



Primary clear cell adenocarcinoma of the lung: a national analysis

Arian Mansur¹, Zain Saleem², Alexandra L. Potter³, Camille Mathey-Andrews³, Priyanka Senthil³, Chi-Fu Jeffrey Yang³

¹Harvard Medical School, Boston, MA, USA; ²University of Ottawa, Ottawa, Ontario, Canada; ³Division of Thoracic Surgery, Department of Surgery, Massachusetts General Hospital, Boston, MA, USA

Contributions: (I) Conception and design: A Mansur, CJ Yang; (II) Administrative support: CJ Yang; (III) Provision of study materials or patients: AL Potter, CJ Yang; (IV) Collection and assembly of data: A Mansur, Z Saleem, AL Potter, C Mathey-Andrews, P Senthil; (V) Data analysis and interpretation: A Mansur, Z Saleem, AL Potter, C Mathey-Andrews, P Senthil; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Arian Mansur, BA. Harvard Medical School, 25 Shattuck Street, Boston, MA 02115, USA. Email: arianmansur@hms.harvard.edu.

Background: Primary clear cell adenocarcinoma of the lung (CCAL) is a rare form of lung cancer with poorly understood clinical features. We sought to investigate the clinicopathological characteristics and independent prognostic factors of primary CCAL.

Methods: Overall survival (OS) of patients with CCAL in the National Cancer Database (NCDB) from 2004 to 2017 was compared to lung adenocarcinoma using Kaplan-Meier analysis, multivariable Cox proportional hazards modeling, and propensity score matching. Independent prognostic indicators for patients with CCAL were determined using multivariable Cox proportional hazards analysis.

Results: A total of 1,396 CCAL and 462,360 lung adenocarcinoma patients were included in our analysis. When compared to patients diagnosed with lung adenocarcinoma, those diagnosed with CCAL were more likely to be younger, white, reside farther from a hospital, have higher Charlson/Deyo comorbidity condition (CDCC) scores, have private insurance, have T1, N0, M0 status. In unadjusted analysis, patients with CCAL had better survival than those with lung adenocarcinoma, although no significant differences in survival were found between the two groups with multivariable Cox proportional hazards and propensity score-matched analyses.

Conclusions: In this national analysis, we found that the clinicopathological characteristics of CCAL are distinct from those of lung adenocarcinoma, but CCAL is not itself an independent predictor of survival after multivariable adjustment or propensity score-matched analysis.

Keywords: Clear cell adenocarcinoma (CCA); lung cancer; adenocarcinoma

Submitted Jan 15, 2023. Accepted for publication Jul 10, 2023. Published online Aug 28, 2023.

doi: 10.21037/jtd-23-76

View this article at: <https://dx.doi.org/10.21037/jtd-23-76>

Introduction

Clear cell adenocarcinoma of the lung (CCAL) is a rare type of lung cancer characterized by an intracellular accumulation of glycogen resulting in a clear cytoplasm (1,2). CCA is often seen in the kidneys (3) and in the female genital tract (4), but primary tumors can be found in the lung (5). CCAL was first described in 1963 by Liebow and Castleman (6) and was later recognized as a distinct histologic subtype of lung cancer in 2004 in the

third edition of the World Health Organization (WHO) classification system for thoracic tumors (7). However, in 2011, the International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society international multidisciplinary classification of lung adenocarcinoma proposed to discontinue CCAL as an adenocarcinoma subtype because of the lack of data showing its clinical significance (8). The WHO classification of thoracic tumors then proceeded to discontinue CCAL as a distinct subtype and instead recognize it as a cytologic

feature (9). Since 2015, however, several studies have found that CCAL appears to have clinicopathological and prognostic features that are different from those of lung adenocarcinoma (10-12).

The purpose of this study is to further elucidate the clinicopathological and prognostic characteristics of patients with primary CCAL using the National Cancer Database (NCDB) in order improve the evidence that was lacking at the time that the determination to exclude CCAL as a separate histology type was made. We present this article in accordance with the STROBE reporting checklist (available at <https://jtd.amegroups.com/article/view/10.21037/jtd-23-76/rc>).

Methods

Data source

The NCDB is a clinical oncology database that is jointly managed by the American College of Surgeons Commission on Cancer and the American Cancer Society. It is estimated that the data provided by the NCDB includes approximately 72% of all newly diagnosed cases of lung cancer in the United States annually (13). The NCDB collects data from over 1,500 cancer centers in the United States and now contains over 30 million patient records. Variables used in the NCDB are available online <https://www.facs.org/quality-programs/cancer-programs/national-cancer-database/puf/> (14).

<https://www.facs.org/quality-programs/cancer-programs/national-cancer-database/puf/> (14).

Study design

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Institutional Review Board of Mass General Brigham (No. 2020P004110) and individual consent for this retrospective analysis was waived. All patients who were diagnosed with CCAL from 2004 to 2017 were identified for inclusion using the International Classification of Diseases for Oncology, 3rd edition (ICD-O-3) histology and topography codes. Using the ICD-O-3, we included tumor histology codes of 8140/3 and 8310/3, which corresponded to adenocarcinoma (not otherwise specified) and CCA (not otherwise specified), respectively. Data from the NCDB is directly abstracted from the pathology report, and the CCA cases in our study represent primary lung cancers that were classified as primary CCA by the pathologist. Years of diagnosis after 2015 were included given that ICD-O-3 is not aligned with the recent WHO classification of Tumors as detailed online <https://seer.cancer.gov/tools/solidtumor/clarifications.html> (15). However, we also performed a sensitivity analysis limited to 2014.

Only patients who were initially diagnosed with a single malignancy of lung adenocarcinoma or CCAL and who were diagnosed and treated at the reporting facility were included in the cohort. Further exclusion criteria included patients who had unknown or missing American Joint Committee on Cancer (AJCC) staging. The primary outcome was overall survival (OS), measured from time of diagnosis to time of death or last follow-up.

Statistical analysis

Patients were grouped according to histological subtype. Baseline characteristics and unadjusted outcomes were compared using the *t*-test or Wilcoxon Rank Sum test, when appropriate, for continuous variables and Pearson χ^2 test or Fisher's Exact test, when appropriate, for discrete variables. Median survival and 5-year survival of the histology groups were analyzed with the log-rank test and Kaplan-Meier product limit approach.

A Cox proportional hazards regression model was used to compare survival between patients of different histologic

Highlight box

Key findings

- In this national analysis, the clinicopathological characteristics of primary clear cell adenocarcinoma of the lung (CCAL) are distinct from those of lung adenocarcinoma, but CCAL is not itself an independent predictor of survival after multivariable adjustment or propensity score-matched analysis.

What is known and what is new?

- CCAL was removed as a distinct histologic subtype of lung cancer by the World Health Organization classification system for thoracic tumors in 2015 because of the lack of evidence on its clinical significance and instead recognized as a cytologic feature.
- Our study helps fill this research gap that led to a change in classification system.

