Lung Cancer 129 (2019) 8–15

Contents lists available at ScienceDirect

Lung Cancer

journal homepage: www.elsevier.com/locate/lungcan

Impact of estrogen monotherapy on survival in women with stage III-IV non-small cell lung cancer

Samuel P. Heilbroner[a,1], Eric P. Xanthopoulos[b,1], Donna Buono[c], Yongmei Huang[c], Daniel Carrier[b], Anand Shah[d], Jerry Kim[e], Michael Corradetti[f], Jason D. Wright[g], Alfred I. Neugut[c], Dawn L. Hershman[c], Simon K. Cheng[b,c,*]

[a] Columbia University College of Physicians & Surgeons, 630 W 168th St, New York, NY, 10032, United States
[b] Department of Radiation Oncology, Columbia University Irving Medical Center, 622 W 168th St, New York, NY, 10032, United States
[c] Herbert Irving Comprehensive Cancer Center, 1130 St Nicholas Ave, New York, NY, 10032, United States
[d] Food and Drug Administration, 10 Exchange Pl #804, Jersey City, NJ, 07302, United States
[e] Columbia Business School, Uris Hall 719, 3022 Broadway, New York, NY, 10027, United States
[f] Duke Cancer Center, 3404 Wake Forest Rd, Raleigh, NC, 27609, United States
[g] Department of Obstetrics and Gynecology, New York Presbyterian Hospital, 161 Fort Washington Avenue, 4th Floor, New York, NY, 10032, United States

ARTICLE INFO

Keywords:
Non-small-cell lung cancer
NSCLC
Estrogen
Outcomes
Locally advanced

ABSTRACT

Objectives: Women with lung cancer have better survival than men. The reasons are unknown, but estrogen is hypothesized to improve survival. Our objective was to examine the association between estrogen monotherapy and cancer-specific and overall survival in elderly women with non-small cell lung cancer (NSCLC).

Materials and methods: We used the SEER-Medicare database to identify women ≥65 years old who were diagnosed with stage III or IV NSCLC. Estrogen monotherapy (EM) was defined as at least one estrogen claim without any progesterone claims 6 months prior to diagnosis. To assess cancer-specific survival and overall survival, we used Kaplan-Meier and multivariate Cox modeling with propensity score adjustments. As an exploratory analysis, we also examined the effect of combined estrogen and progesterone hormonal therapy on survival using Cox modeling.

Results: We identified 6958 women in our initial cohort: 283 used EM (4%) and 6675 (96%) did not. The median follow-up time was 46.5 months in the EM patients and 49.5 months in the non-EM patients. In a Kaplan-Meier analysis, median overall survival was 8.2 months in patients who receive EM and 6.2 months in those who did not (p = 0.004).

In our 1:4 propensity-matched cohort, median follow-up was 46.5 in the EM group and 50.6 in the non-EM group; median overall survival was 8.0 months in the EM group and 6.4 months in the non-EM group (p = 0.02). In a multivariate Cox regression of the matched cohort, EM was significantly associated with overall survival (HR 0.84; 95% CI 0.73 – 0.97). All results were similar for cancer-specific survival. In our exploratory analysis, combined Estrogen-Progesterone did significantly impact overall survival (HR 0.84; 95% CI 0.71–0.99, p = 0.04) but did not appear to effect cancer-specific survival (HR 0.91; 95% CI 0.77–1.09, p = 0.30).

Conclusion: EM was associated with a significant improvement in cancer-specific survival and overall survival in women with late stage NSCLC.

1. Introduction

Multiple studies have explored gender differences in lung cancer incidence. Although differential smoking patterns in men and women contribute to the variations in lung cancer risk, they do not fully explain the difference in incidence [1]. Women are more likely than men to develop adenocarcinomas [2] and develop cancer at a younger age [2]. Hormonal and reproductive factors are postulated to contribute to the incidence of lung cancer, particularly the role of estrogens in cancer development, with inconsistent findings [3–5].

Gender differences also appear to influence lung cancer survival: females with lung cancer live longer than their male counterparts after

* Corresponding author at: Columbia University Irving Medical Center, Department of Radiation Oncology, 622 W 168th St, New York, NY, 10032, United States.
E-mail address: sc2335@cumc.columbia.edu (S.K. Cheng).
[1] Co-first authors contributed equally to this paper.

https://doi.org/10.1016/j.lungcan.2018.12.021
Received 15 May 2018; Received in revised form 13 November 2018; Accepted 23 December 2018

adjustment for other factors [3–5]. However, research on how estrogen effects lung cancer survival remains sparse, and the results mixed. Ganti et al, for example, found that women on hormonal therapy had decreased survival compared with women who had never used it (HR 1.97; 95% CI 1.14–3.39), although this was a retrospective cohort study [6]. In a prospective randomized trial, the Women's Health Initiative (WHI), those in the estrogen-only hormonal therapy arm demonstrated no change in incidence (HR of incidence = 1.17; 95% CI 0.81–1.69) or mortality (HR 1.07; 95% CI 0.66–1.72) from lung cancer [7]. The WHI study also showed that combined estrogen and progesterone hormonal therapy was associated with increased mortality but was not associated with the incidence of lung cancer [8]. Conflicting results have been reported by Katcoff et al; this group demonstrated that hormonal therapy with either estrogen monotherapy (EM) or combination estrogen plus progesterone therapy (EPT) was associated with improved NSCLC survival (HR 0.54; 95% CI 0.37-0.78), especially in women on long-term therapy [9]. In the most recent trial, the California Teachers Study, EM was associated with improved survival (HR 0.69; 95% CI 0.52 – 0.93) while combined EPT use had no effect [10].

We conducted a retrospective cohort study in women with stage III or IV NSCLC to further examine the association between EM and survival.

2. Materials and methods

2.1. Data sources

We conducted our study using the Surveillance, Epidemiology and End Results (SEER) database attached to Medicare claims (SEER-Medicare) [11,12]. SEER provides information on tumor histology, location, stage, treatment, and survival, along with the SEER site of diagnosis, demographic data, and selected census tract-level information for 26% of the United States population. The Medicare database includes Part A (inpatient), Part B (outpatient), and Part D (outpatient drugs) eligibility status, billed claims, and diagnoses. These two files are linked by unique patient identification numbers and provide the ability to determine who has been treated with an EM and the dates of service. We obtained Institutional Review Board approval.

