10-4-2021

Ovarian Cancer Epidemiology, Healthcare Access and Disparities (ORCHiD): Methodology for a Population-Based Study of Black, Hispanic and White Patients with Ovarian Cancer

Tomi Akinyemiju
Duke University

April Deveaux
Duke University

Lauren Wilson
Duke University

Anjali Gupta
Duke University

Ashwini Joshi
Duke University

Ovarian Cancer Epidemiology, Healthcare Access and Disparities (ORCHiD): Methodology for a Population-Based Study of Black, Hispanic and White Patients with Ovarian Cancer

Digital Object Identifier (DOI)
https://doi.org/10.1136/bmjopen-2021-052808

Notes/Citation Information
Published in *BMJ Open*, v. 11, issue 10, e052808.

Authors
Tomi Akinyemiju, April Deveaux, Lauren Wilson, Anjali Gupta, Ashwini Joshi, Malcolm Bevel, Chioma Omeogu, Onyinye Ohamadike, Bin Huang, Maria Pisu, Margaret Liang, Molly McFatrich, Erin Daniell, Laura Jane Fish, Kevin Ward, Maria Schymura, Andrew Berchuck, and Arnold L. Potosky

BMJ Open

Ovarian Cancer Epidemiology, Healthcare Access and Disparities (ORCHiD): methodology for a population-based study of black, Hispanic and white patients with ovarian cancer

Tomi Akinyemiju ,[1,2] April Deveaux ,[1] Lauren Wilson ,[1] Anjali Gupta,[1] Ashwini Joshi ,[1] Malcolm Bevel ,[1] Chioma Omeogu,[1] Onyinye Ohamadike,[1] Bin Huang,[3] Maria Pisu,[4] Margaret Liang,[4,5] Molly McFatrich,[1] Erin Daniell,[1] Laura Jane Fish,[2] Kevin Ward,[6] Maria Schymura,[7] Andrew Berchuck,[8] Arnold L Potosky[9]

To cite: Akinyemiju T, Deveaux A, Wilson L, *et al*. Ovarian Cancer Epidemiology, Healthcare Access and Disparities (ORCHiD): methodology for a population-based study of black, Hispanic and white patients with ovarian cancer. *BMJ Open* 2021;**11**:e052808. doi:10.1136/bmjopen-2021-052808

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-052808).

Received 29 April 2021
Accepted 10 September 2021

For numbered affiliations see end of article.

Correspondence to
Dr Tomi Akinyemiju;
tomi.akinyemiju@duke.edu

ABSTRACT

Introduction Less than 40% of patients with ovarian cancer (OC) in the USA receive stage-appropriate guideline-adherent surgery and chemotherapy. Black patients with cancer report greater depression, pain and fatigue than white patients. Lack of access to healthcare likely contributes to low treatment rates and racial differences in outcomes. The Ovarian Cancer Epidemiology, Healthcare Access and Disparities study aims to characterise healthcare access (HCA) across five specific dimensions—Availability, Affordability, Accessibility, Accommodation and Acceptability—among black, Hispanic and white patients with OC, evaluate the impact of HCA on quality of treatment, supportive care and survival, and explore biological mechanisms that may contribute to OC disparities.

Methods and analysis We will use the Surveillance Epidemiology and Ends Results dataset linked with Medicare claims data from 9744 patients with OC ages 65 years and older. We will recruit 1641 patients with OC (413 black, 299 Hispanic and 929 white) from cancer registries in nine US states. We will examine HCA dimensions in relation to three main outcomes: (1) receipt of quality, guideline adherent initial treatment and supportive care, (2) quality of life based on patient-reported outcomes and (3) survival. We will obtain saliva and vaginal microbiome samples to examine prognostic biomarkers. We will use hierarchical regression models to estimate the impact of HCA dimensions across patient, neighbourhood, provider and hospital levels, with random effects to account for clustering. Multilevel structural equation models will estimate the total, direct and indirect effects of race on treatment mediated through HCA dimensions.

Ethics and dissemination Result dissemination will occur through presentations at national meetings and in collaboration with collaborators, community partners and colleagues across othercancer centres. We will disclose findings to key stakeholders, including scientists, providers and community members. This study has been approved by the Duke Institutional Review Board (Pro00101872). Safety considerations include protection of patient privacy. All disseminated data will be deidentified and summarised.

Strengths and limitations of this study

- Diverse sample population (black, white and Hispanic patients from nine different state cancer registries).
- Evaluation of multiple facets of healthcare access (Accessibility, Acceptability, Affordability, Accommodation and Availability) in a population-based sample across multiple levels of influence with regard to multiple outcomes across the cancer care continuum.
- Incorporation of a cell to society approach by characterisation of measures across multiple levels—biological (saliva and microbiome), patient (sociodemographics, comorbidities, patient-reported outcomes) and social (dimensions of healthcare access).
- Potential for recall bias.
- Inability to make causal inferences.

BACKGROUND

In 2020, there were an estimated 21 750 newly diagnosed cases and an estimated 13 940 deaths attributed to ovarian cancer (OC) in the USA,[1] accounting for 1.2% of all cancers among women and 2.3% of all cancer deaths. Diagnosis at early stages, when the cancer is most amenable to treatment, is difficult due to the lack of easily recognisable symptoms and effective screening tests. Consequently,

up to 75% of patients with OC present with cancer that has spread regionally or has metastasised. For these patients, survival is highly dependent on the quality of treatment. However, data from the National Cancer Institute (NCI) Patterns of Care studies indicate that fewer than half of all patients with OC received guideline-recommended stage-appropriate surgery and multiagent chemotherapy treatment.[2] Lack of quality treatment also contributes to the low 5-year survival rate for OC, which is currently 47%. In addition, striking racial disparities in OC survival persist. Among white women, survival improved from 35% to 47% between 1975–1977 and 2008–2014, but declined from 42% to 39% among black women in the same time period.[3] Lack of access to advanced treatment options has been well documented as a major contributor to lower survival rates and marked disparities in OC survival,[4 5] however, few studies to our knowledge has comprehensively evaluated racial differences across multiple dimensions of healthcare access (HCA) to identify those most impactful for treatment and survival.