What is the implication, and what should change now?

- This national analysis should be taken into consideration when revisiting the role of CCAL and its fate in the classification of lung tumors as it has implications for future research efforts and management of patients.

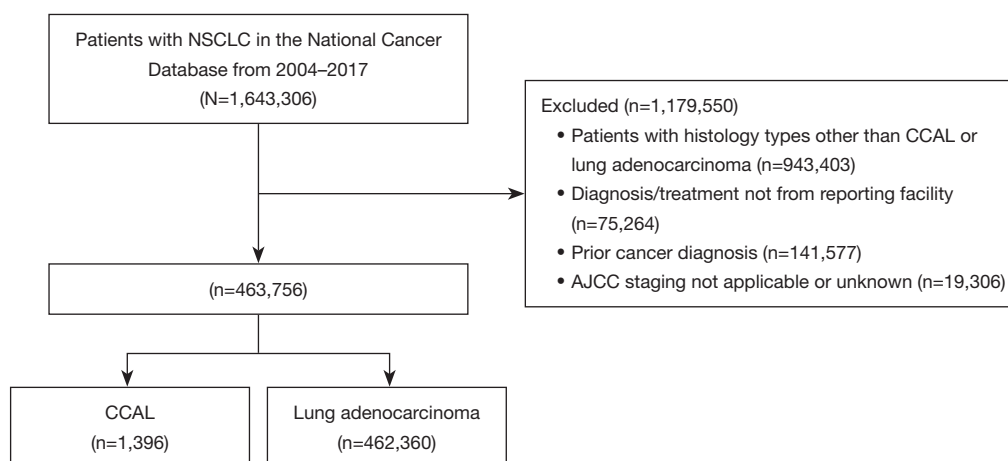


Figure 1 Flow diagram showing schema of study subject selection. NSCLC, non-small cell lung cancer; CCAL, clear cell adenocarcinoma of the lung; AJCC, American Joint Committee on Cancer.

types and to identify predictors of survival in patients with CCAL. Variables in this model included age, sex, race, year of diagnosis, median household income, educational attainment, insurance type, treatment facility type, distance from facility, Charlson/Deyo comorbidity condition (CDCC) score, clinical T status, clinical N status, clinical M status, AJCC stage, tumor size, tumor location, treatment with surgery, chemotherapy, and radiation.

Propensity scores were used to match patients in the CCAL and lung adenocarcinoma groups using similar methods as those previously described (16). Briefly, we first stratified patients into 2 groups (CCAL and lung adenocarcinoma), and then used a logistic regression model to calculate propensity scores. The following covariates were determined *a priori* and were used to calculate these scores: age, sex, race, CDCC score, median census-tract education and income levels, year of diagnosis, T, N, and M status, AJCC stage, tumor size, tumor location, insurance type, grade, distance from facility, tumor location, and treatment with surgery, chemotherapy, and radiation. Moreover, a greedy nearest neighbor algorithm without replacement and with a caliper of 0.01 was used, followed by identifying the most appropriately matched pairs. After matching, the standardized differences were used to evaluate the balance of the match, and Kaplan-Meier analysis was performed to evaluate OS of both groups.

Results

A total of 463,756 patients met study criteria (Figure 1).

Of this cohort, 1,396 patients (0.3%) were diagnosed with CCAL, and 462,360 patients (99.7%) patients were diagnosed with lung adenocarcinoma. The baseline clinicopathological and demographic characteristics are summarized in Table 1. Patients with CCAL were more likely to be younger, white, reside farther from a hospital, have higher CDCC scores, private insurance, earlier year diagnosis, T1, N0, M0 status, more poorly differentiated tumors, earlier stage disease, and were more likely to undergo surgery and less likely to undergo chemotherapy and radiation than patients with lung adenocarcinoma. Similar findings were seen when restricting to cases diagnosed before 2015 (Table S1).

OS stratified by histology type was assessed. The median follow up was 13.8 months (IQR, 3.8–38.0). In unadjusted analysis, CCAL was associated with better survival than lung adenocarcinoma [5-year survival 36% (95% CI: 33–38%) versus 24% (95% CI: 24–24%), log-rank, $P < 0.001$, Figure 2]. In multivariable analysis, there was no significant difference in survival between CCAL and lung adenocarcinoma (adjusted hazard ratio 1.06; 95% CI: 0.93–1.20, $P = 0.38$) (Table 2). Similar findings were seen when restricting to cases diagnosed before 2015 (Figure S1, Table S2).

Propensity-score matching was used to create 2 groups of 520 patients each who had CCAL or lung adenocarcinoma that were well-matched with regard to baseline characteristics (Table 3). All standardized mean differences were less than or equal to 10.6%. There was no significant difference in survival between CCAL and lung

Table 1 Clinicopathologic and demographic characteristics for patients, stratified by lung adenocarcinoma versus clear cell adenocarcinoma of the lung

Patient characteristic	Lung adenocarcinoma (N=462,360)	CCAL (N=1,396)	P value
Age at diagnosis [median (IQR)], years	67.0 (59.0, 75.0)	65.0 (57.0, 72.0)	<0.01
Sex, n (%)			0.62
Male	222,474 (48.1)	681 (48.8)	
Female	239,886 (51.9)	715 (51.2)	
Race, n (%)			<0.01
White	384,206 (83.1)	1,221 (87.5)	
Black	55,543 (12.0)	130 (9.3)	
Other	19,197 (4.2)	31 (2.2)	
Unknown	3,414 (0.7)	14 (1.0)	
Education, n (%)			0.21
17.6%	95,764 (20.7)	269 (19.3)	
10.9–17.5%	121,037 (26.2)	388 (27.8)	
6.3–10.8%	121,931 (26.4)	383 (27.4)	
<6.3%	93,318 (20.2)	264 (18.9)	
Unknown	30,310 (6.6)	92 (6.6)	
CDCC score, n (%)			<0.01
0	279,613 (60.5)	734 (52.6)	
1	121,550 (26.3)	442 (31.7)	
2	41,885 (9.1)	157 (11.2)	
3+	19,312 (4.2)	63 (4.5)	
Year of diagnosis [median (IQR)]	2012 (2008, 2015)	2009 (2006, 2013)	<0.01
Distance from facility [median (IQR)], miles	9.2 (4.1, 22.3)	10.2 (4.3, 24.4)	<0.01
Tumor size [median (IQR)], cm	32.0 (20.0, 50.0)	32.0 (20.0, 54.0)	0.06
Tumor location, n (%)			<0.01
Main bronchus	142,947 (30.9)	475 (34.0)	
RUL	20,012 (4.3)	53 (3.8)	
RML	63,528 (13.7)	187 (13.4)	
RLL	102,638 (22.2)	367 (26.3)	
LUL	50,757 (11.0)	147 (10.5)	
LLL	18,839 (4.1)	46 (3.3)	
Unknown	63,639 (13.8)	121 (8.7)	
Insurance status, n (%)			<0.01
Uninsured	17,119 (3.7)	44 (3.2)	
Private	141,959 (30.7)	502 (36.0)	

