Racial disparities in epithelial ovarian cancer survival: An examination of contributing factors in the Ovarian Cancer in Women of African Ancestry (OCWAA) consortium

Holly R. Harris*, ScD, MPH, Kristin A. Guertin*, PhD, MPH, Tareq F Camacho, MS, Courtney E. Johnson, MS, Anna H. Wu, PhD, MPH, Patricia G. Moorman, PhD, MSPH, Evan Myers, MD, MPH, Traci N. Bethea, PhD, MPA, Elisa V. Bandera, MD, PhD, Charlotte E. Joslin, OD, PhD, Heather M. Ochs-Balcom, PhD, Lauren C. Peres, PhD, MPH, Will T. Rosenow, MS, Veronica W. Setiawan, PhD, Alicia Beeghly-Fadiel, PhD, MPH, Lauren F. Dempsey, MSPH, Lynn Rosenberg[+], ScD, Joellen M. Schildkraut[+], PhD, MPH

*Shared first authorship, contributed equally to this work.

[+]Shared senior authorship, contributed equally to this work.

Table of Contents

Supplemental Methods

Defining Histotypes

Histotype categories were collapsed for analyses into the following groups: high-grade serous, clear cell, and other epithelial ovarian cancer. The high-grade serous category also included carcinosarcoma cases, however we refer to this category as high-grade serous based on the relatively small number of carcinosarcoma cases (n=74 white, 26 black); we combined these two histotypes into one category based on the similar survival in this study and the relatively small numbers of carcinosarcoma cases. The other epithelial category included low-grade serous, endometrioid, mucinous, and other epithelial cancer not otherwise specified. In secondary analyses we considered type I vs type II designation and its impact on survival but in main analyses separated clear cell cases from other type I tumors due to the difference in prevalence by race and the survival disparity that was present within the clear cell histotype.

Defining confounders and mediators

When multiple variables measuring a similar aspect were screened, one variable of the group was selected based on prior research. For education related variables, college graduation status was selected. For BMI related variables, the three-level variable comparing normal, overweight, and obese status was selected. For smoking related variables, the three-level status variable comparing former smoker, current smoker, and never smoker was selected. For postmenopausal hormone usage related variables, the duration variable comparing none, less than 5 years, and at least 5 years was selected. Lastly, the continuous Singh ADI score for the census tract level was selected for ADI related variables. In secondary analyses limited to 1) high grade serous cases and 2) distant stage cases, we rescreened for the mediators for inclusion into the model. We selected variables for inclusion into the models if the tests of the exposure-variable or variable-outcome relationship were significant at the $p<0.10$ level, or in some cases, if only the variable-outcome relationship were significant at the same level. It was

thought to be more likely that these variables fell on the causal pathway and the subsets were not powered to detect an exposure-variable association.

Mediation Assumptions and Interpretation of Effects

Assumptions needed to identify effects under both methods are described next[40,41], using the notation in Figure 1. In addition to standard initial assumptions such as consistency of potential outcomes, under both methods it is sufficient to assume the E-Y relationship is unconfounded conditional on C (i) and the M-Y relationship is unconfounded conditional on C and E (imputation) and E only (Yu) (ii). The imputation method furthermore requires the E-M relationship to be unconfounded (iii), and that no M-Y confounders be affected by E (iv), which is frequently cited as the "cross-world" assumption[42]. Lastly, in order to implement the Yu method, M is assumed not to be causally prior to other mediators (v). Assumptions (i), (ii), and (iii) are presumed to be satisfied by adjusting for C which will account for all or most unmeasured confounding. Assumption (iv) is satisfied by the sequential analysis as joint mediators taken en bloc will have no prior mediators, and assumption (v) is satisfied by not including mediators after M in each step of the sequential analysis.