HCA is defined as five separate but inter-related dimensions: Availability—the type, quality and quantity of healthcare resources; Affordability—the ability to pay for healthcare; Accessibility—the location of healthcare resources in relation to the patient; Accommodation—the organisation of healthcare resources in relation to the patient's constraints and preferences; and Acceptability—the patient's attitude, perceptions and quality of interaction with healthcare providers.[6 7] Prior studies have examined some aspects of HCA dimensions in relation to OC outcomes, such as insurance status and hospital volume[8 9]; for instance, while black patients with breast cancer were more likely to reside in counties with greater availability of healthcare resources (e.g., greater number of oncology hospitals), compared with white patients,[10] they were less likely to receive surgery and experienced lower survival rates. This suggests that availability alone is insufficient for access to care and other dimensions must also be considered. Two HCA dimensions in particular, accommodation and acceptability, are not routinely captured in administrative claims databases and are particularly understudied, although these may be key to understanding how aspects of HCA, such as patient preferences, trust and communication with providers, which interact with other dimensions of HCA, and may be differential by race and impact OC outcomes.

In addition to examining HCA as a social determinant of health, it is also important to evaluate biological mechanisms underlying OC prognosis and disparities. While several aetiological risk factors for OC have been established (including BRCA mutation, parity and oral contraceptive use),[11 12] few studies have characterised biological and genomic factors predicting racial differences in OC prognosis . For instance, there is scientific evidence suggesting an important role for the cervicovaginal microbiome in OC risk and prognosis,[12] and a recent study observed that black and Hispanic women were over-represented among patients that lacked significant lactobacilli and had a higher proportion of anaerobic bacteria.[13] Racial differences in microbiome composition, and those associated with genetic ancestry, may be linked with OC prognosis and partly explain racial disparities observed in OC outcomes[14 15] and deserve further scrutiny.

Therefore, the goal of the Ovarian Cancer Epidemiology, Healthcare Access and Disparities (ORCHiD) study is to address these knowledge gaps by (1) fully characterising HCA dimensions among black, Hispanic and white patients with OC, (2) evaluating the association of HCA dimensions with racial disparities in receipt of quality, guideline-adherent initial treatment, supportive care and survival, (3) and characterising the biological mechanisms underlying OC disparities. The conceptual framework for ORCHiD is presented in figure 1. This comprehensive approach is a necessary first step towards identifying specific interventions focused on important, modifiable social and biological factors to eliminate disparities.

METHODS AND ANALYSIS

Study design and organisation

The ORCHiD study design incorporates two distinct approaches: (1) a retrospective cohort design of black,

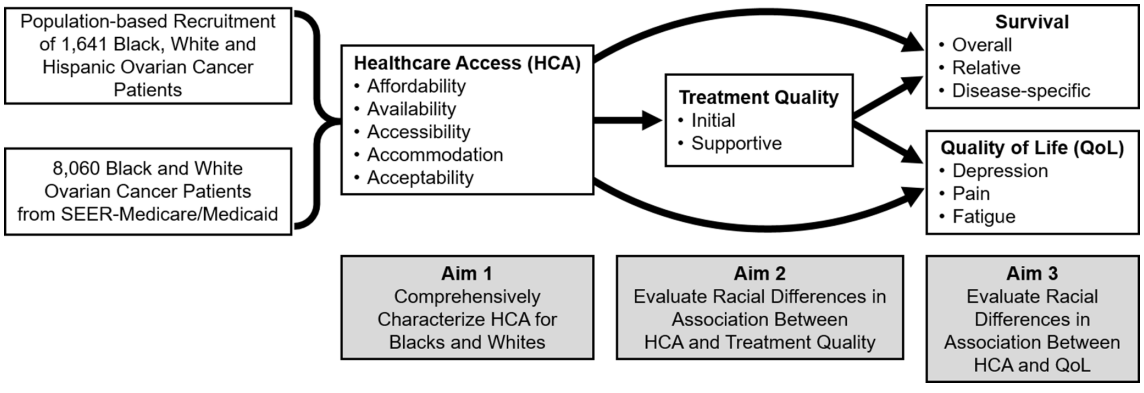

Figure 1 Conceptual overview of ORCHiD aims. ORCHiD, Ovarian Cancer Epidemiology, Healthcare Access and Disparities; SEER, Surveillance Epidemiology and Ends Results.

Hispanic and white patients with OC ages 65–75 years (using secondary data from the Surveillance Epidemiology and Ends Results (SEER)-Medicare linked dataset) and (2) a population-based prospective cohort design with primary survey data of black, Hispanic and white patients with OC ages 20–75 years identified from cancer registries in California, Georgia, Kentucky, New York, North Carolina, Texas, Maryland, Florida and Mississippi. In the SEER-Medicare cohort, we will evaluate the association of three HCA dimensions measurable with administrative data (Affordability, Availability and Accessibility) in relation to guideline adherent initial treatment and supportive care and survival. In the prospective cohort, we will evaluate all five HCA dimensions using a combination of administrative claims data and self-reported measures in relation to guideline adherent treatment and survival, as well as patient-reported outcomes (PROs) such as pain, fatigue and depression.

