American Journal of Epidemiology, 2024, 00, 1–11
https://doi.org/10.1093/aje/kwae033
Advance access publication date April 3, 2024
Study Design

Avanzando Caminos (Leading Pathways): design and procedures of the Hispanic/Latino Cancer Survivorship Study

Frank J. Penedo [1,2,3,*], Patricia I. Moreno [3,4], Magela Pons [4], Paulo S. Pinheiro [4], Michael H. Antoni [1,3], Gilberto Lopes [2], Carmen Calfa [2], Patricia Chalela [5,6], Luz Garcini [7], Chen-Pin Wang [6], Yidong Chen [6], Adolfo Diaz [8], Steve Cole [9,10], Amelie G. Ramirez [5,6]

[1]Department of Psychology, College of Arts and Sciences, University of Miami, Coral Gables, FL 33124, United States
[2]Department of Medicine, Miller School of Medicine, University of Miami, Miami, FL 33136, United States
[3]Cancer Control Research Program, Sylvester Comprehensive Cancer Center, University of Miami, Miami, FL 33136, United States
[4]Department of Public Health Sciences, Miller School of Medicine, University of Miami, Miami, FL 33136, United States
[5]Institute for Health Promotion Research, UT Health San Antonio, San Antonio, TX 78229, United States
[6]Department of Population Health Sciences, Long School of Medicine, UT Health San Antonio, San Antonio, TX 78229, United States
[7]Department of Psychological Sciences, School of Social Sciences, Rice University, Houston, TX 77005, United States
[8]Department of Medicine, Long School of Medicine, UT Health San Antonio, San Antonio, TX 78229, United States
[9]Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, United States
[10]Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, United States
*Corresponding author: Frank J. Penedo, Cancer Control Research Program, Sylvester Comprehensive Cancer Center, University of Miami, 1120 NW 14th Street, Miami, FL 33136 (fpenedo@miami.edu)

Abstract

Avanzando Caminos (Leading Pathways): The Hispanic/Latino Cancer Survivorship Cohort Study aims to examine the influence of sociocultural, medical, stress-related, psychosocial, lifestyle, behavioral, and biological factors on symptom burden, health-related quality of life, and clinical outcomes among Hispanics/Latinos who have been previously treated for cancer. Avanzando Caminos is a prospective, cohort-based study of 3000 Hispanics/Latinos who completed primary cancer treatment within the past 5 years that is representative of the general Hispanic/Latino population in the United States. Participants will complete self-report measures at baseline (time [T] 1), 6 months (T2), 1 year (T3), 2 years (T4), 3 years (T5), 4 years (T6), and 5 years (T7). Blood samples drawn for assessment of leukocyte gene expression, cardiometabolic markers, and genetic admixture will be collected at baseline (T1), 1 year (T3), 3 years (T5), and 5 years (T7). Medical and cancer characteristics and clinical outcomes will be extracted from the electronic medical record and/or state cancer registry at each time point. Data analysis will include general latent variable modeling and latent growth modeling. Avanzando Caminos will fill critical gaps in knowledge in order to guide future secondary and tertiary prevention efforts to mitigate cancer disparities and optimize health-related quality of life among Hispanic/Latino cancer survivors.

Key words: Hispanics; Latinos; cancer; survivorship; posttreatment completion; quality of life.

Introduction

Cancer is the leading cause of death among Hispanics/Latinos (H/Ls) in the United States, accounting for 1 in 5 deaths.[1] At 62.1 million, H/Ls comprise the largest ethnic minority group in the United States (18.5% of the general population), and their numbers are expected to double over the next 4 decades.[2] H/Ls represent individuals from over 15 distinct national origins, and they demonstrate vast variability with respect to socioeconomic status (SES), geographic distribution, genetic admixture, and psychosocial and behavioral determinants of health. Relative to non-H/L Whites, H/Ls have a lower incidence of certain cancers such as lung, breast, and prostate cancer and a higher incidence of infection-related cancers such as stomach, liver, and cervical cancer.[1] Furthermore, existing research examining cancer survivorship among H/Ls suggests that, relative to non-H/L Whites, H/L survivors are more likely to present with advanced disease, to have comorbid conditions, and to experience worse symptom burden and health-related quality of life (HRQoL), and are less likely to return to levels of physical, emotional, and social function that they had prior to cancer treatment.[1,3-7] H/L survivorship may also be further compromised as a result of structural challenges, such as lower SES, lack of health insurance or being underinsured, limited English language proficiency, and barriers to health-care access.[8]

Despite ongoing research, in the absence of a population-based cohort study, the existing evidence base documenting cancer survivorship experience among H/Ls has been limited by multiple challenges, including (1) relatively small samples with short follow-up periods; (2) a narrow focus on common cancers (eg, breast or prostate cancer), with limited attention to cancer

Received: December 20, 2022. **Accepted:** March 29, 2024

sites with high prevalence and mortality among H/Ls; (3) samples restricted to 1 geographic location that lack diversity with respect to H/L country or region of origin and SES; and (4) lack of comprehensive assessments that consider determinants of cancer survivorship outcomes across multiple domains (eg, sociocultural, medical, stress-related, psychosocial, lifestyle, and biological factors).[7,8] The MyHealth Study was a population-based study that aimed to examine differences in HRQoL across racial and ethnic groups and by country of origin with H/L representation (21%).[9] However, that study did not include longitudinal assessments and was limited in its ability to examine within-group differences among H/Ls, particularly among Central and South Americans. The Hispanic Community Health Study/Study of Latinos is a population-based cohort study of US H/Ls that aims to characterize overall health and disease burden, but it has limited cancer-related data.[10,11]

The current study, Avanzando Caminos (Leading Pathways): The Hispanic/Latino Cancer Survivorship Cohort Study, was developed as the first comprehensive cohort study of H/L cancer survivors in the United States. This study is a prospective study of 3000 H/Ls who completed primary cancer treatment within the past 5 years that is representative of the general H/L population in the United States. The overarching goal is to examine the influence of sociocultural, medical, stress-related, psychosocial, lifestyle, behavioral, and biological factors on study outcomes: symptom burden (eg, pain, fatigue, depression, sleep, cognition, toxicities), HRQoL (ie, general and cancer-specific), and disease activity (eg, cancer remission or recurrence, incident secondary cancers, cancer mortality, and all-cause mortality). The aims of Avanzando Caminos are to:

1. Establish an H/L cancer survivorship cohort by recruiting 3000 H/L survivors and conducting comprehensive assessments of sociocultural, medical, stress-related, psychosocial, lifestyle, and biological factors, symptom burden, HRQoL, and disease activity.
2. Examine multiple determinants of symptom burden, HRQoL, and disease activity by determining the extent to which sociocultural, medical, stress-related, psychosocial, lifestyle, behavioral, and biological determinants are associated with symptom burden, HRQoL, and disease status.
3. Evaluate psychosocial mediators by testing whether psychosocial, lifestyle, and behavioral factors mediate the associations of sociocultural and stress-related factors with symptom burden, HRQoL, and disease activity.
4. Evaluate biological mechanisms by testing whether biological factors mediate and moderate associations among sociocultural, stress-related, psychosocial, lifestyle, and behavioral factors and symptom burden, HRQoL, and disease activity.
5. Explore potential moderators by determining whether any of a putative set of moderators (eg, H/L origin, nativity, sex, comorbid conditions, cancer type) modify associations between our determinants and study outcomes.