Table 1 (continued)

Table 1 (continued)

Patient characteristic	Lung adenocarcinoma (N=462,360)	CCAL (N=1,396)	P value
Medicaid	35,089 (7.6)	99 (7.1)	
Medicare	252,466 (54.6)	712 (51.0)	
Other	6,780 (1.5)	17 (1.2)	
Unknown	8,947 (1.9)	22 (1.6)	
Facility type, n (%)			0.45
Community cancer program	32,515 (7.0)	87 (6.2)	
Comprehensive community	188,316 (40.7)	575 (41.2)	
Academic/research program	146,444 (31.7)	460 (33.0)	
Integrated network cancer program	91,615 (19.8)	264 (18.9)	
Unknown	3,470 (0.8)	10 (0.7)	
Median household income, n (%)			0.97
First quartile	90,809 (19.6)	273 (19.6)	
Second quartile	100,587 (21.8)	302 (21.6)	
Third quartile	101,577 (22.0)	302 (21.6)	
Fourth quartile	138,234 (29.9)	426 (30.5)	
Unknown	31,153 (6.7)	93 (6.7)	
Grade/differentiation, n (%)			<0.01
Well differentiated; differentiated, NOS	29,520 (6.4)	52 (3.7)	
Moderately differentiated	92,317 (20.0)	341 (24.4)	
Poorly differentiated; dedifferentiated	124,144 (26.9)	536 (38.4)	
Undifferentiated; anaplastic	2,205 (0.5)	24 (1.7)	
Cell type not determined	214,174 (46.3)	443 (31.7)	
NCDB Analytic Stage Group, n (%)			<0.01
Stage I	106,659 (23.1)	516 (37.0)	
Stage II	31,030 (6.7)	176 (12.6)	
Stage III	86,059 (18.6)	265 (19.0)	
Stage IV	238,612 (51.6)	439 (31.4)	
Clinical T status, n (%)			<0.01
T1a	134,314 (29.0)	488 (35.0)	
T1b	90,897 (19.7)	310 (22.2)	
T1c	54,993 (11.9)	182 (13.0)	
T2a	117,529 (25.4)	299 (21.4)	
Unknown	64,627 (14.0)	117 (8.4)	

Table 1 (continued)

Table 1 (continued)

Patient characteristic	Lung adenocarcinoma (N=462,360)	CCAL (N=1,396)	P value
Clinical N status, n (%)			<0.01
N0	163,585 (35.4)	596 (42.7)	
N1	33,534 (7.3)	92 (6.6)	
N2	130,584 (28.2)	267 (19.1)	
N3	62,407 (13.5)	130 (9.3)	
Unknown	72,250 (15.6)	311 (22.3)	
Clinical M status, n (%)			<.01
M0	219,797 (47.5)	920 (65.9)	
M1	228,244 (49.4)	409 (29.3)	
Unknown	14,319 (3.1)	67 (4.8)	
Treatment, n (%)			<0.01
Surgery	125,133 (27.1)	826 (59.2)	
Chemotherapy	222,452 (48.1)	602 (43.1)	
Radiation	108,132 (23.4)	262 (18.8)	

CCAL, clear cell adenocarcinoma of the lung; IQR, interquartile range; CDCC, Charlson/Deyo comorbidity condition; RUL, right upper lobe; RML, right middle lobe; RLL, right lower lobe; LUL, left upper lobe; LLL, left lower lobe; NOS, not otherwise specified; NCDB, national cancer database.

adenocarcinoma [5-year survival 47% (95% CI: 42–51%) versus 47% (95% CI: 43–51%), log-rank, $P=0.95$, Figure 3]. Similar findings were seen when restricting to cases diagnosed before 2015 (Figure S2).

Multivariable Cox proportional hazards analysis was performed to identify predictors of OS in patients with CCAL. Analysis showed that sex, CDCC score, M status, AJCC stage, and treatment with surgery were independent predictors of survival for patients with CCAL (Table 4). Similar findings were seen when restricting to cases diagnosed before 2015 (Table S3).

Discussion

In this study, we used the NCDB to evaluate the clinicopathological characteristics and independent prognostic factors associated with primary CCAL, while comparing the OS of patients with CCAL to those with lung adenocarcinoma. Although CCAL patients had better survival than patients with lung adenocarcinoma in unadjusted analysis, no significant differences in survival were found following both multivariable and propensity score-matched analysis. There could be several reasons that

could explain the differences between the unadjusted and adjusted analyses. For instance, half of the patients (51.6%) in the lung adenocarcinoma group were diagnosed with stage IV, while in the CCAL group, only a third (31.4%) were diagnosed with stage IV. Moreover, over half of the patients (59.3%) in the CCAL group received surgery while only about a quarter (27.2%) in the lung adenocarcinoma group received surgery.

Several studies have compared prognosis between CCAL and general lung adenocarcinoma with conflicting results. Previous smaller scale studies reported CCAL having either similar prognosis (16) or worse survival (10) when compared to other lung adenocarcinomas. More recently, Ke *et al.* (11) used the Surveillance, Epidemiology, and End Results (SEER) database to compare OS between 1,203 patients with CCAL to 266,652 patients with general lung adenocarcinoma and found that patients with CCAL had better survival, both in unadjusted and adjusted analyses. However, Komiya *et al.* (12) found that CCAL histology was an independent predictor for survival in unadjusted but not in adjusted analysis when compared to patients with general lung adenocarcinomas (1,227 CCAL vs. 233,154 lung adenocarcinoma). Similar to Komiya *et al.*, our

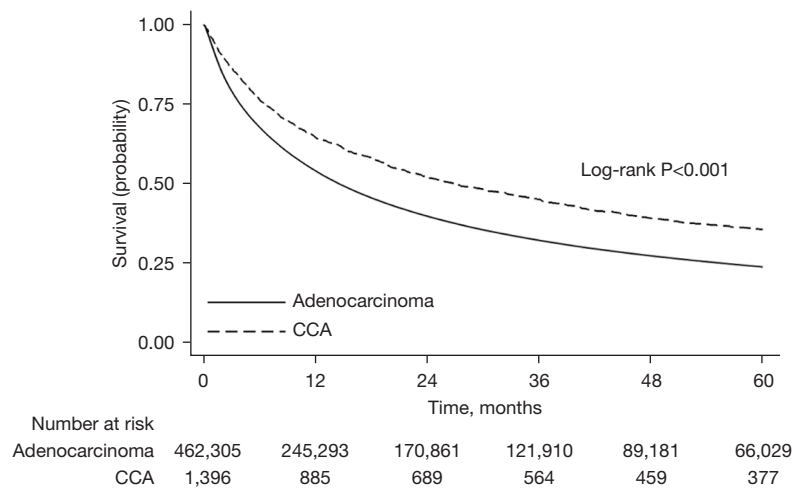


Figure 2 Kaplan-Meier analysis of overall survival for patients with CCA versus patients with lung adenocarcinoma. CCA, clear cell adenocarcinoma.

