

## **Supplemental Text to Describe Methods:**

### *Clinical variables:*

Patient demographics, socioeconomic status, clinical characteristics, diagnostic method, treatment facility, geographic region, tumor, nodes, and metastases (TNM) staging, and treatment types were extracted from the NCDB. Demographic information included age, sex, race, and ethnicity. Socioeconomic status included insurance status and type of insurance, household income, level of education, and geographic classifications of metropolitan, urban, or rural. Household income (median household income for the patient's zip code) and level of education (median percentage of adults ages 25 or older without a high school degree in the patient's zip code) were based on 2016 American Community Survey data. Clinical characteristics included the Charlson/Deyo Comorbidity Score (0, 1, 2, or  $\geq 3$ ) and Model for End-stage Liver Disease (MELD) score, cirrhosis (Ishak fibrosis score 5-6). Diagnostic methods included histology/cytology and clinical diagnosis. Treatment facilities were categorized geographically in the US (Northeast, Midwest, South, West) and were classified as academic ( $> 500$  new cancer diagnoses annually and at least four postgraduate training programs), comprehensive community ( $> 500$  new cancer diagnoses annually), community (100 to 500 new cancer diagnoses annually), and integrated network (no minimum caseload, joint venture with multiple facilities providing integrated cancer care with at least one facility being a hospital and a CoC-accredited cancer program).

### *Statistical analysis:*

Study variables were summarized using standard descriptors such as frequency/percentage for categorical variables and median interquartile range (IQR) or mean standard deviation for continuous variables. The Pearson's chi-square test, Welch's t-test, and Wilcoxon-Mann-Whitney test were used for bivariate comparisons where appropriate. Univariate and multivariable logistic regression modeling was performed to determine factors associated with elevated serum AFP levels. The chained equation approach for multiple imputations was used prior to performing the logistic regression to account for missing data in the NCDB.<sup>1</sup> In order to provide a more intuitive interpretation of the temporal trend of increased AFP, we

estimated the percent of elevated AFP using marginal effects, also known as the average predicted probability, from a logistic regression model.<sup>2</sup> The interaction term between diagnosis year and TNM stage was included in the logistic regression model for pair-wise comparisons in the temporal trend of increased AFP level among different TNM stages. This computation was repeated using the overall sample with complete AFP data and subsample by race/ethnicity. All statistical analyses were performed using R statistical software (version 4.0.5; R Foundation, Vienna, Austria) with two-sided tests and a significance level of 0.05.

## References

1. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2011;30:377-99.
2. Lüdtke D (2021). sjPlot: Data Visualization for Statistics in Social Science. R package version 2.8.8, <https://CRAN.R-project.org/package=sjPlot>.

**Supplemental Table 1. Baseline patient characteristics by AFP**

Characteristic	AFP Level				Elevated AFP		
	Overall N = 133,542	AFP Missing N = 25,189	AFP Reported N = 108,353	p-value	Elevated AFP N = 67,853	Normal AFP N = 40,500	p-value
Age Mean (SD)	64 (11)	65 (11)	64 (10)	< 0.001	64 (10)	64 (10)	< 0.001
Sex				0.011			0.906
Male	101,938 (76%)	19,072 (76%)	82,866 (77%)		51,901 (76%)	30,965 (76%)	
Female	31,604 (24%)	6,117 (24%)	25,487 (24%)		15,952 (24%)	9,535 (24%)	
Race/Ethnicity				< 0.001			< 0.001
White	81,460 (63%)	15,437 (64%)	66,023 (63%)		39,299 (60%)	26,724 (68%)	
Hispanic	16,797 (13%)	3,354 (14%)	13,443 (13%)		8,518 (13%)	4,925 (13%)	
Black	20,442 (16%)	3,528 (15%)	16,914 (16%)		12,188 (19%)	4,726 (12%)	
Asian + Other	11,211 (8.6%)	1,990 (8.2%)	9,221 (8.7%)		6,040 (9.2%)	3,181 (8.0%)	
Insurance Status				< 0.001			< 0.001
Uninsured	7,481 (5.7%)	1,169 (4.8%)	6,312 (5.9%)		4,656 (7.0%)	1,656 (4.1%)	
Private	38,950 (30%)	7,046 (29%)	31,904 (30%)		19,589 (29%)	12,315 (31%)	
Medicaid/Medicare	82,003 (63%)	15,273 (63%)	66,730 (63%)		41,413 (62%)	25,317 (63%)	
Other	2,669 (2.0%)	797 (3.3%)	1,872 (1.8%)		1,163 (1.7%)	709 (1.8%)	
Median Income (quartiles)				0.941			< 0.001
Less than \$40,227	30,767 (25%)	5,854 (26%)	24,913 (25%)		16,604 (27%)	8,309 (23%)	
\$40,227 - \$50,353	27,905 (23%)	5,263 (23%)	22,642 (23%)		14,522 (23%)	8,120 (23%)	
\$50,354 - \$63,332	27,138 (22%)	5,154 (22%)	21,984 (22%)		13,887 (22%)	8,097 (23%)	
\$63,333 +	35,265 (29%)	6,716 (29%)	28,549 (29%)		17,232 (28%)	11,317 (32%)	
No High School Degree (quartiles)				< 0.001			< 0.001
17.6% or More	37,356 (31%)	7,364 (32%)	29,992 (31%)		19,880 (32%)	10,112 (28%)	
10.9%-17.5%	33,156 (27%)	6,416 (28%)	26,740 (27%)		17,176 (28%)	9,564 (27%)	
6.3%-10.8%	29,671 (25%)	5,415 (24%)	24,256 (25%)		15,104 (24%)	9,152 (26%)	
Less than 6.3%	21,124 (17%)	3,832 (17%)	17,292 (18%)		10,223 (16%)	7,069 (20%)	
Geographic Classification				0.218			0.102
Metropolitan	112,597 (87%)	21,252 (87%)	91,345 (87%)		57,465 (87%)	33,880 (87%)	
Urban	15,136 (12%)	2,830 (12%)	12,306 (12%)		7,661 (12%)	4,645 (12%)	
Rural	1,738 (1.3%)	355 (1.5%)	1,383 (1.3%)		840 (1.3%)	543 (1.4%)	
Facility Type				< 0.001			< 0.001
Academic	73,895 (56%)	11,726 (47%)	62,169 (58%)		36,466 (54%)	25,703 (64%)	
Community Cancer Program	6,622 (5.0%)	1,824 (7.3%)	4,798 (4.5%)		3,630 (5.4%)	1,168 (2.9%)	
Comprehensive Community Cancer Program	34,773 (26%)	8,262 (33%)	26,511 (25%)		18,208 (27%)	8,303 (21%)	
Integrated Network	16,659 (13%)	3,071 (12%)	13,588 (13%)		8,789 (13%)	4,799 (12%)	
US Region				< 0.001			< 0.001
Northeast	27,039 (21%)	4,959 (20%)	22,080 (21%)		13,299 (20%)	8,781 (22%)	
Midwest	26,162 (20%)	3,755 (15%)	22,407 (21%)		13,740 (21%)	8,667 (22%)	
<b>Characteristic</b>	<b>AFP Level</b>				<b>Elevated AFP</b>		