The Imputation method decomposes the average total effect conditional on C into an average natural direct effect (NDE) and average natural indirect effect (NIE)[41]. We interpret the NDE as the portion of the total effect which does not operate through the mediator(s) M, and the indirect effect as the portion of the total effect which operates through M. Although a NIE for an individual mediator may be difficult to identify particularly in regards to the violation of the cross-worlds assumption, it is possible to identify individual path-specific effects if assumptions on the joint mediators are met[41]. One such particular effect can be interpreted as the portion of the indirect effect which operates only through M and not through the other mediators which may affect M, as described in the supplement.

The Yu method decomposes the average total effect (ATE) into a direct effect (DE) and an indirect effect (IE). The DE is interpreted as the average exposure effect when the mediator M is intervened by assigning to each case a value randomly sampled from the marginal distribution of M. The indirect effect is the difference between the TE and DE, and interpreted as the change in effect attributable to intervening on the mediator M.

Both of these methods are alternatives to the product-coefficient method, and as such, do not require the rare event assumption to use Cox PH models to calculate effects. Further, sensitivity analyses using Log-normal AFT models were fit which show the convergence of both outcome models and thus the correct specifications of the model.

Imputation

Multilevel Multiple Imputation (MMI)[30] was conducted. As the outcome consisted of censored survival data, we followed the White and Royston procedure, which estimates a cumulative baseline hazard and includes it as an auxiliary variable in the imputation model together with the event indicator[30]. Furthermore, a multilevel extension (JOMO) to multiple imputation using chained equations (MICE) in R[43,44] as implemented to account for differing study designs. The Singh Area Deprivation Index (ADI), was systematically missing for WHI, and diabetes was missing for CCCS. This package allows for multiple imputation to be done by a cluster variable. After stratifying by race, multiple imputation was conducted while clustering by site. This imputation model included all screened mediators, confounders, and outcome related variables. MMI was integrated with the mediation procedures by treating each imputed data set as complete data, calculating the mediation effects for each imputation, and then taking the mean effect to get point estimates. In order to reduce empirically observed variation of the mean effects, we used 25 imputations.

Implementation of the Imputation Method

Details on how to implement imputation method are described in Lange et al.[10], summarized below:

1. Fit an outcome prediction model in the original dataset (D) as a function of confounders, exposure (x) and mediators. Lange[10] suggests fitting an AFT model to the data. As the error term in an AFT model may vary, resulting in different regression such as Weibull, log-logistic, or lognormal, we varied the error distributions to improve model fit according to the AIC criterion as well as to result in the best fit to survival curve using visual plotting. As a result, the model chosen was a lognormal AFT.
2. Create a new variable x* with the same value as the exposure x in D. Then create an identical copy of D (D'). In this copy replace x* with the opposite value of x and set the time-to-event outcome Y to missing.
3. Using the outcome model, impute the missing time-to-event in D' where $x \neq x^*$ by simulating a draw from the distribution implied by the AFT model, given the confounders, exposure x*, and mediators in D'. To avoid extrapolation outside what is supported by the data, any imputed survival time longer than the maximum follow-up (29 years in this case) is artificially censored at the time of maximal follow-up[10]. The censoring indicator for the simulated data is set to indicate an event.
4. Concatenate D and D' into one data set D''. Fit a natural effects model in D" using a Cox PH model as described in Lange[10]. The independent variables in the Cox PH will consist of x, x*, and confounders c. The log hazard takes the following form:

$$\log\ \lambda(t|x_i, x_i^*, c_i) = \log \lambda_0(t) + f(x_i, x_i^*, c_i) \text{ for subjects } i=1,\ldots,2n$$

The beta coefficient, or log hazard, pertaining to x is the natural indirect effect of the exposure, and the beta coefficient for x* is an estimate of the natural direct effect.