Table 2 Multivariable Cox proportional hazards analyses for patients, stratified by lung adenocarcinoma versus clear cell adenocarcinoma of the lung

Variables	Hazard ratio	95% CI	P value
Age (per year)	1.01	1.01, 1.01	<0.01
Female vs. male	0.83	0.81, 0.84	<0.01
Race (ref = white)			
Black	0.92	0.90, 0.95	<0.01
Native American	0.96	0.82, 1.11	0.57
Asian	0.75	0.71, 0.79	<0.01
Year of diagnosis (per year)	0.96	0.96, 0.97	<0.01
Median household income (ref = quartile 1)			
Second quartile	0.98	0.95, 1.00	0.07
Third quartile	0.94	0.92, 0.97	<0.01
Forth quartile	0.89	0.87, 0.92	<0.01
Insurance type (ref = uninsured)			
Private	0.84	0.80, 0.88	<0.01
Medicaid	0.96	0.91, 1.01	0.15
Medicare	0.93	0.89, 0.97	<0.01
Other	0.90	0.83, 0.98	0.02
Education (ref =17.6%)			
10.9–17.5%	1.03	1.00, 1.05	0.02
6.3–10.8%	1.02	0.99, 1.05	0.12
<6.3%	1.01	0.98, 1.05	0.37
Distance from facility (per mile)	1.00	1.00, 1.00	0.08

Table 2 (continued)

Table 2 (continued)

Variables	Hazard ratio	95% CI	P value
Facility type (ref = community cancer program)			<0.01
Comprehensive community clinic	0.95	0.92, 0.97	
Academic/research program	0.84	0.82, 0.87	
Integrated network cancer program	0.93	0.90, 0.96	
CDCC score (ref =0)			<0.01
1	1.16	1.13, 1.18	
2	1.32	1.28, 1.35	
3+	1.56	1.50, 1.63	
Tumor size (per cm)	1.00	1.00, 1.00	<0.01
Tumor location (ref = main bronchus)			
RUL	1.07	1.03, 1.11	<0.01
RML	1.11	1.09, 1.15	<0.01
RLL	1.02	1.00, 1.04	0.03
LUL	1.08	1.05, 1.10	<0.01
LLL	1.14	1.10, 1.19	<0.01
Grade/differentiation (ref = well differentiated)			<0.01
Moderately differentiated	1.23	1.19, 1.27	
Poorly differentiated; dedifferentiated	1.42	1.37, 1.47	
Undifferentiated; anaplastic	1.38	1.27, 1.50	
Clinical T status (ref = T1a)			<0.01
T1b	1.15	1.13, 1.18	
T1c	1.24	1.20, 1.27	
T2a	1.34	1.31, 1.38	
Clinical N status (ref = N0)			<0.01
N1	1.10	1.07, 1.13	
N2	1.21	1.17, 1.25	
N3	1.10	1.04, 1.11	
Clinical M status (ref = M0)			<0.01
M1	0.99	0.99, 0.99	
NCDB Analytic Stage Group (ref = stage I)			<0.01
Stage II	1.87	1.81, 1.94	
Stage III	2.37	2.28, 2.45	
Stage IV	4.14	3.99, 4.29	
Treatment			<0.01
Surgery	0.50	0.49, 0.52	
Chemotherapy	0.50	0.50, 0.51	
Radiation	0.92	0.90, 0.94	
CCAL vs. lung adenocarcinoma	1.06	0.93, 1.20	0.38

CI, confidence interval; CDCC, Charlson/Deyo comorbidity condition; RUL, right upper lobe; RML, right middle lobe; RLL, right lower lobe; LUL; left upper lobe; LLL, left lower lobe; NCDB, national cancer database; CCAL, clear cell adenocarcinoma of the lung.

Table 3 Propensity-matched baseline characteristics stratified by stratified by lung adenocarcinoma versus clear cell adenocarcinoma of the lung

Patient characteristic	Lung adenocarcinoma (N=520)	CCAL (N=520)	Absolute standardized difference (%)	P value
Age at diagnosis [median (IQR)], years	66.0 (57.0, 73.0)	65.0 (57.0, 73.0)	2.9	0.60
Sex, n (%)				0.46
Male	250 (48.1)	238 (45.8)	4.6	
Female	270 (51.9)	282 (54.2)	4.6	
Race				0.81
White, n (%)	463 (89.0)	464 (89.2)	0.6	
Black, n (%)	52 (10.0)	49 (9.4)	1.9	
Other (%)	<10	<10	0.0	
Education, n (%)				0.73
17.6%	113 (21.7)	105 (20.2)	3.9	
10.9–17.5%	160 (30.8)	156 (30.0)	1.7	
6.3–10.8%	138 (26.5)	154 (29.6)	6.8	
<6.3%	109 (21.0)	105 (20.2)	1.9	
CDCC score, n (%)				0.55
0	218 (41.9)	236 (45.4)	7.0	
1	224 (43.1)	206 (39.6)	7.4	
2	65 (12.5)	61 (11.7)	2.5	
3+	13 (2.5)	17 (3.3)	3.8	
Year of diagnosis [median (IQR)]	2010 [2008, 2013]	2010 [2007, 2013]	2.0	0.77
Distance from facility [median (IQR)], miles	11.4 (4.5, 23.8)	10.6 (4.8, 26.0)	5.6	0.44
Tumor size [median (IQR)], cm	30.0 (18.0, 47.5)	31.0 (20.0, 50.0)	10.6	0.08
Tumor location, n (%)				0.69
Main bronchus	179 (34.4)	200 (38.5)	8.4	
RUL	19 (3.7)	20 (3.8)	0.9	
RML	77 (14.8)	71 (13.7)	3.2	
RLL	179 (34.4)	159 (30.6)	8.5	
LUL	51 (9.8)	57 (11.0)	3.6	
LLL	15 (2.9)	13 (2.5)	2.3	
Insurance status				0.87
Uninsured, n (%)	22 (4.2)	18 (3.5)	4.4	
Private, n (%)	193 (37.1)	187 (36.0)	2.4	
Medicaid, n (%)	30 (5.8)	37 (7.1)	5.3	
Medicare, n (%)	266 (51.2)	269 (51.7)	1.2	
Other (%)	<10	<10	0.0	
Facility type, n (%)				0.68
Community cancer program	22 (4.2)	28 (5.4)	4.7	

Table 3 (continued)

Table 3 (continued)