2.2. Patient inclusion criteria

We analyzed women diagnosed with NSCLC between June 1, 2007 and December 31, 2011. The patient inclusion criteria are outlined in Fig. 1. Briefly, we excluded patients with (1) any prior malignancies; (2) autopsy or death certificate as the reporting source of death; (3) age < 65 years or whose reason for Medicare entitlement was not age; (4) > 90 days discordance between SEER and Medicare reported date of death; (5) any health maintenance organization enrollment within one year of diagnosis; or (6) incomplete Part A or B Medicare insurance in the year before or after diagnosis.

We also excluded patients who (7) lacked Part D coverage anytime 6 months prior to diagnosis; [13] (8) were living in a nursing home at the time of diagnosis; or (9) were not stage III or IV at diagnosis (Fig. 1).

We only included patients with non-small-cell disease histology according to the World Health Organization's *International Classification of Diseases for Oncology*, Third Edition (Appendix Table A1, online only). We simultaneously required only primary cancers in the lung and bronchus (SEER site recode 22030) [14,15].

2.3. Demographics, comorbidities and medications

Demographic factors. Patient demographic and tumor factors used in the analysis are listed in Table 1. We also included several other covariates in our analysis: (1) an aggregate socioeconomic status score from education, poverty level, and income information from the 2010 census tract data, as described previously by Du et al. [16] One patient lacked sufficient information in one or more categories and was excluded from the analyses. (2) We used the Klabunde adaptation of the Charlson comorbidity index to assess the prevalence of comorbid disease in our cohort [17–20]. This was calculated using the International Classification of Diseases (ninth revision) codes found on Medicare inpatient and outpatient claims in the 12 months prior to diagnosis [21]. (3) We also included comorbid conditions that are not included in the Charlson score including hip fracture, hyperlipidemia, hypertension, and osteoporosis. We calculated these using the Centers for Medicaid & Medicare Services Chronic Conditions Data Warehouse algorithms [22].

Identifying EM use. Systemic EM use was determined from Medicare Part D claims [23]. Adopting an intent-to-treat analysis, we classified patients as using systemic EM if they had a pharmacy claim submitted within 6 months prior to cancer diagnosis for any systemic estrogen-only brand name drug (Appendix Table A2, online only). We excluded patients who used any progesterone-containing drugs during that same period (Appendix Table A2). Our definition of EM has two potential weaknesses: (1) filling a single prescription for Estrogen does not guarantee sustained usage of the drug and (2) EM use in the pre-diagnosis period does not guarantee use in the post-diagnosis period where its effect on lung cancer biology is best assessed. We addressed these concerns by (1) counting the number of prescriptions for EM filled in the pre-diagnosis period for each patient and assessing the results by histogram and (2) tabulating the number of patients in our EM group who eventually also received estrogen after diagnosis (Table 1), even though part D Medicare coverage was not guaranteed through this period.

Identifying Estrogen + Progesterone Use. As an exploratory analysis, we also set out to determine the effect of combination Estrogen-Progesterone therapy (EPT). Patients were classified as using EPT if they (1) had a pharmacy claim submitted within 6 months prior to cancer diagnosis for any systemic estrogen-progesterone combination brand name drug or (2) had separate claims for pure estrogen and pure progesterone during the same time period.

Identifying Tarceva use. Similar to EM use after diagnosis, the use of Tarceva after diagnosis was tabulated to establish that this drug was generally well balanced across these treatment groups. However, this covariate was not included in our Cox or propensity score models because part D Medicare coverage was not guaranteed after diagnosis.

2.4. Cancer-related factors and treatment

From Medicare claims, we tabulated diagnosis and staging procedures, including positron emission tomography (PET) scan, bone scan, mediastinoscopy (MEED), and fine needle biopsy (FNA) (Appendix **Table A.3**, online only). We classified patients as treated with intravenous chemotherapy if they received it within 6 months of cancer diagnosis (Appendix Table A3) [24].

2.5. Outcome

The primary end point was overall survival (OS). Secondary endpoints were cancer-specific survival (CSS) and toxicity-free survival, including thromboembolic [25] events (TE), cerebrovascular events (CVE), myocardial infarctions (MI), and skeletal-related [26] events (SRE) (Appendix Table A3).

2.6. Statistical analysis

Chi-square analysis. We compared the distribution of patient characteristics between the two treatment groups with the Pearson's chi-square test when all cell counts were > 10. Otherwise, we used a Monte Carlo simulation to estimate p for Fisher's exact test with N = 1000 simulations (seed = 54039).

Kaplan-Meier and log rank survival analysis. We compared survival between women who did and did not take EM using unadjusted