Retrospective cohort study design

We obtained SEER-Medicare data on 9744 black, Hispanic and white patients with OC, stages I–IV, ages 65 years or older, diagnosed with a first primary OC between 2008 and 2015. Patients were excluded if their OC was diagnosed at autopsy or death, or if OC was not either their first or second primary cancer diagnosis. Patients were required to have at least 12 months of continuous enrolment in Medicare fee-for-service parts A and B prior to the SEER diagnosis, and to have at least one Medicare inpatient, outpatient or carrier claim with a diagnosis code for OC (International Classification of Disease ICD-9-CM and ICD-10-CM diagnosis codes 183.0 or C569) within 2 months of the SEER diagnosis; the date of the earliest claim will serve as the patient's exact OC diagnosis date as SEER provides only the month and year of diagnosis. Patients were required to have continuous fee-for-service Medicare enrolment in the 12 months following their diagnosis date, or until death. The SEER dataset is routinely linked with Medicare claims data for patients with cancer ages 65+, and includes all SEER data items (sociodemographics, cancer stage, subtype, date of initial diagnosis, vital status and survival time), Medicare claims data from the Center for Medicare and Medicaid Services (CMS)[16] and linked external datasets. The SEER-Medicare cohort is linked with external data from the US Census Bureau American Community Survey (ACS), the Area Healthcare Resource file and the NCI hospital files to obtain data at the patient, neighbourhood, physician and hospital levels. The ACS census tract and ZIP code file was used to link region-specific socioeconomic data to the claims records; the Hospital file was used to determine whether a treatment facility was affiliated with an NCI designated cancer centre, whether a facility was ever non-accredited, total number of beds, number of Medicare/Medicaid certified beds, participation in Medicare/Medicaid, urban vs rural location, or teaching status as a measure of academic standing. SEER-Medicare provides the Hospital file linked to their identifiers. We then evaluated the three HCA dimensions—Availability, Affordability and Accommodation—that can be feasibly measured in SEER-Medicare.

Prospective cohort study design

We will recruit 1641 black, Hispanic and white patients with OC ages 20–75 years from nine population-based cancer registries including California, Georgia, Kentucky, New York, North Carolina, Texas, Maryland, Florida and Mississippi (figure 2). These registries cover a diverse range of patients, including typically difficult-to-reach populations, such as Appalachian whites in Kentucky and rural blacks in the Mississippi Delta. We chose to focus on this age group in the prospective cohort (20–75 years) because the median age at diagnosis for OC is 63 years, and 55% of cases are younger than 65 years at diagnosis, enabling us to examine OC outcomes among patients not captured in SEER-Medicare. Sociodemographic and clinical data from registry records will be used to identify patients eligible for the study based on: black, Hispanic or white race; between the ages of 20 and 75 at initial diagnosis; a resident of the registry geographical region at the time of diagnosis and initial contact; at least 6 months postdiagnosis; first primary clinically confirmed OC. To ensure that we have an adequate number of racially diverse participants, we will include patients diagnosed with any stage and subtype of OC and adjust for these variables in statistical models. We will also attempt to recruit all black patients with OC in each registry and use a stratified random sampling approach to select white patients age matched to black patients on ages 20–64 and 65–75 years. We will attempt to recruit all Hispanic patients in two registries—Texas and California—that have high Hispanic populations.

Prospective cohort recruitment

For each of the target state registries, registry staff will review patient data/records to confirm study eligibility. For eligible patients, a letter describing the study will be sent to the treating physician (when required by the registry). The letter will ask the physician to contact the registry if there are reasons the subject should not be contacted. If no reasons are reported or there is no response from the physician within 2–3 weeks, the registries will send eligible participants a letter informing them of their eligibility for the research study along with a preaddressed postage paid response card which may be used to select whether or not they want their names to be released to the ORCHiD study team. In some registries, if a response card is not received within 2 weeks, phone calls will be made by registry staff in an attempt to get verbal permission and only those who agree to have their names released will be forwarded to the ORCHiD study team for recruitment. At regular intervals, the ORCHiD research team will receive a list of eligible participants from participating registries. ORCHiD study staff will mail out a study packet containing an introduction letter, the study brochure, any registry-required documentation,

