Cancer Causes & Control (2019) 30:1231–1241
https://doi.org/10.1007/s10552-019-01229-4

ORIGINAL PAPER

Race/ethnicity and lung cancer survival in the United States: a meta-analysis

Madelyn Klugman[1] · **Xiaonan Xue**[1] · **H. Dean Hosgood III**[1]

Received: 7 March 2019 / Accepted: 4 September 2019 / Published online: 14 September 2019

Abstract

Purpose Lung cancer mortality has been shown to vary by race and ethnicity in cancer registries; however, studies often do not account for smoking status. We sought to summarize the independent contribution of race and ethnicity to survival in US lung cancer patients, accounting for important variables including smoking status.

Methods PubMed was used to identify 1,877 potentially eligible studies of which 27 were included. Studies were excluded if they did not account for age, race and/or ethnicity, and smoking status. Fixed- and random-effects meta-analyses were conducted using the reported adjusted hazard ratios (HR) of Hispanic ethnicity and Asian and African-American race compared to Non-Hispanic whites (NHWs) on overall survival in lung cancer.

Results Hispanic ethnicity and Asian race were associated with decreased adjusted risk of death (Hispanic: $N_{studies} = 5$, $N_{subjects} = 108{,}810$, HR = 0.95, 95% CI 0.90–1.00; Asian: $N_{studies} = 6$, $N_{subjects} = 128{,}950$, HR = 0.86, 95% CI 0.81–0.90). The results were similar when excluding studies of solely never-smokers. There was no significant difference in survival between African-American and white race after adjustment ($N_{studies} = 10$, $N_{subjects} = 131{,}378$, HR = 0.98, 95% CI 0.96–1.01). Other prognostic factors were female gender (HR = 0.88, 95% CI 0.87–0.89), unmarried status (HR = 1.08, 95% CI 1.04–1.11), ever-smoking status (HR = 1.11, 95% CI 1.08–1.15), having comorbidities (HR = 1.39, 95% CI 1.24–1.56), and treatment receipt (surgery: HR = 0.33, 95% CI 0.32–0.34; radiation: HR = 0.87, 95% CI 0.85–0.88; chemotherapy: HR = 0.64, 95% CI 0.63–0.65).

Conclusions Even after adjustment for clinical factors and smoking status, Hispanics and Asians experienced improved survival compared to NHWs. Future studies are needed to elucidate the drivers of these survival disparities.

Keywords Hispanics · African-Americans · Ethnicity · Prognostic · Smoking · Review

Introduction

Lung cancer is the leading cause of cancer death for both men and women in the United States (US), accounting for over 150,000 deaths in 2017 [1]. Ethnic minorities experience heterogeneous mortality rates compared to whites, with some groups higher and some groups lower. According to data from the Surveillance, Epidemiology, and End Results (SEER) database, African-American (AA) patients have the highest rate of lung cancer mortality, although the disparity is decreasing, possibly because of decreasing smoking initiation by AA adolescents [2]. AA patients are more likely to present at an advanced stage [1, 3] and are less likely to undergo definitive therapy (treatments with intent of cure) [3, 4]. After adjustment for age, sex, stage at diagnosis and histology, AA non-small cell lung cancer (NSCLC) patients have significantly lower survival, and Hispanics have significantly greater survival than white patients [5]. Asian-Americans have the lowest rate of lung cancer mortality in the United States [3]. An important limitation of the large-scale cancer registries is that many lack robust covariate data, most notably smoking [5, 6]. Cigarette smoking is the primary risk factor for developing lung cancer in the US and may be a prognostic marker, as it has been shown to influence treatment outcomes [7, 8]. Other potential confounders

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10552-019-01229-4) contains supplementary material, which is available to authorized users.

✉ Madelyn Klugman
madelyn.klugman@einsteinmed.org

[1] Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, USA

of the relationship between race/ethnicity and lung cancer survival that may not be available in large-scale registries include comorbidities (smoking-related and smoking-independent), socioeconomic status, and social support [5]. Institutional series often offer greater covariate information but many have small sample sizes. To better understand the differences in survival by race/ethnicity and robustly summarize the existing literature, we conducted a meta-analysis assessing the independent contributions of race/ethnicity to survival as well as the potential demographic (e.g., gender, marital status) and molecular drivers of lung cancer survival.

Materials and methods

Search strategy

The PubMed database was searched to identify peer-reviewed manuscripts published through June 30, 2018, written in English and not using animal models. Combinations of search terms in the following categories were applied: "lung cancer," "survival," "United States," and "minorities" (see Online Resource 1 for a detailed list of search terms). All of the racial and ethnic group classifications included in the United States Census [9] were incorporated into the keyword search, as well as other common terms. The goal of our PubMed search was to ensure that we considered a wide array of studies related to our a priori hypothesis. We then applied rigorous eligibility criteria to select only studies that accounted for important variables that we hypothesized a priori to be related to lung cancer survival.

Eligibility criteria

Studies investigating the causes of lung cancer survival disparities in United States minority populations were eligible for this review. The following inclusion criteria were used in the initial selection of articles: (i) primary data analysis used; (ii) study population consisted of incident primary lung cancer patients in the US, (iii) that included non-Caucasian and/or Hispanic participants; and (iv) survival statistics were reported for at least one sociodemographic, genetic, or environmental exposure variable. During manuscript review, we applied additional inclusion criteria: (v) the survival (overall or lung cancer-specific) statistics were in the form of adjusted hazard ratios (HR) and associated 95% confidence intervals (CI) in a multivariable model that (vi) accounted for at minimum age, race and/or ethnicity, and smoking status. We chose to exclude studies that did not account for smoking status because we believed a priori that smoking status may confound the relationship between race/ethnicity and survival in lung cancer. Smoking status has been shown to influence response to lung cancer treatments, including chemotherapy, radiation, and surgery [7, 8]. By including only studies that accounted for smoking status, we would be better able to tease apart the associations with survival attributed to smoking versus race/ethnicity. Accounting for covariates could be in the form of restriction, stratification, or inclusion as a covariate in the regression model. Common reasons for exclusion included presentation of univariate Cox regression results only, lack of clarity about which covariates were included in the final model, no adjustment for or stratification by smoking status, and publication of HR and p value but no 95% CI. The extraction was independently performed by two authors who resolved any disagreements in person by consensus.