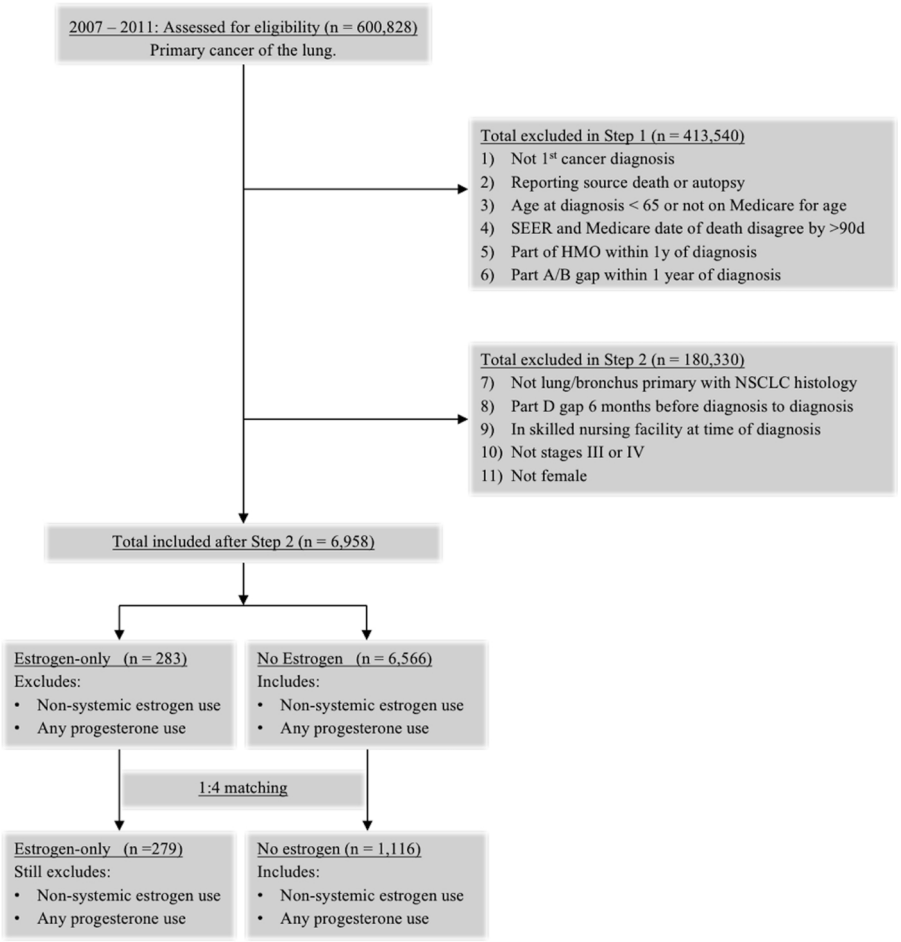

Fig. 1. Cohort Consort Diagram. Consort diagram of patient selection to evaluate the benefits of estrogen use in women with stage III or IV non-small-cell lung cancer diagnosed in the SEER-Medicare database from 2007 – 2011. Abbreviations: HMO = health maintenance organization; NSCLC = non-small cell lung cancer.

Kaplan-Meier curves and the log-rank tests. Median survival was also reported using Kaplan-Meier curves. For the OS analysis, women alive at the end of the study period were censored and contributed to the time interval from their date of diagnosis to the end of the study on December 31, 2013. Likewise, for the CSS analysis, patients alive at the end of the study period or who died from causes other than their lung cancer were censored. Median follow-up time was calculated using Kaplan-Meier curves and reverse censoring.

Propensity scoring. Propensity score methods adjusted for potential selection bias [27,28]. The propensity score was calculated using a logistic regression model which fit first- and second-order interactions on 12 covariates using backward elimination with a threshold p-value ≤ 0.10. The 12 covariates fed into the regression model included: (1) age, (2) year of diagnosis, (3) race, (4) marital status, (5) state of residence, (6) socioeconomic status, (7) Charlson score, (8) homebound status, and a history of (9) hip fracture, (10) hyperlipidemia, (11) hypertension, and (12) osteoporosis. Patients were randomly sorted and then matched without replacement to a nearest neighbor with match caliper = 0.01. We assessed the quality of our propensity score model by checking the area under the curve ($c = 0.86 > 0.70$) and goodness of fit ($p = 0.80 > 0.05$). As described immediately below, we included the propensity score in some of our Cox models to adjust for potential confounding from selection bias.

Cox regression adjustment. To build each of our Cox regression models, we started with a pool of 9 baseline covariates (Table 2) that could not be included in our propensity score model because their values were measured at or after diagnosis. These covariates were: (1) EM, (2) histology, (3) stage, (4) tumor site/location, (5) staging PET scan, (6) staging bone scan, (7) staging MEED, (8) staging FNA, and (9) receipt of chemotherapy. Staging studies and receipt of chemotherapy were modeled as time dependent variables to avoid immortality bias. First we performed a propensity score unadjusted Cox regression on all eligible patients (n = 6958) both to (a) identify individual factors associated with survival differences between the treatment groups, and (b) adjust for potential confounders. Then we performed two propensity score adjusted Cox regressions to reduce the risk of confounding, fitting additional models on: (i) all eligible patients (n = 6958) after stratifying them by propensity score deciles, and (ii) a subset of patients obtained by propensity score matching 1 control patient to 4 treated patients (n = 1395). The use of EM and Tarceva after diagnosis were not included as covariates because some patients dropped their part D coverage after diagnosis.

In the matched sub-analysis, we did not Cox regress on covariates that were initially part of the propensity score that became well balanced after matching. We used the Schoenfeld residuals test to verify the underlying proportional hazards assumption for each of our Cox models.

Toxicity-free survival secondary endpoint. We reported toxicity-free survival while adjusting for the competing risk of death using a multivariate Cox regression model in the propensity score matched cohort described above [29].

Effect of Estrogen + Progesterone Use. As an exploratory analysis, we set out to examine the effect of combination EPT use on survival by comparing it to EM and all other patients. Because three treatment groups were included in this analysis, no propensity score methods were applied. Instead, hormone use (either EM, EPT, or neither) was

Table 1
Patient Characteristics. Patient characteristics of all patients and a sub-set of propensity score matched patients with one experimental patient matched to four controls. Abbreviations: FNA = fine needle aspiration; MEED = mediastinoscopy; PET = positron emission tomography.

