ORIGINAL ARTICLE

Racial Disparities in Cancer Guideline-Concordant Treatment Using Surveillance, Epidemiology, and End Results Data for Patients With NSCLC

Eric Ababio Anyimadu, MS,[a] Jacklyn M. Engelbart, MD,[b,c] Jason Semprini, PhD,[d] Amanda Kahl, MPH,[b] Cameron Trentz, MS,[a] John M. Buatti, MD,[e] Thomas L. Casavant, PhD,[a] Mary E. Charlton, PhD,[b] Guadalupe Canahuate, PhD[a,*]

[a]*Department of Electrical and Computer Engineering, University of Iowa, Iowa City, Iowa*
[b]*Department of Epidemiology, University of Iowa, Iowa City, Iowa*
[c]*Department of Surgery, University of Iowa Hospitals and Clinics, Iowa City, Iowa*
[d]*Department of Public Health, Des Moines University, Des Moines, Iowa*
[e]*Department of Radiation Oncology, University of Iowa, Iowa City, Iowa*

Received 25 September 2023; revised 5 October 2024; accepted 12 October 2024
Available online - 19 October 2024

ABSTRACT

Introduction: Despite efforts to achieve health care equality, racial/ethnic disparities persist in lung cancer survival in the United States, with non-Hispanic Black patients experiencing higher mortality compared with non-Hispanic Whites. Previous research often focused on single treatments, overlooking the broad range of options available. We aimed to highlight disparities in survival and receipt of comprehensive lung cancer treatment by developing a guideline-concordant initial treatment (GCIT) indicator based on disease stage and recommended treatment.

Methods: Using data of the Surveillance, Epidemiology, and End Results on 377,370 patients with NSCLC, we derived a GCIT indicator based on National Comprehensive Cancer Network guidelines. Observed probabilities and logistic regression models adjusted for age, disease stage, and race were used to assess racial disparities in treatment and survival, with the Kaplan-Meier method evaluating survival rates. Racial/ethnic groups analyzed included non-Hispanic White, non-Hispanic Black, Asian/Pacific Islander, Hispanic, and American Indian/Alaska Native.

Results: Non-Hispanic Black patients had lower odds of receiving GCIT (OR = 0.80; 95% confidence interval [CI]: 0.78–0.82) and surviving 2 years after diagnosis (OR = 0.80; 95% CI: 0.78–0.82). Non-Hispanic Asians had the highest odds of receiving GCIT (OR = 1.02; 95% CI: 0.99–1.05). Patients receiving GCIT had improved survival, with early stage patients experiencing median survival of 67 to 102 months, compared with 11 to 17 months for those without GCIT.

Conclusion: Receiving GCIT considerably improves survival across all races, though disparities in receipt are observed. Interventions are needed to ensure equitable access to guideline-concordant care and reduce survival disparities for patients.

Keywords: Disparities; Non–small cell lung cancer; Guideline-concordant initial treatment; Survival

Introduction

Lung cancer is the primary cause of cancer-related deaths in the United States with a 5-year relative survival rate of 22.9% reported from 2012 to 2018.[1,2] For NSCLC that accounts for 84% of all lung cancer

***Corresponding author.**

Address for correspondence: Guadalupe Canahuate, PhD, Department of Electrical and Computer Engineering, University of Iowa, 4016 Seamans Center for the Engineering Arts and Sciences, Iowa City, Iowa 52242. E-mail: guadalupe-canahuate@uiowa.edu

Cite this article as: Anyimadu EA, Engelbart JM, Semprini J, et al. Racial disparities in cancer guideline-concordant treatment using surveillance, Epidemiology, and End Results Data for patients with NSCLC. *JTO Clin Res Rep* 2025;6:100747

ISSN: 2666-3643

https://doi.org/10.1016/j.jtocrr.2024.100747

diagnoses, the 5-year relative survival rate is 28% for the same period.[3]

Despite the general decline in lung cancer-related mortality in recent years due in part to lifestyle changes, improved screening, and treatment,[3–5] racial disparities occur such that non-Hispanic Black (NHB) patients experience worse survival outcomes compared with non-Hispanic White (NHW) patients.[6,7] Particularly, the 5-year survival rate of NHB patients with NSCLC is at least 2% lower than that of NHW patients.[8] The source of this disparity in survival has largely been attributed to discrepancies in access to stage-dependent treatment in several studies.[8–11] However, much of the research on treatment disparities has primarily emphasized single-treatment modalities, overlooking the broad spectrum of recommended treatment options available for lung cancer therapy.

Furthermore, owing to the growing knowledge among physicians regarding lung cancer, there are continuous advancements and emerging treatment recommendations.[12] Consequently, evaluating disparities in treatment becomes a multifaceted task, as treatment recommendations undergo constant changes and evolution across stage and time periods.

The National Comprehensive Cancer Network (NCCN) provides recommendations on the therapeutic care that patients with cancer should receive at initial diagnosis based on the stage or extent of disease and histology, including grade.[12,13] The standard course of treatment for NSCLC entails surgery, chemotherapy, and radiation therapy alone or in combination, and treatment that aligns with NCCN-recommended treatment is called a guideline-concordant treatment.[12]

Some studies have investigated treatment guidelines and variables to identify a guideline-concordant indicator. This indicator has been used to quantify racial disparities in treatment, revealing a historical trend where NHB patients are less likely relative to NHW patients to receive recommended treatment.[11,14,15]

Therefore, in this paper, we derive a similar indicator called guideline-concordant initial treatment (GCIT) using data from the Surveillance, Epidemiology, and End Results (SEER) database and evaluate racial disparities in receiving guideline-concordant treatment and 2-year survival after disease diagnosis. This indicator is extracted based on the NCCN guidelines published between 2003 and 2018, the American Joint Committee on Cancer (AJCC) staging variables (T, N, and M stages of the disease), and first course of therapy received by the patient (including treatment variables surgery, radiation therapy, and chemotherapy) available in the SEER database.

The SEER database is supported by the Surveillance Research Program in National Cancer Institute's Division of Cancer Control and Population Sciences.[16,17] SEER gathers population-based information from 21 different registries across the United States, and approximately 48% of the U.S. population are represented. The SEER database is considered the accepted standard for information on cancer diagnosis and treatment and is regularly used in quality control, data monitoring, and review of programs to guarantee the accuracy and reliability of data.[2]

In summary, this paper makes two relevant contributions. First, we investigate the racial discrepancies in receiving GCIT among patients with NSCLC over a substantial length of time. Second, we analyze the relationship between receiving GCIT and disparities in survival rates among the racial/ethnic groupings among patients with NSCLC and draw attention to the existence of disparities using cancer registry data to help improve patient outcomes.