Patient characteristic	Lung adenocarcinoma (N=520)	CCAL (N=520)	Absolute standardized difference (%)	P value
Comprehensive community	238 (45.8)	223 (42.9)	5.8	
Academic/research program	175 (33.7)	177 (34.0)	0.8	
Integrated network cancer program	85 (16.3)	92 (17.7)	3.4	
Median household income, n (%)				0.72
First quartile	105 (20.2)	101 (19.4)	2.0	
Second quartile	127 (24.4)	131 (25.2)	1.8	
Third quartile	130 (25.0)	117 (22.5)	6.0	
Fourth quartile	158 (30.4)	171 (32.9)	5.3	
Grade/differentiation, n (%)				0.79
Well differentiated; differentiated, NOS	23 (4.4)	30 (5.8)	4.7	
Moderately differentiated	189 (36.3)	188 (36.2)	0.4	
Poorly differentiated; dedifferentiated	294 (56.5)	287 (55.2)	2.7	
Undifferentiated; anaplastic	14 (2.7)	15 (2.9)	1.4	
NCDB Analytic Stage Group, n (%)				0.99
Stage I	265 (51.0)	270 (51.9)	2.0	
Stage II	88 (16.9)	86 (16.5)	1.1	
Stage III	101 (19.4)	100 (19.2)	0.5	
Stage IV	66 (12.7)	64 (12.3)	0.9	
Clinical T status, n (%)				0.98
T1a	208 (40.0)	212 (40.8)	1.6	
T1b	145 (27.9)	144 (27.7)	0.4	
T1c	88 (16.9)	83 (16.0)	2.7	
T2a	79 (15.2)	81 (15.6)	1.0	
Clinical N status, n (%)				0.99
N0	365 (70.2)	369 (71.0)	0.0	
N1	46 (8.8)	46 (8.8)	0.0	
N2	90 (17.3)	87 (16.7)	1.4	
N3	19 (3.7)	18 (3.5)	0.7	
Clinical M status, n (%)				0.85
M0	460 (88.5)	458 (88.1)	0.1	
M1	60 (11.5)	62 (11.9)	0.1	
Treatment, n (%)				
Surgery	424 (81.5)	425 (81.7)	0.4	0.94
Chemotherapy	206 (39.6)	196 (37.7)	3.9	0.52
Radiation	107 (20.6)	107 (20.6)	0.0	1.00

CCAL, clear cell adenocarcinoma of the lung; IQR, interquartile range; CDCC, Charlson/Deyo comorbidity condition; RUL, right upper lobe; RML, right middle lobe; RLL, right lower lobe; LUL, left upper lobe; LLL, left lower lobe; NOS, not otherwise specified; NCDB, national cancer database.

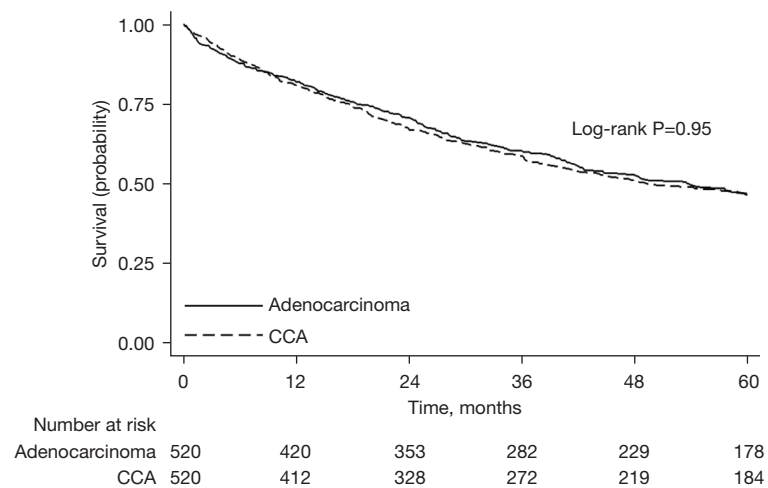


Figure 3 Kaplan-Meier analysis of overall survival for propensity score-matched patients, stratified by lung adenocarcinoma versus clear cell adenocarcinoma of the lung. CCA, clear cell adenocarcinoma.

Table 4 Independent predictors of overall survival after Cox proportional hazards adjustment for patients with clear cell adenocarcinoma of the lung

Variables	Hazard ratio	95% CI	P value
Age (per year)	1.01	0.99, 1.03	0.16
Female vs. male	0.69	0.51, 0.93	0.02
Race (ref = white)			
Black	1.24	0.72, 2.14	0.49
Native American	0.00	0.00, N/A	1.00
Asian	0.95	0.19, 4.82	0.95
Year of diagnosis (per year)	0.97	0.92, 1.02	0.19
Median household income (ref = quartile 1)			
Second quartile	0.72	0.44, 1.16	0.18
Third quartile	0.71	0.43, 1.18	0.19
Forth quartile	1.28	0.71, 2.32	0.41
Insurance type (ref = uninsured)			
Private	0.42	0.13, 1.31	0.14
Medicaid	0.55	0.16, 1.81	0.32
Medicare	0.53	0.17, 1.66	0.28
Other	0.71	0.16, 3.14	0.65
Education (ref = 17.6%)			
10.9–17.5%	1.36	0.89, 2.09	0.16
6.3–10.8%	0.92	0.57, 1.48	0.73
<6.3%	0.63	0.35, 1.15	0.14
Distance from facility (per mile)	1.00	0.99, 1.00	0.67
Facility type (ref = community cancer program)			

Table 4 (continued)

Table 4 (continued)

Variables	Hazard ratio	95% CI	P value
Comprehensive community clinic	1.53	0.73, 3.19	0.26
Academic/research program	1.27	0.59, 2.71	0.54
Integrated network cancer program	1.57	0.71, 3.46	0.27
CDCC score (ref =0)			
1	1.10	0.78, 1.53	0.59
2	1.85	1.19, 2.86	<0.01
3+	0.90	0.40, 1.99	0.79
Tumor size (per cm)	1.01	1.00, 1.01	0.24
Tumor location (ref = main bronchus)			
RUL	2.03	0.91, 4.55	0.08
RML	1.31	0.83, 2.06	0.25
RLL	1.06	0.73, 1.55	0.75
LUL	1.54	0.95, 2.49	0.07
LLL	0.99	0.47, 2.06	0.97
Grade/differentiation (ref = well differentiated)			
Moderately differentiated	1.14	0.59, 2.23	0.69
Poorly differentiated; dedifferentiated	1.06	0.56, 1.99	0.85
Undifferentiated; anaplastic	0.90	0.29, 2.82	0.86
Clinical T status (ref = T1a)			
T1b	1.22	0.81, 1.84	0.34
T1c	1.59	0.90, 2.82	0.11
T2a	1.35	0.66, 2.74	0.41
Clinical N status (ref = N0)			
N1	0.90	0.54, 1.51	0.70
N2	0.76	0.35, 1.64	0.5
N3	0.99	0.57, 1.74	0.99
Clinical M status (ref = M0)			<0.01
M1	0.98	0.97, 0.99	
NCDB Analytic Stage Group (ref = stage I)			
Stage II	1.20	0.70, 2.06	0.51
Stage III	3.04	1.82, 5.08	<0.01
Stage IV	7.86	3.98, 15.52	<0.01
Treatment			
Surgery	0.51	0.29, 0.90	0.02
Chemotherapy	0.67	0.44, 1.03	0.07
Radiation	0.94	0.60, 1.46	0.78

CI, confidence interval; CDCC, Charlson/Deyo comorbidity condition; RUL, right upper lobe; RML, right middle lobe; RLL, right lower lobe; LUL; left upper lobe; LLL, left lower lobe; NCDB, national cancer database.

findings indicate that CCA histology is not an independent prognostic indicator after performing multivariable Cox proportional hazards and propensity-score matching analyses but was associated with improved survival in unadjusted analyses. These differences in findings may be attributed to the fact that our study used a different national database and additionally adjusted for prognostic variables such as sex, comorbidities, insurance status, education, income, Tumor, Node, Metastasis (TNM) staging, and use of chemotherapy, which were not controlled for in the past studies.