Characteristic	1a Estrogen unmatched	1b No Estrogen unmatched	1c p	2a Estrogen 1:4 match	2b No Estrogen 1:4 match	2c p
Age†			< .01			0.28
65 - 74	162 (57%)	3087 (46%)		159 (57%)	622 (56%)	
75 - 84	96 (34%)	2816 (42%)		95 (34%)	420 (38%)	
85 +	25 (9%)	772 (12%)		25 (9%)	74 (7%)	
Year of diagnosis†			0.45			0.23
2007	29 (10%)	729 (11%)		27 (10%)	147 (13%)	
2008	65 (23%)	1472 (22%)		64 (23%)	247 (22%)	
2009	72 (25%)	1536 (23%)		72 (26%)	253 (23%)	
2010	48 (17%)	1422 (21%)		47 (17%)	228 (20%)	
2011	69 (24%)	1516 (23%)		69 (25%)	241 (22%)	
Race†			< .01			1.00
White	264 (93%)	5424 (81%)		260 (93%)	1040 (93%)	
Black	14 (5%)	702 (11%)		14 (5%)	55 (5%)	
Hispanic	*	96 (1%)		*	*	
Other	*	453 (7%)		*	17 (2%)	
Married*†	104 (37%)	1977 (30%)	0.01	102 (37%)	393 (35%)	0.67
State†			< .01			0.89
California	114 (40%)	2057 (31%)		114 (41%)	503 (45%)	
Connecticut	*	248 (4%)		*	19 (2%)	
Georgia	41 (15%)	832 (13%)		40 (14%)	171 (15%)	
Hawaii	*	56 (1%)		*	*	
Iowa	11 (4%)	464 (7%)		11 (4%)	37 (3%)	
Kentucky	27 (10%)	722 (11%)		27 (10%)	113 (10%)	
Louisiana	23 (8%)	460 (7%)		21 (8%)	72 (7%)	
Michigan	19 (7%)	475 (7%)		19 (7%)	62 (6%)	
New Jersey	*	833 (13%)		*	32 (3%)	
New Mexico	*	105 (2%)		*	32 (3%)	
Utah	*	61 (1%)		*	11 (1%)	
Washington	20 (7%)	362 (5%)		20 (7%)	62 (6%)	
Socioeconomic rank†			0.59			0.69
1 st quintile	48 (17%)	1341 (20%)		46 (17%)	209 (19%)	
2nd quintile	65 (23%)	1404 (21%)		65 (23%)	250 (22%)	
3rd quintile	41 (15%)	1082 (16%)		40 (14%)	186 (17%)	
4th quintile	73 (26%)	1616 (24%)		73 (26%)	259 (23%)	
5th quintile	56 (20%)	1231 (18%)		55 (20%)	211 (19%)	
Unknown	0 (0%)	*		0 (0%)	*	
Charlson score†			< .01			0.57
0	102 (36%)	2174 (33%)		102 (37%)	371 (33%)	
1-2	153 (54%)	3316 (50%)		149 (53%)	632 (57%)	
> 2	28 (10%)	1185 (18%)		28 (10%)	113 (10%)	
Homebound†	*	97 (2%)	0.04	*	16 (1%)	0.42
Stage‡			0.38			0.41
III	116 (41%)	2564 (38%)		114 (41%)	426 (38%)	
IV	167 (59%)	4111 (62%)		165 (59%)	690 (62%)	
Histology‡			0.11			0.65
Adenocarcinoma	147 (52%)	3441 (52%)		145 (52%)	593 (53%)	
Squamous	58 (21%)	1589 (24%)		56 (20%)	238 (21%)	
Large Cell	13 (5%)	167 (3%)		13 (5%)	36 (3%)	
Other	65 (23%)	1478 (22%)		65 (23%)	249 (22%)	
Location‡			0.03			0.04
Upper lobe	138 (49%)	3133 (47%)		137 (49%)	491 (44%)	
Middle lobe	19 (7%)	249 (4%)		18 (7%)	43 (4%)	
Lower lobe	75 (27%)	1779 (27%)		75 (27%)	324 (29%)	
Other Site	51 (18%)	1514 (23%)		49 (18%)	258 (23%)	
Staging studies						
PET‡	184 (65%)	3827 (57%)	0.01	181 (65%)	656 (59%)	0.06
FNA‡	172 (61%)	3363 (50%)	< .01	170 (61%)	587 (53%)	0.01
Bone scan‡	71 (25%)	1712 (26%)	0.83	70 (25%)	277 (25%)	0.93
MEED‡	33 (12%)	448 (7%)	< .01	32 (12%)	79 (7%)	0.02
Post-diagnosis treatment						
Chemotherapy‡ (iv)	144 (51%)	2953 (44%)	0.03	141 (51%)	549 (49%)	0.69
Tarceva** (oral)	39 (14%)	970 (15%)	0.73	39 (14%)	172 (15%)	0.55
Estrogen**	145 (51%)	131 (2%)	< .01	142 (51%)	32 (3%)	< .01
Pre-diagnosis comorbidities						
Hip fracture†	*	204 (3%)	0.14	*	27 (2%)	0.15
Hyperlipidemia†	197 (70%)	4620 (69%)	0.89	195 (70%)	794 (71%)	0.68
Hypertension†	222 (78%)	5477 (82%)	0.12	218 (78%)	895 (80%)	0.44
Osteoporosis†	66 (23%)	1612 (24%)	0.75	64 (23%)	241 (22%)	0.63

* Low value cells masked to protect patient privacy per the National Cancer Institute's SEER-Medicare data user agreement.
** During patient selection, Part D coverage only required in the 6 months prior to diagnosis to preserve study power.
† In subset analysis, pre-diagnosis characteristic balanced across cohorts using 1:4 propensity score matching.
‡ Later Cox regression modeling used to address any potential imbalances in these unmatched characteristics.

Table 2
Cox Regression Analysis. Uni- and multivariable Cox survival analysis. 1) Univariate Cox regression on estrogen use. 2) Multivariable Cox regression that only uses the 9 covariates that could not be included in the propensity score analysis. 3) Multivariable Cox regression on our propensity score matched cohort. 4) Multivariable Cox regression using the 9 covariates described in 2 as well as propensity score decile. Abbreviations: FNA = fine needle aspiration; MEED = mediastinoscopy; PET = positron emission tomography.