Data abstraction

Results related to study design, subject population, outcome variable, and adjustment variables were extracted from each study. The adjusted HRs and 95% CIs were extracted for the following covariates: gender (female vs. male), race [African-American vs. white or Non-Hispanic white (NHW), Asian vs. white or NHW], ethnicity (Hispanic vs. NHW or Non-Hispanic], smoking status (ever vs. never or current vs. former/never), treatment (receipt of chemotherapy, radiation, and/or surgery), comorbidities/performance status, and marital status. Of note, one study compared non-whites versus whites and was considered in the African-American model because 98% of the subjects in the non-white group were African-American [10]. In addition, another study compared Asians to Non-Asians [11], which was comprised of 89% Caucasians; this study was included in the Asians versus white model. When multiple models were reported, we chose the maximally adjusted model. When the reference group varied by study (e.g., male vs. female), the reciprocals of the hazard ratio and 95% CI were extracted from the study with the less commonly used reference group. For articles with overlapping patient populations (i.e., [12, 13], and [14, 15]) or articles with overlapping models (i.e., [13]), all unique HRs of interest were extracted. When similar HRs were reported in those studies with overlapping populations, the HR was extracted from the one with the larger population (i.e., [12, 15]). If dummy variables for race/ethnicity were further stratified by foreign versus US-born, the hazard ratio for the dummy variable of larger sample size was chosen.

Statistical analyses

All statistical analyses were performed using STATA (v 15.1, STATA Corporation, College Station, TX, US). The HRs and 95% CIs were pooled for each covariate and summary-effect estimates were calculated. All of the HRs accounted for age, sex, smoking status, and stage, and

many also accounted for treatment, histology, and measures of socioeconomic status. These HRs were then used to estimate the meta-analyses HRs. We chose to employ fixed-effects models for the main analysis due to (a) wide variation in sample size across studies, as these models are less influenced by studies of lower sample size, and (b) low heterogeneity was assumed given the strict covariate criteria required for inclusion in the study [16, 17]. Random-effects models were also employed for comparison. Graphical presentations in the form of forest plots were generated. Heterogeneity among studies was determined using the I^2 test of homogeneity and the Cochran Q statistic. We characterized I^2 values of 25, 50, and 75% as low, moderate and high heterogeneity, respectively [18].

Sub-analyses were also performed to assess if the difference in survival between race and ethnicity groups, when accounting for at least age, sex, and smoking status, varied by lung cancer presentation (NSCLC only, early stage vs. late stage, smokers vs. non-smokers). Three additional sub-analyses assessed if the difference in survival between race and ethnicity groups remains the same with the addition of other potentially meaningful covariates. These analyses included studies that accounted for socioeconomic status (marital status, measures of poverty, or insurance), histology (e.g., adenocarcinoma, squamous cell carcinoma), and treatment (chemotherapy, radiation, surgery, or a combination), in addition to age, sex, race/ethnicity and smoking status. Publication bias was assessed via funnel plots and the Begg's test (with continuity correction).

Results

The initial keyword searches yielded 1,877 manuscripts of which 1,023 titles were selected for abstract review (Fig. 1). Subsequently, 660 full manuscripts were assessed, and 208 articles were selected for further evaluation. 27 studies met our final inclusion criteria [10–15, 19–39]. Table 1 summarizes the study populations of the included articles. The studies were conducted in populations spanning the US. Nine studies reported the presence of Hispanics in their study population, 24 reported the presence of African-Americans, and 11 reported the presence of Asian-Americans. Per eligibility criteria, all included studies accounted for age, race and/or ethnicity, and smoking status; all included studies also happened to account for stage in the models. A majority of studies accounted for socioeconomic status ($N_{studies} = 16$) and histology ($N_{studies} = 19$), and nearly all studies accounted for treatment ($N_{studies} = 23$).

Hispanic ethnicity as compared to whites (or Non-Hispanics) was found to be associated with a 5% decreased risk of death ($N_{studies} = 5$, $N_{subjects} = 108{,}810$, HR = 0.95, 95% CI 0.90–1.00), with moderate heterogeneity present ($I^2 = 57.4\%$) (Fig. 2). The findings were similar and heterogeneity was minimal when excluding studies that

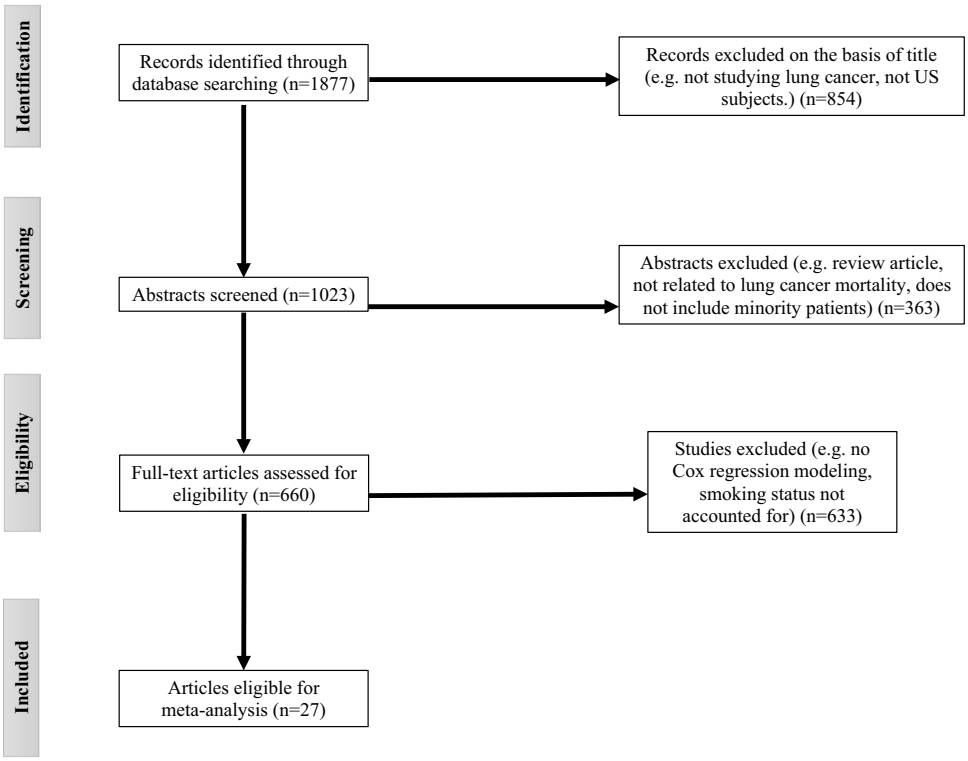


Fig. 1 Flowchart presenting the selection of the eligible studies related to race/ethnicity and survival in US lung cancer patients

Table 1 Characteristics of studies included in meta-analysis of sociodemographics and lung cancer survival in the United States