This study has several important limitations. First, it is a retrospective cohort analysis and there is always a chance of inherent unmeasured confounding present in the study. Second, in a rare tumor type of lung cancer that is often considered a cytologic feature, the NCDB data may have cases where tumors should have been categorized as CCAL but were misclassified. Third, the NCDB does not have performance status and pulmonary function data. Fourth, the two groups analyzed in our study had unequal sample sizes. Fifth, we note that the study period analyzed was from 2004–2017, during which AJCC staging guidelines changed; the cases used in this study had their stage classified by the AJCC guideline that was available at the time of their diagnosis and was not reclassified according to the 8th edition.

Conclusions

In conclusion, in this national analysis, CCAL was found to be associated with different clinicopathological characteristics, more early-stage disease, and better survival when compared to lung adenocarcinoma in unadjusted analysis. However, no significant differences in survival between the two groups were found following both multivariable and propensity score-matched analysis. Given that CCAL was found to have distinct clinicopathological features from lung adenocarcinoma, more efforts, especially prospective, multi-institutional studies, should be made to further elucidate the diagnostic and prognostic significance of CCAL in order to guide further management of the disease.

Acknowledgments

The NCDB states: “The data used in this study are derived from a de-identified NCDB file. The American College of Surgeons and the Commission on Cancer have not verified and are not responsible for the analytic or statistical

methodology employed, or the conclusions drawn from these data by the investigator.”

Funding: None.

Footnote

Reporting Checklist: The authors have completed the STROBE reporting checklist. Available at <https://jtd.amegroups.com/article/view/10.21037/jtd-23-76/rc>

Peer Review File: Available at <https://jtd.amegroups.com/article/view/10.21037/jtd-23-76/prf>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://jtd.amegroups.com/article/view/10.21037/jtd-23-76/coif>). CJY serves as an unpaid editorial board member of *Journal of Thoracic Disease* from February 2023 to January 2025. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Institutional Review Board of Mass General Brigham (No. 2020P004110) and individual consent for this retrospective analysis was waived.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Gaffey MJ, Mills SE, Ritter JH. Clear cell tumors of the lower respiratory tract. *Semin Diagn Pathol* 1997;14:222-32.
2. Guo Y, Shrestha A, Maskey N, et al. Recent Trends in the Incidence of Clear Cell Adenocarcinoma and Survival Outcomes: A SEER Analysis. *Front Endocrinol (Lausanne)*

- 2022;13:762589.
3. Jonasch E, Walker CL, Rathmell WK. Clear cell renal cell carcinoma ontogeny and mechanisms of lethality. *Nat Rev Nephrol* 2021;17:245-61.
 4. Offman SL, Longacre TA. Clear cell carcinoma of the female genital tract (not everything is as clear as it seems). *Adv Anat Pathol* 2012;19:296-312.
 5. Shen L, Lin J, Ren Z, et al. Clear cell tumor of the lung could be aggressive: a case report and review of the literature. *J Cardiothorac Surg* 2020;15:177.
 6. Liebow AA, Castleman B. Benign "clear cell tumors" of the lung. *Am J Pathol* 1963;43:13-4.
 7. Travis WD, Brambilla E, Muller-Hermelink HK, et al. Pathology & genetics tumours of the lung, pleura, thymus and heart. World Health Organization classification of tumours. Lyon: IARC Press; 2004.
 8. Travis WD, Brambilla E, Noguchi M, et al. International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. *J Thorac Oncol* 2011;6:244-85.
 9. Travis WD, Brambilla E, Nicholson AG, et al. The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. *J Thorac Oncol* 2015;10:1243-60.
 10. Gu C, Pan X, Wang R, et al. Analysis of mutational and clinicopathologic characteristics of lung adenocarcinoma with clear cell component. *Oncotarget* 2016;7:24596-603.
 11. Ke SJ, Wang P, Xu B. Clear cell adenocarcinoma of the lung: a population-based study. *Cancer Manag Res* 2019;11:1003-12.
 12. Komiya T, Guddati AK, Nakanishi Y. Clear cell adenocarcinoma of the lung: a SEER analysis. *Transl Lung Cancer Res* 2019;8:187-91.
 13. Mallin K, Browner A, Palis B, et al. Incident Cases Captured in the National Cancer Database Compared with Those in U.S. Population Based Central Cancer Registries in 2012-2014. *Ann Surg Oncol* 2019;26:1604-12.
 14. American College of Surgeons. Participant User Files. Available online: <https://www.facs.org/quality-programs/cancer-programs/national-cancer-database/puf/>. Accessed October 30th, 2022.
 15. Histology Coding Clarifications. 2020. Available online: <https://seer.cancer.gov/tools/solidtumor/clarifications.html>. Accessed January 9th, 2023.
 16. Yang CF, Kumar A, Gulack BC, et al. Long-term outcomes after lobectomy for non-small cell lung cancer when unsuspected pN2 disease is found: A National Cancer Data Base analysis. *J Thorac Cardiovasc Surg* 2016;151:1380-8.