Methods

Data

This retrospective, population-based study evaluates NSCLC cases diagnosed between 2004 and 2018 with data sourced from the SEER database. The SEER registries included in this study are from Greater California, Los Angeles, San Francisco-Oakland SMSA, Connecticut, Metropolitan Detroit, Greater Georgia, Metropolitan Atlanta, Rural Georgia, Hawaii, Iowa, Kentucky, Louisiana, New Jersey, New Mexico, San Jose-Monterey, Seattle - Puget Sound, and Utah.[18]

To reduce treatment variability, only patients with first-primary NSCLC (International Classification of Diseases for Oncology, third edition site codes C34.0–C34.3, C34.8, C34.9) were considered. Patients with occult or 0 (in situ) stages were excluded from the analyses. In addition, patients for which outcome variables could not be determined due to missing data (i.e., missing staging variables) were also excluded. Patients of unknown race or ethnicity were similarly not included in the study.

We used the SEER race/ethnicity variable to identify the following 5 mutually exclusive groups: NHW, NHB, Hispanic all races (H), non-Hispanic American Indian/Alaska Native (NHAIAN), and non-Hispanic Asian/Pacific Islander (NHAPI).

Outcome Formulation

Two binary outcome variables were used for the analyses: the GCIT indicator and 2-year survival after diagnosis.

The GCIT variable was coded using the AJCC's T, N, and M stages. The T stage indicates the extent to which the primary tumor has grown whereas the N stage corresponds to whether the primary tumor has spread to

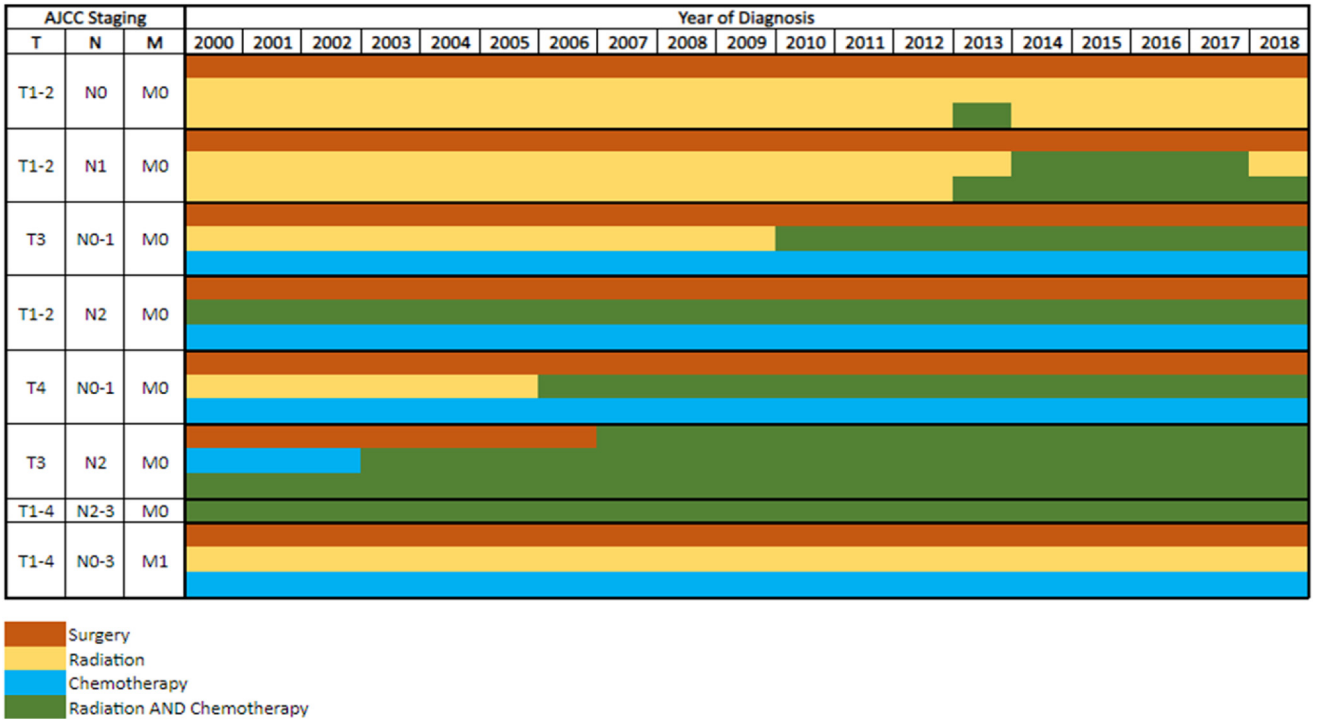

Figure 1. AJCC stage stratified-NCCN-recommended treatment for the period 2000 to 2018. Patients who were diagnosed in a particular year and received treatment based on their AJCC TNM stage and the corresponding NCCN-recommended treatment for that year are considered to be recipients of GCIT. AJCC, American Joint Committee on Cancer; GCIT, guideline-concordant initial treatment; NCCN, National Comprehensive Cancer Network.

nearby lymph nodes. The M stage also signifies whether the tumor has spread to various regions within the body through metastasis.

The classification of patients based on the TNM staging was done using the year of diagnosis and the corresponding AJCC version for that year. For cases diagnosed between 2004 and 2009, the AJCC sixth edition TNM staging variables were used, whereas those diagnosed between 2010 and 2015 were staged using the AJCC seventh edition TNM variables. Furthermore, the AJCC seventh edition variables were used for cases diagnosed between 2016 and 2017. Finally, for cases diagnosed in 2018, AJCC eighth edition variables were used.

The second factor used in defining the GCIT indicator was the NCCN-recommended treatment guidelines for the different AJCC stages and the corresponding year.

Because it is not possible to separate the No/Unknown categories for the treatment variables, we adopt the same notation for GCIT and use Yes when SEER abstractors were able to ascertain that the patient received GCIT and No/Unknown, otherwise.

Figure 1 provides in further detail the NCCN-recommended treatment considering the AJCC staging per year of diagnosis. These guidelines serve as the criteria for assigning a GCIT status to a given patient. Each row represents a TNM stage combination and each column a different year. The colors represent the following different treatments: surgery, radiation, chemotherapy, and concurrent chemotherapy (radiation and chemotherapy). For each patient, we identify the year of diagnosis, staging, surgery performed, radiotherapy, and chemotherapy variables in the SEER data to indicate whether the patient received GCIT. As can be found, the treatment recommendations change over time. For early stages (T1–2, N0, M0—first row in Fig. 1), the recommended treatment is either surgery or radiotherapy alone, and in 2013, it also included concurrent chemotherapy (radiotherapy and chemotherapy together) as a recommended alternative. For subsequent years, chemotherapy was usually recommended for the more advanced stages. For metastatic stages (M1—last row in Fig. 1), patients receiving either surgery, radiotherapy, or chemotherapy treatments (alone or in combination) would be compliant with the published guidelines for all years included in this study. Surgery recommendation targets interventions at the primary cancer site, including tumor resection or destruction. Nevertheless, more details about the specific procedures performed such as pleurodesis on a patient are not included in the data.

The 2-year survival outcome variable was derived using vital status data and survival/follow-up time at a

24-month minimum period for each patient. Living patients with less than 24-month follow-up were not included in the survival analysis.