References	Study period	State	# lung cancer cases	%				Accounted for[a]				
				Hispanic[b]	AA[c]	Asian	Female	Smoking	Stage	SES[d]	Histology	Treatment
Ademuyiwa et al. [19]	2002–2006	IN	203	NR	5	NR	34	✓	✓			✓
Aldrich et al. [20]	2002–2011	SE US[e]	501	0	70	0	45	✓	✓	✓		
Biswas et al. [21]	2001–2010	NC	569	0	25	0	40	✓	✓	✓	✓	✓
Biswas et al. [22]	2001–2010	NC	2,351	0	30	0	37	✓	✓	✓	✓	
Brasky et al. [23]	2000–2009	WA	785	6% non-white			56	✓	✓		✓	✓
Brzezniak et al. [24]	2003–2013	US	4,751	2.4	11	9.9	36	✓	✓	✓	✓	✓
Clément-Duchêne et al. [13]	2003–2005	US	3,410	4.3	10.7	4.1	40	✓	✓	✓	✓	✓
Elchoufani et al. [25]	2001–2010	NC	2,351	0	70	0	37	✓	✓	✓	✓	✓
Enewold et al. [26]	NR[f]	MD	242	0	29	0	45	✓	✓		✓	
Enewold et al. [27]	1998–2003	MD	334	0	24	0	50	✓	✓		✓	✓
Gomez et al. [28]	1998–2008	CA	462	40	12	45	100	✓	✓	✓	✓	✓
Jones et al. [29]	2002–2012	SE US	286	0	100	0	40	✓	✓	✓		✓
Katcoff et al. [30]	2001–2005	MI	485	0	23.1	0	100	✓	✓	✓		✓
Mulligan et al. [31]	1990–2000	DC	886	0	20	0	33	✓	✓			✓
Ou et al. [32]	1991–2005	CA	1,124	0	0	100	36	✓	✓	✓	✓	✓
Ou et al. [33]	1991–2005	CA	4,782	7	3	4	47	✓	✓	✓	✓	✓
Ou et al. [11]	2001–2010	CA	20,140	8	3	6	45	✓	✓	✓	✓	✓
Pennella et al. [34]	2002–2006	US	434	7	15	9	46	✓	✓	✓	✓	✓
Pine et al. [35]	1984–2004	MD	173 [g]	0	100	0	45	✓	✓			
Sausville et al. [36]	2002–2010	SE US	395	0	67	0	45	✓	✓			✓
Tannenbaum et al. [15]	1996–2007	FL	98,541	5	7	1	44	✓	✓	✓	✓	✓
Van Dyke et al. [37]	2001–2005	MI	238	NR	19	NR	100	✓	✓		✓	✓
Vyfhuis et al. [38]	2000–2013	MD	355	1	42	1	43	✓	✓	✓		✓
Wan et al. [39]	NR	US	442	0	NR	1	50	✓	✓		✓	✓
Williams et al. [12]	2001–2010	US	18,466	0	15	0	NR	✓	✓		✓	✓
Xu et al. [10]	1998–2010	US	224	0	20	1	48	✓	✓		✓	✓
Yang et al. [14]	1998–2002	FL	76,086	5.7	6.7	NR	45	✓	✓	✓	✓	✓

[a]Adjusted for this variable in Cox regression or reported strata-specific Cox models

[b]Race (White, African-American, and Asian) and ethnicity (Hispanic vs. Non-Hispanic) reporting varied. Most studies categorized subjects into mutually exclusive groups, but some studies reported race and ethnicity separately. Therefore, there may be some overlap in some categories (for example, a subject may be both African-American race and Hispanic ethnicity)

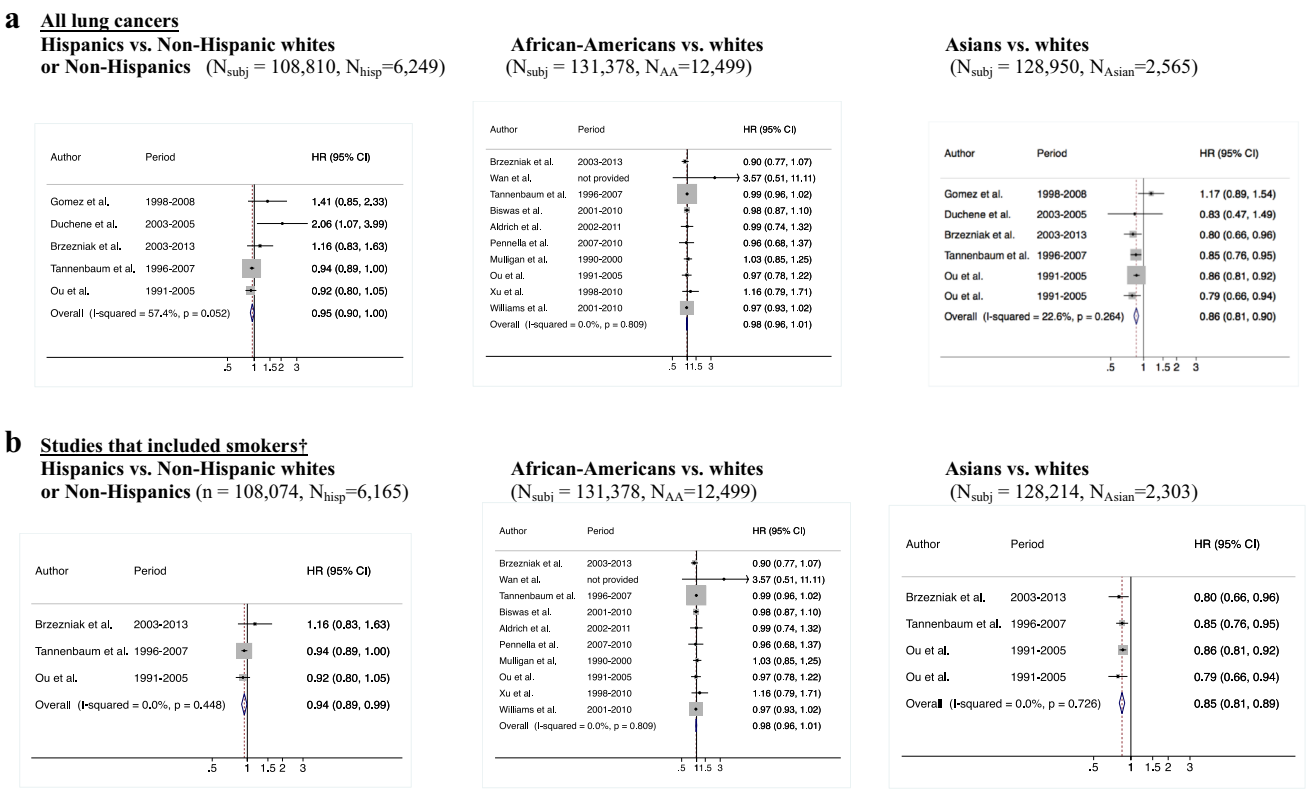
[c]African-American

[d]Socioeconomic status

[e]Southeastern US

[f]Not reported

[g]These authors presented stratified results by race; these subjects are from the African-American model



†These studies excluded studies that contained only never-smokers. The included studies consisted of a mix of smokers and never-smokers.