Our research team has extensive expertise in cancer survivorship among H/Ls, assessment of patient-reported outcomes, recruitment of diverse individuals with cancer, longitudinal cohort study design, cancer registry–based recruitment, community-based participatory research, leukocyte gene expression, medical oncology, comprehensive survivorship care, and advanced statistical analysis.

Avanzando Caminos will test determinants of symptom burden, HRQoL, and disease activity using our integrated model of biopsychosocial determinants of health in H/L cancer survivors (Figure 1). This model demonstrates the conceptual organization of our proposed aims and constructs of interest. Findings from the current study will provide novel, critically needed information to inform and guide future secondary and tertiary prevention among H/L cancer survivors by (1) characterizing diverse factors that influence trajectories and outcomes; (2) documenting differential impacts of our determinants across H/Ls with significant diversity with respect to country/region of origin, SES, and urbanicity (urban vs rural residence); and (3) identifying modifiable psychosocial, behavioral, and biological mechanisms that can be targeted via psychosocial interventions or clinical management to improve survivorship outcomes.

Methods

Cohort

A total of 3000 H/Ls (1500 men and 1500 women) of diverse H/L background and cancer type will be recruited from 2 major US geographic regions (1500 from South Florida and 1500 from South Texas). Participants will have completed primary treatment for breast, colorectal, kidney, liver, lung, prostate, stomach, or cervical cancer within the past 5 years. The sample will include individuals of Mexican, Central American, South American, Puerto Rican, Dominican, and Cuban descent in a distribution that approximates the general US H/L population (Table 1).

This study received institutional review board approval on December 24, 2020. The first participant enrollment occurred on August 19, 2021, and it is anticipated that the study will conclude by February 28, 2027.

Inclusion and exclusion criteria

Inclusion criteria are: (1) age 18 years or older; (2) completion of primary cancer treatment for a primary tumor with or without neoadjuvant or current adjuvant treatment within the past 5 years; (3) a primary stage I-III diagnosis of cancer of the breast (female), colon/rectum, kidney or renal pelvis, liver and intrahepatic bile duct, lung and bronchus, prostate (male), stomach, or uterine cervix (female) as confirmed by electronic medical record (EMR) and/or cancer registry; and (4) willingness to attend study visits. Exclusion criteria are: (1) severe cognitive impairment; (2) inpatient psychiatric treatment in the past 6 months; and (3) stage IV cancer diagnosis.

Recruitment

Potential participants will primarily be recruited from the University of Miami Sylvester Comprehensive Cancer Center and the UT Health San Antonio Mays Cancer Center and its affiliated institutions (70%), with additional recruitment through the Florida Cancer Data System (FCDS) and the Texas Cancer Registry TCR (30%). Recruitment will be stratified by geographic region, cancer type, and sex assigned at birth for comparability with the registry data, in order to ensure (1) an equal sex distribution across geographic regions, (2) inclusion of less studied cancers (eg, stomach, liver), and (3) proportions of cancers reflecting the registry H/L rates of the FCDS and TCR. Table 2 shows the distribution of the sample by cancer type and sex. Cancer types were selected on the basis of the Sylvester and Mays Cancer Center catchment area data.

At the Sylvester Comprehensive Cancer Center, both the tumor registry and the Consent to Contact registry (a repository of patients who have agreed to be contacted for research purposes) will be used to generate lists of potentially eligible patients' names in weekly requested authorized reports based on H/L ethnicity in

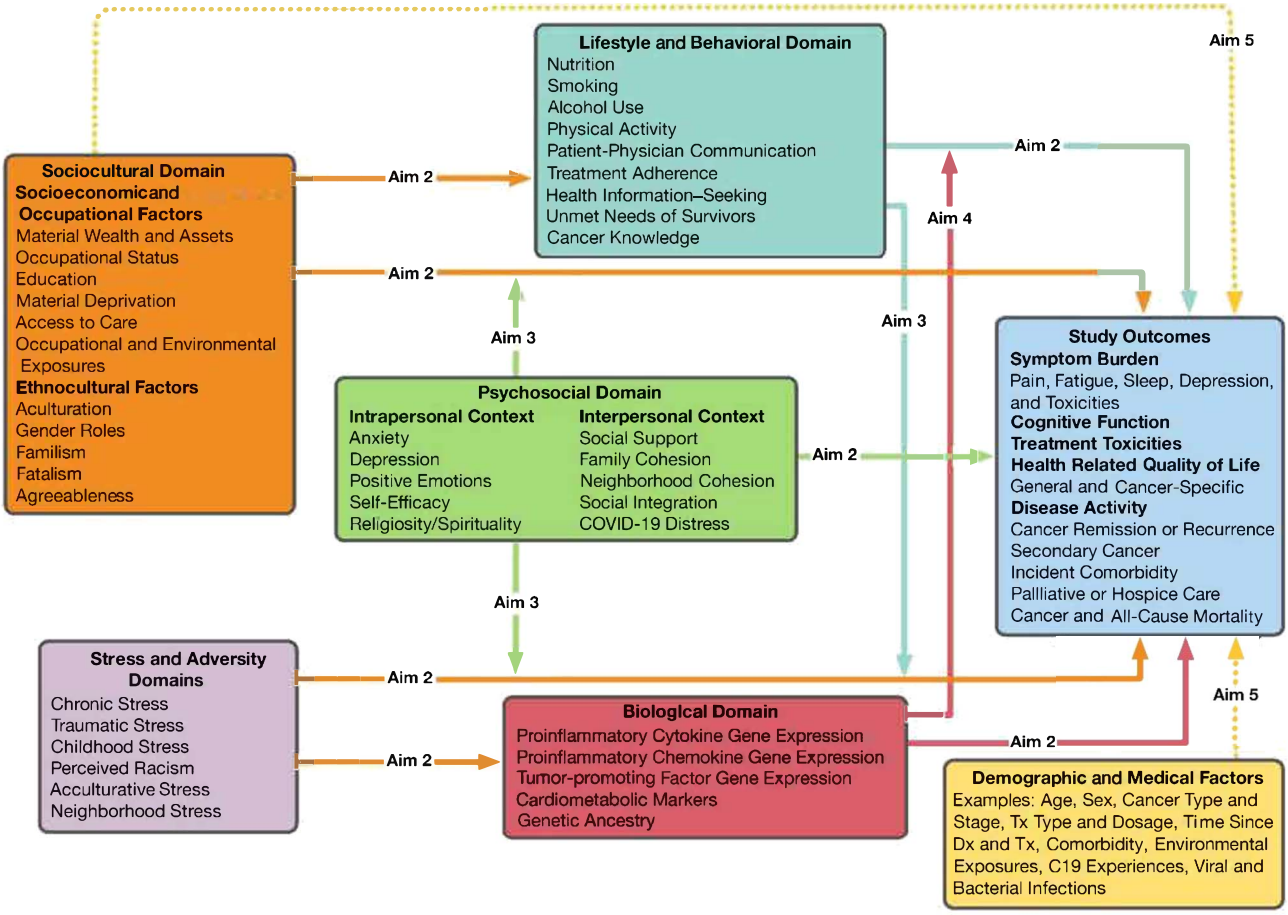

Figure 1. The authors' integrated model of biopsychosocial determinants of health outcomes among Hispanic/Latino cancer survivors in 2 US locations (South Florida and South Texas). The model forms the basis of the conceptual organization of the cohort measures, proposed aims, and hypothesized relationships in Avanzando Caminos (Leading Pathways): The Hispanic/Latino Cancer Survivorship Cohort Study. Pathways leading to boxed domains represent proposed direct associations, while pathways leading to lines represent mediation or moderation pathways. Dotted paths represent select tests of moderation. C19, COVID-19; Dx, diagnosis; Tx, treatment.

