%)	<0.001
\$50,255–\$55,092	4,571 (29%)	24,622 (19%)	<0.001
\$55,093–\$62,512	3,695 (23%)	25,356 (20%)	<0.001
\$62,513–\$71,724	2,313 (15%)	26,943 (20%)	<0.001
\$71,725–\$110,970	1,975 (12%)	27,008 (21%)	<0.001

Table 2

Disparity in restricted mean survival time truncated at month 60 since diagnosis by stage at diagnosis.

	Estimate (C.I.)
Stage	
I	-3.2 (-4.4, -1.7)
II	-2.4 (-3.6, -1.0)
III	-0.8 (-2.1, 0.6)
IV	-2.8 (-3.9, -1.5)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Disparity in restricted mean survival time truncated at month 60 since diagnosis before and after the hypothetical intervention on stage at diagnosis overall and by age and gender. In parenthesis the number of black and white patients contributing to the estimation of the quantities.

	Estimate (C.I.)
OVERALL (nblack=13,168, nwhite=73,779)	
Disparity	-4.0 (-4.6, -3.3)
Residual Disparity	-2.6 (-3.4, -1.7)
Percent Disparity Reduction	35% (11%, 58%)
AGE<50 (nblack=1,198, nwhite= 2,910)	
Disparity	-6.0 (-7.6, -4.3)
Residual Disparity	-4.1 (-6.0, -2.3)
Percent Disparity Reduction	30% (-6.9%, 67%)
50<=AGE<65 (nblack=4,105, nwhite=14,498)	
Disparity	-5.1 (-6.2, -4.0)
Residual Disparity	-3.8 (-5.2, -2.3)
Percent Disparity Reduction	26% (-5.6%, 57%)
AGE>65 (nblack=7,865, nwhite=56,371)	
Disparity	-2.4 (-3.4, -1.5)
Residual Disparity	-1.2 (-2.4, -0.4)
Percent Disparity Reduction	50% (1.0%, 100%)
MALE (nblack=6,235, nwhite=34,669)	
Disparity	-4.0 (-5.0, -3.0)
Residual Disparity	-2.9 (-4.1, -1.7)
Percent Disparity Reduction	26% (-8.1%, 61%)
FEMALES (nblack=6,933, nwhite=39,110)	
Disparity	-3.8 (-4.8, -2.8)
Residual Disparity	-2.1 (-3.3, -0.9)
Percent Disparity Reduction	44% (10%, 78%)