Statistical Analysis

To determine the crude or unadjusted OR for GCIT across the various race/ethnic groups, we used the observed probabilities within the data set. In addition, we stratified the ORs based on disease stage, distinguishing between early stage (I and II) and late stage (III and IV).

In analyzing the 2-year survival outcome, we trained a logistic regression model to characterize the outcome based on the following adjustment predictors: age, sex, disease stage, and race/ethnicity. A standard 80/20 split was used to create the train and test sets from the original data set. To evaluate the impact of GCIT on 2-year survival, two variations of the model were developed. One included the GCIT indicator and the other excluded it as a predictor. The categorical variables (stage, sex, race, and GCIT) were transformed into binary representations using one-hot encoding. In addition, for each predictor, a category was selected and dropped to serve as a reference. Stage: IV, sex: male, race: NHW, and GCIT: no were used as references for the predictors.

We report the accuracy, area under the curve (AUC), and F1 score for each race/ethnicity and provide a plot of the OR for all covariates for both models. A confidence interval (CI) of 95% was adopted for reporting all the ORs.

Furthermore, to find the relationship between receiving GCIT and survival, we compared the trends in the Kaplan-Meier survival curves for the different racial/ethnic groups given the GCIT status of the treatment received. p values less than 0.01 were recognized to be statistically significant in this study. All analyses were conducted using Python version 3.9, with the NumPy and Pandas libraries.

Results

A total of 377,370 patients with NSCLC met the inclusion criteria for our analyses. Table 1 presents the distribution of patients by selected characteristics, treatment variables, and outcomes categorized by race/ethnicity. Most patients in the data set were NHW (74.6%), followed by NHB (12.1%), NHAPI (7.0%), and H (5.9%), whereas NHAIAN (0.4%) were the smallest proportion. In comparison to the other race/ethnic groups, NHB patients emerged as the youngest cohort at the time of diagnosis, with 48.7% falling below the age of 65 years and only 18.0% surpassing 75 years. Conversely, NHW and NHAPI patients represented relatively older groups, as 28.5% of NHW and 29.1% of NHAPI individuals were aged above 75 years. Moreover, NHB individuals were the highest proportion of late-stage diagnoses (74.0%), with 25.6% at stage III and 48.4% at stage IV. This was followed by NHAPI (72.4%) individuals with 21.1% at stage III and 51.3% at stage IV, whereas H (71.3%) had 22.2% at stage III and 49.1% at stage IV. NHW patients had the lowest proportion diagnosed at late stages (65.9%), with 23.5% at stage III and 42.4% at stage IV.

Other attributes such as diagnostic confirmation, tumor location, and laterality have no noticeable racial differences in their distribution.

Racial disparities in the distribution of treatment modalities among the cohort are suggested. Notably, NHW had the highest percentage (29.6%) of surgery performed whereas NHB had the lowest at 20.4%. NHB received more chemotherapy (at least 44.7%) and radiation (at least 44.1%) than NHW (at least 42.3%) for chemotherapy and (at least 40.7%) for radiation. NHW had the smallest proportion of patients receiving chemotherapy, whereas the smallest racial/ethnic grouping receiving radiation were H (at least 34.9%).

Considering the derived outcome variable GCIT and 2-year survival in all stages, the NHW group had the most patients receiving GCIT at 65.1% followed by NHAPI and NHAIAN with 63.9% and 63.0%, respectively, whereas NHB and H had the lowest at 61.1% and 58.9%, respectively.

Moreover, comparably more NHAPI and NHW patients survived 2 years after diagnosis (34.2% and 31.5%, respectively), followed by H and NHAIAN with 29.5% and 29.4%, respectively, whereas NHB had the lowest with only 26.7% patients surviving 2 years after diagnosis.

Stage-Stratified Outcomes

Figure 2 reveals the distribution of patients who received GCIT (Fig. 2*A*) and patients who survived 2 years after their diagnosis (Fig. 2*B*) stratified by stage for each race/ethnicity.

As can be found in Figure 2*A*, in the early stages, a greater proportion of patients of all races are known to receive GCIT as compared with the late stages. In the later stages, the percentage of patients known to receive GCIT is close to 55% for all racial/ethnic groups. Notably, the percentage of NHB patients known to have received GCIT in the early stages (stages I and II) was the lowest among all races, whereas H had the lowest percentages for the later stages (stages III and IV). Figure 2*B* reveals the 2-year survival rates stratified by stage. As expected, the survival rates decreased for more advanced stages. For all stages, NHB had the worst 2-year survival observed probability for all stages

Table 1. Distribution of Cohort by Selected Attributes Across Race/Ethnicity

Race/Ethnicity		NHW, n (%)	NHB, n (%)	H, n (%)	NHAIAN, n (%)	NHAPI, n (%)	All, n (%)	*p* Value
Count(%)		281,388 (74.6)	45,840 (12.1)	22,385 (5.9)	1372 (0.4)	26,385 (7.0)	377,370 (100)	-
Age at diagnosis (y)	<65	94,828 (33.7)	22,324 (48.7)	8372 (37.4)	538 (39.2)	9578 (36.3)	129,438 (34.3)	<0.001
	65–75	106,365 (37.8)	15,265 (33.3)	7924 (35.4)	519 (37.8)	9129 (34.6)	143,023 (37.9)	
	>75	80,196 (28.5)	8251 (18.0)	6089 (27.2)	316 (23.0)	7678 (29.1)	104,909 (27.8)	
Stage	I	72,598 (25.8)	8618 (18.8)	4790 (21.4)	317 (23.1)	5514 (20.9)	91,701 (24.3)	<0.001
	II	23,355 (8.3)	3300 (7.2)	1634 (7.3)	111 (8.1)	1768 (6.7)	30,190 (8.0)	
	III	66,126 (23.5)	11,735 (25.6)	4969 (22.2)	340 (24.8)	5567 (21.1)	88,682 (23.5)	
	IV	119,309 (42.4)	22,187 (48.4)	10,991 (49.1)	604 (44.0)	13,536 (51.3)	166,798 (44.2)	
Diagnostic confirmation	Clinical diagnosis only	1407 (0.5)	183 (0.4)	90 (0.4)	4 (0.3)	79 (0.3)	1887 (0.5)	NS
	Microscopically confirmed	268,726 (95.5)	44,098 (96.2)	21,713 (97.0)	1308 (95.3)	25,752 (97.6)	361,520 (95.8)	
	Radiography without microscopic confirm	10,693 (3.8)	1467 (3.2)	560 (2.5)	60 (4.4)	528 (2.0)	13,208 (3.5)	
	Unknown	563 (0.2)	92 (0.2)	22 (0.1)	0 (0.0)	26 (0.1)	755 (0.2)	
Laterality	Bilateral	2814 (1.0)	504 (1.1)	336 (1.5)	8 (0.6)	290 (1.1)	3774 (1.0)	NS
	Left - origin of primary	113,681 (40.4)	17,832 (38.9)	8909 (39.8)	534 (38.9)	10,554 (40.0)	151,325 (40.1)	
	Right - origin of primary	162,079 (57.6)	27,046 (59.0)	12,782 (57.1)	811 (59.1)	15,251 (57.8)	218,497 (57.9)	
	Other	2814 (1.0)	458 (1.0)	358 (1.6)	19 (1.4)	290 (1.1)	3774 (1.0)	
Treatment variables								
Chemotherapy	No/Unknown	162,361 (57.7)	25,350 (55.3)	12,692 (56.7)	792 (57.7)	13,536 (51.3)	214,724 (56.9)	<0.001
	Yes	119,027 (42.3)	20,490 (44.7)	9693 (43.3)	580 (42.3)	12,849 (48.7)	162,646 (43.1)	
Radiation administered	No/Unknown	166,863 (59.3)	25,625 (55.9)	14,573 (65.1)	811 (59.1)	16,860 (63.9)	224,913 (59.6)	<0.001
	Yes	114,525 (40.7)	20,215 (44.1)	7812 (34.9)	561 (40.9)	9525 (36.1)	152,457 (40.4)	
Surgery performed	No/Unknown	198,097 (70.4)	36,489 (79.6)	16,431 (73.4)	1018 (74.2)	19,261 (73.0)	271,329 (71.9)	<0.001
	Yes	83,291 (29.6)	9351 (20.4)	5954 (26.6)	354 (25.8)	7124 (27.0)	106,041 (28.1)	
Outcome indicators								
2-y survival after diagnosis	No	169,114 (60.1)	29,750 (64.9)	13,140 (58.7)	849 (61.9)	13,799 (52.3)	226,422 (60.0)	<0.001
	Unknown	23,637 (8.4)	3851 (8.4)	2641 (11.8)	119 (8.7)	3562 (13.5)	33,963 (9.0)	
	Yes	88,637 (31.5)	12,239 (26.7)	6604 (29.5)	403 (29.4)	9024 (34.2)	116,985 (31.0)	
Guideline-concordant initial treatment	No/Unknown	98,204 (34.9)	17,832 (38.9)	9200 (41.1)	508 (37.0)	9525 (36.1)	135,098 (35.8)	<0.001
	Yes	183,184 (65.1)	28,008 (61.1)	13,185 (58.9)	864 (63.0)	16,860 (63.9)	242,272 (64.2)	