the EMR and clustered by cancer type (breast, colorectal, lung, liver, prostate, melanoma, etc) based on *International Classification of Diseases, Tenth Revision*, codes. Staff will review lists to identify participants to be recruited. Furthermore, staff will access electronic data warehouse systems to prescreen potential participants (eg, *International Classification of Diseases, Tenth Revision*, diagnosis, completion of primary treatment, H/L ethnicity) when a physician or clinic staff refers a patient. Once a patient is identified as eligible, staff will contact the physician and medical team of record and obtain approval to make contact. These approaches include contact preferences to guide communication method (eg, letter, call, email, text). We will also use the

Table 1. Distribution of the Avanzando Caminos (Leading Pathways) cohort of cancer survivors by Hispanic/Latino origin relative to the overall US Hispanic/Latino population and the catchment area for each participating cancer center (UT Health San Antonio Mays Cancer Center and University of Miami Sylvester Comprehensive Cancer Center).

H/L origin	Cohort sample		Approximate proportion of US H/Ls, %[a]	Proportion of participants by cancer center, %	
	No.	%		UT Health San Antonio Mays Cancer Center	University of Miami Sylvester Comprehensive Cancer Center
Mexican	1500	50	60	90	6
Central/South American	675	23	17	<5	35
Cuban	525	18	4	<5	44
Puerto Rican or Dominican	300	15	13	<5	15

Abbreviation: H/L, Hispanic/Latino.
[a]Pew Research Center calculations based on the 2010 and 2021 American Community Surveys (US Census Bureau).[12]

Table 2. Distribution of the Avanzando Caminos (Leading Pathways) cohort of Hispanic/Latino cancer survivors by cancer type and sex.

Cancer type	No. of participants		
	Male	Female	Total
Stomach	100	90	190
Colon/rectal	285	270	555
Liver	130	90	220
Lung	260	230	490
Breast	N/A	525	525
Cervical	N/A	170	170
Kidney	155	125	280
Prostate	570	N/A	570
Total	1500	1500	3000

Abbreviation: N/A, not applicable.

EMR-integrated patient portal to communicate with eligible participants who consent to be contacted. At the Mays Cancer Center and its affiliate site, the University of Texas Education and Research Center at Laredo, medical records and cancer registry data sets will be queried weekly to identify potentially eligible participants based on H/L ethnicity in the EMR without direct provider referral, per institutional policy. At the Mays Cancer Center affiliate university health system, potentially eligible participants will be identified and referred directly by the physician and medical team who will guide the communication method.

Potential participants identified from the FCDS and TCR will be mailed a packet including a patient contact letter and a patient response form and a letter from the respective department of health. If, after 3 weeks from sending the initial packet, there is no response from a potential participant, a second mailing will be sent, with the addition of a telephone opt-out card. The telephone opt-out card explains to the patient that if no response is received, the study investigator and/or a member of the study team will attempt to contact them via a telephone call, text, or email to introduce the study. If there is no response to the second mailing and the telephone opt-out card after 3 weeks, a telephone call will be attempted by the study staff.

Community engagement and cultural sensitivity

Additional community-based recruitment will include advertisements in predominately H/L news, radio, and television media. Both the Sylvester and Mays cancer centers have longstanding partnerships with several community organizations and will establish a community advisory board of stakeholders at each recruitment site that meets twice a year to provide feedback on study procedures and materials and to disseminate information about the study to community physicians and health service agencies. Study flyers and brochures are designed to represent each cancer center's catchment area H/L population (the Mays sample is comprised primarily of individuals of Mexican and Central American heritage, and the Sylvester sample is comprised primarily of individuals of Caribbean and South American heritage). Both sites have fully bilingual staff members that undergo culturally specific trainings that emphasize common H/L cultural values, such as alocentrismo, familismo, respeto, machismo/marianismo, and simpatía.

Screening

Once a potential participant expresses interest in participation, study staff will verify self-identification as H/L and fluency in spoken English or Spanish via a telephone screen. Patients who confirm H/L identity and demonstrate fluency in English or Spanish are scheduled for an in-person assessment, where cognitive impairment will be assessed prior to obtaining informed consent using a score of 7 or below on the Short Portable Mental Status Questionnaire.[13]

Assessments

Participants will complete self-report measures at each time point: baseline (time [T] 1), 6 months (T2), 1 year (T3), 2 years (T4), 3 years (T5), 4 years (T6), and 5 years (T7). Certified phlebotomists will conduct peripheral blood draws to assess biomarkers (ie, leukocyte gene expression, cardiometabolic markers, and genetic admixture) at every other time point: baseline (T1), 1 year (T3), 3 years (T5), and 5 years (T7). Genetic admixture is only assessed once at baseline, while leukocyte gene expression and cardiometabolic markers are assessed repeatedly. Highly trained staff take clinical measurements of blood pressure (diastolic and systolic) and waist circumference at baseline (T1), 1 year (T3), 3 years (T5), and 5 years (T7). Medical and cancer characteristics and clinical outcomes will be extracted from the EMR and/or state cancer registry at each time point. Gender identity, such as cis-woman/man, trans-woman/man, or nonbinary, will be assessed via self-report. Participants who are not affiliated with the Sylvester and Mays cancer centers will be asked to complete a third-party health-records release form to access their community medical records in order to capture any missing clinical or medical data (eg, comorbidity). Rurality will be defined as residence in a nonmetropolitan county with fewer than 50 000 people, per US Census Bureau and Office of Management and Budget standards. Variables on which data were collected and their corresponding measures and data sources are listed in Table 3.

All assessments that include peripheral blood draws (ie, baseline [T1], 1 year [T3], 3 years [T5], and 5 years [T7]) will be conducted in person in the laboratory. In order to accommodate a diversity of educational attainment and literacy in this population, bilingual staff will verbally administer self-report measures to participants and directly enter participant responses into REDCap (a web-based data management system). Assessments that do not include peripheral blood draws (ie, 6 months [T2], 2 years [T4], and 4 years [T6]) will be conducted either in person as described above or via teleconference or phone. Figure 2 illustrates the participant and data collection workflow. Assessments will last 2-3 hours with breaks, and participants will be compensated $50 per visit ($350 total). Parking and transportation reimbursements are available to participants for all in-person visits.

Retention strategies

Strategies for optimizing engagement and retention include reminder calls, text messages, and emails prior to upcoming assessments (tailored based on participants' preferences) and provision of participants with study contact phone numbers and email addresses to facilitate communication and accessibility. Participants are offered flexible scheduling for assessments to accommodate their availability, including evening hours and weekends. Birthday and holiday cards are sent to participants throughout the duration of their participation. Participants who cannot be reached via phone or email are mailed a follow-up letter in an attempt to reengage them. Participants who express a desire to withdraw from the study are contacted by a study coordinator to understand this desire and discuss possible accommodations that may enable continued participation.

Table 3. Study variables and measures to be included in Avanzando Caminos (Leading Pathways): The Hispanic/Latino Cancer Survivorship Cohort Study.