5. Repeat step 10 described in Part A. to get point estimates and confidence intervals for the NDE and NIE.

Implementation of the Yu Method

Details on the original algorithms are provided in Yu, et al.[11,40]. Let y_i represent the observed time for subject i (censored or event time), $\mathbf{c_i}$ be a vector of confounders, $\mathbf{m_i} = (m_{1i},\ldots,m_{pi})$ be the vector of mediators for subject i, and x_i be the a binary exposure level for subject i. Two prediction models for outcome survival time were considered. Firstly, a Cox Proportional Hazards (Cox PH) model was used to predict the outcome, where the log hazards takes the following form:

$$\log\ \lambda(t|x_i,m_i,c_i) = \log \lambda_0(t)+f(x_i,m_{1i},\ldots,m_{pi},c_i),\text{for subjects } i=1,\ldots,n,$$

f is a linear function of the arguments, and $\lambda_0(t)$ is a baseline hazard function.

Secondly, and Accelerated Failure Time (AFT) model with the following form was considered:

$$\log y_i = f(x_i,m_{1i},\ldots,m_{pi},c_i) + \sigma\varepsilon_i$$

where f is a linear function, σ is a scaling parameter, and ε_i is a random error term.

Yu et al.'s original algorithm assumes the distribution of mediators is multivariate normal ($M|x,c \sim N(0,\Sigma)$) and draws random values for the mediators from this distribution. We propose to broaden this assumption by the application of method of chained equations (MICE) to estimate the total effects for binary exposure as follows. Under this extension, the values to draw are viewed as missing data to be imputed, with the benefit that categorical imputation is allowed and multivariate distribution is not restricted to normality.

Step a. Estimation of total effects

1. Assuming the original dataset (D) has no missing values, create two copies of D: D1 and D0. In D1 replace all values of the X exposure variable to the comparison level (x=1), and in D0 replace the all values with the baseline level (x = 0). In both datasets, set the mediator variables m to missing, and leave the confounder variables c as in the original.
2. Create a concatenated data set D+D0+D1. For this concatenated dataset, impute the missing mediation variables using MICE. The MICE imputation model will impute missing mediators by randomly sampling from an approximation of the distribution of M|x,c. Only the mediators, the exposure, and confounders are included into the imputation model without specifying other auxiliary and extraneous variables.
3. In D0, D1, estimate the predicted value $f(x_i,m_{1i},\ldots,m_{pi}, c_i)$ for each case using the Cox PH or AFT model predictions estimated from the original data D.
4. The Average Total Effect ($\log \lambda_{TE}$) is estimated by taking the difference between D1 and D0 of the sample means f_i.

Step b. Estimation of direct effects

5. Concatenate D1 and D0 from step 1 into a data set **D***
6. Randomly shuffle, or permute, the mediator(s) of interest to different cases in **D***. This simulates a random intervention.
7. Estimated predicted values in the shuffled data as in Step 3.
8. The average direct effect ($\log \lambda_{DE}$) is estimated by taking the difference in sample means of f_i between cases with x = 0 and x = 1 in **D***.
9. The indirect effect is estimated by the difference between the total effect and direct effect.

Step c. Repetition of steps a. and b.

10. As in a multiple imputation method, repeat steps 1-4 on separately imputed data sets and get a combined point estimate by taking the mean effects across the imputations. In this case we used 25 imputations. To get 95% confidence intervals use bootstrapping methods which combine with multiple imputation and which are described in Schomaker et al[29].