Note: p value NS implies the category did not meet the frequency threshold and *p* values were not computed.

H, Hispanic all races; NHAIAN, non-Hispanic American Indian/Alaska Native; NHAPI, non-Hispanic Asian/Pacific Islander; NHB, non-Hispanic Black; NHW, non-Hispanic White; NS, not significant.

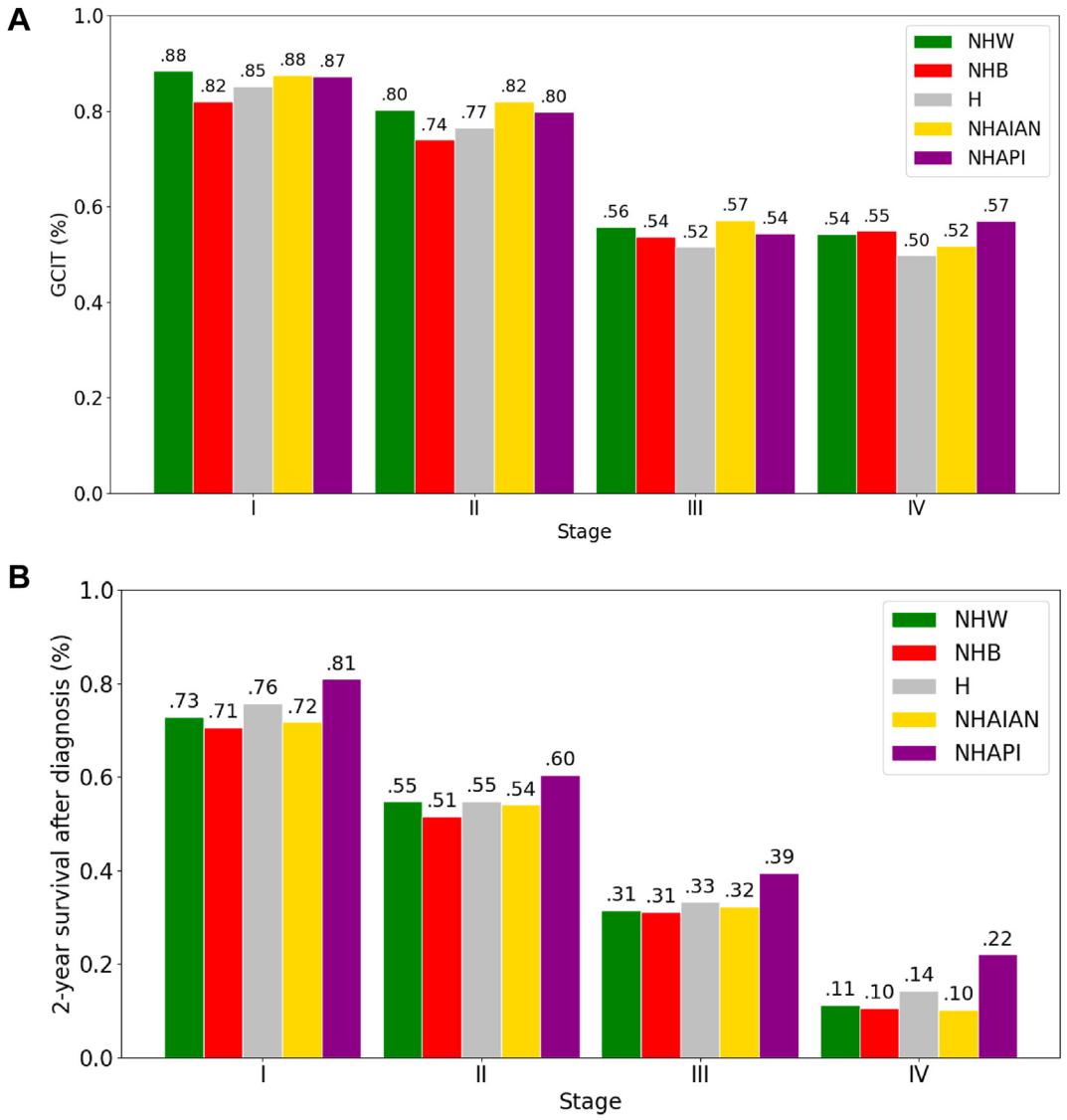

Figure 2. Percentage of patients by stage who received (*A*) GCIT and (*B*) 2-year survival post-diagnosis, categorized by race/ethnicity. GCIT, guideline-concordant initial treatment.

whereas NHAPI had the highest 2-year survival rate. Supplementary Table A1 provides further details on the distribution of outcome variables GCIT and 2-year survival by race/ethnicity and stage.