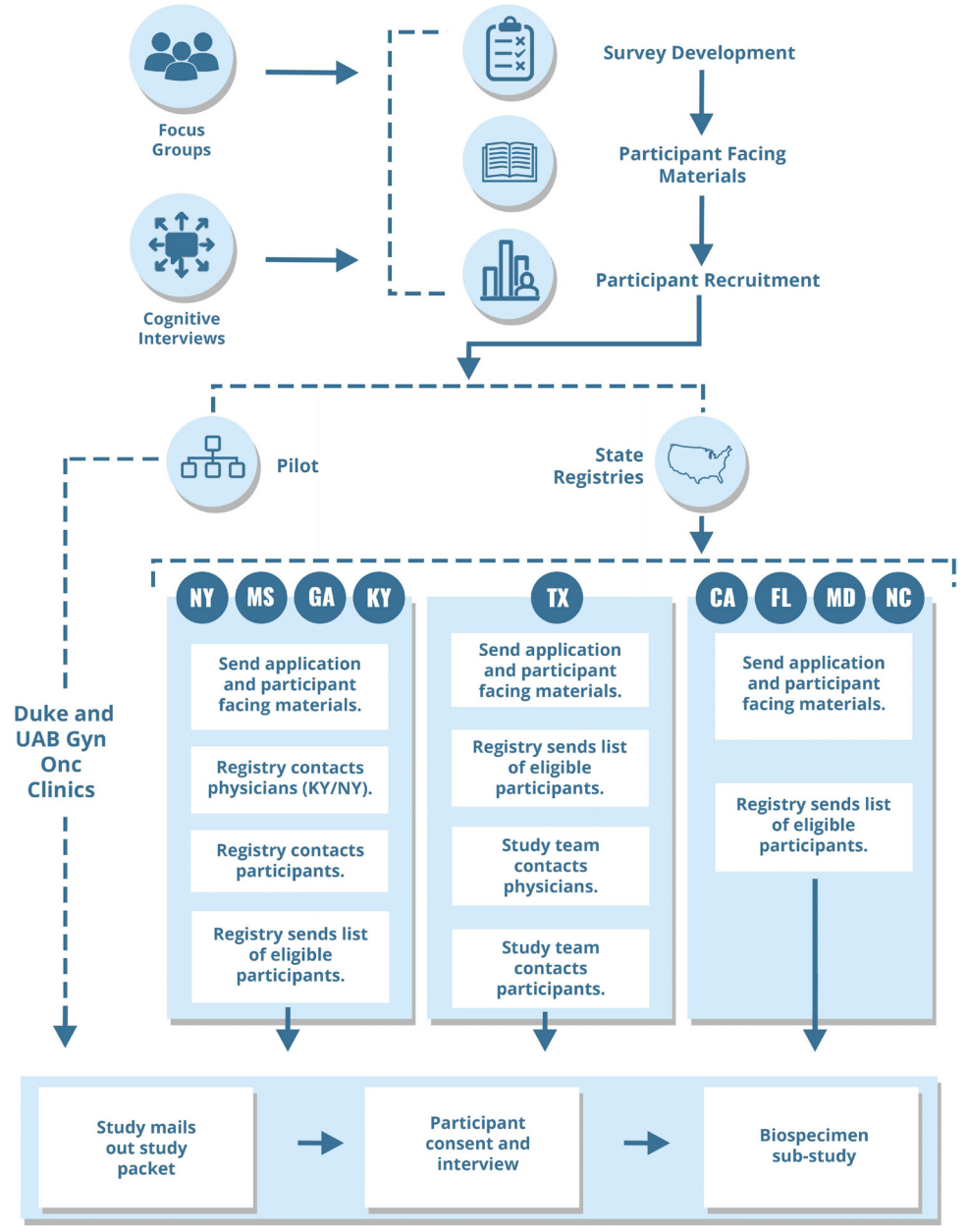

Figure 2 Prospective primary cohort study schema. UAB, University of Alabama at Birmingham.

the study consent form and a prepaid envelope. Between 2 and 3 weeks after the study packet is mailed out, the study interviewers will begin making phone calls to participants to enrol them into the study, answer any questions, conduct informed consent and complete the survey. Phone contact will follow a schedule of attempts during day, evening, and weekend hours. The ORCHiD study was granted a waiver of documentation of consent, therefore, verbal consent obtained during the phone call and documented by the interviewer on an institutional review board (IRB)-approved phone script enables the survey interview to occur during the same phone call. In addition, while not required, participants will be encouraged to also provide documentation of consent via an electronic consent form on REDCap or by signing and returning the mailed paper consent form. The study survey is expected to last 55–65 minutes and can be completed at one time during the call or rescheduled, if needed. On completion of the survey, the biospecimen substudy component will be introduced to participants and interested participants will be administered a biospecimen consent and then subsequently mailed the biospecimen collection kit. An incentive of a US$25 gift card will be provided for completing the study survey, and an additional US$20 gift card for returning the biospecimen kit.

Study measures and data collection

To create the unique richness and depth of data for ORCHiD, we will use a variety of data sources as described below to characterise measures at the patient, neighbourhood, physician and hospital levels.

Table 1 Healthcare access dimensions, component measures and geographic level of assessment

Dimension	Measure	Level of assessment	Data source
Affordability Price, willingness and ability to pay for healthcare services	Income	Individual	Survey*
	Insurance type	Individual	Survey
	Insurance co-pay	Individual	Survey
	Out-of-pocket cost	Individual	Survey
	Census tract Median Household income, poverty	Neighbourhood (census tract)	US census
	% Uninsured	Neighbourhood (census tract)	US census
Availability Type, quality and volume of healthcare services in relation to patient need	Usual healthcare provider	Individual	Survey
	No of hospitals	Neighbourhood (county)	Area resource file
	No of Gyn-oncologists, Ob-Gyn	Neighbourhood (county)	Area resource file
	Physician specialty	Physician	SEER-medicare
	Physician ovarian cancer patient volume	Physician	SEER-medicare
	Hospital teaching status	Hospital	Area Resource File
	Hospital quality metrics	Hospital	SEER-medicare
Accessibility Location of healthcare services in relation to patients	Travel time (minutes)	Individual	Survey
	Mode of transport	Individual	Survey
	Distance to hospital (miles)	Geographic	Survey+Geographic Information System (GIS)
	Rural/urban location	Patient, hospital	Survey+area resource file
Accommodation Organisation of healthcare services and resources in relation to patients' ability to accommodate such services	Hospital bed size	Hospital	Area resource file
	Hospital average wait times	Hospital	Hospital information
	Access to support services	Individual	Survey
Acceptability Patient attitude to personal and practice characteristics of healthcare provider	Trust (oncologist, primary care provider)	Individual	Survey
	Comfort (oncologist)	Individual	Survey
	Race concordance (oncologist)	Individual, physician	Survey
	Reputation/credentials (oncologist)	Individual	Survey
	Transportation/distance to facility	Individual	Survey
	Time away from work/family	Individual	Survey

*Participant surveys conducted 6–12 months postdiagnosis for primary ovarian cancer diagnosis to establish healthcare access characteristics around the time of diagnosis that may influence choice of healthcare facility and provider.
SEER, Surveillance Epidemiology and Ends Results.

Measuring HCA dimensions

The Penchansky and Thomas framework[6 7] provides five dimensions of HCA. As OC outcomes are particularly impacted by access to care, we will use this framework as a comprehensive measure of access using variables relevant to OC that can be measured at either the patient, neighbourhood, provider or health system levels using data sources as outlined in table 1. As part of formative work for ORCHiD, we conducted a comprehensive review of the literature to identify existing instruments to characterise HCA across all five dimensions. However, we noted significant challenges: (1) none of the existing instruments comprehensively assessed all five dimensions; most surveys covered aspects of one specific dimension for example, financial toxicity as part of the Affordability component, patient–provider communication as part of the Acceptability component, but these were not comprehensive even within each dimension; (2) none of the existing instruments appeared to have considered the varying experiences of different racial groups in developing the items, making it unclear if the questions would capture the experiences of black and Hispanic patients, the major focus of ORCHiD and (3) combining questions or subscales from various surveys into one HCA instrument