Supplemental Table 1. Missing Data Patterns by Study Before Imputation

	Black women N (%) Missing	White women N (%) Missing
Area Deprivation Index (ADI)[a]		
AACES	40 (6.9%)	--
BWHS	7 (9.1%)	--
LACOCS	8 (6.5%)	81 (7.1%)
CCCCS	0 (0.0%)	0 (0.0%)
MEC	29 (31.9%)	98 (67.6%)
NCOCS	12 (10.3%)	56 (6.8%)
WHI	47 (100.0%)	952 (100.0%)
Education		
AACES	0 (0.0%)	--
BWHS	0 (0.0%)	--
LACOCS	0 (0.0%)	0 (0.0%)
CCCCS	0 (0.0%)	0 (0.0%)
MEC	1 (1.1%)	1 (0.7%)
NCOCS	1 (0.9%)	0 (0.0%)
WHI	0 (0.0%)	7 (0.7%)
Nulliparity		
AACES	0 (0.0%)	--
BWHS	0 (0.0%)	--
LACOCS	0 (0.0%)	0 (0.0%)
CCCCS	0 (0.0%)	0 (0.0%)
MEC	0 (0.0%)	0 (0.0%)
NCOCS	1 (0.9%)	0 (0.0%)
WHI	0 (0.0%)	1 (0.1%)
Smoking Status		
AACES	0 (0.0%)	--
BWHS	0 (0.0%)	--
LACOCS	0 (0.0%)	0 (0.0%)
CCCCS	0 (0.0%)	11 (5.3%)
MEC	2 (2.2%)	3 (2.1%)
NCOCS	1 (0.9%)	0 (0.0%)
WHI	0 (0.0%)	9 (0.9%)
Body Mass Index (BMI)		
AACES	4 (0.7%)	--
BWHS	0 (0.0%)	--
LACOCS	0 (0.0%)	0 (0.0%)
CCCCS	0 (0.0%)	1 (0.5%)
MEC	1 (1.1%)	0 (0.0%)
NCOCS	2 (1.7%)	21 (2.6%)
WHI	0 (0.0%)	3 (0.3%)
Diabetes		
AACES	0 (0.0%)	--
BWHS	0 (0.0%)	--
LACOCS	0 (0.0%)	0 (0.0%)
CCCCS	39 (100.0%)	208 (100.0%)
MEC	0 (0.0%)	0 (0.0%)
NCOCS	0 (0.0%)	0 (0.0%)
WHI	0 (0.0%)	0 (0.0%)
Postmenopausal Hormone Duration		
AACES	4 (0.7%)	--
BWHS	0 (0.0%)	--
LACOCS	0 (0.0%)	0 (0.0%)
CCCCS	1 (2.6%)	7 (3.4%)
MEC	7 (7.7%)	6 (4.1%)
NCOCS	3 (2.6%)	39 (4.8%)
WHI	0 (0.0%)	0 (0.0%)
Stage		
AACES	36 (6.2%)	--

BWHS	13 (16.9%)	--
LACOCS	2 (1.6%)	8 (0.7%)
CCCCS	1 (2.6%)	0 (0.0%)
MEC	6 (6.6%)	2 (1.4%)
NCOCS	2 (1.7%)	15 (1.8%)
WHI	0 (0.0%)	3 (0.3%)

SD: Standard deviation; AACES: African-American Cancer Epidemiology Study; BWHS, Black Women's Health Study; CCCCS, Cook County Case-Control Study; LACOCS, Los Angeles County Ovarian Cancer Study; MEC, Multiethnic Cohort Study; NCOCS, North Carolina Ovarian Cancer Study; WHI, Women's Health Initiative; BMI: body mass index; ADI, Area Deprivation Index

[a]ADI is the Singh score at census tract level

Supplemental Table 2. Sequential Multiple Mediation of the Effect of Race on Ovarian Cancer Survival using an imputation method[10]: Area Deprivation Index as a Mediator[a]

Step[b]	Mediator(s) Added	Path Specific IE HR (95% CI) [c]	Indirect Effect HR (95% CI)[d]	Direct Effect HR (95% CI)[e]	Total Effect HR (95% CI)[f]	% Mediated[g]	Path Specific % Mediated[h]
1	College Area Deprivation Index	1.06 (1.00, 1.14)	1.06 (1.00, 1.14)	1.23 (1.18, 1.29)	1.31 (1.26, 1.38)	21.5 (10.9)	21.5 (10.9)
2	Nulliparity	1.03 (0.93, 1.09)	1.09 (1.01, 1.15)	1.23 (1.17, 1.29)	1.34 (1.26, 1.38)	29.8 (10.9)	8.7 (14.9)
3	Smoking	0.99 (0.92, 1.08)	1.08 (1.01, 1.14)	1.23 (1.17, 1.29)	1.32 (1.26, 1.38)	25.7 (10.8)	-5.5 (15.1)
4	BMI Diabetes Diabetes/Race Interaction	1.01 (0.94, 1.10)	1.09 (1.03, 1.16)	1.21 (1.16, 1.27)	1.32 (1.26, 1.38)	30.2 (10.3)	3.3 (14.6)
5	PMH Duration PMH Duration/Race Interaction PMH Duration/Age Interaction	1.04 (0.95, 1.12)	1.13 (1.06, 1.21)	1.14 (1.09, 1.19)	1.29 (1.23, 1.35)	47.8 (10.8)	13.1 (16.1)
6	Histotype Stage	1.04 (0.96, 1.12)	1.17 (1.11, 1.24)	1.09 (1.05, 1.14)	1.28 (1.23, 1.34)	63.5 (8.8)	14.8 (15.5)