Covariate	1 Univariate HR (95% CI, p)	2 Propensity Score Unadjusted, Multivariable HR (95% CI, p)	3 Multivariable HR (95% CI, p) matched	4 Multivariable HR (95% CI, p) ranked
Estrogen Mono. (Y vs N)	0.78 (0.69 - 0.89, < 0.001)	0.83 (0.73 - 0.95, 0.0058)	0.84 (0.73 - 0.98, 0.022)	0.86 (0.75 - 0.98, 0.025)
Histology	–	–	–	–
Adenocarcinoma	Reference	Reference	Reference	Reference
Squamous	1.07 (1.01 - 1.14, 0.027)	1.18 (1.11 - 1.26, < 0.001)	1.32 (1.14 - 1.52, < 0.001)	1.18 (1.11 - 1.26, < 0.001)
Large Cell	1.51 (1.29 - 1.76, < 0.001)	1.61 (1.38 - 1.88, < 0.001)	2.01 (1.49 - 2.72, < 0.001)	1.6 (1.37 - 1.87, < 0.001)
Other	1.28 (1.2 - 1.37, < 0.001)	1.32 (1.23 - 1.41, < 0.001)	1.45 (1.25 - 1.69, < 0.001)	1.31 (1.23 - 1.41, < 0.001)
Stage (4 vs 3)	1.11 (0.98 - 1.25, 0.113)	1.01 (0.89 - 1.14, 0.8643)	1.1 (0.82 - 1.48, 0.526)	1.01 (0.9 - 1.15, 0.835)
Site	–	–	–	–
Upper lobe	Reference	Reference	Reference	Reference
Middle lobe	1.01 (0.89 - 1.16, 0.84)	0.95 (0.83 - 1.09, 0.4843)	0.89 (0.66 - 1.19, 0.430)	0.95 (0.83 - 1.09, 0.454)
Lower lobe	1.04 (0.98 - 1.1, 0.22)	1.05 (0.98 - 1.11, 0.1472)	1.12 (0.98 - 1.28, 0.110)	1.05 (0.98 - 1.11, 0.145)
Other	1.49 (1.4 - 1.59, < 0.001)	1.32 (1.24 - 1.41, < 0.001)	1.4 (1.2 - 1.62, < 0.001)	1.33 (1.24 - 1.42, < 0.001)
Staging studies	–	–	–	–
PET† (Y vs N)	0.51 (0.48 - 0.53, < 0.001)	0.58 (0.55 - 0.61, < 0.001)	0.52 (0.46 - 0.59, < 0.001)	0.57 (0.54 - 0.61, < 0.001)
Bone Scan† (Y vs N)	1.46 (1.38 - 1.55, < 0.001)	1.33 (1.26 - 1.41, < 0.001)	1.4 (1.22 - 1.59, < 0.001)	1.33 (1.25 - 1.4, < 0.001)
MEED† (Y vs N)	0.56 (0.5 - 0.62, < 0.001)	0.68 (0.61 - 0.76, < 0.001)	0.79 (0.63 - 0.99, 0.040)	0.68 (0.61 - 0.76, < 0.001)
FNA† (Y vs N)	0.91 (0.86 - 0.96, < 0.001)	0.97 (0.92 - 1.02, 0.1771)	1.13 (1.01 - 1.28, 0.037)	0.97 (0.92 - 1.02, 0.208)
Chemotherapy† (Y vs N)	0.71 (0.68 - 0.75, < 0.001)	0.77 (0.73 - 0.81, < 0.001)	0.71 (0.62 - 0.8, < 0.001)	0.78 (0.73 - 0.82, < 0.001)
Propensity score decile	–	–	–	–
1st decile	Reference	Not in model	Not in model	Reference
2nd decile	0.93 (0.83 - 1.04, 0.188)	Not in model	Not in model	0.99 (0.89 - 1.11, 0.872)
3rd decile	0.87 (0.78 - 0.97, 0.010)	Not in model	Not in model	0.94 (0.84 - 1.05, 0.300)
4th decile	0.93 (0.83 - 1.04, 0.202)	Not in model	Not in model	1.04 (0.93 - 1.16, 0.477)
5th decile	0.93 (0.84 - 1.04, 0.213)	Not in model	Not in model	1.04 (0.93 - 1.17, 0.457)
6th decile	0.94 (0.85 - 1.05, 0.311)	Not in model	Not in model	1.12 (1 - 1.25, 0.046)
7th decile	0.86 (0.77 - 0.96, 0.006)	Not in model	Not in model	0.98 (0.88 - 1.09, 0.708)
8th decile	0.82 (0.74 - 0.92, < 0.001)	Not in model	Not in model	0.94 (0.84 - 1.05, 0.246)
9th decile	0.79 (0.7 - 0.88, < 0.001)	Not in model	Not in model	0.92 (0.83 - 1.03, 0.171)
10th decile	0.71 (0.63 - 0.79, < 0.001)	Not in model	Not in model	0.88 (0.79 - 0.99, 0.032)

† Time-dependent adjustment applied during regression.

entered into a new Cox model with all covariates that were used previously in either the calculation of a propensity score or in the Cox models.

All analyses were performed with SAS 9.4 (SAS, Cary, NC) using two-tailed testing, with $\alpha = 0.05$.

3. Results

3.1. Characteristics

A total of 6958 patients satisfied the inclusion criteria for our study. Of these, 283 (4%) women received EM and 6675 (96%) women did not (Table 1). In our 1:4 matched sub-analysis of 1395 patients, 279 received EM and 1116 did not (Table 1). In our exploratory analysis, an additional 169 received combination EPT in the entire cohort.

After adjustment, the study groups were well-balanced across all pre-diagnosis characteristics included in the propensity score model: the adjusted cohorts remained unbalanced on some covariates that were not included in the propensity scoring regression: use of FNA staging and use of mediastinoscopy staging ($p < 0.05$ for both, Table 1). We adjusted for these imbalances by including these covariates in our Cox regression.

Of the patients who receive EM in the pre-diagnosis period, 76% filled at least two prescriptions (Fig. 2). At least 51% received EM after diagnosis (Table 1), although the exact number could not be assessed because some patients had dropped their part D coverage. Tarceva use appeared to be well balanced between the two groups, but cannot be fully assessed because part D coverage was not guaranteed.

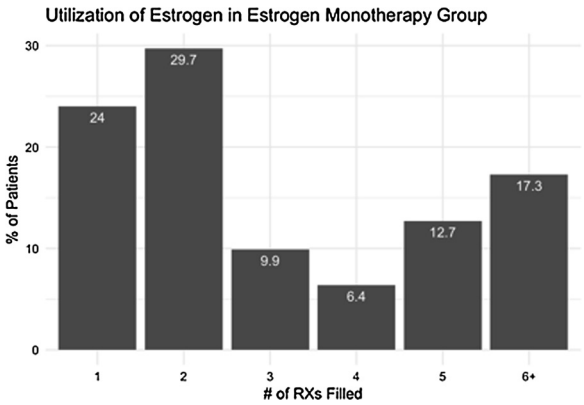

Fig. 2. Utilization of Estrogen in Estrogen Monotherapy Group. Percent of estrogen monotherapy patients who filled 1, 2, 3, 4, 5, or 6+ Estrogen prescriptions in the 6 months before diagnosis.