Fig. 2 Summary risk estimates for death associated with race/ethnicity in United States lung cancer patients, with (**a**) and without (**b**) never-smoking papers. The solid diamond symbols represent each study's published adjusted hazard ratio (HR), with the size proportional to the number of cases and horizontal lines representing the 95% CIs; summary ORs and 95% CIs were calculated using fixed-effects models

were comprised of only never-smokers (studies containing a mix of never-smokers and smokers were still included) ($N_{studies} = 3$, $N_{subjects} = 108{,}074$, HR = 0.94, 95% CI 0.89–0.99, $I^2 < 0.1\%$). The decreased risk of death in Hispanics ranged from 5 to 8% in subgroup analyses with minimal heterogeneity (Table 2). However, Hispanic ethnicity was a negative prognostic marker when analyses were restricted to studies of early stage and never-smokers only (early stage: $N_{studies} = 1$, $N_{Hispanic}/N_{total} = 36/1954$, HR = 2.64, 95% CI 1.06–6.62; never-smokers: $N_{studies} = 2$, $N_{Hispanic}/N_{total} = 84/73$, HR = 1.62, 95% CI 1.08–2.42). All of the Hispanic survival estimates used in the overall analyses accounted for SES, histology, and treatment, so those subgroup analyses have identical hazard ratios and heterogeneity estimates. There was evidence of publication bias for Hispanics ($p_{Begg's} = 0.03$) and examination of the funnel plot demonstrated right skew (Online Resource 3). Publication bias was no longer evident with restriction to models with ever-smokers only ($p_{Begg's} > 0.99$).

African-American race was not significantly associated with survival (vs. whites, $N_{studies} = 10$, $N_{subjects} = 131{,}378$, HR = 0.98, 95% CI 0.96–1.01), with minimal heterogeneity ($I^2 < 0.1\%$) (Fig. 2). Of note, none of the studies in the main analysis of African-American versus white race examined solely never-smokers, so results are unchanged in subgroup analyses excluding studies that modeled survival in never-smokers only. These findings remained true in subgroup analyses, except in one study of never-smokers only that reported 75% increased risk of death in African-Americans (HR = 1.75, 95% CI 0.96–3.18) (Table 2). There was no evidence of publication bias ($p_{Begg's} = 0.28$) (Online Resource 3). Among individual studies publishing results in uniquely stratified models, e.g., types of insurance, African-American race was still not a predictor of survival. One study found that African-American versus white race was not a significant independent predictor of survival among each stratum of health insurance [25], and another among each stratum of main treatment modality [12]. Asian race (vs. Non-Asian, Non-Hispanic white, or Caucasian race) was also significantly associated with increased survival across all studies ($N_{studies} = 6$, $N_{subjects} = 128{,}950$, HR = 0.86, 95% CI 0.81–0.90) with low heterogeneity ($I^2 = 22.6\%$). Results were

Table 2 Subgroup analysis of race/ethnicity (accounting for age, sex, and smoking status) in relation to survival in US lung cancer patients

Subgroup	Number of studies	Number of subjects	Heterogeneity [I^2 (%); p value]	HR (95% CI)
Hispanic[a]		$N_{Hispanic}/N_{total}$		
NSCLC only[b]	3	5,986/103,754	47.8; 0.15	0.95 (0.90–1.01)
Early-stage cases only[c]	1	36/1,954	n/a	2.64 (1.06–6.62)
Late-stage cases only[d]	2	288/6,994	0.0; 0.74	0.92 (0.81–1.06)
Never/non-smokers only	2	84/736	0.0; 0.37	1.62 (1.08–2.42)
Accounted for SES[e]	5	6,249/108,810	57.4; 0.05	0.95 (0.90–1.00)
Accounted for histology[f]	5	6,249/108,810	57.4; 0.05	0.95 (0.90–1.00)
Accounted for treatment[g]	5	6,249/108,810	57.4; 0.05	0.95 (0.90–1.00)
African-American[h]		$N_{AfricanAmerican}/N_{total}$		
NSCLC only[b]	8	11,255/124,245	0.0; 0.63	0.98 (0.96–1.01)
Early-stage cases only[c]	3	3,199/20,989	0.0; 0.71	0.97 (0.93–1.02)
Late-stage cases only[d]	3	414/7,428	0.0; 0.75	0.92 (0.80–1.06)
Never/non-smokers only	1	27/274	n/a	1.75 (0.96–3.18)
Accounted for SES[e]	6	9,001/131,500	0.0; 0.94	0.99 (0.96–1.02)
Accounted for histology[f]	8	11,532/129,991	0.0; 0.65	0.98 (0.96–1.00)
Accounted for treatment[g]	7	8,694/111,525	0.0; 0.57	0.99 (0.96–1.02)
Asian[i]		N_{Asian}/N_{total}		
NSCLC only[b]	4	2,379/123,894	44.8; 0.14	0.86 (0.82–0.91)
Early-stage cases only[c]	1	189/1,954	n/a	0.67 (0.41–1.09)
Late-stage cases only[d]	2	373/6,994	0.0; 0.64	0.81 (0.70–0.93)
Never/non-smokers only	3	581/6,575	55.0; 0.11	0.90 (0.79–1.02)
Accounted for SES[e]	6	2,565/128,950	22.6; 0.26	0.86 (0.81–0.90)
Accounted for histology[f]	6	2,565/128,950	22.6; 0.26	0.86 (0.81–0.90)
Accounted for treatment[g]	6	2,565/128,950	22.6; 0.26	0.86 (0.81–0.90)

[a]Compared to Caucasian/whites, Non-Hispanic whites (NHW), or Non-Hispanics

[b]Non-small cell lung cancer

[c]Studies restricted to stages I–IIIA (or a subset) only. All NSCLC

[d]Studies restricted to stage IIIB or IV NSCLC or extensive disease SCLC

[e]Defined as measures of poverty, marital status, and/or insurance type

[f]Must include at least two histologies more specific than subtype (e.g., adenocarcinoma, squamous cell carcinoma; these are more specific than SCLC or NSCLC)

[g]Radiation, surgery, and/or chemotherapy

[h]Compared to Caucasian/whites or Non-Hispanic whites. In one study [10], "non-white" described a sample of 98% African-Americans

[i]Compared to Caucasian/whites, NHWs or, in one study [11], Non-Asians (89% NHW)

similar in subgroup analyses for the Asian HR (Table 2). There was no evidence of publication bias ($p_{Begg's} > 0.99$) (Online Resource 3). The results were similar in random-effects models for each race/ethnicity category (Online Resource 2).