Determinant	Variable	Measure	Source(s) of data
Socioeconomic and occupational factors	Material wealth and assets	Home ownership, automobile ownership	Self-report
	Occupational status	Type of job	Self-report
	Education	Total number of years of education, years of US education, degrees	Self-report
	Material deprivation	Difficulty with food purchases, housing expenses	Self-report
	Occupational and environmental exposures	Exposures to metals, dust, fibers, chemicals, fumes, radiation, pesticides, and herbicides; length of exposure[14]	Self-report
	Access to care	NCI HINTS[15]	Self-report
Ethnocultural factors	Acculturation	Short Acculturation Scale for Hispanics[16]	Self-report
	Gender roles	Marianismo Scale[17,18] and Masculinity Scale[19]	Self-report
	Familism	Familism Scale[20]	Self-report
	Fatalism	Fatalism Scale[21]	Self-report
	Agreeableness	Simpatía Scale[22,23]	Self-report
Stress and adversity determinants	Chronic stress	Chronic Burden Scale[24]	Self-report
	Traumatic stress	Traumatic Stress Schedule[25]	Self-report
	Childhood stress	Adverse Childhood Experiences Scale[26]	Self-report
	Perceived racism	Brief Perceived Ethnic Discrimination Questionnaire-Community Scale[27]	Self-report
	Acculturative stress	Abbreviated Hispanic Stress Inventory-Immigrant Scale[28]	Self-report
	Neighborhood stress	Neighborhood Problems Scale[29]	Self-report
Psychosocial determinants	Anxiety	PROMIS-Ca Anxiety Short Form[30]	Self-report
	Depression	PROMIS-Ca Depression Short Form[30]	Self-report
	Positive emotions	PROMIS Positive Affect Short Form[31]	Self-report
	Optimism	Life Orientation Test-Revised[32]	Self-report
	Self-efficacy	PROMIS Self-Efficacy Short Form[33]	Self-report
	Spirituality	Functional Assessment in Chronic Illness Therapy—Spiritual Well-Being Scale (FACIT-Sp)[34]	Self-report
	Religiosity	Duke University Religion Index[35,36]	Self-report
	Social support	Interpersonal Support Evaluation List[37,38]	Self-report
	Family cohesion	Family Environment Scale, cohesion and conflict subscales[39]	Self-report
	Neighborhood cohesion	Neighborhood Social Cohesion Scale[40]	Self-report
	Social integration	Social Network Index[41]	Self-report
	COVID-19 distress	COVID-19 Practical and Psychosocial Experiences Questionnaire[42]	Self-report
Lifestyle and behavioral determinants	Nutrition	Food Propensity Questionnaire[43]	Self-report
	Smoking	Smoking Tobacco Use Questionnaire[44]	Self-report
	Alcohol use	Alcohol use disorder risk (NIAAA)[45]	Self-report
	Physical activity	Global Physical Activity Questionnaire[46]	Self-report
	Patient-physician communication	Perceived Efficacy in Patient-Physician Interactions (PEPPI) Scale[47]	Self-report
	Treatment adherence	Skipped medications, appointments, and treatments	Self-report
	Health information–seeking	NCI HINTS[15]	Self-report
	Unmet needs of survivors	Supportive Care Needs Survey[48]	Self-report
	Cancer knowledge	NCI HINTS[15]	Self-report
Biological determinants	Proinflammatory cytokine gene expression	*IL1A*, *IL1B*, *IL6*, *IL8/CXCL8*, *TNF*, and *PTGS2/COX-2* genes	Extracted RNA
	Proinflammatory chemokine gene expression	*CCL2*, *CCL3*, *CCL7*, *CCL20*, *CCL3L1*, *CCL4L2*, and *CXCR7* genes	Extracted RNA
	Tumor-promoting factor gene expression	*MMP9* and *LMNA* genes in circulating peripheral blood mononuclear cells[49,50] Protocol for monocyte attainment, handling, transportation, profiling and genetic analysis[51-57]	Extracted RNA
	Cardiometabolic markers	HDL and LDL cholesterol, triglycerides, fasting glucose, hemoglobin A1c, and metabolic syndrome components using Adult Treatment Panel III[58] Blood pressure (diastolic, systolic) Waist circumference	Peripheral blood clinical measurement
	Genetic ancestry	Genetic markers (ancestry-informative markers) and sequencing protocol used to assess Hispanic/Latino ancestry admixture at baseline only[59-65]	Extracted DNA

(Table continues)

Table 3. Continued.

Determinant	Variable	Measure	Source(s) of data
Study outcomes	Symptom burden	PROMIS-29 Profile (version 2.1)[66]	Self-report
	Cognitive function	PROMIS Cognitive Function Short Form[67]	Self-report
	Treatment toxicities	NCI PRO-CTCAE[68]	Self-report
	General HRQoL	Functional Assessment of Cancer Therapy-General (FACT-G)[69]	Self-report
	Cancer site–specific HRQoL	Cancer site-specific FACT symptom burden subscales[70]	Self-report
	Disease activity	Cancer remission, cancer recurrence, cancer progression, secondary cancer/s, incident comorbidity, palliative care, hospice care, cancer mortality, and all-cause mortality	Electronic medical records; Florida Cancer Data System; Texas Cancer Registry; health-care providers

Abbreviations: Ca, cancer; FACT, Functional Assessment of Cancer Therapy; HDL, high-density lipoprotein; HINTS, Health Information National Trends Survey; HRQoL, health-related quality of life; LDL, low-density lipoprotein; NIAAA, National Institute on Alcohol Abuse and Alcoholism; NCI, National Cancer Institute; PRO-CTCAE, Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events; PROMIS, Patient-Reported Outcomes Measurement Information System.

Data analyses

Data will be inspected and subjected to quality control procedures. Non–normally distributed variables will be transformed by Box-Cox power transformation prior to parametric statistical analyses as appropriate, or analyzed with nonparametric statistics. To identify underlying constructs in multivariate measures of the 5 determinants (sociocultural, medical, stress-related, psychosocial, lifestyle, behavioral, and biological) of the 3 study outcomes, we will use general latent variable modeling (GLVM)[71] with maximum likelihood robust variance estimation. GLVM is an advanced generalized structural equation modeling technique, and it is well suited for analysis of these study data because it accommodates diverse distributions of multivariate measures of each determinant as well as the multilevel structure (repeated measures nested within patients) in the longidtudinal study design. In addition, GLVM is appealing for its ability to characterize heterogeneity in the determinants using both discrete and continuous latent variables. In the primary analyses, we will identify the construct(s) of each determinant using a separate GLVM. The (distinct) trajectory profiles (classes) associated with each determinant derived from GLVM will emerge as the construct(s) associated with each determinant; and the patient-specific propensity scores of trajectory profiles will be used to characterize each patient's sociocultural, medical, stress-related, psychosocial, lifestyle, behavioral, and biological determinants. For a consistent interpretation of the construct(s) underlying each determinant over time, we will also derive constructs/scores of each determinant based on GLVM using baseline measurements. Scores for other time points will then be derived using empirical Bayes estimation conditioned on GLVM derived from baseline and observed data at each follow-up. Multiple group comparisons within GLVM will test measurement and structural invariance across groups (eg, H/L groups).