BMI: body mass index; CI: Confidence Interval; HR: hazard ratio; IE: indirect effect; PMH: post-menopausal hormone.
[a] Confounders adjusted for in the model include year of diagnosis, age at diagnosis, site, and area deprivation index (ADI). Bold numbers reflect statistically significant values.
[b] Sequential order in which the selected mediators were added to the model. Model also includes variables in all preceding steps.
[c] Path specific indirect effect is the indirect effect through the new variable(s) added only and not through other mediators or through the joint pathways of the new variable(s) with prior mediators.
[d] The indirect effect is the cumulative indirect effect of the variables selected as mediators in a given step and all preceding steps.
[e] The direct effect is the remaining direct effect through all other non-mediated pathways.
[f] Total effect is calculated as log(Indirect Effect HR) + log(Direct Effect HR).
[g] Percent (%) mediated is calculated as log(indirect effect HR)/log(total effect HR).
[h] Path Specific (%) mediated is calculated as log(path specific indirect effect HR)/log(total effect HR).

Supplemental Table 3. Sequential Multiple Mediation of the Effect of Race on Ovarian Cancer Survival using an imputation method[10]: Log-Normal AFT Estimation[a]

Step[b]	Mediator(s) Added	Path Specific IE HR (95% CI)[c]	Indirect Effect HR (95% CI)[d]	Direct Effect HR (95% CI)[e]	Total Effect HR (95% CI)[f]	% Mediated[g]	Path Specific % Mediated[h]
1	College	1.01 (0.92, 1.09)	1.01 (0.92, 1.09)	0.75 (0.71, 0.80)	0.76 (0.69, 0.82)	-5.6 (15.6)	-5.6 (15.6)
2	Nulliparity	0.97 (0.88, 1.11)	0.98 (0.92, 1.08)	0.76 (0.71, 0.80)	0.74 (0.69, 0.82)	6.0 (15.2)	10.2 (21.4)
3	Smoking	1.01 (0.89, 1.12)	0.99 (0.92, 1.08)	0.76 (0.71, 0.80)	0.75 (0.69, 0.82)	2.8 (15.2)	-5.3 (21.7)
4	BMI Diabetes Diabetes/Race Interaction	0.97 (0.87, 1.09)	0.96 (0.89, 1.05)	0.77 (0.73, 0.82)	0.74 (0.69, 0.81)	12.3 (13.8)	7.7 (20.4)
5	PMH Duration PMH Duration/Race Interaction PMH Duration/Age Interaction	0.96 (0.86, 1.08)	0.92 (0.85, 1.01)	0.84 (0.78, 0.89)	0.77 (0.71, 0.84)	29.7 (14.4)	13.2 (23.5)
6	Histotype Stage	0.95 (0.85, 1.06)	0.88 (0.81, 0.95)	0.88 (0.83, 0.93)	0.77 (0.71, 0.83)	50.0 (10.8)	20.0 (21.3)