Other potential drivers beyond ethnicity were identified. Female gender was found to be significantly independently associated with improved survival ($N_{studies} = 12$, $N_{subjects} = 110{,}813$, HR = 0.88, 95% CI 0.87–0.89) with low heterogeneity ($I^2 = 36.7\%$) (Table 3). These findings were similar with low heterogeneity when restricting the analysis to studies of NSCLC cases. Subjects with high comorbidities or lower performance status experienced a 39% increased risk of death compared to those without comorbidities or with the highest performance status (HR = 1.39, 95% CI 1.24–1.56, $N_{studies} = 4$), with minimal heterogeneity ($I^2 < 0.1\%$). Not being married was associated with an 8% increased risk of death (HR = 1.08, 95% CI 1.04–1.11); however, there was high heterogeneity ($I^2 = 86.1\%$). Subjects who reported being an ever or current smoker experienced an 11% increased risk of death (HR = 1.11, 95% CI 2.08–1.15), with moderate heterogeneity ($I^2 = 56.0\%$). Receipt of surgery was associated with a two-thirds reduction in risk of death and was the treatment associated with the greatest survival benefit, with low heterogeneity (HR = 0.33, 95% CI 0.32–0.34, $I^2 = 31.0\%$) (Table 3).

Table 3 Analysis of factors other than race and ethnicity in relation to survival in US lung cancer patients (fixed-effects model)

Covariate	Number of studies	Number of subjects	Heterogeneity [I^2 (%); p value]	HR (95% CI)
Female	12	110,813	36.7; 0.10	0.88 (0.87–0.89)
NSCLC only	8	32,102	0.0; 0.76	0.84 (0.82–0.87)
Early-stage cases only[a]	2	2,523	73.5; 0.05	0.93 (0.76–1.12)
Late-stage cases only[b]	2	6,994	0.0; 0.41	0.83 (0.78–0.89)
Never/non-smokers only[d]	2	2,113	0.0; 0.82	0.79 (0.71–0.88)
Ever-smokers only	1	18,301	n/a	0.85 (0.82–0.88)
Accounted for SES[e]	7	108,885	60.0; 0.02	0.88 (0.87–0.89)
Accounted for histology[f]	6	108,608	43.4; 0.09	0.88 (0.87–0.89)
Accounted for treatment	9	107,788	38.4; 0.11	0.88 (0.86–0.89)
Ever/current smoker[c]	9	112,389	56.0; 0.02	1.11 (1.08–1.15)
NSCLC only	5	25,750	18.8; 0.30	1.06 (1.01–1.12)
Early-stage cases only	2	2,523	0.0; 0.69	1.09 (0.76–1.56)
Late-stage cases only	2	6,994	0.0; 0.78	1.34 (1.14–1.57)
Accounted for SES	6	111,520	69.5; 0.01	1.11 (1.08–1.15)
Accounted for histology	8	112,186	59.8; 0.02	1.11 (1.08–1.15)
Accounted for treatment	8	110,038	57.9; 0.02	1.12 (1.08–1.16)
Receipt of surgery[d]	5	120,303	31.0; .22	0.33 (0.32–0.34)
Receipt of radiation[d]	5	195,503	85.5; < 0.001	0.87 (0.85–0.88)
Receipt of chemotherapy[e]	5	101,637	98.5; < 0.001	0.64 (0.63–0.65)
Comorbidities[f]	4	5,452	46.1; 0.14	1.39 (1.24–1.56)
Not married[g]	2	24,922	86.1; 0.007	1.08 (1.04–1.11)

The bold represents the overall analysis of the variable (e.g. Female versus Male)

[a]Studies restricted to stages I–IIIA (or a subset) only. All non-small cell lung cancer (NSCLC)

[b]Studies restricted to stage IIIB or IV NSCLC or extensive disease SCLC

[c]Ever or current smoker versus never/non-current smoker. Some studies additionally adjusted for smoking pack-years

[d]One extensive disease small cell lung cancer-only article was excluded from this analysis due to substantial heterogeneity and limited clinical use of surgery/radiation in this disease type and stage

[e]Receipt of any chemotherapy or consolidation chemotherapy

[f]High versus low number of comorbidities or performance status

[g]Single, separated, divorced, or widowed versus married

Radiation and chemotherapy were also protective markers, although less so and with high heterogeneity (radiation: HR = 0.87, 95% CI 0.85–0.88, $I^2 = 85.5\%$; chemotherapy: HR = 0.64, 95% CI 0.63–0.65, $I^2 = 98.5\%$).

Of note, eight [10, 26, 27, 29, 35–37, 39] of the 27 studies presented unique genetic/molecular markers associated with lung cancer survival. Six studies stratified their genetic analysis by race (African-American vs white) [26, 27, 29, 35–37]. Four of them identified biomarkers/variants that were significantly associated with survival in one but not the other race; these markers included serum interleukins 10 and 12 [27] and variants in the matrix metalloproteinase–1 gene [27], in eicosanoid pathway genes [36] and in the mannose-binding lectin 2 gene [35]. None of the studies investigated the same markers.

Discussion

While multiple institutions and population registries in the US have assessed the role of race and ethnicity in survival in lung cancer, this is the first report, to the best of our knowledge, to summarize this relationship in a meta-analysis. We found that Hispanic ethnicity and Asian race are associated with superior survival in US lung cancer patients. Compared to whites, African-American race is not associated with better or worse survival. These findings generally remained when analyzing only early-stage cases, late-stage cases, and never-smokers; studies excluding never-smokers; and studies that adjusted for socioeconomic status, histology, and treatment. Female gender

was associated with improved survival in lung cancer after accounting for clinical and social factors, and these findings are similar to those reported in a prior meta-analysis [40]. Comorbidities, smoking, and marital status were found to be significantly associated with mortality.

The Hispanic paradox refers to findings showing similar or better health outcomes in Hispanics than in non-Hispanic whites in the US, despite lower average socioeconomic status [41, 42]. This has been most consistently demonstrated in CVD: Hispanics have lower rates of coronary heart disease, CVD mortality, and overall mortality despite a greater prevalence of CVD risk factors. The Hispanic paradox has also shown in NSCLC in large-scale studies [5], but these findings are not adjusted for smoking. In the present meta-analysis that included over 100,000 subjects with lung cancer, of which over 6,000 identified as Hispanic, and accounted for smoking and other important covariates, we found evidence of the Hispanic paradox. There was stronger evidence of the paradox when excluding analyses of solely never-smokers. While the drivers of this paradox are still not fully elucidated, there are several theories: data artifacts, migration effects, and cultural effects [43]. The first theory relates to issues of misclassification of Hispanics on death certificates; however, it has been shown that the net ascertainment of Hispanic origin is 5% higher on survey records than on death certificates, which is not enough to explain the approximately 20% age-adjusted mortality advantage in Hispanics [44]. Hypotheses that there is self-selection of healthy Hispanic immigrants to the US, and that sick Hispanics return to their home countries when near death, have been demonstrated to exert minimal influence on mortality differences [45–47]. The role of Hispanic culture on health has been shown based on comparison of US-born and foreign-born Hispanics, the latter with generally superior outcomes [48]. It has been proposed that with those who are less acculturated are more likely to have healthier diets [49, 50]; a better social support network, particularly among those of lower socioeconomic status [51]; decreased smoking [52, 53]; and decreased alcohol use [54]. However, studies have shown that the CVD mortality benefit is still present even after adjusting for smoking and other sociodemographic factors [45, 55]. Soneji et al. [56] found that Hispanics had better overall survival in early-stage lung cancer relative to Non-Hispanic Whites (adjusted for clinical and demographic factors but not smoking) that was largely driven by a decreased risk of death from other causes than lung cancer. They hypothesized the improved survival could be related to decreased smoking-related comorbidity. Of note, the term "Hispanic paradox" does not take into account heterogeneity among Hispanics: studies have shown differences in disease risk factors and mortality by country or region of origin [57, 58].