Cite this article as: Mansur A, Saleem Z, Potter AL, Mathey-Andrews C, Senthil P, Yang CJ. Primary clear cell adenocarcinoma of the lung: a national analysis. *J Thorac Dis* 2023;15(8):4248-4261. doi: 10.21037/jtd-23-76

Table S1 Clinicopathologic and demographic characteristics for patients, stratified by lung adenocarcinoma versus CCAL, for cases diagnosed before 2015

Patient characteristic	Lung adenocarcinoma (N=338,385)	CCAL (N=1,183)	P value
Age at diagnosis (median, IQR), years	67.0 (58.0, 75.0)	65.0 (57.0, 72.0)	<0.01
Sex, n (%)			0.48
Male	164,148 (48.5)	586 (49.5)	
Female	174,237 (51.5)	597 (50.5)	
Race, n (%)			<0.01
White	283,049 (83.6)	1,043 (88.2)	
Black	39,980 (11.8)	104 (8.8)	
Other	12,749 (3.8)	26 (2.2)	
Unknown	2,607 (0.8)	10 (0.8)	
Education, n (%)			0.44
17.6%	71,828 (21.2)	233 (19.7)	
10.9%-17.5%	90,374 (26.7)	331 (28.0)	
6.3%-10.8%	90,938 (26.9)	325 (27.5)	
<6.3%	70,017 (20.7)	233 (19.7)	
Unknown	15,228 (4.5)	61 (5.2)	
CDCC Score, n (%)			<0.01
0	203,766 (60.2)	623 (52.7)	
1	92,355 (27.3)	380 (32.1)	
2	30,455 (9.0)	136 (11.5)	
3+	11,809 (3.5)	44 (3.7)	
Year of Diagnosis (median, IQR)	2010 (2007, 2012)	2008 (2006, 2011)	<0.01
Distance from Facility (median, IQR), miles	8.9 (4.0, 21.8)	10.2 (4.3, 24.5)	<0.01
Tumor Size (median, IQR), cm	32.0 (20.0, 50.0)	33.0 (20.0, 55.0)	0.09
Tumor Location, n (%)			<0.01
Main bronchus	103,984 (30.7)	404 (34.2)	
RUL	14,584 (4.3)	42 (3.6)	
RML	45,844 (13.5)	152 (12.8)	
RLL	75,010 (22.2)	308 (26.0)	
LUL	36,412 (10.8)	129 (10.9)	
LLL	14,158 (4.2)	40 (3.4)	
Unknown	48,393 (14.3)	108 (9.1)	
Insurance Status, n (%)			<0.01
Uninsured	13,743 (4.1)	38 (3.2)	
Private	106,978 (31.6)	437 (36.9)	
Medicaid	24,026 (7.1)	77 (6.5)	
Medicare	181,972 (53.8)	597 (50.5)	
Other	4,445 (1.3)	13 (1.1)	
Unknown	7,221 (2.1)	21 (1.8)	
Facility Type, n (%)			0.32
Community cancer program	23,913 (7.1)	72 (6.1)	
Comprehensive community	138,066 (40.8)	491 (41.5)	
Academic/research program	105,527 (31.2)	387 (32.7)	
Integrated network cancer program	68,216 (20.2)	224 (18.9)	
Unknown	2,663 (0.8)	9 (0.8)	
Median Household Income, n (%)			0.78
First quartile	68,274 (20.2)	229 (19.4)	
Second quartile	75,205 (22.2)	271 (22.9)	
Third quartile	75,847 (22.4)	255 (21.6)	
Fourth quartile	103,172 (30.5)	366 (30.9)	
Unknown	15,887 (4.7)	62 (5.2)	
Grade/Differentiation, n (%)			<0.01
Well differentiated; differentiated, NOS	20,634 (6.1)	44 (3.7)	
Moderately differentiated	70,860 (20.9)	287 (24.3)	
Poorly differentiated; dedifferentiated	96,280 (28.5)	471 (39.8)	
Undifferentiated; anaplastic	1,801 (0.5)	20 (1.7)	
Cell type not determined	148,810 (44.0)	361 (30.5)	
NCCDB Analytic Stage Group, n (%)			<0.01
Stage I	76,727 (22.7)	445 (37.6)	
Stage II	22,875 (6.8)	146 (12.3)	
Stage III	66,851 (19.8)	235 (19.9)	
Stage IV	171,932 (50.8)	357 (30.2)	
Clinical T status, n (%)			<0.01
T1a	95,875 (28.3)	412 (34.8)	
T1b	67,277 (19.9)	265 (22.4)	
T1c	38,562 (11.4)	158 (13.4)	
T2a	88,504 (26.2)	248 (21.0)	
Unknown	46,547 (13.8)	95 (8.0)	
Clinical N status, n (%)			<0.01
N0	112,621 (33.3)	484 (40.9)	
N1	24,741 (7.3)	75 (6.3)	
N2	94,661 (28.0)	219 (18.5)	
N3	42,588 (12.6)	104 (8.8)	
Unknown	63,744 (18.8)	301 (25.4)	
Clinical M status, n (%)			<0.01
M0	164,010 (48.5)	791 (66.9)	
M1	162,988 (48.2)	332 (28.1)	
Unknown	11,387 (3.4)	60 (5.1)	
Treatment, n (%)			<0.01
Surgery	95,145 (28.1)	723 (61.1)	
Chemotherapy	165,706 (49.0)	513 (43.4)	
Radiation	77,529 (22.9)	221 (18.7)	

Table S2 Multivariate Cox proportional hazards analyses for patients, stratified by lung adenocarcinoma versus CCAL, for cases diagnosed before 2015

Variables	Hazard ratio	95% CI	P
Age (per year)	1.01	1.01, 1.01	<0.01
Female vs. male	0.82	0.81, 0.84	<0.01
Race (ref = white)			
Black	0.92	0.89, 0.95	<0.01
Native American	0.93	0.79, 1.10	0.40
Asian	0.74	0.70, 0.78	<0.01
Year of diagnosis (per year)	0.97	0.97, 0.97	<0.01
Median household income (ref = quartile 1)			
Second quartile	0.98	0.97, 1.01	0.23
Third quartile	0.96	0.93, 0.99	<0.01
Forth quartile	0.91	0.88, 0.94	<0.01
Insurance type (ref = uninsured)			
Private	0.85	0.81, 0.89	<0.01
Medicaid	0.97	0.92, 1.03	0.38
Medicare	0.95	0.90, 0.99	0.03
Other	0.92	0.84, 1.01	0.08
Education (ref = 17.6%)			
10.9%-17.5%	1.02	1.00, 1.05	0.10
6.3%-10.8%	1.01	0.98, 1.04	0.40
<6.3%	1.00	0.97, 1.04	0.86
Distance from facility (per mile)	1.00	1.00, 1.00	0.02
Facility type (ref = community cancer program)			<0.01
Comprehensive community clinic	0.95	0.92, 0.98	
Academic/research program	0.86	0.83, 0.89	
Integrated network cancer program	0.97	0.90, 0.97	
CDCC score (ref = 0)			<0.01
1	1.16	1.13, 1.18	
2	1.32	1.28, 1.36	
3+	1.57	1.49, 1.65	
Tumor size (per cm)	1.00	1.00, 1.00	<0.01
Tumor location (ref = Main bronchus)			
RUL	1.08	1.03, 1.12	<0.01
RML	1.14	1.11, 1.17	<0.01
RLL	1.02	1.00, 1.05	0.05
LUL	1.09	1.06, 1.12	<0.01
LLL	1.13	1.08, 1.18	<0.01
Grade/Differentiation (ref = well differentiated)			<0.01
Moderately differentiated	1.24	1.20, 1.29	
Poorly differentiated; dedifferentiated	1.44	1.39, 1.50	
Undifferentiated; anaplastic	1.44	1.32, 1.57	
Clinical T status (ref = T1a)			<0.01
T1b	1.16	1.13, 1.19	
T1c	1.24	1.20, 1.28	
T2a	1.36	1.31, 1.40	
Clinical N status (ref = N0)			<0.01
N1	1.10	1.07, 1.13	
N2	1.22	1.18, 1.27	
N3	1.10	1.06, 1.13	
Clinical M status (ref = M0)			<0.01
M1	0.99	0.99, 0.99	
Analytic Stage Group (ref = stage I)			<0.01
Stage II	1.84	1.77, 1.91	
Stage III	2.29	2.21, 2.37	
Stage IV	4.03	3.88, 4.19	
Treatment			<0.01
Surgery	0.50	0.48, 0.51	
Chemotherapy	0.52	0.51, 0.53	
Radiation	0.92	0.90, 0.94	
CCAL v Lung adenocarcinoma	1.10	0.96, 1.24	0.20