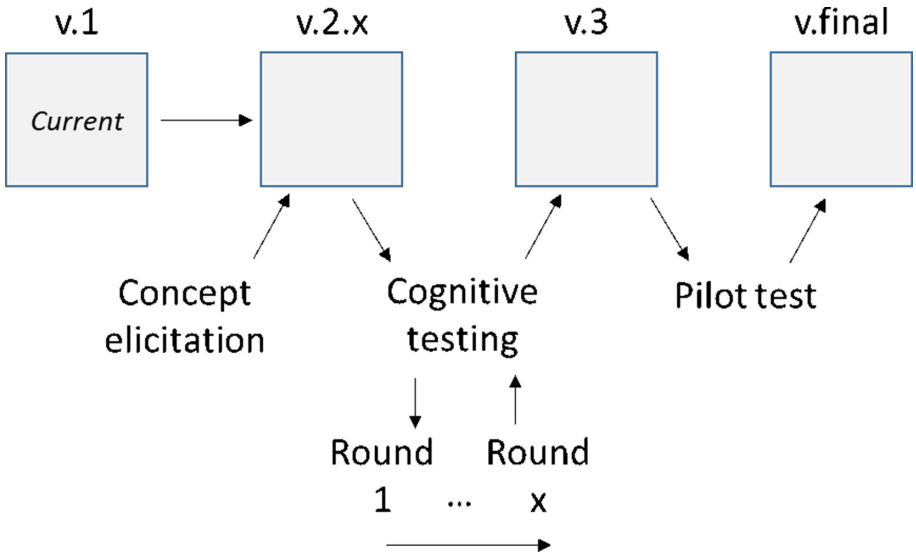

Figure 3 Conceptual framework for qualitative study and survey development.

would likely result in limited validity with poor psychometric properties. Therefore, we addressed these critical limitations by conducting an initial qualitative study that included focus group discussions and cognitive interviews with cancer survivors to provide content validation data for the HCA survey to be used in our prospective cohort. Additionally, the original framework was developed in the context of primary care. Data gained from focus groups identified how each HCA dimension applied to a socioeconomic and racially diverse group of patients with cancer, allowing us to obtain measures of access specific to cancer care in each of the dimensions. The process of concept elicitation with focus groups, cognitive interviews and pilot testing is described below and summarised in figure 3.

1. Draft survey: We began drafting a questionnaire by pulling survey items from existing, validated surveys that addressed the topics of interest. Validated questionnaires are instruments that have undergone evaluation to ensure the instrument measures what it is intended to measure.
2. Concept elicitation: To guide the development of survey questions that were not captured by existing instruments, we conducted focus group discussions with 32 female cancer survivors (63% black and 18% Hispanic) using a semistructured interview guide. The focus groups solicited input regarding their perceptions of each dimension of HCA, experiences accessing care after their cancer diagnosis, and facilitators and barriers to receiving quality treatment. Information from the focus groups were synthesised and analysed and used to guide survey development. Focus group participants were provided a US$25 gift card for participation.
3. Cognitive testing: We conducted cognitive testing of the HCA survey items with in-depth interviews among seven patients with OC diagnosed at Duke University (three black, one Hispanic and three white patients). Interview participants were asked to answer the questionnaire, think about their responses (i.e., retrospective probing) and describe their thought processes after responding. The interviewer probed on (1) comprehension (level of understanding), (2) clarity (level of straightforward meaning), (3) knowledge and memory (ease of recall of information needed to respond) and (4) judgement (ease of fitting personal experiences to the measure response options). Debriefing forms were completed for each interview and used by the study team to review findings and modify questions as needed. Participants were provided a US$25 gift card for participation.
4. Pilot testing: With the ORCHiD survey at near final version, having been edited for length, clarity and content, we conducted a pilot study of 50 patients with OC recruited from the Duke Cancer Center and the University of Alabama at Birmingham Comprehensive Cancer Center. The full-length survey was administered by phone (preferred option) or self-administered electronically. A small proportion of patients preferred to complete the paper version and were mailed a shortened version of the full survey to enhance completion. The biospecimen substudy component was also tested in this phase. Feedback from pilot study participants regarding the survey was incorporated into the draft version before being finalised for the main study. Participants were provided a US$25 gift card incentive for completing the survey and an additional US$20 gift card for providing biospecimen.

Measuring guideline-recommended treatment

In both the retrospective and prospective cohorts, we will use previously well-validated treatment algorithms combining registry and claims data to assess receipt of guideline-adherent OC treatments.[17–19] We will compare treatments received to the treatments recommended in professional guidelines published by the National Comprehensive Cancer Network.[20] We will also select

well-defined supportive care quality measures from the CMS Quality Payment Programme, including medications for treatment of pain, depression or anxiety. We will evaluate care quality for comorbid conditions, measured by hospitalisations and emergency room visits for potentially avoidable admissions related to suboptimal management of multimorbidity. Concordance between recommended treatment and treatment outcomes in this cohort will be assessed from multiple overlapping data sources: (1) patient-reported treatment information from the survey (in prospective cohort); there is evidence that patients can reliably report receipt of recent radiation and chemotherapy, although not the specific agents used,[21–24] (2) cancer registry data on surgery and chemotherapy for all participants; these data have been shown to have a level of reliability similar to claims data,[25] (3) claims data on treatment through linkages with Medicare or Medicaid claims data and (4) a manual review of medical records for patients in the Kentucky registry as further validation. These multiple layers of treatment information in the primary survey cohort will minimise the likelihood of misclassification and also enable us to assess the validity of treatment information obtained from survey data.

Measuring PROs and other patient variables

We will measure patient-reported symptoms using the Patient-Reported Outcomes Measurement Information System (PROMIS) four-item short forms in the prospective cohort study. PROMIS item banks have been extensively tested,[26 27] and short forms are reliable and valid for use in the general US population, in cancer-specific populations[28] and in non-white and low-income subgroups.[29–32] We will assess four symptom domains widely recognised as prevalent and manageable in cancer survivors: fatigue, depression, anxiety and sleep disturbance. Self-efficacy for healthcare will also be assessed with a PROMIS measure.[33 34] We will use two out of five possible subdomains from this item bank: confidence in managing symptoms and confidence in managing medications and also assess functional status. Registry-derived data including sociodemographics, month and year of diagnosis, cancer site, histologic type, stage at diagnosis, survival time and vital status, and first course of therapy will be obtained for all patients in both cohorts. Since comorbidity status is a significant predictor of poor survival with marked racial differences, we will assess comorbidities in both cohorts. In the retrospective SEER-Medicare cohort, we use claims data to create a comorbidity index,[35] and in the prospective cohort, we ask respondents about common and serious comorbid conditions adapted from the Charlson index. We also include measures of select risk factors (e.g., smoking, body mass index) that are both associated with poor survival and also potentially modifiable targets for improving outcomes.[36 37]

Biospecimen collection

Following completion of the survey in the prospective cohort, participants will be asked if they are interested in participating in future research. If they decline, no further studies will be introduced, however, if they agree, the participants will be offered the opportunity to participate in the biospecimen substudy. Interested participants will be asked to provide self-collected cervicovaginal microbiome specimen using the DNA Genotek OMR-130 kit (or similar) and saliva sample using DNA Genotek OGR-500 kit (or similar). Samples will be shipped back to the study team at Duke for processing and analysis, and molecular data will be linked to clinical data obtained in the ORCHiD study for future analysis.