Outcome Evaluation

Figure 3 reveals the race/ethnicity OR for GCIT for (*A*) early stage patients, (*B*) late stage patients, and (*C*) all stages combined. Compared with NHW, NHB patients have at least 39% lower odds of receiving GCIT (OR = 0.61; 95% CI: 0.56–0.66) in the early stages and 6% lower odds (OR = 0.94; 95% CI: 0.91–0.96) in the late stages. In the late stages, H has the lowest odds of receiving GCIT (OR = 0.88; 95% CI: 0.84–0.91) among all races. Overall, for all stages combined, compared with NHW, NHB patients have the lowest odds (OR = 0.80; 95% CI: 0.78–0.82) of receiving GCIT whereas NHAPI patients have the highest (OR = 1.02; 95% CI: 0.99–1.05). Respectively, H and NHAIAN followed with (OR = 0.82; 95% CI: 0.79–0.84) and (OR = 0.92; 95% CI: 0.80–1.04).

Figure 4 reveals the ORs derived from the logistic regression models for 2-year survival using age, race, sex, and AJCC stage as predictors when GCIT is included (Fig. 4*A*) and when it is excluded (Fig. 4*B*) as a predictor in the model. The strongest predictor for survival in both models is AJCC staging (ORs: GCIT included [19.214–4.083], GCIT excluded [25.246–3.879]), followed by GCIT (ORs: 3.651; 95% CI: 3.569–3.735).

Table 2 compares the performance metrics over the test data overall and stratified by race/ethnicity for both models (including/excluding GCIT). As can be found, the inclusion of GCIT improves all the performance metrics

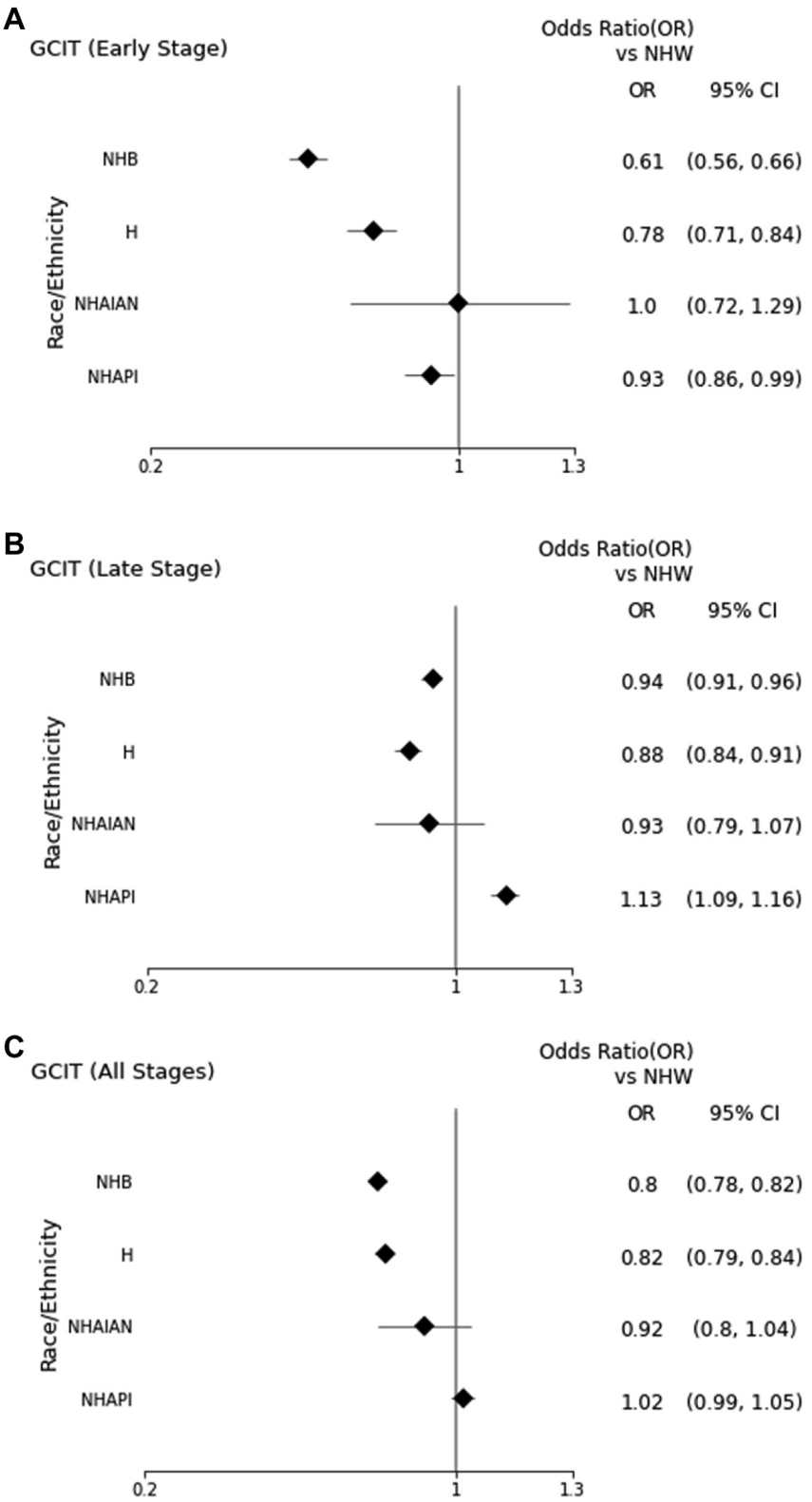

Figure 3. ORs for receiving GCIT among race/ethnic groups relative to NHW in (*A*) early stage, (*B*) late stage, and (*C*) all stages. GCIT, guideline-concordant initial treatment; NHW, non-Hispanic White.

both overall and for each individual race/ethnic group. The AUC improved from 82% to 84% when GCIT was included in the model. Similarly, the model accuracy and F1 score improved from 0.792 to 0.806 and 0.636 to 0.662, respectively, when the GCIT indicator was included in the model.