BMI: body mass index; CI: Confidence Interval; HR: hazard ratio; IE: indirect effect; PMH: post-menopausal hormone.
[a] Confounders adjusted for in the model include year of diagnosis, age at diagnosis, site, and area deprivation index (ADI).
Bold numbers reflect statistically significant values.
[b] Sequential order in which the selected mediators were added to the model. Model also includes variables in all preceding steps.
[c] Path specific indirect effect is the indirect effect through the new variable(s) added only and not through other mediators or through the joint pathways of the new variable(s) with prior mediators.
[d] The indirect effect is the cumulative indirect effect of the variables selected as mediators in a given step and all preceding steps.
[e] The direct effect is the remaining direct effect through all other non-mediated pathways.
[f] Total effect is calculated as log(Indirect Effect HR) + log(Direct Effect HR).
[g] Percent (%) mediated is calculated as log(indirect effect HR)/log(total effect HR).
[h] Path Specific (%) mediated is calculated as log(path specific indirect effect HR)/log(total effect HR).

Supplemental Table 4. Average Path Specific Indirect Effects (IE) and Path Specific Percent Mediated (PM) Considering All Possible Orderings of the Six Steps[a]

	Path Specific IE HR[b]			Path Specific % Mediated[c]		
	Mean	**Median**	**(Minimum, Maximum)**	**Mean**	**Median**	**(Minimum, Maximum)**
College	0.998	0.997	(0.991, 1.005)	-2.6	-3.2	(-6.8, 1.3)
Nulliparity	1.002	1.002	(0.995, 1.011)	-0.6	-0.7	(-4.3, 4.2)
Smoking Status	1.000	1.000	(0.994, 1.008)	-1.6	-2.0	(-5.0, 2.5)
BMI & Diabetes	1.020	1.020	(1.008, 1.029)	8.4	8.9	(2.9, 12.8)
PMH Duration	1.027	1.026	(1.019, 1.035)	12.4	12.0	(9.0, 17.4)
Stage & Histotype	1.038	1.036	(1.023, 1.051)	17.2	16.8	(10.7, 22.8)

IE, Indirect Effect; HR, Hazard Ratio

[a] 100 Imputations of ordering combinations of the six steps in models examining the sequential multiple mediation of the effect of race on ovarian cancer survival using an imputation method[10]

[b] Path specific indirect effect is the indirect effect through the new variable(s) added only and not through other mediators or through the joint pathways of the new variable(s) with prior mediators.

[c] Path Specific (%) mediated is calculated as log(path specific indirect effect HR)/log(total effect HR).

Supplemental Table 5. Sequential Multiple Mediation of the Effect of Race on Ovarian Cancer Survival using an imputation method[10]: High-Grade Serous Cases (N = 2730)[a]

Step[b]	Mediator(s) Added	Path Specific IE HR (95% CI)[c]	Indirect Effect HR (95% CI)[d]	Direct Effect HR (95% CI)[e]	Total Effect HR (95% CI)[f]	% Mediated (SE)[g]	Path Specific % Mediated (SE)[h]
1	College	0.99 (0.93, 1.07)	0.99 (0.93, 1.07)	1.27 (1.20, 1.32)	1.26 (1.17, 1.34)	-5.3 (18.1)	-5.3 (18.1)
2	Smoking	1.00 (0.90, 1.11)	0.99 (0.93, 1.08)	1.26 (1.19, 1.32)	1.25 (1.17, 1.34)	-5.1 (18.1)	-1.2 (24.9)
3	BMI Diabetes	1.05 (0.93, 1.15)	1.04 (0.96, 1.12)	1.20 (1.15, 1.27)	1.25 (1.17, 1.34)	17.3 (16.5)	20.6 (23.9)
4	PMH Duration PMH Duration/Age	1.02 (0.92, 1.14)	1.06 (0.99, 1.15)	1.18 (1.12, 1.23)	1.25 (1.17, 1.34)	26.1 (15.2)	7.2 (24.7)
5	Prior Breast Cancer	0.99 (0.90, 1.11)	1.06 (0.99, 1.15)	1.17 (1.12, 1.23)	1.24 (1.17, 1.34)	24.8 (15.1)	-3.5 (25.6)
6	Stage	1.03 (0.92, 1.13)	1.09 (1.01, 1.17)	1.14 (1.09, 1.20)	1.24 (1.17, 1.33)	37.8 (13.5)	10.0 (24.4)

BMI: body mass index; CI: Confidence Interval; SE: standard error; HR: hazard ratio; IE: indirect effect; PMH: post-menopausal hormone.

