Check for updates

Impact of gender and mutational differences in hormone receptor expressing non-small cell lung cancer

Robert Hsu[1,2]*, Denaly Chen[1,2], Bing Xia[1,2†], Rebecca Feldman[3], Wendy Cozen[4], Luis E. Raez[5], Hossein Borghaei[6], Chul Kim[7], Misako Nagasaka[4], Hirva Mamdani[8], Ari M. Vanderwalde[3], Gilberto Lopes[9], Mark A. Socinski[10], Antoinette J. Wozniak[11†], Alexander I. Spira[12], Stephen V. Liu[7] and Jorge J. Nieva[1,2]

[1]Department of Internal Medicine, Division of Medical Oncology, University of Southern California, Los Angeles, CA, United States, [2]Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA, United States, [3]Caris Life Sciences, Phoenix, AZ, United States, [4]Division of Hematology-Oncology, Department of Medicine, University of California Irvine School of Medicine, Chao Family Comprehensive Cancer Center, Orange, CA, United States, [5]Thoracic Oncology Program, Memorial Cancer Institute/Florida Atlantic University, Pembroke Pines, FL, United States, [6]Department of Hematology-Oncology, Fox Chase Cancer Center, Philadelphia, PA, United States, [7]Lombardi Comprehensive Cancer Center, MedStar Georgetown University Hospital, Washington, DC, United States, [8]Department of Oncology, Barbara Ann Karmanos Cancer Institute, Detroit, MI, United States, [9]Department of Medical Oncology, Sylvester Comprehensive Cancer Center at the University of Miami, Miami, FL, United States, [10]AdventHealth Cancer Institute, Orlando, FL, United States, [11]Hillman Cancer Center, Department of Medicine, Division of Hematology/Oncology, University of Pittsburgh Medical Center, Pittsburgh, PA, United States, [12]US Oncology Research, Virginia Cancer Specialists, Fairfax, VA, United States

Background: The incidence of lung cancer in the US has been decreasing but a bigger decline has been observed in men despite similar declines in tobacco use between men and women. Multiple theories have been proposed, including exposure to exogenous estrogens. Our study seeks to understand the relationship between hormone receptors (HR), gender, and the genomic landscape of non-small lung cancer (NSCLC).

Methods: 3,256 NSCLC tumor samples submitted for molecular profiling between 2013-2018 were retrospectively identified and assessed for HR expression. Hormone receptor (HR+) was defined as $\geq$ 1% nuclear staining of estrogen receptor-alpha (ER-a) or progesterone receptor (PR) by immunohistochemistry. DNA sequencing by NGS included cases sequenced by the Illumina MiSeq hot spot 47 gene panel (n=2753) and Illumina NextSeq 592 gene panel (n=503). An adjusted p-value (q-value) <0.05 was determined significant.

Results: HR+ was identified in 18.3% of NSCLC. HR+ occurred more commonly in women compared to men (19.6% vs 11.4%, p <0.0001, q <0.0001). *EGFR* mutations occurred more commonly in HR+ NSCLC than HR- NSCLC (20.2% vs. 14.6%, p = 0.002, q=0.007). Overall, men with *EGFR* mutations were affected by HR status with a higher prevalence in HR+ NSCLC while such differences were not seen in women. However, in women ages $\leq$45, there was a trend towards

greater prevalence HR+ NSCLC (25.25% vs. 11.32%, q= 0.0942) and 10/25 (40.0%) of HR+ cases in young women were found to be *EGFR* mutated. *KRAS* mutations and ALK+ IHC expression occurred more in HR+ NSCLC whereas *TP53* mutations occurred more in HR- NSCLC.

Conclusions: Women were more likely to have HR+ NSCLC than men and *EGFR* and *KRAS* mutations occurred more commonly in HR+ NSCLC. Additional studies with more strict inclusion criteria for HR+ are warranted to see if there is benefit to targeting HR in these subgroups.

Background

Lung cancer is the most common cause of cancer-related deaths in the United States; cigarette smoking is a major risk factor (1). The general incidence of lung cancer has been decreasing in both men and women largely due to decrease in the incidence of smoking, but there has been a minimal decline in the incidence of lung cancer in women (2, 3). Although smoking behaviors are similar between men and women today, historically men had higher prevalence of smoking than women resulting in higher incidence rates of lung cancer. As smoking declined in men, their rates of lung cancer declined precipitously. However, women have not experienced a decline of the same magnitude (4).