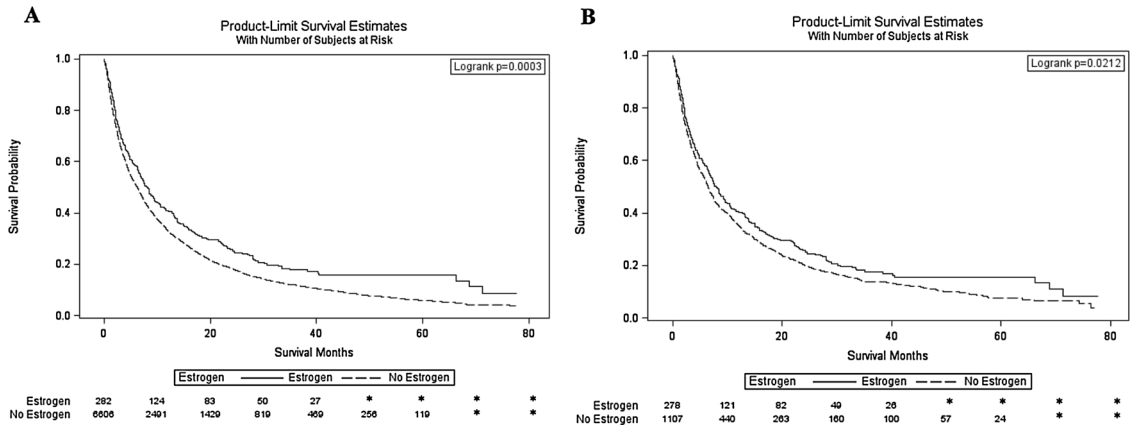

Fig. 3. Kaplan-Meier Analysis of Overall Survival. Kaplan-Meier analysis of (A) overall survival (OS) in all eligible patients and (B) OS in sub-cohort with one experimental estrogen-patient propensity score-matched to four controls. *s indicate hidden value to protect patient anonymity.

3.2. Overall and cancer-specific survival

In all eligible patients (n = 6958), the median follow-up time was 46.5 months (IQR 31.6–61.6 months) in the 283 EM patients and 49.5 months in the 589 surviving non-EM patients (IQR 35.5–62.5 months; p = 0.54). In the matched sub-analysis (n = 2115), the median follow-up was 46.5 months (IQR 31.6–61.5 months) for the 279 EM patients and 50.6 months (range 36.6–63.6 months; p = 0.39) for the 1116 non-EM patients.

Using Kaplan-Meier analysis (Fig. 3), median OS was 8.2 versus 6.2 (log-rank p = 0.0003) months in all eligible, and 8.0 versus 6.4 months (log-rank p = 0.0212) in matched EM versus non-EM patients. Statistics and plots for CSS were comparable and are presented in the appendix (Appendix Table A4, Figure A1).

Results were nearly identical using several Cox models (Table 2) including: (1) a univariate Cox regression (2) propensity score unadjusted, multivariable Cox regression across all eligible patients, (3) multivariable Cox regression on a subset of 1:4 propensity score matched patients, and (4) propensity score-adjusted, multivariable Cox regression across all eligible patients (HR 0.84; CI 0.73 - 0.97 for the propensity score matched model). Table 2 also lists the other covariates that were associated with OS. EM was also associated with CSS (HR 0.83; 95% CI 0.71 - 0.97 for the propensity score matched model) and the entire output is provided in the appendix (Appendix Table A5).

3.3. Exploratory analysis of estrogen-progesterone therapy

In our exploratory analysis using multivariable Cox regression analysis, combined Estrogen-Progesterone did significantly impact overall survival (HR 0.84; 95% CI 0.71 - 0.99, p = 0.04) but did not appear to effect cancer-specific survival (HR 0.91; 95% CI 0.77–1.09, p = 0.30). Full output from the cox-models is shown in the appendix (Appendix Tables A6 and A7).

3.4. Toxicity-free survival

In the multivariable Cox regression analysis of our matched cohort, systemic EM significantly reduced the rates of SREs (HR 0.80; 95% CI 0.65 - 0.98; p = 0.03) between the two groups. EM had no effect on the rates of TEs (HR 1.01; 95% CI 0.85–1.19; p = 0.94), MIs (HR 0.97 95% CI 0.82–1.15; p = 0.72), or CVEs (HR 0.99; 95% CI 0.83–1.18; p = 0.92).

4. Discussion

We found an association between EM in the 6 months prior to diagnosis and improved CSS and OS in women with stage III or IV NSCLC. There was also an association between EM and a decreased risk of SREs but no association with the risk of TEs, MIs, or CVEs. These results were the same, regardless of how our statistical analysis was performed.

Prior studies showed that gender differences are associated with lung cancer survival [3–5]. While prior studies were mixed, multiple series have shown that EM has shown more promise to improve [9] – or at least not affect [7] – survival outcomes in women with NSCLC. Indeed, the WHI randomized controlled trial found no difference in mortality [7] and a more recent study by Katcoff actually showed a benefit [9]. Data for combined EPT use remains less conclusive, with literature reporting the gamut of worse [8] to improved [9] survival outcomes. All of these studies were limited by small sample sizes and a small number of covariates; limitations which we attempted to address in this study. In addition, the clinical significance of this study's observed OS increase (1.5–2 months) with EM is on the same order of benefit as showed for other repurposed medications, such as metformin [30,31] and denosumab [32,33], which are currently being studied in phase III clinical trials in NSCLC.