[a] Confounders adjusted for in the model include year of diagnosis, age at diagnosis, site, and area deprivation index (ADI). Bold numbers reflect statistically significant values.

[b] Sequential order in which the selected mediators were added to the model. Model also includes variables in all preceding steps.

[c] Path specific indirect effect is the indirect effect through the new variable(s) added only and not through other mediators or through the joint pathways of the new variable(s) with prior mediators.

[d] The indirect effect is the cumulative indirect effect of the variables selected as mediators in a given step and all preceding steps.

[e] The direct effect is the remaining direct effect through all other non-mediated pathways.

[f] Total effect is calculated as log(Indirect Effect HR) + log(Direct Effect HR).

[g] Percent (%) mediated is calculated as log(indirect effect HR)/log(total effect HR).

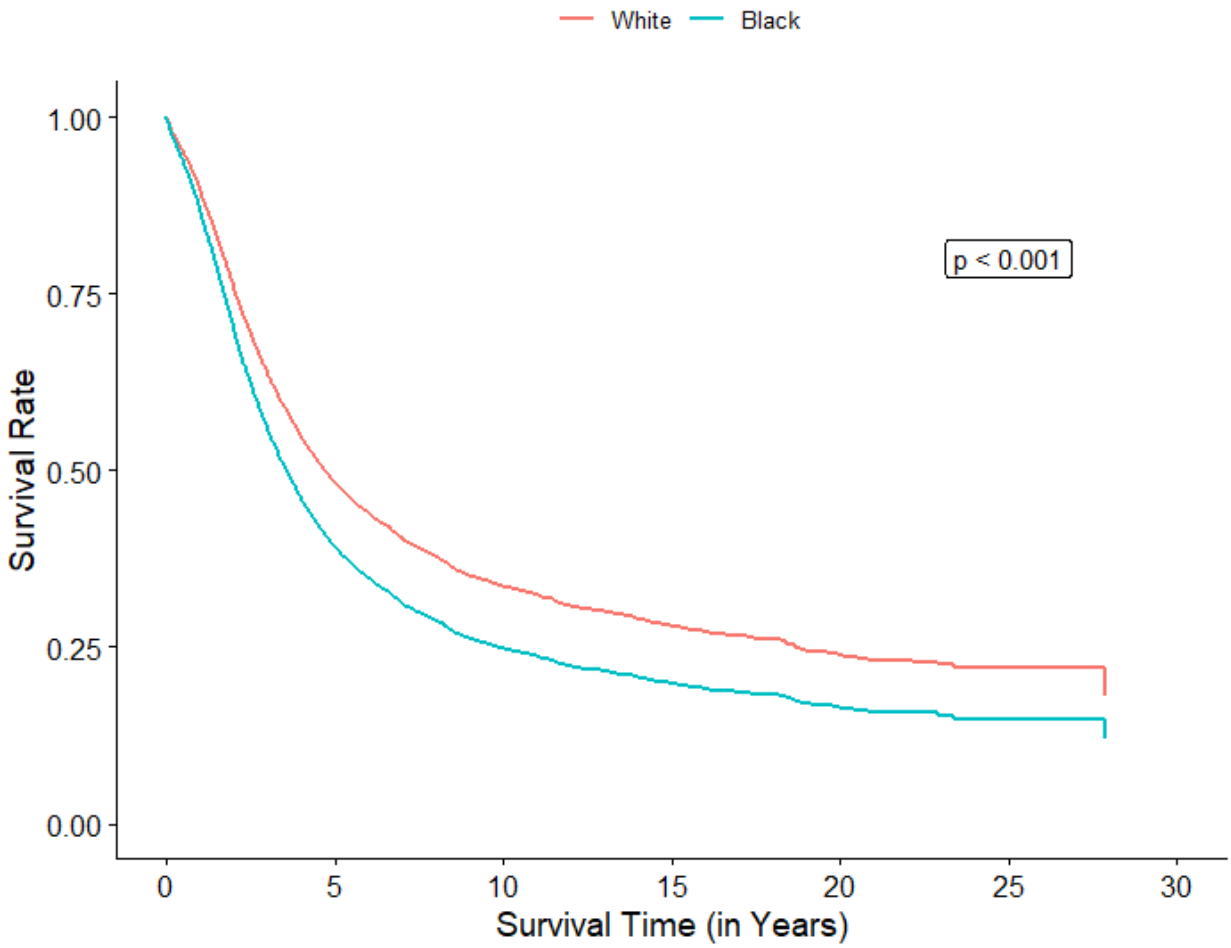
[h] Path Specific (%) mediated is calculated as log(path specific indirect effect HR)/log(total effect HR).