Patient and public involvement

Essentially, the goal of ORCHiD is to address these knowledge gaps by (1) fully characterising HCA dimensions among black, Hispanic, and white patients with OC, (2) evaluating the association of HCA dimensions with racial disparities in receipt of quality, guideline-adherent initial treatment, supportive care and survival and (3) and characterising the biological mechanisms underlying OC disparities. Patients, cancer survivors and community members were involved in the study design, and contributed important insights into the recruitment and conduct of this research study. The study protocol was presented to the Duke Cancer Institute Community Advisory Council for feedback, and information obtained during the focus group with cancer survivors and the pilot study of ovarian cancer survivors were incorporated into the study design. Feedback on each step of the study design, pilot and ongoing recruitment activities are disseminated back to patients and survivors, with a plan to provide ongoing updates.

Analytical plan

Descriptive statistics will be conducted for all variables in each study cohort overall and by race, region and age group. We will compare categorical variables by group using χ^2 tests, and continuous variables using t-tests or non-parametric tests as appropriate. Hierarchical multiple regression analysis will be used to evaluate the association between each main exposures of interest (HCA dimensions) and study outcomes (guideline-adherent treatment, PROs, survival), accounting for the clustering of patients within neighbourhoods (e.g., county), physicians and hospitals, and adjusting for a wide range of relevant study covariates. Variable collinearity will be assessed using the variance inflation factor (VIF) method. Covariates with VIF values of 10 or higher will be dropped from the multivariable models. Additionally, the planned factor analysis will be used to create composite scores for the HCA dimensions; these scores will be used in lieu of adjustment for each individual variable, which should help to reduce issues of collinearity between variables of interest. Non-response from the survey and its characteristics will be recorded and evaluated to identify potential biases. We will conduct multiple imputation, for instance, full conditional specification approach, to handle missing data. If missing data are deemed to be not

missing at random, non-standard conditionally specified models will be used.

Sample size and power analysis

To estimate the sample size for the primary survey cohort, we focused our power calculation on the outcome for Aim 2: receipt of guideline-adherent treatment. A 10%–14% difference in the receipt of guideline-adherent treatment was observed between black and white patients ages 65+ in a population-based cohort,[2 38] with larger differences observed in those less than age 65.[4 39] We will recruit all potential black patients with OC from participating registries, age group-matched white patients and all potential Hispanic patients with OC from Texas and California registries. We will, therefore, have a total sample size of 9744 patients in the SEER-Medicare cohort, and 1641 patients (413 blacks, 299 Hispanics and 929 whites) in the primary survey cohort. Accounting for intraclass correlation coefficient (conservatively assumed to not exceed 0.1) within the multilevel statistical model, and assuming 60% of study subjects reside in neighbourhoods with high HCA availability, we will have 80% power to detect racial differences in treatment as small as 6% using SEER-Medicare data and as small as 14% using the primary survey cohort data. Based on similar assumptions, we will be well powered to detect hazard ratios as small as 1.09 in SEER-Medicare data, and as small as 1.21 in the primary survey cohort. We will also have 80% power to detect at least a 1.5-point difference (or 0.15*SD) in PROMIS scores.

DISCUSSION

The ORCHiD study will generate novel and valuable data on a diverse patient population to identify salient factors associated with disparities in initial treatment, supportive care and quality of life among patients with OC. This study incorporates a cell to society approach by characterising measures across multiple levels—biological (saliva and microbiome), patient (sociodemographics, comorbidities, PROs) and social (HCA dimensions)—in relation to OC outcomes and racial disparities, making our study uniquely comprehensive in scope and addressing several major limitations in the current literature. HCA is fundamental to receipt of quality, guideline-adherent treatment, which is critical to survival. Understanding racial differences in salient HCA dimensions that incorporate the experiences of black and Hispanic patients will provide unique insights.

Measures of HCA are derived from the Penchansky and Thomas framework that was published in 1981. This comprehensive study provided a framework for dimensions of HCA and specified five in particular (Availability, Affordability, Accommodation, Acceptability and Accessibility). [40] Data collected will allow us to identify specific interventions to address barriers to HCA among cancer patients in order to ensure that all patients receive quality and timely treatment.

The first scientific papers for ORCHiD will describe results of the qualitative analysis of HCA focus groups with diverse cancer survivors and present empirical data on racial differences in affordability, availability and accessibility HCA dimensions on quality of initial treatment from the retrospective SEER-Medicare cohort study. Subsequent papers will describe HCA in relation to supportive care and quality of life, and report on all five HCA dimensions from the prospective cohort.

The study design for ORCHiD is motivated primarily by an urgent need to close the treatment and survival gap between black and white patients with OC—most, if not all of the benefits of improved treatment strategies over the past several decades have been experienced by white patients, while survival rates have declined for black women. Importantly, when black and white women who received complete guideline-adherent treatment are compared, racial disparities are no longer apparent.[41 42] There is a unique opportunity to address this gap by understanding the drivers of poor access to care among black and Hispanic women and developing targeted interventions to address them. As with most observational studies that rely on self-report, there are several inherent limitations to our study design, such as the potential for recall bias, inability to make causal inferences and potential for differential misclassification of exposures. However, our study design addresses these challenges by: (1) collecting data from several distinct sources, including standardised cancer registry data and external secondary data, (2) applying advanced statistical methodological approaches to account for the multilevel clustered nature of the dataset and to mitigate selection bias, (3) conducting extensive prestudy activities to ensure that our recruitment materials are compelling among black and Hispanic, hard-to-reach populations and (4) using existing validated survey measures on outcomes of interest, such as PROs.