To test direct, indirect/mediation, and interaction/moderation effects on temporal changes of symptom burden and HRQoL outcomes, separate latent growth models (LGMs) are proposed to model the trajectory of each outcome, and the effects of underlying constructs (derived from GLVM) on outcome trajectory parameters (eg, slope). The LGM was chosen over the mixed-effects model because it is more general, especially when normal transformation presents challenges. Model fit and refinement of GLVM and LGM is guided by bayesian information criterion and

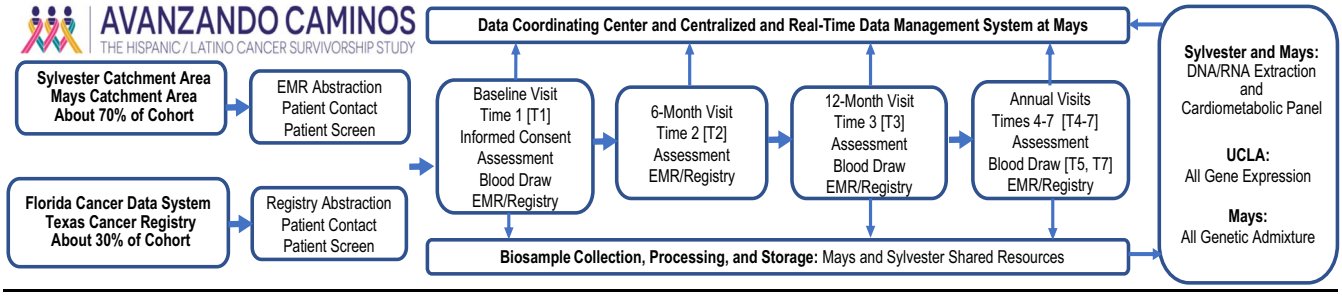

Figure 2. Participant and data collection workflow in Avanzando Caminos (Leading Pathways): The Hispanic/Latino Cancer Survivorship Cohort Study. Recruitment will take place via 3 sources: the University of Miami Sylvester Comprehensive Cancer Center (ie, Sylvester), the UT Health San Antonio Mays Cancer Center (ie, Mays), and state cancer registries (the Florida Cancer Data System and the Texas Cancer Registry). The assessment schedule reflects 7 data collection points that include baseline, 6-month follow-up, 12-month follow-up, and annual follow-up thereafter. Biosample collection occurs at baseline and at the 12-, 36-, and 60-month follow-ups. All data are captured in real time and centralized in a web-based data management system. All biospecimen data are collected, processed, and stored at the Sylvester and Mays cancer centers' shared resources. The initial 100 samples collected for leukocyte gene expression and genetic admixture are analyzed in year 2 of the UH3 phase. The remaining samples are batched and assayed during the UG3 phase. All cardiometabolic panels and RNA and DNA extraction procedures are conducted at Sylvester and Mays upon collection. Batched samples for leukocyte gene expression and genetic admixture are sent to the University of California, Los Angeles (UCLA) and Mays, respectively. EMR, electronic medical record; UT, University of Texas.

LGM diagnostics.[72] Domain experts will evaluate the plausibility of model interpretations. Because our models will capitalize on multiple mediation pathways in our outcomes, joint modeling of multiple mediators as prescribed in VanderWeele and Vansteelandt[73] will be used to formally assess mediation effects associated with the sociocultural, medical, stress-related, psychosocial, lifestyle, behavioral, and biological factors that are derived from GLVM. Due to the time-to-event nature of disease activity outcomes (eg, recurrence, mortality), the overall effects of each determinant and the associated mediation or moderation effects will be assessed on the log hazard ratio scale via Cox regression models. The Hosmer-Lemeshow test and the C statistic will be used to evaluate model fit. To control the influence of covariates in relation to all analyses, we will include demographic variables (eg, age) as predictors in all analyses. For clinical covariates such as cancer type, stage, etc, we will conduct variable selection using theory-guided hypotheses and statistical methods such as the likelihood ratio test or the elastic net regularization method.[74]

Missing data

Observed data likelihood will be used for deriving primary inference. We will also derive estimates based on multiple imputations. To gain further insights, we will examine missing-data patterns using latent class analyses, where missing data due to death (for nonmortality outcomes), nonresponse, and loss to follow-up are differentiated by separate indicators. Predictors associated with missing-data patterns are incorporated into GLVM and LGM analyses as auxiliary variables. All approaches above rely on the "missing at random" assumption to ensure consistent estimates. Sensitivity analyses will be conducted to assess the impacts of "missing at random" violation.

Application to analysis of determinants of primary outcomes (aim 2)

To address aim 2, we will conduct LGM analyses to evaluate the effectiveness of determinant scores derived from GLVM in association with the trajectory of each repeatedly measured outcome, and Cox regression analyses for the time-to-event outcome. Using the effect of lifestyle on repeated measurements of HRQoL as an example, we will first conduct GLVM to identify distinct lifestyle profiles (based on repeated measures of multiple lifestyle variables) and derive patient-level lifestyle scores using the empirical Bayes method (eg, the posterior probability of being in the poorer lifestyle profile or the indicator of the most likely lifestyle profile conditioned on observed individual data and GLVM estimates). This will be followed by LGM analyses of repeated measures of HRQoL over time, where the overall temporal pattern will be characterized by suitable trajectory parameters (eg, intercept, linear, quadratic, etc), and each trajectory parameter is further modeled in terms of patient-level lifestyle determinant scores, covariates (eg, sex), and potentially random effects accounting for unobserved variation between H/Ls. This will be followed by testing to determine whether there exists at least 1 underlying construct/score associated with a determinant that is a significant predictor for trajectory parameters of symptom burden or HRQoL. For disease activity, Cox regression analyses will be conducted to examine the associations between incidence of each disease activity and the underlying constructs of each determinant. Time-varying covariates will be adjusted for, and the proportional hazards assumption will be verified via Schoenfeld residuals.

Mediation and moderation (aims 3-5)

To test whether psychosocial, lifestyle, and behavioral factors jointly mediate the association between the sociocultural determinant and HRQoL, LGM trajectory parameters of HRQoL will be modeled in terms of sociocultural determinant scores as well as psychosocial and lifestyle (mediators) determinants. The indirect effects of sociocultural scores mediated via psychosocial and lifestyle scores will be calculated on the basis of VanderWeele and Vansteelandt equations.[73] To test whether biological factors jointly mediate the association between stress and disease activity, the log-transformed hazard function of the disease activity outcome will be modeled in terms of inflammatory, prometastatic, and glucocorticoid resistance gene expression profile scores derived from GLVM, and the mediation and moderation effects will also be assessed on the log-hazard scale. The moderation effects of several factors on the associations between a determinant and a study outcome will be tested by LGMs, with outcome trajectory parameters being modeled in terms of moderators, constructs of determinants, and interactions of moderators with determinants. The moderation effects of biological factors on the associations between determinants and a disease activity outcome will be assessed by Cox regression model, with the log-hazard function being modeled in terms of moderators, constructs of determinants, and interactions between moderators with determinants.

Power estimate

Based on recruitment and a 10% attrition rate per follow-up, 2446, 1596, 902, and 350 participants are expected to have a minimum of 2, 3, 4, and 5 years of follow-up data, respectively. We are assuming that (1) at least 1 construct of each determinant will be associated with each outcome with an effect size greater than 0.20; (2) for each pair of determinants on a mediation pathway (eg, interpersonal resources mediate the effect of stress), the effect size of this association will be greater than 0.21; (3) the temporal correlation of each outcome will be greater than 0.4; and (4) the variance of inflation due to missing data and collinearity will be less than 1.5. Based on simulations emulating the design matrix consistent with the recruitment plan and follow-up, the estimated levels of statistical power needed to verify aims 2, 3, and 4 will be 0.841, 0.822, and 0.816, respectively, conditioned on an overall type 1 error adjusting for multiple comparisons. The effect size in our calculation corresponds to the difference in slopes for each outcome between high and low risk/score determinant strata with diverse distributions (0.1-0.9), where the variance of slopes is calculated on the basis of the design matrix (7 repeated measures), assuming varied autocorrelation.