The mechanism by which estrogen may contribute to improved lung cancer survival are unknown. Women with lung cancer are at increased risk for premature death due to cardiovascular events or other complications; treatment with EM may ameliorate this risk. However, we did not find significant reductions in the risk of TEs, MIs, or CVEs between the two study arms. These results may be explained by the short life expectancy and high rate of cancer death among these patients. EM did significantly reduce the rate of SREs in our cohort. Estrogen's osteotrophic effects are well known. However, an improvement in the rates of SREs is unlikely to singularly explain our findings. This is supported by the fact that EM was also associated with a significant improvement in CSS. In general, there was very few discrepancies between overall and cancer-specific survival analysis, suggesting that estrogens effect on survival may be directly related to cancer physiology.

Contrary to our findings, cell biological and mouse data demonstrate that estrogen may be pro-mitogenic [9,34–36]. But pro-proliferative estrogen receptors do not operate in a vacuum; they interact with many other parallel and downstream pathways in still poorly understood ways. For example, progesterone action – mediated by the progesterone receptor, PR – may stimulate tissue differentiation and inhibit cell proliferation [35,37]. This study suggests that the cumulative impact of the total, not-yet-fully-elucidated estrogen cascade may produce a net lung tumor suppression effect. Some possibilities include direct anti-tumor effects or acting as a sensitizer for later chemotherapy or radiation.

In defining our study, we deliberately defined EM in the six-month window prior to cancer diagnosis to eliminate immortality bias [38]. Briefly, immortality bias occurs when retrospectively studying an intervention that occurs after diagnosis. In these instances, there will be a

false correlation between survival and the intervention, regardless of its actual efficacy.

However, any bias that our definition of EM use would cause would likely tilt towards the null hypothesis: we would expect *not* to see a survival benefit if too many patients did not continue to use EM after diagnosis or if estrogen quickly dissipated after use. But our results overcame this countervailing hurdle and rejected the null hypothesis. Unfortunately, it is not possible to know how many women remained on estrogen therapy after diagnosis in our analysis, since we only required part D coverage until diagnosis.

We intentionally limited our study to women who used systemic estrogen only and excluded women who concurrently used any progesterone therapy. We did this because allowing estrogen use with or without progesterone use would complicate, or even confound, efforts to deconstruct what drove any survival differences between the control and study groups. In a brief exploratory analysis, the effects of combined estrogen-progesterone therapy were more mixed with a significant improvement in overall-survival but not cancer-specific survival (Appendix, Tables A6 and A7). This is an interesting avenue for further research.

Our study has several limitations: (1) It is always possible that selection bias may confound the results. The SEER database does not record performance status; smoking status; weight loss; tumor molecular marker phenotype; radiation fraction size, total dose, or location irradiated; or surgical margin status. These potential confounders can drive selection bias even after compensating for them using propensity score analysis and multivariate Cox regressions. One particularly important potential confounder is the fact that EM use may also simply reflect better access to healthcare or increased health awareness. We indirectly addressed this limitation by including a composite socioeconomic rank - as described by Du et al [16] – in our model. (2) Another limitation was our inability to confirm adherence to EM therapy or assess for a dose-dependent effect on survival. However, lack of adherence would likely bias our results towards the null hypothesis; but our findings rejected the null hypothesis despite this bias. (3) Finally, our cohort only included patients $\geq$65 years of age; accordingly, we could not assess the effect of EM in younger patients. In light of these limitations, interpretation of our data should be viewed as hypothesis generating. Never the less, our data does suggest that EM use may be associated with CSS and OS in women with stage III or IV NSCLC.

Conflicts of interest statement

None declared.

Acknowledgments

We thank Barry Neustein for his generosity and support for the lung cancer research program in Radiation Oncology at Columbia University. We thank Dr Tom K Hei, PhD, Vice-Chair and Professor of Radiation Oncology and Environmental Health Sciences, Columbia University for comments that greatly improved the manuscript. This study used the linked SEER-Medicare database. The interpretation and reporting of these data are the sole responsibility of the authors. The authors acknowledge the efforts of the National Cancer Institute, the Office of Research, Development, and Information, Centers for Medicare & Medicaid Services; Information Management Services, Inc.; and the SEER program tumor registries in the creation of the SEER-Medicare database.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.lungcan.2018.12.021.