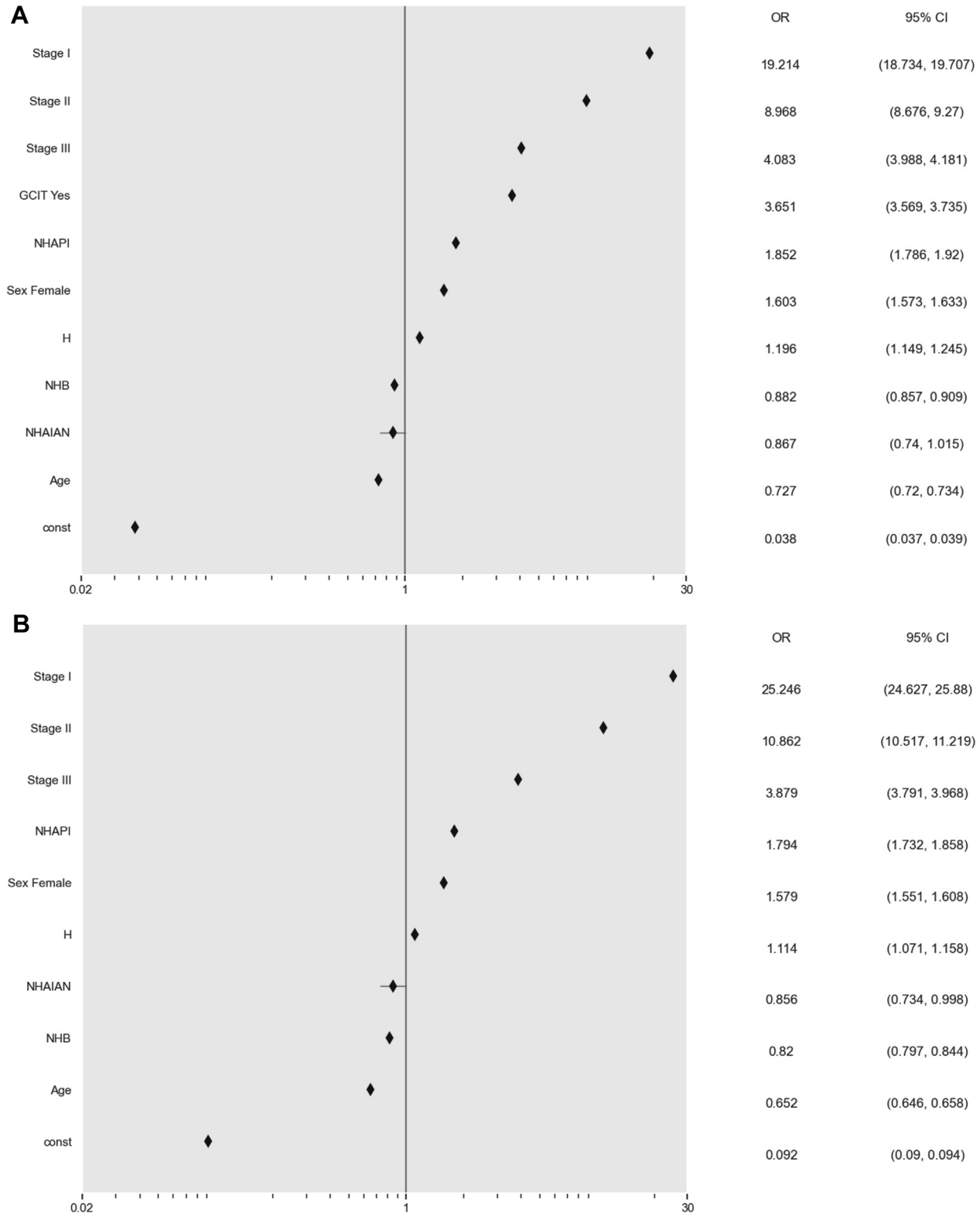

Figure 4. ORs for predictors in the two-year survival logistic model across all stages, (*A*) including GCIT as a predictor and (*B*) excluding GCIT as a predictor. GCIT, guideline-concordant initial treatment. A. Including GCIT as a predictor in the model. B. Excluding GCIT as a predictor from the model

Figure 5 reveals the Kaplan-Meier curves comparing the 5-year overall survival for NHW and each of the other racial/ethnic groups for early stage (stages I and II) stratified by GCIT. One plot is found for each comparison with NHW and the NHB (Fig. 5*A*), H (Fig. 5*B*), NHAIAN (Fig. 5*C*), and NHAPI (Fig. 5*D*) patients. Two curves are found

Table 2. Performance Metrics of 2-Year Survival Models Across Racial/Ethnic Groups

Race/ Ethnicity	Accuracy			F1 score			AUC		
	GCIT Excluded	+GCIT	% Improvement	GCIT Excluded	+GCIT	% Improvement	GCIT Excluded	+GCIT	% Improvement
NHW	0.794	0.808	1.8	0.647	0.672	3.9	0.825	0.845	2.4
NHB	0.795	0.809	1.8	0.562	0.591	5.2	0.807	0.828	2.6
H	0.795	0.810	1.9	0.617	0.653	5.8	0.812	0.833	2.6
NHAIAN	0.767	0.792	3.3	0.558	0.591	5.9	0.806	0.833	3.3
NHAPI	0.766	0.772	0.8	0.637	0.661	3.8	0.800	0.822	2.7
Overall	0.792	0.806	1.8	0.636	0.662	4.1	0.821	0.842	2.6

Note: The table highlights the effect of incorporating the GCIT indicator as a predictor on the accuracy, F1 score, and AUC of 2-year survival models across various racial/ethnic groups. The results reveal a consistent improvement in model performance with the inclusion of the GCIT indicator, underscoring its value as a predictive factor across all groups.

AUC, area under the curve; GCIT, guideline-concordant initial treatment; H, Hispanic all races; NHAIAN, non-Hispanic American Indian/Alaska Native; NHAPI, non-Hispanic Asian/Pacific Islander; NHB, non-Hispanic Black; NHW, non-Hispanic White.