	<b>Overall</b> <i>N = 133,542</i>	<b>AFP Missing</b> <i>N = 25,189</i>	<b>AFP Reported</b> <i>N = 108,353</i>	<b>p-value</b>	<b>Elevated AFP</b> <i>N = 67,853</i>	<b>Normal AFP</b> <i>N = 40,500</i>	<b>p-value</b>
South	53,104 (40%)	11,780 (47%)	41,324 (39%)		26,767 (40%)	14,557 (36%)	
West	25,644 (19%)	4,389 (18%)	21,255 (20%)		13,287 (20%)	7,968 (20%)	
Charlson/Deyo Comorbidity Score				< 0.001			< 0.001
0 or 1	97,227 (73%)	19,896 (79%)	77,331 (71%)		49,524 (73%)	27,807 (69%)	
2	12,409 (9.3%)	2,034 (8.1%)	10,375 (9.6%)		6,005 (8.9%)	4,370 (11%)	
3 or More	23,906 (18%)	3,259 (13%)	20,647 (19%)		12,324 (18%)	8,323 (21%)	
Diagnosis Method				< 0.001			< 0.001
Cytology or Histology	77,198 (58%)	17,638 (70%)	59,560 (55%)		35,981 (53%)	23,579 (58%)	
Clinical Diagnosis	56,344 (42%)	7,551 (30%)	48,793 (45%)		31,872 (47%)	16,921 (42%)	
MELD Median [IQR]	12 [8, 20]	13 [8, 21]	12 [8, 19]	< 0.001	12 [9, 20]	11 [8, 17]	< 0.001
Cirrhosis				< 0.001			< 0.001
No	6,673 (21%)	847 (25%)	5,826 (21%)		3,074 (20%)	2,752 (22%)	
Yes	24,736 (79%)	2,536 (75%)	22,200 (79%)		12,597 (80%)	9,603 (78%)	
Tumor Size (cm) Median [IQR]	4.2 [2.5, 7.4]	4.2 [2.5, 7.4]	4.2 [2.5, 7.4]	0.994	5.0 [2.9, 8.4]	3.40 [2.2, 5.6]	< 0.001
TNM Stage				< 0.001			< 0.001
Stage 1	44,757 (38%)	8,304 (42%)	36,453 (38%)		17,741 (29%)	18,712 (51%)	
Stage 2	24,306 (21%)	3,463 (18%)	20,843 (22%)		11,921 (20%)	8,922 (24%)	
Stage 3	25,120 (22%)	3,554 (18%)	21,566 (22%)		16,320 (27%)	5,246 (14%)	
Stage 4	22,342 (19%)	4,249 (22%)	18,093 (19%)		14,448 (24%)	3,645 (10%)	
Treatment				< 0.001			< 0.001
No Treatment	28,147 (22%)	7,007 (29%)	21,140 (20%)		15,601 (24%)	5,539 (14%)	
Curative	33,921 (26%)	6,882 (29%)	27,039 (26%)		12,515 (19%)	14,524 (37%)	
Non-curative Treatment	67,665 (52%)	10,256 (43%)	57,409 (54%)		38,172 (58%)	19,237 (49%)	