Some theories postulate that women have a higher sensitivity to adverse biological effects of smoking including a prevalence of tumor protein *p53* (*TP53*)/Kirsten rat sarcoma viral oncogene homolog (*KRAS*) co-mutations, higher levels of polycyclic aromatic hydrocarbons (PAH)-DNA adducts at any given level of smoking, and higher CYP1A1 expression (which encodes an enzyme used in the metabolism of PAHs) (5, 6). Women may have a higher exposure to passive smoking, which is a known risk factor for lung cancer (7). In addition, adenocarcinoma is more common in women especially never-smokers and the risk of adenocarcinoma decreases more slowly than other histologies (8). Consistent with this observation, women have more *EGFR* mutations than men (9). Other non-smoking risk factors for lung cancer include passive smoking (10), viral infections such as human papilloma virus (HPV) (11), low body mass index (BMI) (12), diet (significantly lower grain and carbohydrate consumption in patients with epidermal growth factor receptor (*EGFR*) mutations) (13), socioeconomic status (14), and exposures to arsenic, asbestos and radon (15).

The effect of estrogen on lung cancer pathogenesis is complex and not well understood, and current data linking the effect of estrogen and hormone replacement therapy (HRT) on the incidence of lung cancer is conflicting. Estrogen has two major receptors implicated in carcinogenesis of non-small cell lung cancer (NSCLC): estrogen receptor alpha (ER-a) and estrogen receptor beta (ER-b), both with high affinity for estradiol (16). It has been shown that ER-a, in the presence of estrogen, activates transcription, whereas ER-b inhibits transcription in presence of estrogen (16). Regarding the mechanism of estrogen and the carcinogenesis of lung cancer, some studies show that blocking ER can inhibit proliferation of NSCLC in mice while others have shown estrogen can reduce inflammatory cytokines which reduce the risk of NSCLC. Some have proposed that there may be a protective role with HRT in smokers due to the anti-inflammatory properties of estrogen by neutralizing the extra inflammation induced by smoking (14).

The role of progesterone receptors (PR) in the pathogenesis of NSCLC is unclear and there have been mixed results on the prognostic implications. Ishibashi et al. first looked at PR in NSCLC and showed that PR+ NSCLC was inversely associated with tumor node metastasis (TNM) stage and histology with better clinical outcomes in patients with PR+ status (17). A later study further showed that PR expression in tumor-surrounding stromal cells is associated with improved disease-specific survival and positive PR expression in tumor epithelial cells is associated with poor disease-specific survival in females (18). Yet, Raso et al. did not show any correlation between PR and patient clinicopathologic characteristics, which included histology, gender, tobacco history, and staging (19). We sought in our retrospective study to understand the relationship between HR status, gender, and the genomic landscape in NSCLC.

Methods

Tumor samples

The study included NSCLC tumor samples submitted to Caris Life Sciences (Phoenix, AZ) for analysis. This study was conducted in accordance with guidelines of the Declaration of Helsinki, Belmont report, and U.S. Common rule. In keeping with 45 CFR 46.101(b) (4), this study was performed utilizing retrospective, deidentified clinical data. Therefore, this study was considered IRB exempt and patient consent was not required.

Immunohistochemistry

Immunohistochemistry (IHC) was performed on formalin-fixed paraffin-embedded (FFPE) sections of glass slides. Slides were stained using automated staining techniques, per the manufacturer's instructions, and were optimized and validated per CLIA/CAO and ISO requirements. HR-positive (HR+) status was defined as ≥ 1+ and ≥ 1% nuclear staining of ER-a (SP1, Ventana) and/or PR (IE2, Ventana) by immunohistochemistry. ALK IHC status was determined using the Ventana ALK CDx Assay (D5F3,Ventana); ALK positivity was defined as 3+ in >1% of cells (20).

Next-generation sequencing

NGS was performed on genomic DNA isolated from FFPE tumor samples using the NextSeq platform (Illumina, Inc., San Diego, CA). Cases were either sequenced by the Illumina MiSeq hot spot 47 gene panel (n=2753) and Illumina NextSeq 592 gene panel (n=503). For tumors tested with MiSeq, specific regions of the genome were amplified using the Illumina TruSeq Amplicon Cancer Hotspot panel (21, 22). For NextSeq, a custom-designed SureSelect XT assay was used to enrich 592 whole-gene targets (Agilent Technologies, Santa Clara, CA) (23). All variants were detected with > 99% confidence based on allele frequency and amplicon coverage, with an average sequencing depth of coverage of > 500 and an analytic sensitivity of 5%. Prior to molecular testing, tumor enrichment was achieved by harvesting targeted tissue using manual microdissection techniques. Genetic variants were interpreted by molecular geneticists and categorized as "pathogenic," "presumed pathogenic," "pathogenic variant", "variant of unknown significance," "presumed benign" or "benign" according to the American College of Medical Genetics and Genomics (ACMG) standards. "Pathogenic", "presumed pathogenic", and "pathogenic variants" were counted as mutations whereas "benign", "presumed benign", and "variants of unknown significance" were excluded. Pan wild type tumors were defined as tumors that did not contain a "pathogenic," "presumed pathogenic," or "pathogenic variant" mutation.