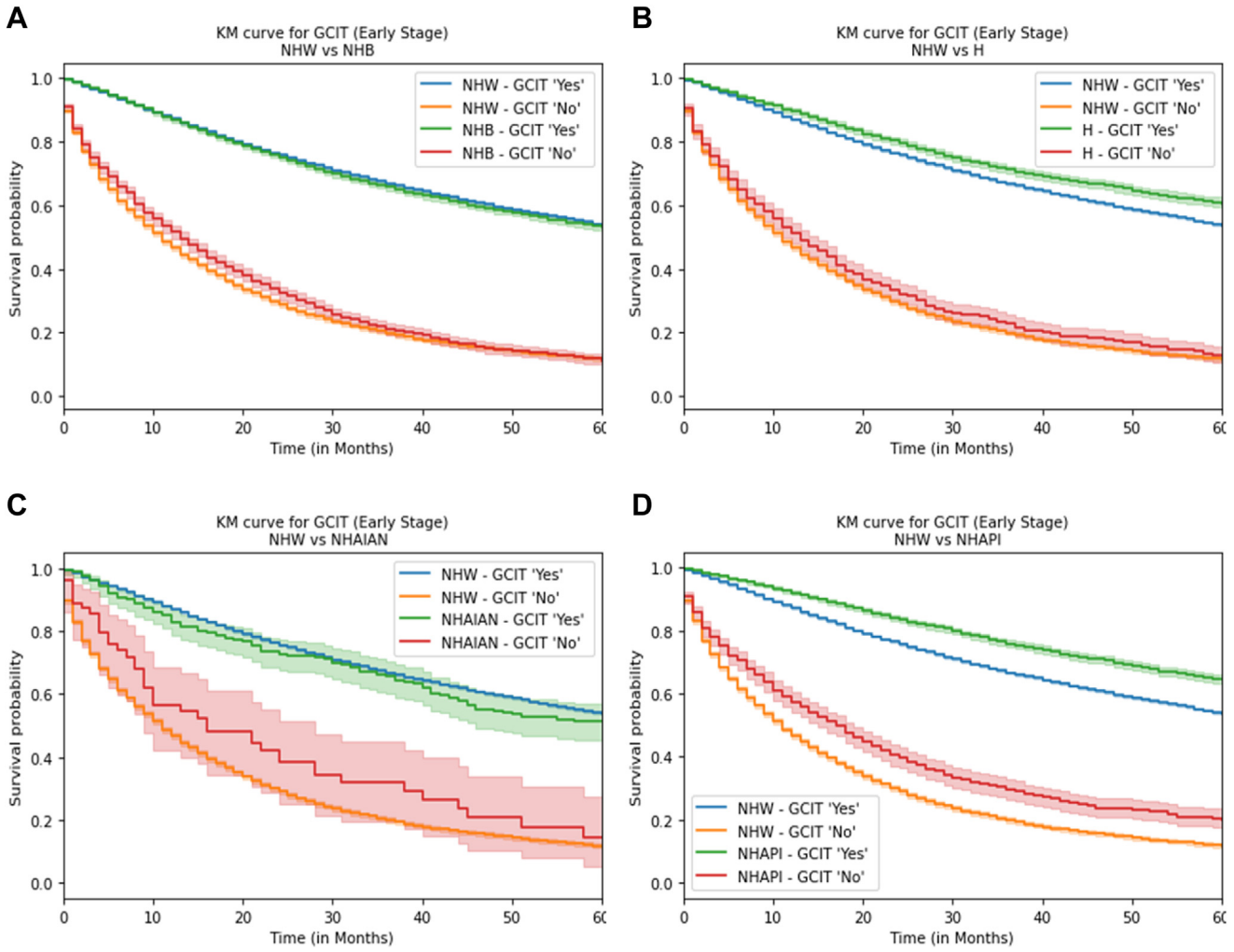

Figure 5. Kaplan-Meier overall survival curves comparing NHW with each racial/ethnic group: (*A*) NHB, (*B*) H, (*C*) NHAIAN, and (*D*) NHAPI, stratified by GCIT indicator for early stage. GCIT, guideline-concordant initial treatment; H, Hispanic all races; NHAIAN, non-Hispanic American Indian/Alaska Native; NHAPI, non-Hispanic Asian/Pacific Islander; NHB, non-Hispanic Black; NHW, non-Hispanic White.

for each race/ethnic group corresponding to those who received GCIT and those who did not.

Notably, patients known to receive GCIT had greater probability of survival as compared with patients who did not. In the early stage, NHW and NHB who received GCIT had almost identical survival curves (median survival: 67–69 mo, p value: 0.95). For patients who did not receive GCIT, NHB performed better than NHW (median survival: 13 versus 11 mo, p value: <0.005). For H patients known to receive GCIT, survival is better than for NHW (median survival: 88 mo, p value: <0.005) and similar to NHB who did not received GCIT (median survival: 13 mo, p value: 0.01). Relative to other races, NHAPI patients had the highest median survival for patients known to receive GCIT (median survival: 102 mo, p value: <0.005) and patients who did not (median survival: 17 mo, p value: <0.005). The survival curves for NHAIAN patients were not significantly different than NHW for GCIT received (median survival: 62 mo, p value: 0.28) and for GCIT not received (median survival: 16 mo, p value: 0.17).

Similar survival relationships as found in the early stage are observed in the late stage. Details on the late-stage median survival rates are provided in the Supplementary Data (Supplementary Fig. A2 and Supplementary Table A3).

Discussion

In this study, we evaluated the racial/ethnic disparities that exist particularly between NHW and NHB patients by considering the associated differences in the likelihood of receiving the standard treatment recommended by the NCCN and its relationship with survival for NSCLC. We derived a GCIT indicator using the disease stage and treatment variables available in the SEER data set to represent receipt of NCCN-recommended treatment. The GCIT indicator accounts for published treatment guidelines over time and satisfies the proportional hazards assumption needed in survival models, such as the Cox model.

Considering the overall results of this study, among all the racial/ethnic groups, most patients in the early stage received GCIT (83%–89% in stage I and 74%–82% in stage II), whereas in later or more advanced stages, patients were less likely to receive GCIT (52%–58% in stage III and 50%–57% in stage IV).

When evaluating the receipt of GCIT among different racial/ethnic groups, it is found that NHB patients exhibit a 20% lower likelihood of receiving GCIT compared with NHW patients across all stages. Notably, in the early stage where most patients undergo GCIT, NHB individuals are 39% less likely to receive such treatment. Even in the late stage, where GCIT is less often administered overall, NHB patients still have 6% lower odds compared with NHW patients.

The prediction performance of a logistic regression model trained to predict 2-year survival adjusted for age, stage, sex, and race improves profundly when the GCIT indicator is included as a predictor in the model. As found by the logistic regression model results, although AJCC staging remains the strongest predictor of survival overall, GCIT is also a strong predictor even when accounting for other factors. Across all races, patients who received GCIT have better survival than patients who did not. Specifically, for patients with early stage diagnosis, the median survival is 62 to 69 months for patients who received GCIT in contrast to 11 to 17 months for patients who did not receive GCIT. For late stages, the median survival for patients who received GCIT is 10 to 15 months versus only 2 to 3 months for patients who did not receive GCIT (Supplementary Table A3). It is also noteworthy that the survival rates are comparable between NHW and NHB known to have received GCIT intervention, in both early and late stages. This finding is encouraging and should motivate future research to identify the factors contributing to different treatment rates.

Note that as we underscore the significance of GCIT concerning survival outcomes, we also recognize the potential influence of external factors, beyond the scope of our study, which may contribute to survival differences. Therefore, these estimates related to GCIT and survival should not be considered causal. We cannot definitively say that receipt of GCIT improves survival by approximately 50 months. There are certain unobserved or confounding biases that may vary by race/ethnicity that affect these relationships. Nevertheless, our results suggest that observed racial/ethnic disparities between NHW and NHB and NHW and H patients with NSCLC dissipate accounting for GCIT status.

It is important to acknowledge the following data limitations in this study. First, the GCIT indicator is applied in a time-sensitive way to patients with defined stages of disease identified in the SEER database. The GCIT indicator was simplified as a binary indicator given the lack of variables that could inform a more granular classification which might reflect a more accurate burden of disease within stages. Moreover, missing data in SEER, either for existing variables or the lack of specific variables needed to capture treatment considerations, prevented the determination of whether the patient had received GCIT in some cases. Later stages (stages III and IV) had the highest percentage of patients for whom GCIT could not be ascertained. Although data limitation is a contributing factor, the specificity of the NCCN guidelines could also be a factor. Stage III, for example, was the only stage where certain substages

were not recommended to receive surgery per NCCN guidelines. For many substages of stage III cases, radiation was not recommended to be given alone and was to be given in combination with chemotherapy. For other stages, this was not the case. For the stage III cases that could not be marked as receiving GCIT, but did receive some form of treatment, 56% received radiation alone.