MELD: Model for End-Stage Liver Disease; TNM: tumor, nodes, metastases

**Supplemental Table 2. Factors associated with elevated AFP**

Characteristic	Univariate		Multivariable	
	Odds Ratio	p-value	Odds Ratio	p-value
Age (per 10 years)	0.94 (0.93 to 0.95)	< 0.001	0.96 (0.95 to 0.98)	< 0.001
Female Sex (Ref: Male)	1.00 (0.97 to 1.03)	0.900	1.11 (1.08 to 1.15)	< 0.001
Race/Ethnicity (Ref: White)				
Black	1.74 (1.67 to 1.80)	< 0.001	1.59 (1.53 to 1.66)	< 0.001
Hispanic	1.17 (1.13 to 1.22)	< 0.001	1.11 (1.06 to 1.16)	< 0.001
Asian + Other	1.29 (1.22 to 1.34)	< 0.001	1.30 (1.24 to 1.36)	< 0.001
Insurance Status (Ref: No insurance)				
Private Insurance	0.57 (0.54 to 0.61)	< 0.001	0.81 (0.76 to 0.87)	< 0.001
Medicaid/Medicare	0.59 (0.55 to 0.62)	< 0.001	0.81 (0.76 to 0.86)	< 0.001
Other Insurance	0.59 (0.53 to 0.65)	< 0.001	0.83 (0.74 to 0.92)	0.001
Median Income (quartiles) (Ref: <\$40,227)				
\$40,227 - \$50,353	0.90 (0.86 to 0.92)	< 0.001	0.99 (0.95 to 1.03)	0.721
\$50,354 - \$63,332	0.86 (0.82 to 0.88)	< 0.001	0.98 (0.93 to 1.02)	0.352
\$63,333 +	0.77 (0.74 to 0.79)	< 0.001	0.95 (0.90 to 0.99)	0.025
No High School Degree (quartiles) (Ref: >17.5%)				
10.9%-17.5%	0.91 (0.88 to 0.94)	< 0.001	0.98 (0.95 to 1.02)	0.297
6.3%-10.8%	0.83 (0.81 to 0.86)	< 0.001	0.97 (0.93 to 1.02)	0.155
Less than 6.3%	0.74 (0.71 to 0.77)	< 0.001	0.92 (0.88 to 0.98)	0.002
Geographic Classification (Ref: Metropolitan)				
Urban	0.97 (0.94 to 1.01)	0.123	0.98 (0.94 to 1.03)	0.445
Rural	0.92 (0.83 to 1.03)	0.115	0.88 (0.79 to 1.00)	0.025
Facility Type (Ref: Academic)				
Community Cancer Program	2.17 (2.01 to 2.30)	< 0.001	1.74 (1.62 to 1.86)	< 0.001
Comprehensive Community Cancer Program	1.54 (1.50 to 1.59)	< 0.001	1.40 (1.37 to 1.46)	< 0.001
Integrated Network	1.29 (1.24 to 1.34)	< 0.001	1.25 (1.20 to 1.30)	< 0.001
US Region (Ref: Northeast)				
Midwest	1.05 (1.01 to 1.09)	0.019	1.02 (0.98 to 1.06)	0.267
South	1.21 (1.17 to 1.25)	< 0.001	1.07 (1.03 to 1.11)	< 0.001
West	1.10 (1.06 to 1.14)	< 0.001	1.02 (0.97 to 1.06)	0.368
Charlson/Deyo Comorbidity Score (Ref: 0-1)				
2	0.77 (0.74 to 0.80)	< 0.001	0.86 (0.82 to 0.89)	< 0.001
3 or More	0.83 (0.81 to 0.86)	< 0.001	0.85 (0.82 to 0.88)	< 0.001
MELD (per 10 unit)	1.17 (1.15 to 1.19)	< 0.001	1.09 (1.07 to 1.10)	< 0.001
Tumor Size (cm)	1.10 (1.09 to 1.10)	< 0.001	1.04 (1.03 to 1.04)	< 0.001
Year of Diagnosis	0.94 (0.93 to 0.94)	< 0.001	0.93 (0.93 to 0.94)	< 0.001
Diagnosis Method (Ref: Cytology or Histology)				
Clinical Diagnosis	1.23 (1.20 to 1.27)	< 0.001	1.45 (1.41 to 1.49)	< 0.001
Cirrhosis (Ref: No cirrhosis)	1.05 (1.03 to 1.09)	0.001	1.05 (1.01 to 1.08)	0.009
TNM Composite Stage (Ref: Stage 1)				
Stage 2	1.36 (1.32 to 1.41)	< 0.001	1.38 (1.33 to 1.42)	< 0.001
Stage 3	3.06 (2.96 to 3.17)	< 0.001	2.52 (2.43 to 2.62)	< 0.001
Stage 4	3.78 (3.62 to 3.91)	< 0.001	3.10 (2.97 to 3.23)	< 0.001

MELD: Model for End-Stage Liver Disease; TNM: tumor, nodes, metastases