Discussion

The goal of Avanzando Caminos (Leading Pathways): The Hispanic/Latino Cancer Survivorship Cohort Study is to determine the influence of sociocultural, medical, stress-related, psychosocial, lifestyle, behavioral, and biological factors on symptom burden, HRQoL, and disease activity among H/Ls who completed primary cancer treatment. We hypothesize that less favorable sociocultural factors and greater stress and adversity will be significantly associated with poorer lifestyle and behavioral factors, greater proinflammatory/prometastatic gene expression, and more cardiometabolic comorbidity and with greater symptom burden, worse HRQoL, and greater disease activity among H/L cancer survivors. Alternatively, we hypothesize that greater psychosocial

resources, more favorable lifestyle and behavioral factors, and better gene expression and cardiometabolic profiles will be significantly associated with lower symptom burden, better HRQoL, and less disease activity among H/L cancer survivors. We also hypothesize that (1) associations of sociocultural and stress-related factors with symptom burden, HRQoL, and disease activity will be mediated by psychosocial factors and that (2) associations of sociocultural, stress-related, and psychosocial factors with symptom burden, HRQoL, and disease activity will be mediated by lifestyle and behavioral factors (Figure 1). Furthermore, we hypothesize that associations of sociocultural, stress-related, psychosocial, lifestyle, and behavioral factors with symptom burden, HRQoL, and disease activity will be mediated by gene expression profiles and cardiometabolic profiles. Lastly, we expect that the relationships described above may be modified by factors such H/L country/region of origin, nativity, genetic admixture, sex, comorbidity, and cancer type.

Strengths of this study include (1) the development and testing of our integrated model of biopsychosocial determinants of health in H/L cancer survivors; (2) a well-characterized cohort of H/L survivors who are diverse with respect to country/region of origin, SES, and urbanicity (urban vs rural residence), allowing us to examine how survivorship experiences and outcomes vary across these domains; and (3) comprehensive assessment of novel sociocultural and psychosocial factors, leukocyte gene expression, and cardiometabolic markers in the context of H/L survivorship for identification of potentially modifiable factors that may influence outcomes—resulting in a rich repository of data and biospecimens. Limitations include enrollment in 2 distinct geographic regions of the United States (South Texas and South Florida) and primary reliance on clinic- and registry-based recruitment, which limit generalizability.

Lessons Learned

- Assessments are needed across multiple levels of influence, including sociocultural, medical, stress, psychosocial, lifestyle, behavioral, and biological factors, to effectively characterize determinants of quality of life, symptom burden, disease activity, and clinical outcomes among diverse Hispanic/Latino cancer survivors.
- Culturally informed and targeted strategies are critical to optimizing participant recruitment, engagement, and retention, particularly among diverse, underrepresented populations.
- Robust statistical analyses are ideally suited for handling complex data, including data with a multilevel structure of repeated measures nested within patients, as well as both discrete and continuous latent predictor variables. General latent variable models and latent growth models will be particularly useful in our research on diverse Hispanic/Latino cancer survivors.
- Results will help guide targeted patient-, community-, and system-level interventions to reduce the disparate burden of cancer observed in the Hispanic/Latino community.

Conclusions

To our knowledge, the proposed study is the first comprehensive cohort study of H/L cancer survivors in the United States that will address a crucial gap in the literature and contribute to our understanding of how multiple determinants contribute to outcomes (including symptoms, HRQoL, and clinical outcomes) in diverse H/L cancer survivors.

Funding

This work was supported by a National Cancer Institute grant (UG3CA260317) awarded to F.J.P. and A.G.R. In addition, P.I.M. was supported by a National Cancer Institute Career Development Award (K01CA258955).

Conflict of interest

The authors have no conflicts of interest to disclose.

Disclaimer

The content of this article is solely the responsibility of the authors and does not represent the official views of the National Institutes of Health.

Data availability

Data-sharing is not applicable to this article, as no new data were generated or analyzed in this protocol.

References

1. American Cancer Society. *Cancer Facts & Figures for Hispanic/Latino People 2021-2023*. American Cancer Society; 2021.
2. Population Estimates Program, US Census Bureau. Population and housing unit estimates. 2021. Accessed June 1, 2022. https://www.census.gov/programs-surveys/popest.html
3. Bustillo NE, McGinty HL, Dahn JR, et al. Fatalism, medical mistrust, and pretreatment health-related quality of life in ethnically diverse prostate cancer patients. *Psychooncology*. 2017;26(3):323-329. https://doi.org/10.1002/pon.4030
4. Yanez B, Thompson EH, Stanton AL. Quality of life among Latina breast cancer patients: a systematic review of the literature. *J Cancer Surviv*. 2011;5(2):191-207. https://doi.org/10.1007/s11764-011-0171-0
5. Luckett T, Goldstein D, Butow PN, et al. Psychological morbidity and quality of life of ethnic minority patients with cancer: a systematic review and meta-analysis. *Lancet Oncol*. 2011;12(13):1240-1248. https://doi.org/10.1016/S1470-2045(11)70212-1
6. Ayanian JZ, Zaslavsky AM, Guadagnoli E, et al. Patients' perceptions of quality of care for colorectal cancer by race, ethnicity, and language. *J Clin Oncol*. 2005;23(27):6576-6586. https://doi.org/10.1200/JCO.2005.06.102
7. Yanez B, McGinty HL, Buitrago D, et al. Cancer outcomes in Hispanics/Latinos in the United States: an integrative review and conceptual model of determinants of health. *J Lat Psychol*. 2016;4(2):114-129. https://doi.org/10.1037/lat0000055
8. Kronenfeld JP, Graves KD, Penedo FJ, et al. Overcoming disparities in cancer: a need for meaningful reform for Hispanic and Latino cancer survivors. *Oncologist*. 2021;26(6):443-452. https://doi.org/10.1002/onco.13729