Furthermore, chemotherapy is frequently given outside of a hospital setting, and treatment in those settings is difficult to capture by cancer registries. It is possible that patients received chemotherapy outside of a hospital setting in combination with their radiation treatment, and it was not captured by cancer registries. For the SEER registry data, when the chemotherapy and radiation variables are recorded as Yes, then we are fairly certain that the patient received the given treatment. Nevertheless, when these variables are recorded as No/Unknown, it is possible that the patient had received the treatment, but the information was not available in their record. Furthermore, the SEER registry does not include targeted therapies and other factors that might affect treatment such as social determinants of health and other comorbidities, some of which have been found to vary across races.[14,17]

Despite all these limitations, deriving this GCIT indicator using SEER data is not without merit. The probabilities observed in this study for the receipt of GCIT are consistent with what has been observed in other studies that used other sources of data.[14,15] This implies that even when the data are not perfect, it is still able to capture important elements in GCIT.

Ethnic/racial disparities are a complex issue. From the analysis presented in this paper, there does not seem to be a single major source of disparity, but rather small disparities with an additive effect. NHB present at younger ages and later stages compared with NHW. Early stages had considerable better survival than late stages. Patients receiving GCIT have better survival than those who do not. Later stages are less likely to receive GCIT, and NHB, specifically, are less likely to receive GCIT than NHW. All these contribute to a 2-year survival observed probability of 36.5% for NHW versus 31.2% for NHB.

In conclusion, we were able to derive a GCIT indicator from staging and treatment variables (surgery, radiotherapy, and chemotherapy) from SEER data for NSCLC. The observed probabilities of the GCIT indicator across the racial/ethnic groupings are consistent with the likelihoods observed using other sources of data. Results reveal that the GCIT indicator is an effective way of correlating stage and recommended treatment over time. Racial disparities are apparent for both treatment and survival outcomes within this analysis, particularly for NHB when compared with NHW. Although efforts addressing social and structural determinants of health and health inequities remain paramount, our study reveals that greater delivery of GCIT for lung cancer could advance health equity and future research, either by implementing randomized clinical trials or leveraging natural experiments; however, to estimate the causal impact of GCIT on all-cause mortality remains warranted.

CRediT Authorship Contribution Statement

Eric Ababio Anyimadu: Conceptualization, Data curation, Methodology, Formal analysis, Investigation, Visualization, Writing - review & editing.

Jacklyn M. Engelbart: Conceptualization, Data curation, Methodology, Investigation, Writing - review & editing.

Jason Semprini: Conceptualization, Data curation, Methodology, Formal analysis, Investigation, Writing - review & editing.

Amanda Kahl: Conceptualization, Data curation, Methodology, Formal analysis, Investigation, Writing - review & editing.

Cameron Trentz: Data curation, Methodology, Investigation, Writing - review & editing.

John M. Buatti: Conceptualization, Data curation, Methodology, Formal analysis, Investigation, Visualization, Writing - review & editing.

Thomas L. Casavant: Conceptualization, Methodology, Investigation, Writing - review & editing.

Mary E. Charlton: Conceptualization, Data curation, Methodology, Investigation, Writing - review & editing.

Guadalupe Canahuate: Conceptualization, Data curation, Methodology, Formal analysis, Investigation, Visualization, Writing - review & editing.

Disclosure

The authors declare no conflict of interest.

Acknowledgments

This work received support from The University of Iowa Jumpstarting Tomorrow Program, the National Insitute of Health (NIH)/National Cancer Institute (NCI) contract number: HHSN261201800012I/HHSN26100001; NIH/NCI P30 CA086862) and NIH-NIDCD grants DC002842, DC012049, and DC017955.

Supplementary Data

Note: To access the supplementary material accompanying this article, visit the online version of the *JTO Clinical and Research Reports* at www.jtocrr.org and at https://doi.org/10.1016/j.jtocrr.2024.100747.

References

1. Torre LA, Siegel RL, Jemal A. Lung cancer statistics. *Adv Exp Med Biol*. 2016;893:1–19.
2. National Cancer Institute SEER Program. Cancer stat facts: lung and bronchus cancer. https://seer.cancer.gov/statfacts/html/lungb.html. Accessed April 10, 2023.
3. American Cancer Society. Key statistics for lung cancer: how common is lung cancer? http://www.cancer.org/cancer/lung-cancer/about/key-statistics.html. Accessed April 10, 2023.
4. Jeon J, Holford TR, Levy DT, et al. Smoking and lung cancer mortality in the United States from 2015 to 2065: a comparative modeling approach. *Ann Intern Med*. 2018;169:684–693.
5. Jani C, Marshall DC, Singh H, et al. Lung cancer mortality in Europe and the USA between 2000 and 2017: an observational analysis. *ERJ Open Res*. 2021;7:00311–2021.
6. Soneji S, Tanner NT, Silvestri GA, Lathan CS, Black W. Racial and ethnic disparities in early-stage lung cancer survival. *Chest*. 2017;152:587–597.
7. Jones CC, Mercaldo SF, Blume JD, et al. Racial disparities in lung cancer survival: the contribution of stage, treatment, and ancestry. *J Thorac Oncol*. 2018;13:1464–1473.
8. Driessen EJ, Aarts MJ, Bootsma GP, van Loon JG, Janssen-Heijnen ML. Trends in treatment and relative survival among non-small cell lung cancer patients in the Netherlands (1990-2014): disparities between younger and older patients. *Lung Cancer*. 2017;108:198–204.
9. Osuoha CA, Callahan KE, Ponce CP, Pinheiro PS. Disparities in lung cancer survival and receipt of surgical treatment. *Lung Cancer*. 2018;122:54–59.
10. DeSantis CE, Miller KD, Goding Sauer A, Jemal A, Siegel RL. Cancer statistics for African Americans, 2019. *CA Cancer J Clin*. 2019;69:211–233.
11. Tannenbaum SL, Koru-Sengul T, Zhao W, Miao F, Byrne MM. Survival disparities in non-small cell lung cancer by race, ethnicity, and socioeconomic status. *Cancer J*. 2014;20:237–245.
12. Harrison S, Judd J, Chin S, Ragin C. Disparities in lung cancer treatment. *Curr Oncol Rep*. 2022;24:241–248.
13. Ettinger DS, Wood DE, Aisner DL, et al. NCCN guidelines insights: non-small cell lung cancer, version 2. *J Natl Compr Canc Netw*. 2021;19:254–266.
14. Blom EF, Ten Haaf K, Arenberg DA, de Koning HJ. Disparities in receiving guideline-concordant treatment for lung cancer in the United States. *Ann Am Thorac Soc*. 2020;17:186–194.
15. Rosell R, Felip E. Role of multimodality treatment for lung cancer. *Semin Surg Oncol*. 2000;18:143–151.
16. Duggan MA, Anderson WF, Altekruse S, Penberthy L, Sherman ME. The Surveillance, Epidemiology, and End Results (SEER) program and pathology: toward strengthening the critical relationship. *Am J Surg Pathol*. 2016;40:e94–e102.
17. National Cancer Institute SEER Program. SEER is an authoritative source for cancer statistics in the United States. www.seer.cancer.gov. Accessed April 10, 2023.
18. National Cancer Institute. SEER Research Plus data description cases diagnosed in 1975-2018. https://seer.cancer.gov/data-software/documentation/seerstat/nov2020/TextData.FileDescription.pdf. Accessed April 10, 2023.