Regarding the Hispanic paradox in lung cancer, there is some evidence that foreign-born Hispanics have decreased lung cancer mortality compared to US-born Hispanics [28, 59]. This may point to differences in acculturation as mentioned above or perhaps different environmental exposures in Hispanic countries and the United States. Smoking habits are particularly important when considering acculturation's role in lung cancer mortality. One study showed that Hispanics are over five times more likely to be light intermittent smokers than Non-Hispanic whites, although it also found African-Americans to be nearly four times as likely to be light intermittent smokers as well [60]. These results indicate that smoking may capture some but not all of the differences. While all of the studies accounted for smoking, smoking status of ever versus never (the most common metric in this meta-analysis) may not capture dose and frequency differences in Hispanic and non-Hispanic smokers. Environmental exposures leading to different mutational profiles in Hispanics may also explain some of the survival advantage. Importantly, none of the included studies evaluated genetics in Hispanics alone. This may be due to low sample size and supports the importance of the development of a Hispanic/Latino Lung Cancer Registry to identify mutation profiles for NSCLC in this population [61]. This registry has demonstrated an increased frequency of *EGFR* mutations relative to Caucasians, similar to findings of East Asian lung cancer patients, as well as a decreased frequency of *TP53* mutations. The increased frequency of *EGFR* may be due to increased wood smoke exposure, decreased smoking, and increased tuberculosis exposure [41].

We also demonstrated that when adjusting for clinical and social factors, African-American and white subjects with lung cancer had similar survival. This suggests that the increased unadjusted mortality of African-Americans compared to whites can be explained by known prognostic factors (e.g., stage, treatment receipt) and social factors, particularly smoking status. Importantly, many of the studies in the meta-analysis had patient populations with similar socioeconomic status or health insurance (e.g., Veterans Health Administration systems). Of note, there have been a few studies in which accounting for socioeconomic status, but not smoking, has demonstrated decreased risk of adverse outcomes in African-Americans versus whites. This includes decreased risk of in-patient mortality after surgery [62] and, in stage III NSCLC patients, increased overall survival [63]. Asian-Americans with lung cancer were found to have superior survival in the meta-analysis and in subgroup analyses. Of note, one study of Asian-American patients [32] found that gender and smoking status were not significant predictors of death. Some theories of why Asians may have superior survival include high rates of never-smokers with lung cancer and higher prevalence of favorable and actionable genetic mutations including EGFR [11]. This is likely

secondary to environmental tobacco smoke exposure [64] and household coal use [65, 66].

Our study has a number of strengths and limitations. We are the first to assess by meta-analysis the independent (accounting for age and smoking) effect of race and ethnicity in survival from lung cancer in a large sample of US patients. Our studies were selected in a systematic and objective manner. We found that the overwhelming majority of studies assessing the role of race and ethnicity in survival in lung cancer did not account for smoking, so excluding these articles substantially reduced smoking's confounding effect and heterogeneity of results. However, using a binary variable for smoking status does not account for as detailed of a smoking history as more refined exposure assessments such as pack-years of smoking. Unfortunately, the vast majority of studies did not account for or report an adjusted HR for pack-years of smoking and this level of analysis was not feasible. Another strength of our study is that we assessed the robustness of our findings using sensitivity and subgroup analyses. If the overall results were heterogeneous, subgroup analyses aimed to explain and reduce heterogeneity were further conducted. A limitation of our study includes low number of studies with such robust statistics available for extraction. Because many of our analyses had fewer than ten studies, interpretation of statistical tests for publication bias, either using Begg's or Egger's tests, should be limited due to low power [67]. However, results were generally consistent across the studies, particularly in subgroup analysis, lending confidence to our findings. In addition, our findings were similar in fixed- and random-effects analyses. Furthermore, there was evidence of publication bias in the reporting of the HR of Hispanic ethnicity. Notably, inspection of the funnel plot demonstrated high right skew among the small-scale studies, suggesting potentially unpublished studies supporting our main findings. This may be indicative of a file drawer effect due to low sample sizes of Hispanics typically observed in studies and thus limited power for publication.

In summary, racial/ethnic differences in overall survival from lung cancer in Americans are present; notably, Hispanics and Asians had improved survival relative to NHWs. Future studies are needed in light of our findings. There are several factors that may influence lung cancer survival whose independent effects should be investigated: environmental exposures other than tobacco, use of targeted therapies, and utilization of lung cancer screening. Furthermore, given that some racial/ethnic groups were independent predictors of survival, factors related to survival differences (including nationality in Hispanics) should be examined in those populations. Comorbidities, measures of social support including marital status, and particularly smoking (ideally utilizing pack-year data to avoid residual confounding) should be considered as covariates in analyses of future studies as we have shown an association with mortality in this review. In our study, we found a particularly elevated HR of having comorbidities compared to no comorbidities even after accounting for smoking, suggesting that differences in smoking-related and smoking-unrelated comorbidities by race/ethnicity could at least partially explain racial/ethnic differences in survival in lung cancer. In addition, while there were no significant differences in survival in African-Americans versus NHWs after controlling for age and smoking; these results should be replicated in study populations with a greater range in access to health care. The ultimate goal will be to develop interventions that will successfully mitigate drivers of survival disparities.

Acknowledgments This work was supported by NIH/National Center for Advancing Translational Science (NCATS) Einstein-Montefiore CTSA Grant Number ULITR001073.