Tumor mutational burden

TMB was measured (592 genes and 1.4 megabases [MB] sequenced per tumor) by counting all non-synonymous missense mutations found per tumor that had not been previously described as germline alterations. TMB analysis was available only for those tumors that were tested with the Illumnia NextSeq 592 gene panel NGS testing.

Statistical analyses

Standard descriptive statistics were used for this retrospective analysis. For dichotomous outcomes, Fisher's exact test was performed. For comparison of TMB, student's t-test was performed. Given the nature of multiple comparisons, p-values with multiple comparisons were further corrected using the Benjamini-Hochberg method and an adjusted p-value (q-value) of <0.05 was considered a significant difference. However, due to the exploratory nature of the investigation, multivariate analysis was not performed, only univariate analysis was performed. Statistical analyses were conducted using R (version 3.5.0) and Prism Graphpad (version 10.0.0).

Results

Baseline characteristics

3,256 NSCLC tumor samples submitted for molecular profiling between 2013-2018 were retrospectively identified and assessed for HR expression. There was nearly an even split between males (n=1629) and female (n=1627) NSCLC tumor samples. The mean age at collection of the sample was 64 years with an average age of 63.75 years in females and 65.08 years in males. By age group, 4.66% (n=152) samples were in patients ages ≤ 45, 42.62% (n=1388) samples were in patients ages 46-64, and 52.70% (n=1716) samples were in patients ages ≥ 65. (Table 1).

In terms of prevalence of mutations, TP53 was most commonly seen in 51.35% of tumor samples followed by $KRAS$ (26.44%) and $EGFR$ mutations (15.48%). In females, TP53 mutations were seen in 47.88% of patients followed by $KRAS$ mutations in 30.91% of patients and $EGFR$ mutations in 20.71% of patients. In males, TP53 mutations were seen in 54.82% of patients followed by $KRAS$ mutations in 21.98% of patients, and then $EGFR$ mutations in 10.25% of patients. (Table 1).

The overall mean TMB was 11.01 mutations/Mb among the 503 patients with TMB tested; the mean TMB was 10.58 mutations/Mb in females (n= 243) and 11.43 mutations/Mb in males (n=260). (Table 1).

Hormone receptor positivity in NSCLC

Hormone receptor positivity (HR+) was identified in 504/3256 (18.3%) of NSCLC tumors. By gender, HR+ occurred more commonly in women compared to men (19.6% vs 11.4%; p<0.0001, q<0.0001). (Table 2A) When stratified by age, women age≥65 were more likely than men age≥65 to have HR+ NSCLC (160/815, 19.6% vs. 95/901, 10.5%; p<0.0001, q<0.0001). In young patients (age ≤ 45), there was a trend towards increased likelihood in women (25/99, 25.3%) compared to men (6/53, 11.3%) (p= 0.0565, q = 0.0942). Among HR+ patients, women had a significantly greater prevalence of ER+ cases (255/318, 80.19% vs.132/186, 70.97%, p = 0.0216, q = 0.0432) while males trended towards having a greater prevalence of PR+ cases (71/172, 38.17% vs. 31.76%, p= 0.1457, q = 0.1457). By estrogen and progesterone receptor positivity, there was a trend towards women having a greater but not statistically significant percentage of ER-a+/PR- prevalence (217/318, 68.24%) compared to men (115/186, 61.83%), (p=0.1457, q= 0.2186) while men trended towards a higher

TABLE 1 Baseline Characteristics of study population.