By using high-quality data and a comprehensive cell to society approach, we will not only characterise racial disparities in OC care and survival, but we will also determine the role of differential access to care and biological mechanisms that sustain these disparities, enabling us to assess the relative importance of each factor independently and jointly. Our data will provide useful information regarding which HCA dimensions and biological factors are most relevant and highlight specific areas where tailored clinical and public health strategies are needed to improve quality care for all patients with OC. While we are unable to infer causality, our data will provide relevant information as a guide to investigating specific healthcare policies, such as value-based payment models or collaborative care models, regarding strategies that may promote the quality and value of healthcare services while reducing disparities in health outcomes.

Ethics and dissemination

Duke University's Institutional Review Board reviewed and approved the ethics associated with this study

(Pro00101872). As it involves collection of personal health information, we have a detailed security plan that involves deidentification of survey answers and biological samples when possible. Sensitive information is kept on encrypted and firewalled databases. This study has considerable potential for contributing high-impact findings, many of which might have direct public health benefit. There are two parts to our plan: First, the study's investigators are very active in national and international forums dealing with their respective areas of expertise. Through the leading roles they play at such gatherings, they will present the findings from this new study on a regular basis and discuss the public health implications as well as implications for future research. Second, cancer clinicians and scientists are among the key target audiences for the proposed study. The study measurements represent the cutting edge of implementation science, oncology and cancer disparities research. We will continue to foster wide interest in these areas by keeping the practitioners apprised of our findings and their relevance. We plan reach out to other cancer centres to describe our research findings, and to target national meetings such as AcademyHealth, American Assoication for Cancer Research, American Society for Preventive Oncology, American Society of Clinical Oncology, and Society for Epidemiologic Research, among others. We will disseminate findings to key stakeholders, including, scientists and providers, and to community members through presentations of findings to the Community Advisory Board, at local, state and local conferences, and actively seek out opportunities to engage community members. We will also disseminate findings to our study participants via a newsletter at the end of the study.

Author affiliations
[1]Department of Population Health Sciences, Duke University School of Medicine, Durham, North Carolina, USA
[2]Duke University School of Medicine, Duke Cancer Institute, Durham, North Carolina, USA
[3]Department of Biostatistics and Kentucky Cancer Registry, University of Kentucky, Lexington, Kentucky, USA
[4]Division of Preventive Medicine, The University of Alabama, Birmingham, Alabama, USA
[5]Division of Hematology and Supportive Care, University of Alabama, Birmingham, Alabama, USA
[6]Georgia Cancer Registry, Emory University, Atlanta, Georgia, USA
[7]New York State Cancer Registry, New York State Department of Health, Albany, New York, USA
[8]Division of Gynecologic Oncology, Duke University School of Medicine, Durham, North Carolina, USA
[9]Georgetown University Medical Center, Washington, District of Columbia, USA

Acknowledgements We thank all the cancer patients, survivors and patient navigators who participated in the ORCHiD study for their vital contribution in advancing the science of cancer disparities in the United States.

Contributors TA conceptualised and led the study; LW conducted statistical analysis and data management; AJ, AG, AD, MB, CO and OO contributed to data collection, project management and analysis; BH, MP, ALP, ML, MM, ED, LJF, KW, MS and AB all contributed to study design, data collection and interpretation of results. All authors have reviewed and approved the final version of the manuscript. All included authors provided written informed consent.

Funding This research was funded by the National Institutes of Health/National Cancer Institute (Grant Number R37CA233777 and 5R37CA23377).

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Tomi Akinyemiju http://orcid.org/0000-0002-1412-3234
April Deveaux http://orcid.org/0000-0002-1185-7750
Lauren Wilson http://orcid.org/0000-0002-5953-2293
Ashwini Joshi http://orcid.org/0000-0002-9287-6370
Malcolm Bevel http://orcid.org/0000-0002-9907-0089

REFERENCES

11 NationalCancerInstitute. NCI SEER cancer statistics: cancer STAT facts: ovarian cancer, 2020. Available: https://seer.cancer.gov/statfacts/html/ovary.html
2 Warren JL, Harlan LC, Trimble EL, *et al*. Trends in the receipt of guideline care and survival for women with ovarian cancer: a population-based study. *Gynecol Oncol* 2017;145:486–92.
3 AmericanCancerSociety. *ACS cancer facts and figures 2017*. Atlanta, GA: American Cancer Society, 2017. https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2017/cancer-facts-and-figures-2017.pdf
4 Bristow RE, Powell MA, Al-Hammadi N, *et al*. Disparities in ovarian cancer care quality and survival according to race and socioeconomic status. *J Natl Cancer Inst* 2013;105:823–32.
5 Aranda MA, McGory M, Sekeris E, *et al*. Do racial/ethnic disparities exist in the utilization of high-volume surgeons for women with ovarian cancer? *Gynecol Oncol* 2008;111:166–72.
6 Penchansky R, Thomas JW. The concept of access: definition and relationship to consumer satisfaction. *Med Care* 1981;19:127–40.
7 Thomas JW, Penchansky R. Relating satisfaction with access to utilization of services. *Med Care* 1984;22:553–68.
8 Bristow RE, Chang J, Ziogas A, *et al*. High-volume ovarian cancer care: survival impact and disparities in access for advanced-stage disease. *Gynecol Oncol* 2014;132:403–10.
9 Long B, Chang J, Ziogas A, *et al*. Impact of race, socioeconomic status, and the health care system on the treatment of advanced-stage ovarian cancer in California. *Am J Obstet Gynecol* 2015;212:468.e1–468.e9.
10 Akinyemiju T, Moore JX, Ojesina AI, *et al*. Racial disparities in individual breast cancer outcomes by hormone-receptor subtype, area-level socio-economic status and healthcare resources. *Breast Cancer Res Treat* 2016;157:575–86.
11 Fortner RT, Poole EM, Wentzensen NA, *et al*. Ovarian cancer risk factors by tumor aggressiveness: an analysis from the ovarian cancer cohort Consortium. *Int J Cancer* 2019;145:58–69.
12 Nené NR, Reisel D, Leimbach A, *et al*. Association between the cervicovaginal microbiome, BRCA1 mutation status, and risk of ovarian cancer: a case-control study. *Lancet Oncol* 2019;20:1171–82.
13 Ravel J, Gajer P, Abdo Z, *et al*. Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci U S A* 2011;108 Suppl 1:4680–7.
14 Champer M, Wong AM, Champer J, *et al*. The role of the vaginal microbiome in gynaecological cancer. *BJOG* 2018;125:309–15.
15 Chase D, Goulder A, Zenhausern F, *et al*. The vaginal and gastrointestinal microbiomes in gynecologic cancers: a review of applications in etiology, symptoms and treatment. *Gynecol Oncol* 2015;138:190–200.
16 Warren JL, Klabunde CN, Schrag D, *et al*. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population. *Med Care* 2002;40:IV-3-18.
17 Lund JL, Stürmer T, Harlan LC, *et al*. Identifying specific chemotherapeutic agents in Medicare data: a validation study. *Med Care* 2013;51:e27–34.