Author contributions MK, XX, and HDH conceived and designed the analysis; MK analyzed the data; MK, XX, and HDH contributed to the interpretation of the results and manuscript preparation.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

References

1. Siegel RL, Miller KD, Jemal A (2017) Cancer statistics, 2017. CA Cancer J Clin 67:7–30
2. American Cancer Society (2019) Cancer facts & figures for African Americans 2019–2021. American Cancer Society, Atlanta
3. Aizer AA, Wilhite TJ, Chen MH et al (2014) Lack of reduction in racial disparities in cancer-specific mortality over a 20-year period. Cancer 120:1532–1539
4. Shugarman LR, Mack K, Sorbero ME et al (2009) Race and sex differences in the receipt of timely and appropriate lung cancer treatment. Med Care 47:774–781
5. Saeed AM, Toonkel R, Glassberg MK et al (2012) The influence of Hispanic ethnicity on nonsmall cell lung cancer histology and patient survival: an analysis of the survival, epidemiology, and end results database. Cancer 118:4495–4501
6. Siegel DA, Henley SJ, Wike JM et al (2018) Capture of tobacco use among population-based registries: findings from 10 National Program of Cancer Registries states. Cancer 124:2381–2389
7. Tammemagi CM, Neslund-Dudas C, Simoff M, Kvale P (2004) Smoking and lung cancer survival: the role of comorbidity and treatment. Chest 125:27–37
8. Leduc C, Antoni D, Charloux A, Falcoz PE, Quoix E (2017) Comorbidities in the management of patients with lung cancer. Eur Respir J 49:1601721
9. US Census Bureau, US Department of Commerce (2017) Race & ethnicity
10. Xu T, Wei Q, Lopez Guerra JL et al (2012) HSPB1 gene polymorphisms predict risk of mortality for US patients after radio(chemo) therapy for non-small cell lung cancer. Int J Radiat Oncol Biol Phys 84:e229–e235
11. Ou SH, Ziogas A, Zell JA (2009) Asian ethnicity is a favorable prognostic factor for overall survival in non-small cell lung cancer (NSCLC) and is independent of smoking status. J Thorac Oncol 4:1083–1093

12. Williams CD, Salama JK, Moghanaki D, Karas TZ, Kelley MJ (2016) Impact of race on treatment and survival among U.S. veterans with early-stage lung cancer. J Thorac Oncol 11:1672–1681
13. Clement-Duchene C, Stock S, Xu X et al (2016) Survival among never-smokers with lung cancer in the cancer care outcomes research and surveillance study. Ann Am Thorac Soc 13:58–66
14. Yang R, Cheung MC, Byrne MM et al (2010) Do racial or socioeconomic disparities exist in lung cancer treatment? Cancer 116:2437–2447
15. Tannenbaum SL, Koru-Sengul T, Zhao W, Miao F, Byrne MM (2014) Survival disparities in non-small cell lung cancer by race, ethnicity, and socioeconomic status. Cancer J 20:237–245
16. Nikolakopoulou A (2014) Demystifying fixed and random effects meta-analysis. Evid Based Mental Health 17:53–57
17. Barili F, Parolari A, Kappetein PA, Freemantle N (2018) Statistical primer: heterogeneity, random- or fixed-effects model analyses? Interact Cardiovasc Thorac Surg 27:317–321
18. Higgins JP, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. BMJ 327:557–560
19. Ademuyiwa FO, Johnson CS, White AS et al (2007) Prognostic factors in stage III non-small-cell lung cancer. Clin Lung Cancer 8:478–482
20. Aldrich MC, Grogan EL, Munro HM, Signorello LB, Blot WJ (2013) Stage-adjusted lung cancer survival does not differ between low-income Blacks and Whites. J Thorac Oncol 8:1248–1254
21. Biswas T, Walker P, Podder T, Efird JT (2015) Effect of race and insurance on the outcome of stage i non-small cell lung cancer. Anticancer Res 35:4243–4249
22. Biswas T, Walker P, Podder T, Rosenman J, Efird J (2014) Important prognostic factors for lung cancer in tobacco predominant Eastern North Carolina: study based on a single cancer registry. Lung Cancer 84:116–120
23. Brasky TM, Baik CS, Slatore CG, Alvarado M, White E (2012) Prediagnostic nonsteroidal anti-inflammatory drug use and lung cancer survival in the VITAL study. J Thorac Oncol 7:1503–1512
24. Brzezniak C, Satram-Hoang S, Goertz HP et al (2015) Survival and racial differences of non-small cell lung cancer in the United States Military. J Gen Intern Med 30:1406–1412
25. Elchoufani SE, Efird JT, O'Neal WT, Davies SW, Landrine H, Biswas T (2013) The relation of race and type of health insurance to long-term risk of mortality among lung cancer patients in rural Eastern North Carolina. N C Med J 74:464–469
26. Enewold L, Mechanic LE, Bowman ED, Platz EA, Alberg AJ (2012) Association of matrix metalloproteinase-1 polymorphisms with risk of COPD and lung cancer and survival in lung cancer. Anticancer Res 32:3917–3922
27. Enewold L, Mechanic LE, Bowman ED et al (2009) Serum concentrations of cytokines and lung cancer survival in African Americans and Caucasians. Cancer Epidemiol Biomark Prev 18:215–222
28. Gomez SL, Chang ET, Shema SJ et al (2011) Survival following non-small cell lung cancer among Asian/Pacific Islander, Latina, and Non-Hispanic white women who have never smoked. Cancer Epidemiol Biomark Prev 20:545–554
29. Jones CC, Bush WS, Crawford DC et al (2017) Germline genetic variants and lung cancer survival in African Americans. Cancer Epidemiol Biomark Prev 26:1288–1295
30. Katcoff H, Wenzlaff AS, Schwartz AG (2014) Survival in women with NSCLC: the role of reproductive history and hormone use. J Thorac Oncol 9:355–361
31. Mulligan CR, Meram AD, Proctor CD, Wu H, Zhu K, Marrogi AJ (2006) Unlimited access to care: effect on racial disparity and prognostic factors in lung cancer. Cancer Epidemiol Biomark Prev 15:25–31
32. Ou SH, Ziogas A, Zell JA (2010) A comparison study of clinicopathologic characteristics of Southern California Asian American Non-small Cell Lung Cancer (NSCLC) patients by smoking status. J Thorac Oncol 5:158–168
33. Ou SH, Ziogas A, Zell JA (2009) Prognostic factors for survival in extensive stage small cell lung cancer (ED-SCLC): the importance of smoking history, socioeconomic and marital statuses, and ethnicity. J Thorac Oncol 4:37–43
34. Pennella E, Obasaju CK, Pohl G et al (2013) Prospective observational comparison of clinical outcomes between African-American and Caucasian patients receiving second-line treatment with pemetrexed for advanced non-small-cell lung cancer. Clin Lung Cancer 14:726–735
35. Pine SR, Mechanic LE, Ambs S et al (2007) Lung cancer survival and functional polymorphisms in MBL2, an innate-immunity gene. J Natl Cancer Inst 99:1401–1409
36. Sausville LN, Jones CC, Aldrich MC, Blot WJ, Pozzi A, Williams SM (2017) Genetic variation in the eicosanoid pathway is associated with non-small-cell lung cancer (NSCLC) survival. PLoS ONE 12:e0180471
37. Van Dyke AL, Cote ML, Prysak GM et al (2008) COX-2/EGFR expression and survival among women with adenocarcinoma of the lung. Carcinogenesis 29:1781–1787
38. Vyfhuis MAL, Bhooshan N, Molitoris J et al (2017) Clinical outcomes of black vs. non-black patients with locally advanced non-small cell lung cancer. Lung Cancer. 114:44–49
39. Wan YW, Beer DG, Guo NL (2012) Signaling pathway-based identification of extensive prognostic gene signatures for lung adenocarcinoma. Lung Cancer 76:98–105
40. Nakamura H, Ando K, Shinmyo T et al (2011) Female gender is an independent prognostic factor in non-small-cell lung cancer: a meta-analysis. Ann Thorac Cardiovasc Surg 17:469–480
41. Arrieta O, Ramirez-Tirado LA, Baez-Saldana R, Pena-Curiel O, Soca-Chafre G, Macedo-Perez EO (2015) Different mutation profiles and clinical characteristics among Hispanic patients with non-small cell lung cancer could explain the "Hispanic paradox". Lung Cancer 90:161–166
42. Ruiz JM, Steffen P, Smith TB (2013) Hispanic mortality paradox: a systematic review and meta-analysis of the longitudinal literature. Am J Public Health 103:e52–e60
43. Palloni A, Arias E (2004) Paradox lost: explaining the Hispanic adult mortality advantage. Demography 41:385–415
44. Arias E, Eschbach K, Schauman WS, Backlund EL, Sorlie PD (2010) The Hispanic mortality advantage and ethnic misclassification on US death certificates. Am J Public Health 100(Suppl 1):S171–S177
45. Medina-Inojosa J, Jean N, Cortes-Bergoderi M, Lopez-Jimenez F (2014) The Hispanic paradox in cardiovascular disease and total mortality. Prog Cardiovasc Dis 57:286–292
46. Abraido-Lanza AF, Dohrenwend BP, Ng-Mak DS, Turner JB (1999) The Latino mortality paradox: a test of the "salmon bias" and healthy migrant hypotheses. Am J Public Health 89:1543–1548
47. Turra CM, Elo IT (2008) The Impact of Salmon Bias on the Hispanic mortality advantage: new evidence from social security data. Popul Res Policy Rev 27:515–530
48. Argeseanu Cunningham S, Ruben JD, Narayan KM (2008) Health of foreign-born people in the United States: a review. Health Place 14:623–635
49. Bolstad AL, Bungum T (2013) Diet, acculturation, and BMI in Hispanics living in southern Nevada. Am J Health Behav 37:218–226
50. Yoshida Y, Scribner R, Chen L, Broyles S, Phillippi S, Tseng T-S (2017) Role of age and acculturation in diet quality among Mexican Americans: findings from the National Health and Nutrition Examination Survey, 1999–2012. Prev Chronic Dis 14:E59
51. Almeida J, Molnar BE, Kawachi I, Subramanian SV (2009) Ethnicity and nativity status as determinants of perceived