References

[1] E.A. Zang, E.L. Wynder, Differences in lung cancer risk between men and women: examination of the evidence, J. Natl. Cancer Inst. 88 (1996) 183–192.
[2] J.S. Donington, Y.L. Colson, Sex and gender differences in non-small cell lung cancer, Semin. Thorac. Cardiovasc. Surg. 23 (2011) 137–145.
[3] A.L. Visbal, B.A. Williams, F.C. Nichols 3rdet al., Gender differences in non-small-cell lung cancer survival: an analysis of 4,618 patients diagnosed between 1997 and 2002, Ann. Thorac. Surg. 78 (2004) 209–215 discussion 215.
[4] C. Alexiou, C.V. Onyeaka, D. Beggs, et al., Do women live longer following lung resection for carcinoma? Eur. J. Cardiothorac. Surg. 21 (2002) 319–325.
[5] S. Kligerman, C. White, Epidemiology of lung cancer in women: risk factors, survival, and screening, AJR Am. J. Roentgenol. 196 (2011) 287–295.
[6] A.K. Ganti, A.E. Sahmoun, A.W. Panwalkar, K.K. Tendulkar, A. Potti, Hormone replacement therapy is associated with decreased survival in women with lung cancer, J. Clin. Oncol. 24 (2006) 59–63.
[7] R.T. Chlebowski, G.L. Anderson, J.E. Manson, et al., Lung cancer among postmenopausal women treated with estrogen alone in the women's health initiative randomized trial, J. Natl. Cancer Inst. 102 (2010) 1413–1421.
[8] R.T. Chlebowski, A.G. Schwartz, H. Wakelee, et al., Oestrogen plus progestin and lung cancer in postmenopausal women (Women's Health Initiative trial): a post-hoc analysis of a randomised controlled trial, Lancet 374 (2009) 1243–1251.
[9] H. Katcoff, A.S. Wenzlaff, A.G. Schwartz, Survival in women with NSCLC: the role of reproductive history and hormone use, J. Thorac. Oncol. 9 (2014) 355–361.
[10] J. Clague, P. Reynolds, K.D. Henderson, et al., Menopausal hormone therapy and lung cancer-specific mortality following diagnosis: the California teachers Study, PLoS One 9 (2014) e103735.
[11] J.L. Warren, C.N. Klabunde, D. Schrag, P.B. Bach, G.F. Riley, Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population, Med. Care 40 (2002) IV-3-18.
[12] A.L. Potosky, G.F. Riley, J.D. Lubitz, R.M. Mentnech, L.G. Kessler, Potential for cancer related health services research using a linked Medicare-tumor registry database, Med. Care 31 (1993) 732–748.
[13] C.N. Klabunde, L.C. Harlan, J.L. Warren, Data sources for measuring comorbidity: a comparison of hospital records and medicare claims for cancer patients, Med. Care 44 (2006) 921–928.
[14] J.A. Ferguson, C.J. Suelzer, G.J. Eckert, X.H. Zhou, R.S. Dittus, Risk factors for delirium tremens development, J. Gen. Intern. Med. 11 (1996) 410–414.
[15] T. Palmstierna, A model for predicting alcohol withdrawal delirium, Psychiatr. Serv. 52 (2001) 820–823.
[16] X.L. Du, S. Fang, S.W. Vernon, et al., Racial disparities and socioeconomic status in association with survival in a large population-based cohort of elderly patients with colon cancer, Cancer 110 (2007) 660–669.
[17] https://healthcaredelivery.cancer.gov/seermedicare/program/comorbidity.html. NCIDoCCPSS-MCoCWa.
[18] M.E. Charlson, F.L. Sax, C.R. MacKenzie, S.D. Fields, R.L. Braham, R.G. Douglas Jr., Assessing illness severity: does clinical judgment work? J. Chronic Dis. 39 (1986) 439–452.
[19] C.N. Klabunde, J.M. Legler, J.L. Warren, L.M. Baldwin, D. Schrag, A refined comorbidity measurement algorithm for claims-based studies of breast, prostate, colorectal, and lung cancer patients, Ann. Epidemiol. 17 (2007) 584–590.
[20] C.N. Klabunde, A.L. Potosky, J.M. Legler, J.L. Warren, Development of a comorbidity index using physician claims data, J. Clin. Epidemiol. 53 (2000) 1258–1267.
[21] E.M. Cassidy, I. O'Sullivan, P. Bradshaw, T. Islam, C. Onovo, Symptom-triggered benzodiazepine therapy for alcohol withdrawal syndrome in the emergency department: a comparison with the standard fixed dose benzodiazepine regimen, Emerg. Med. J. 29 (2012) 802–804.
[22] S. Naveau, V. Giraud, E. Borotto, A. Aubert, F. Capron, J.C. Chaput, Excess weight risk factor for alcoholic liver disease, Hepatology 25 (1997) 108–111.
[23] K.A. Martin, R.L. Barbieri, T.W. Post (Ed.), Preparations for Menopausal Hormone Therapy, vol. 2016, UpToDate, UpToDate, Waltham, MA, 2016.
[24] J.L. Warren, L.C. Harlan, A. Fahey, et al., Utility of the SEER-Medicare data to identify chemotherapy use, Med. Care 40 (2002) IV-55-61.
[25] S. Shantakumar, A. Connelly-Frost, M.G. Kobayashi, R. Allis, L. Li, Older soft tissue sarcoma patients experience increased rates of venous thromboembolic events: a retrospective cohort study of SEER-Medicare data, Clin. Sarcoma Res. 5 (2015) 18.
[26] A. Aly, E. Onukwugha, C. Woods, et al., Measurement of skeletal related events in SEER-Medicare: a comparison of claims-based methods, BMC Med. Res. Methodol. 15 (2015) 65.
[27] L.H. Curtis, B.G. Hammill, E.L. Eisenstein, J.M. Kramer, K.J. Anstrom, Using inverse probability-weighted estimators in comparative effectiveness analyses with observational databases, Med. Care 45 (2007) S103–107.
[28] J.D. Wright, W.M. Burke, A.I. Tergas, et al., Comparative effectiveness of minimally invasive hysterectomy for endometrial cancer, J. Clin. Oncol. 34 (2016) 1087–1096.
[29] L.S. Scrucca, F. A Aversa, Competing risk analysis using R: an easy guide for clinicians, Bone Marrow Transplant. 40 (2007) 381–387.
[30] G. Gupta, T. de Jesus Andreoli Pinto, D.K. Chellappan, A. Mishra, H. Malipeddi, K. Dua, A clinical update on metformin and lung cancer in diabetic patients, Panminerva Med. 60 (2018) 70–75.
[31] J.J. Lin, E.J. Gallagher, K. Sigel, et al., Survival of patients with stage IV lung cancer with diabetes treated with metformin, Am. J. Respir. Crit. Care Med. 191 (2015) 448–454.
[32] J. De Castro, R. Garcia, P. Garrido, et al., Therapeutic potential of denosumab in

patients with lung Cancer: beyond prevention of skeletal complications, Clin. Lung Cancer 16 (2015) 431–446.
[33] G.V. Scagliotti, V. Hirsh, S. Siena, et al., Overall survival improvement in patients with lung cancer and bone metastases treated with denosumab versus zoledronic acid: subgroup analysis from a randomized phase 3 study, J. Thorac. Oncol. 7 (2012) 1823–1829.
[34] J.M. Siegfried, P.A. Hershberger, L.P. Stabile, Estrogen receptor signaling in lung cancer, Semin. Oncol. 36 (2009) 524–531.
[35] N. Kazmi, D.C. Marquez-Garban, L. Aivazyan, et al., The role of estrogen, progesterone and aromatase in human non-small-cell lung cancer, Lung Cancer Manag. 1 (2012) 259–272.
[36] S. Chakraborty, A.K. Ganti, A. Marr, S.K. Batra, Lung cancer in women: role of estrogens, Expert Rev. Respir. Med. 4 (2010) 509–518.
[37] H. Ishibashi, T. Suzuki, S. Suzuki, et al., Progesterone receptor in non-small cell lung cancer–a potent prognostic factor and possible target for endocrine therapy, Cancer Res. 65 (2005) 6450–6458.
[38] J.H.M. Crowley, Covariance analysis of heart transplant survival data, J. Am. Stat. Assoc. 72 (1977) 27–36.