Table S3 Independent predictors of overall survival after Cox proportional hazards adjustment for patients with CCAL for cases diagnosed before 2015

Variables	Hazard ratio	95% CI	P
Age (per year)	1.01	0.99, 1.03	0.19
Female v male	0.68	0.50, 0.93	0.02
Race (ref = white)			
Black	1.31	0.74, 2.33	0.35
Native American	0.00	N/A, N/A	N/A
Asian	1.38	0.29, 6.55	0.69
Year of diagnosis (per year)	1.00	0.95, 1.07	0.92
Median household income (ref = quartile 1)			
Second quartile	0.59	0.36, 0.98	0.04
Third quartile	0.61	0.35, 1.04	0.07
Forth quartile	1.03	0.56, 1.90	0.92
Insurance type (ref = uninsured)			
Private	0.44	0.12, 1.59	0.21
Medicaid	0.62	0.16, 2.46	0.49
Medicare	0.56	0.15, 2.07	0.39
Other	1.21	0.24, 6.17	0.82
Education (ref = 17.6%)			
10.9%-17.5%	1.56	0.99, 2.46	0.05
6.3%-10.8%	1.40	0.84, 2.33	0.19
<6.3%	0.89	0.47, 1.66	0.70
Distance from facility (per mile)	1.00	0.99, 1.00	0.57
Facility type (ref = community cancer program)			
Comprehensive community clinic	1.01	0.45, 2.22	0.99
Academic/research program	0.80	0.35, 1.82	0.60
Integrated network cancer program	1.02	0.44, 2.39	0.96
CDCC score (ref = 0)			
1	1.39	0.97, 1.99	0.07
2	2.02	1.28, 3.17	<0.01
3+	1.10	0.48, 2.50	0.84
Tumor size (per cm)	1.01	1.00, 1.02	0.03
Tumor location (ref = Main bronchus)			
RUL	2.50	1.09, 5.70	0.03
RML	1.58	0.98, 2.55	0.06
RLL	1.27	0.85, 1.89	0.25
LUL	1.64	0.99, 2.72	0.05
LLL	0.92	0.44, 1.95	0.83
Grade/Differentiation (ref = well differentiated)			
Moderately differentiated	0.96	0.46, 1.97	0.90
Poorly differentiated; dedifferentiated	0.96	0.48, 1.94	0.91
Undifferentiated; anaplastic	0.61	0.19, 1.97	0.40
Clinical T status (ref = T1a)			
T1b	1.17	0.77, 1.77	0.47
T1c	1.20	0.66, 2.20	0.55
T2a	1.26	0.60, 2.62	0.54
Clinical N status (ref = N0)			
N1	1.07	0.62, 1.82	0.82
N2	0.47	0.19, 1.15	0.10
N3	0.94	0.53, 1.65	0.82
Clinical M status (ref = M0)			<0.01
M1	0.98	0.97, 0.99	
Analytic Stage Group (ref = stage I)			
Stage II	1.14	0.65, 1.98	0.65
Stage III	3.15	1.86, 5.33	<0.01
Stage IV	8.04	3.89, 16.61	<0.01
Treatment			
Surgery	0.35	0.19, 0.63	<0.01
Chemotherapy	0.62	0.40, 0.96	0.03
Radiation	0.77	0.49, 1.24	0.29

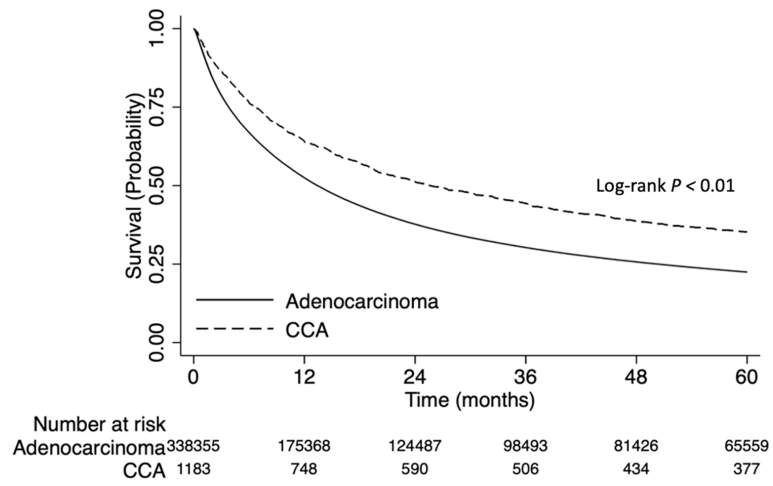


Figure S1 Kaplan-Meier analysis of overall survival for patients with CCAL *vs.* patients with lung adenocarcinoma for cases diagnosed before 2015.

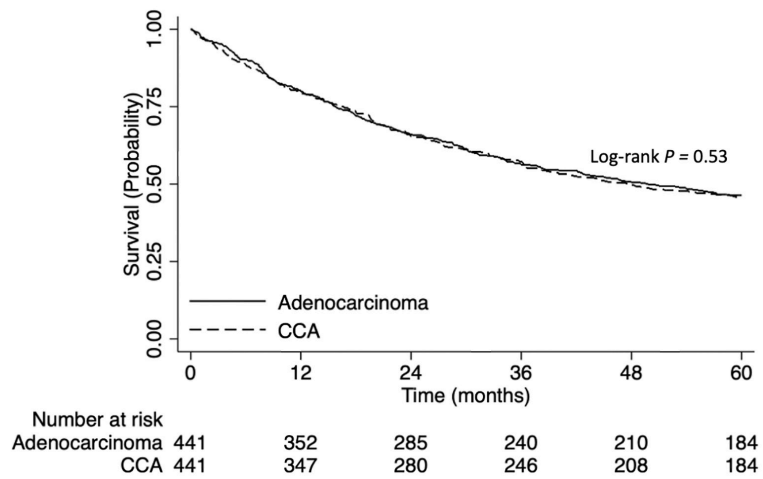


Figure S2 Kaplan-Meier analysis of overall survival for propensity score-matched patients, stratified by lung adenocarcinoma versus CCA, for cases diagnosed before 2015.