social support: testing the concept of familism. Soc Sci Med 68:1852–1858
52. Kaplan RC, Bangdiwala SI, Barnhart JM et al (2014) Smoking among U.S. Hispanic/Latino adults: the Hispanic community health study/study of Latinos. Am J Prev Med 46:496–506
53. Kondo KK, Rossi JS, Schwartz SJ, Zamboanga BL, Scalf CD (2016) Acculturation and cigarette smoking in Hispanic women: a meta-analysis. J Ethn Subst Abuse 15:46–72
54. Lui PP, Zamboanga BL (2018) A critical review and meta-analysis of the associations between acculturation and alcohol use outcomes among Hispanic Americans. Alcohol Clin Exp Res 42:1841–1862
55. Willey JZ, Rodriguez CJ, Moon YP et al (2012) Coronary death and myocardial infarction among Hispanics in the Northern Manhattan study: exploring the Hispanic paradox. Ann Epidemiol 22:303–309
56. Soneji S, Tanner NT, Silvestri GA, Lathan CS, Black W (2017) Racial and ethnic disparities in early-stage lung cancer survival. Chest 152:587–597
57. Schneiderman N, Chirinos DA, Avilés-Santa ML, Heiss G (2014) Challenges in preventing heart disease in hispanics: early lessons learned from the Hispanic Community Health Study/Study of Latinos (HCHS/SOL). Prog Cardiovasc Dis 57:253–261
58. Fenelon A, Chinn JJ, Anderson RN (2017) A comprehensive analysis of the mortality experience of hispanic subgroups in the United States: variation by age, country of origin, and nativity. SSM Popul Health 3:245–254
59. Patel MI, Schupp CW, Gomez SL, Chang ET, Wakelee HA (2013) How do social factors explain outcomes in non-small-cell lung cancer among Hispanics in California? Explaining the Hispanic paradox. J Clin Oncol 31:3572–3578
60. Reyes-Guzman CM, Pfeiffer RM, Lubin J et al (2017) Determinants of light and intermittent smoking in the United States: results from three pooled national health surveys. Cancer Epidemiol Biomark Prev 26:228–239
61. Watza D, Schwartz AG (2017) Addressing Underrepresented populations in lung cancer research: the Hispanic/Latino lung cancer registry identifies distinct mutation profiles for NSCLC. J Thorac Oncol 12:1744–1745
62. LaPar DJ, Bhamidipati CM, Harris DA et al (2011) Gender, race, and socioeconomic status affects outcomes after lung cancer resections in the United States. Ann Thorac Surg 92:434–439
63. Vyfhuis MAL, Bentzen SM, Molitoris JK et al (2019) Patterns of care and survival in Stage III NSCLC among black and latino patients compared with white patients. Clin Lung Cancer 20(248–57):e4
64. Thun MJ, Hannan LM, Adams-Campbell LL et al (2008) Lung cancer occurrence in never-smokers: an analysis of 13 cohorts and 22 cancer registry studies. PLoS Med 5:e185
65. Lan Q, He X, Shen M et al (2008) Variation in lung cancer risk by smoky coal subtype in Xuanwei, China. Int J Cancer 123:2164–2169
66. Hosgood HD 3rd, Wei H, Sapkota A et al (2011) Household coal use and lung cancer: systematic review and meta-analysis of case–control studies, with an emphasis on geographic variation. Int J Epidemiol 40:719–728
67. Ioannidis JP, Trikalinos TA (2007) The appropriateness of asymmetry tests for publication bias in meta-analyses: a large survey. CMAJ 176:1091–1096