9. Reeve BB, Graves KD, Lin L, et al. Health-related quality of life by race, ethnicity, and country of origin among cancer survivors. *JNCI J Natl Cancer Inst*. 2023;115(3):258-267. https://doi.org/10.1093/jnci/djac230
10. Sorlie PD, Avilés-Santa LM, Wassertheil-Smoller S, et al. Design and implementation of the Hispanic Community Health Study/Study of Latinos. *Ann Epidemiol*. 2010;20(8):629-641. https://doi.org/10.1016/j.annepidem.2010.03.015
11. LaVange LM, Kalsbeek WD, Sorlie PD, et al. Sample design and cohort selection in the Hispanic Community Health Study/Study of Latinos. *Ann Epidemiol*. 2010;20(8):642-649. https://doi.org/10.1016/j.annepidem.2010.05.006
12. US Census Bureau. American Community Survey data. Revised August 16, 2023. Accessed December 20, 2022. https://www.census.gov/programs-surveys/acs/data.html
13. Pfeiffer E. A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. *J Am Geriatr Soc*. 1975;23(10):433-441. https://doi.org/10.1111/j.1532-5415.1975.tb00927.x
14. Rapisarda V, Ledda C, Ferrante M, et al. Blood pressure and occupational exposure to noise and lead (Pb): a cross-sectional study. *Toxicol Ind Health*. 2016;32(10):1729-1736. https://doi.org/10.1177/0748233715576616
15. National Cancer Institute. Health Information National Trends Survey. Survey instruments. 2022. Accessed June 1, 2022. https://hints.cancer.gov/data/survey-instruments.aspx
16. Marin G, Sabogal F, Marin BV, et al. Development of a short acculturation scale for Hispanics. *Hisp J Behav Sci*. 1987;9(2):183-205. https://doi.org/10.1177/07399863870092005
17. Castillo LG, Perez FV, Castillo R, et al. Construction and initial validation of the Marianismo Beliefs Scale. *Couns Psychol Q*. 2010;23(2):163-175. https://doi.org/10.1080/09515071003776036
18. Kulis S, Marsiglia FF, Lingard EC, et al. Gender identity and substance use among students in two high schools in Monterrey, Mexico. *Drug Alcohol Depend*. 2008;95(3):258-268. https://doi.org/10.1016/j.drugalcdep.2008.01.019
19. Arciniega GM, Anderson TC, Tovar-Blank ZG, et al. Toward a fuller conception of machismo: development of a traditional machismo and caballerismo scale. *J Couns Psychol*. 2008;55(1):19-33. https://doi.org/10.1037/0022-0167.55.1.19
20. Sabogal F, Marín G, Otero-Sabogal R, et al. Hispanic familism and acculturation: what changes and what doesn't? *Hisp J Behav Sci*. 1987;9(4):397-412. https://doi.org/10.1177/07399863870094003
21. Cuéllar I, Arnold B, González G. Cognitive referents of acculturation: assessment of cultural constructs in Mexican Americans. *J Community Psychol*. 1995;23(4):339-356. https://doi.org/10.1002/1520-6629(199510)23:4<339::AID-JCOP2290230406>3.0.CO;2-7
22. Gamble WC, Modry-Mandell K. Family relations and the adjustment of young children of Mexican descent: do family cultural values moderate these associations? *Soc Dev*. 2008;17(2):358-379. https://doi.org/10.1111/J.1467-9507.2007.00429.X
23. Griffith JD, Joe GW, Chatham LR, et al. The development and validation of a simpatia scale for Hispanics entering drug treatment. *Hisp J Behav Sci*. 1998;20(4):468-482. https://doi.org/10.1177/07399863980204004
24. Bromberger JT, Matthews KA. A longitudinal study of the effects of pessimism, trait anxiety, and life stress on depressive symptoms in middle-aged women. *Psychol Aging*. 1996;11(2):207-213. https://doi.org/10.1037//0882-7974.11.2.207
25. Norris FH. Screening for traumatic stress: a scale for use in the general population. *J Appl Soc Psychol*. 1990;20(20):1704-1715. https://doi.org/10.1111/j.1559-1816.1990.tb01505.x
26. Felitti VJ, Anda RF, Nordenberg D, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The Adverse Childhood Experiences (ACE) Study. *Am J Prev Med*. 1998;14(4):245-258. https://doi.org/10.1016/s0749-3797(98)00017-8
27. Brondolo E, Kelly KP, Coakley V, et al. The Perceived Ethnic Discrimination Questionnaire: development and preliminary validation of a community version. *J Appl Soc Psychol*. 2005;35(2):335-365. https://doi.org/10.1111/j.1559-1816.2005.tb02124.x
28. Cavazos-Rehg PA, Zayas LH, Walker MS, et al. Evaluating an abbreviated version of the Hispanic Stress Inventory for immigrants. *Hisp J Behav Sci*. 2006;28(4):498-515. https://doi.org/10.1177/0739986306291740
29. Mujahid MS, Diez Roux AV, Morenoff JD, et al. Assessing the measurement properties of neighborhood scales: from psychometrics to ecometrics. *Am J Epidemiol*. 2007;165(8):858-867. https://doi.org/10.1093/aje/kwm040
30. Pilkonis PA, Choi SW, Reise SP, et al. Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS®): depression, anxiety, and anger. *Assessment*. 2011;18(3):263-283. https://doi.org/10.1177/1073191111411667
31. Salsman JM, Lai J-S, Hendrie HC, et al. Assessing psychological well-being: self-report instruments for the NIH Toolbox. *Qual Life Res*. 2014;23(1):205-215. https://doi.org/10.1007/s11136-013-0452-3
32. Scheier MF, Carver CS, Bridges MW. Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the life orientation test. *J Pers Soc Psychol*. 1994;67(6):1063-1078. https://doi.org/10.1037/0022-3514.67.6.1063
33. Kupst MJ, Butt Z, Stoney CM, et al. Assessment of stress and self-efficacy for the NIH Toolbox for neurological and behavioral function. *Anxiety Stress Coping*. 2015;28(5):531-544. https://doi.org/10.1080/10615806.2014.994204
34. Peterman AH, Fitchett G, Brady MJ, et al. Measuring spiritual well-being in people with cancer: the Functional Assessment of Chronic Illness Therapy—Spiritual Well-Being Scale (FACIT-Sp). *Ann Behav Med*. 2002;24(1):49-58. https://doi.org/10.1207/S15324796ABM2401_06
35. Storch EA, Roberti JW, Heidgerken AD, et al. The Duke Religion Index: a psychometric investigation. *Pastor Psychol*. 2004;53(2):175-181. https://doi.org/10.1023/B:PASP.0000046828.94211.53
36. Koenig HG, Büssing A. The Duke University Religion Index (DUREL): a five-item measure for use in epidemiological studies. *Religions*. 2010;1(1):78-85. https://doi.org/10.3390/rel1010078
37. Cohen S, Mermelstein R, Kamarck T, Hoberman HM. Measuring the functional components of social support. In: Sarason IG, Sarason BR, eds. *Social Support: Theory, Research and Applications*. (NATO ASI Series, Series D: Behavioural and Social Sciences, no. 24). Springer-Verlag; 1985:73-94. https://doi.org/10.1007/978-94-009-5115-0_5
38. Gallo LC, de los Monteros KE, Ferent V, et al. Education, psychosocial resources, and metabolic syndrome variables in Latinas. *Ann Behav Med*. 2007;34(1):14-25. https://doi.org/10.1007/BF02879917
39. Moos RH, Moos BS. A typology of family social environments. *Fam Process*. 1976;15(4):357-371. https://doi.org/10.1111/j.1545-5300.1976.00357.x
40. Dupuis M, Baggio S, Gmel G. Validation of a brief form of the perceived neighborhood social cohesion questionnaire. *J Health Psychol*. 2017;22(2):218-227. https://doi.org/10.1177/1359105315600234