Supplemental Table 6. Sequential Multiple Mediation of the Effect of Race on Ovarian Cancer Survival using an imputation method[10]: Women with Distant Stage Cancer at Diagnosis (N = 3036)[a]

Step[b]	Mediator(s) Added	Path Specific IE HR (95% CI)[c]	Indirect Effect HR[d] (95% CI)	Direct Effect HR (95% CI)[e]	Total Effect HR (95% CI)[f]	% Mediated[g]	Path Specific % Mediated[h]
1	College	1.00 (0.94, 1.07)	1.00 (0.94, 1.07)	1.27 (1.22, 1.34)	1.28 (1.22, 1.34)	1.3 (14.2)	1.3 (14.2)
2	Smoking Status	1.00 (0.92, 1.10)	1.01 (0.94, 1.08)	1.27 (1.21, 1.33)	1.28 (1.20, 1.37)	2.3 (14.0)	0.1 (19.7)
3	BMI Diabetes Diabetes/Race	1.05 (0.95, 1.15)	1.05 (0.99, 1.13)	1.21 (1.16, 1.28)	1.28 (1.20, 1.37)	19.6 (12.6)	16.3 (19.1)
4	PMH Duration PMH Duration/Age	1.02 (0.92, 1.12)	1.07 (1.00, 1.25)	1.20 (1.14, 1.25)	1.28 (1.20, 1.36)	28.0 (12.2)	7.7 (20.0)

BMI: body mass index; CI: Confidence Interval; HR: hazard ratio; IE: indirect effect; PMH: post-menopausal hormone.
[a] Confounders adjusted for in the model include year of diagnosis, age at diagnosis, site, and area deprivation index (ADI). Bold numbers reflect statistically significant values.
[b] Sequential order in which the selected mediators were added to the model. Model also includes variables in all preceding steps.
[c] Path specific indirect effect is the indirect effect through the new variable(s) added only and not through other mediators or through the joint pathways of the new variable(s) with prior mediators.
[d] The indirect effect is the cumulative indirect effect of the variables selected as mediators in a given step and all preceding steps.
[e] The direct effect is the remaining direct effect through all other non-mediated pathways.
[f] Total effect is calculated as log(Indirect Effect HR) + log(Direct Effect HR).
[g] Percent (%) mediated is calculated as log(indirect effect HR)/log(total effect HR).
[h] Path Specific (%) mediated is calculated as log(path specific indirect effect HR)/log(total effect HR).

Supplemental Figure 1. Cox Proportional Hazards Curves: epithelial ovarian cancer survival by Race, Adjusted for Age at Diagnosis, Year of Diagnosis, and Site (N = 1,074 Black women and 3,263 White women)



Note: p-value for difference in survival rate is from ANVOA for race term in Cox PH model including race, site, age at diagnosis, and year of diagnosis