18 Livaudais-Toman J, Egorova N, Franco R, *et al*. A validation study of administrative claims data to measure ovarian cancer recurrence and secondary debulking surgery. *EGEMS* 2016;4:22.
19 Bristow RE, Chang J, Ziogas A, *et al*. Sociodemographic disparities in advanced ovarian cancer survival and adherence to treatment guidelines. *Obstet Gynecol* 2015;125:833–42.
20 National Comprehensive Cancer Network. *NCCN guidelines for treatment of cancer*. Fort Washington, PA: National Comprehensive Cancer Network, 2017.
21 Liu Y, Diamant AL, Thind A, *et al*. Validity of self-reports of breast cancer treatment in low-income, medically underserved women with breast cancer. *Breast Cancer Res Treat* 2010;119:745–51.
22 Maunsell E, Drolet M, Ouhoummane N, *et al*. Breast cancer survivors accurately reported key treatment and prognostic characteristics. *J Clin Epidemiol* 2005;58:364–9.
23 Clegg LX, Potosky AL, Harlan LC, *et al*. Comparison of self-reported initial treatment with medical records: results from the prostate cancer outcomes study. *Am J Epidemiol* 2001;154:582–7.
24 Zhu K, McKnight B, Stergachis A, *et al*. Comparison of self-report data and medical records data: results from a case-control study on prostate cancer. *Int J Epidemiol* 1999;28:409–17.
25 Noone A-M, Lund JL, Mariotto A, *et al*. Comparison of SEER treatment data with Medicare claims. *Med Care* 2016;54:e55–64.
26 Cella D, Yount S, Rothrock N, *et al*. The patient-reported outcomes measurement information system (PROMIS): progress of an NIH roadmap cooperative group during its first two years. *Med Care* 2007;45:S3–11.
27 DeWalt DA, Rothrock N, Yount S, *et al*. Evaluation of item candidates: the PROMIS qualitative item review. *Med Care* 2007;45:S12–21.
28 Cella D, Riley W, Stone A, *et al*. The patient-reported outcomes measurement information system (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. *J Clin Epidemiol* 2010;63:1179–94.
29 Teresi JA, Ocepek-Welikson K, Kleinman M, *et al*. Measurement equivalence of the patient reported outcomes measurement information system® (PROMIS®) anxiety short forms in ethnically diverse groups. *Psychol Test Assess Model* 2016;58:183–219.
30 Teresi JA, Ocepek-Welikson K, Kleinman M, *et al*. Psychometric properties and performance of the patient reported outcomes measurement information system® (PROMIS®) depression short forms in ethnically diverse groups. *Psychol Test Assess Model* 2016;58:141–81.
31 Fieo R, Ocepek-Welikson K, Kleinman M, *et al*. Measurement equivalence of the patient reported outcomes measurement information system® (PROMIS®) applied cognition - general concerns, short forms in ethnically diverse groups. *Psychol Test Assess Model* 2016;58:255–307.
32 Jensen RE, Potosky AL, Reeve BB, *et al*. Validation of the PROMIS physical function measures in a diverse us population-based cohort of cancer patients. *Qual Life Res* 2015;24:2333–44.
33 Hanmer J, Jensen RE, Rothrock N. A reporting checklist for HealthMeasures' patient-reported outcomes: ASCQ-Me, Neuro-QoL, NIH Toolbox, and PROMIS. *J Patient Rep Outcomes* 2020;4.
34 Gruber-Baldini AL, Velozo C, Romero S, *et al*. Validation of the PROMIS® measures of self-efficacy for managing chronic conditions. *Qual Life Res* 2017;26:1915–24.
35 Elixhauser A, Steiner C, Harris DR, *et al*. Comorbidity measures for use with administrative data. *Med Care* 1998;36:8–27.
36 Haley SM, McHorney CA, Ware JE. Evaluation of the MOS SF-36 physical functioning scale (PF-10): I. Unidimensionality and reproducibility of the Rasch item scale. *J Clin Epidemiol* 1994;47:671–84.
37 US Centers for Disease Control, National Center for Health Statistics. National health interview survey, 2016. Available: https://www.cdc.gov/nchs/nhis/index.htm
38 Howell EA, Egorova N, Hayes MP, *et al*. Racial disparities in the treatment of advanced epithelial ovarian cancer. *Obstet Gynecol* 2013;122:1025–32.
39 Sakhuja S, Yun H, Pisu M, *et al*. Availability of healthcare resources and epithelial ovarian cancer stage of diagnosis and mortality among blacks and whites. *J Ovarian Res* 2017;10:57.
40 Saurman E. Improving access: modifying Penchansky and Thomas's theory of access. *J Health Serv Res Policy* 2016;21:36–9.
41 Terplan M, Temkin S, Tergas A, *et al*. Does equal treatment yield equal outcomes? the impact of race on survival in epithelial ovarian cancer. *Gynecol Oncol* 2008;111:173–8.
42 Farley JH, Tian C, Rose GS, *et al*. Race does not impact outcome for advanced ovarian cancer patients treated with cisplatin/paclitaxel: an analysis of gynecologic Oncology Group trials. *Cancer* 2009;115:4210–7.