41. Cohen S, Doyle WJ, Skoner DP, et al. Social ties and susceptibility to the common cold. *JAMA*. 1997;277(24):1940-1944. https://doi.org/10.1001/jama.1997.03540480040036
42. Penedo FJ, Cohen L, Bower J, Antoni M. COVID-19: impact of the pandemic and HRQoL in cancer patients and survivors. April 24, 2020. Updated June 4, 2021. Accessed June 1, 2022. https://doi.org/10.17605/OSF.IO/2N7SY
43. Hispanic Community Health Study/Study of Latinos. HCHS/SOL Food Propensity Questionnaire. (Version B). Collaborative Studies Coordinating Center, University of North Carolina at Chapel Hill; 2012. Accessed June 1, 2022. https://sites.cscc.unc.edu/hchs/system/files/forms/COM08RETFoodPropensityQxVerBFPE02082012.pdf
44. National Center for Health Statistics. Smoking and Tobacco Use – SMQ (NHANES 2015 questionnaire). Centers for Disease Control and Prevention; 2015. Accessed June 1, 2022. https://wwwn.cdc.gov/Nchs/Data/Nhanes/2015-2016/Questionnaires/SMQ_I.pdf
45. National Institute on Alcohol Abuse and Alcoholism. Health professionals & communities. Accessed June 1, 2022. https://www.niaaa.nih.gov/health-professionals-communities
46. Chu AHY, Ng SHX, Koh D, et al. Reliability and validity of the self- and interviewer-administered versions of the Global Physical Activity Questionnaire (GPAQ). *PloS One*. 2015;10(9):e0136944. https://doi.org/10.1371/journal.pone.0136944
47. Maly RC, Frank JC, Marshall GN, et al. Perceived Efficacy in Patient-Physician Interactions (PEPPI): validation of an instrument in older persons. *J Am Geriatr Soc*. 1998;46(7):889-894. https://doi.org/10.1111/j.1532-5415.1998.tb02725.x
48. Bonevski B, Sanson-Fisher R, Girgis A, et al. Evaluation of an instrument to assess the needs of patients with cancer. *Cancer*. 2000;88(1):217-225. https://doi.org/10.1002/(SICI)1097-0142(20000101)88:1<217::AID-CNCR29>3.0.CO;2-Y
49. Cole SW. Chronic inflammation and breast cancer recurrence. *J Clin Oncol*. 2009;27(21):3418-3419. https://doi.org/10.1200/JCO.2009.21.9782
50. Cole SW. Social regulation of human gene expression. *Curr Dir Psychol Sci*. 2009;18(3):132-137. https://doi.org/10.1111/j.1467-8721.2009.01623.x
51. Miltenyi Biotech. CD14 MicroBeads, human. Accessed June 1, 2022. https://www.miltenyibiotec.com/US-en/products/cd14-microbeads-human.html#130-050-201
52. QIAGEN. RNeasy Kits. Accessed June 1, 2022. https://www.qiagen.com/us/products/discovery-and-translational-research/dna-rna-purification/rna-purification/total-rna/rneasy-kits/
53. Miller GE, Chen E, Fok AK, et al. Low early-life social class leaves a biological residue manifested by decreased glucocorticoid and increased proinflammatory signaling. *Proc Natl Acad Sci U S A*. 2009;106(34):14716-14721. https://doi.org/10.1073/pnas.0902971106
54. Miller G, Chen E. Unfavorable socioeconomic conditions in early life presage expression of proinflammatory phenotype in adolescence. *Psychosom Med*. 2007;69(5):402-409. https://doi.org/10.1097/PSY.0b013e318068fcf9
55. Miller GE, Murphy MLM, Cashman R, et al. Greater inflammatory activity and blunted glucocorticoid signaling in monocytes of chronically stressed caregivers. *Brain Behav Immun*. 2014;41:191-199. https://doi.org/10.1016/j.bbi.2014.05.016
56. Cole SW, Galic Z, Zack JA. Controlling false-negative errors in microarray differential expression analysis: a PRIM approach. *Bioinformatics*. 2003;19(14):1808-1816. https://doi.org/10.1093/bioinformatics/btg242
57. Cole SW, Yan W, Galic Z, et al. Expression-based monitoring of transcription factor activity: the TELiS database. *Bioinformatics*. 2005;21(6):803-810. https://doi.org/10.1093/bioinformatics/bti038
58. National Cholesterol Education Program. ATP III Guidelines At-A-Glance Quick Desk Reference. (NIH publication no. 01-3305). National Heart, Lung, and Blood Institute; 2001. Accessed June 1, 2022. https://www.nhlbi.nih.gov/files/docs/guidelines/atglance.pdf
59. Wang L-J, Zhang CW, Su SC, et al. An ancestry informative marker panel design for individual ancestry estimation of Hispanic population using whole exome sequencing data. *BMC Genomics*. 2019;20(suppl 12):1007. https://doi.org/10.1186/s12864-019-6333-6
60. Galanter JM, Fernandez-Lopez JC, Gignoux CR, et al. Development of a panel of genome-wide ancestry informative markers to study admixture throughout the Americas. *PLoS Genet*. 2012;8(3):e1002554. https://doi.org/10.1371/journal.pgen.1002554
61. Phillips C, Salas A, Sánchez JJ, et al. Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs. *Forensic Sci Int Genet*. 2007;1(3-4):273-280. https://doi.org/10.1016/j.fsigen.2007.06.008
62. Phillips C, Parson W, Lundsberg B, et al. Building a forensic ancestry panel from the ground up: the EUROFORGEN Global AIM-SNP set. *Forensic Sci Int Genet*. 2014;11:13-25. https://doi.org/10.1016/j.fsigen.2014.02.012
63. McKenna A, Hanna M, Banks E, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297-1303. https://doi.org/10.1101/gr.107524.110
64. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res*. 2012;22(3):568-576. https://doi.org/10.1101/gr.129684.111
65. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155(2):945-959. https://doi.org/10.1093/genetics/155.2.945
66. Hays R, Spritzer K, Schalet B, et al. PROMIS®-29 v2.0 profile physical and mental health summary scores. *Qual Life Res*. 2018;27(7):1885-1891. https://doi.org/10.1007/s11136-018-1842-3
67. Lai J-S, Wagner LI, Jacobsen PB, et al. Self-reported cognitive concerns and abilities: two sides of one coin? *Psychooncology*. 2014;23(10):1133-1141. https://doi.org/10.1002/pon.3522
68. Basch E, Dueck AC, Rogak LJ, et al. Feasibility of implementing the Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events in a multicenter trial: NCCTG N1048. *J Clin Oncol*. 2018;36(31):JCO2018788620. https://doi.org/10.1200/JCO.2018.78.8620
69. Winstead-Fry P, Schultz A. Psychometric analysis of the Functional Assessment of Cancer Therapy-General (FACT-G) scale in a rural sample. *Cancer*. 1997;79(12):2446-2452. https://doi.org/10.1002/(SICI)1097-0142(19970615)79:12<2446::AID-CNCR23>3.0.CO;2-Q
70. Webster K, Cella D, Yost K. The Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System: properties, applications, and interpretation. *Health Qual Life Outcomes*. 2003;1:79. https://doi.org/10.1186/1477-7525-1-79
71. Muthén BO. Beyond SEM: general latent variable modeling. *Behav Ther*. 2002;29(1):81-117. https://doi.org/10.2333/bhmk.29.81
72. Wang CP, Brown CH, Bandeen-Roche K. Residual diagnostics for growth mixture models: examining the impact of a preventive intervention on multiple trajectories of aggressive

behavior. *J Am Stat Assoc*. 2005;100(471):1054-1076. https://doi.org/10.1198/016214505000000501

73. VanderWeele T, Vansteelandt S. Mediation analysis with multiple mediators. *Epidemiol Methods*. 2014;2(1):95-115. https://doi.org/10.1515/em-2012-0010
74. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J R Stat Soc Ser B Stat Methodol*. 2005;67(2):301-320. https://doi.org/10.1111/j.1467-9868.2005.00503.x