Total	3256		
Female n (%)	1627 (49.96)		
Male n (%)	1629 (50.04)		
Age at Collection of Sample	Overall	Female	Male
Mean (SD)	64.41 (10.90)	63.75 (11.37)	65.08 (10.30)
Age ≤45 n (%)	152 (4.66)	99 (6.08)	53 (3.25)
Age 46-64 n (%)	1388 (42.62)	713 (43.82)	675 (41.44)
Age ≥ 65 n (%)	1716 (52.70)	815 (50.09)	901 (55.31)
HR+	n	%	
Total	504	18.31	
Female	318	19.55	
Male	186	11.42	
Prevalence of Mutations	Overall	Female	Male
EGFR n (%)	504 (15.48)	337 (20.71)	167 (10.25)
KRAS n (%)	861 (26.44)	503 (30.91)	358 (21.98)
TP53 n (%)	1672 (51.35)	779 (47.88)	893 (54.82)
ERBB2 n (%)	25 (0.77)	12 (0.74)	13 (0.80)
BRAF n (%)	134 (4.12)	84 (5.16)	50 (3.07)
ALK+ IHC n/total tested (%)	31/1052 (2.95)	23/502 (4.58)	8/550 (1.45)
TMB	Overall n=503	Female n=243	Male n= 260
	11.01 (8.85)	10.58 (7.44)	11.43 (9.99)

prevalence of ER-a-/PR+ cases (54/186, 29.03% vs. 63/318, 19.81%; p = 0.0216, q = 0.0648) (Table 2A, B).

EGFR in HR+ NSCLC

EGFR mutations were observed in 102/504 (20.2%) HR+ NSCLC tumors. *EGFR* mutations occurred more commonly in HR+ NSCLC than HR- NSCLC (102/504, 20.2% vs. 402/2752, 14.6%; p= 0.002, q = 0.007) (Table 3A, Figure 1). When stratified by gender, men with *EGFR* mutations were affected by HR status with a higher prevalence in HR+ (33/186, 17.7% vs. 134/1443, 9.3%; p = 0.0008, q=0.0056) while there was nearly equal incidence of EGFR mutations in HR+ and HR- females (HR+: 69/318, 21.7% vs. HR-: 268/1309, 20.47%; p =0.6436, q=0.7509).

When further stratified by age, men age ≥ 65 with HR+ NSCLC had significantly greater prevalence of *EGFR* mutations (14/95, 14.74%) compared to HR- NSCLC (p= 0.0124, q=0.0289) while women age ≥ 65 with HR+ NSCLC had similar prevalence of *EGFR* mutations (39/171, 24.38%) compared to HR- NSCLC (121/644,

20.15%) (p=0.2357, q= 0.4125) (Table 3A). There was a small sample size of HR+ young patients ≤ 45 years with *EGFR* mutations being tested (*n*= 31), but 10/25 (40.0%) young females had *EGFR* mutations. When we examined estrogen and progesterone receptor positivity in *EGFR* mutants, we observed a similar trend with about 60-65% of both females and males having ER-a+/PR- subtype. However, in *EGFR* wild type patients, females had a greater prevalence of ER-a+ cases (200/249, 80.32% vs. 106/153, 69.28%; p = 0.0157, q= 0.0314). (Table 3B).

In HR+ NSCLC, there was no significant difference in the prevalence of *EGFR* mutations (69/318, 21.70% vs. 33/186, 17.74%,; p = 0.3032, q=0.3826) in females. *EGFR* exon 19 deletions were the most common subtype observed at (42/102, 41.18%) followed by *EGFR* L858R mutations (27/102, 26.47%), and exon 20 insertions (9/102, 8.82%). There was also a small percentage of secondary T790M mutations (8/102, 7.84%). Males and females had a similar distribution amongst the *EGFR* mutation subtypes, but it is worth noting that 8 of the 9 *EGFR* Exon 20 insertions in HR+ NSCLC occurred in women. With regards to *EGFR TP53* comutations, there was a suggestive trend towards higher prevalence in women (42/69, 60.87% vs. 15/33, 45.45%, p= 0.2007; q= 0.2007) in HR+ NSCLC. Meanwhile, males had a greater prevalence of *EGFR TP53* comutations in HR- NSCLC compared to HR+ NSCLC (91/134, 67.91% vs. 15/33, 45.45%, p = 0.0254, q= 0.0762). (Table 4).

Other mutations in HR+ NSCLC

There was a significantly higher prevalence of *TP53* mutations in HR- NSCLC vs. HR+ NSCLC (53.89% vs. 42.54%; p<0.0001, q <0.0001); larger differences were seen in males (HR-: 57.72% vs. HR+ 40.22%; p<0.0001, q<0.0001) compared to females (HR-: 49.69% vs. HR+: 43.91%, p=0.0676, q = 0.1082) (Table 5, Figure 2A).

There was a higher prevalence of *KRAS* mutations in HR+ NSCLC (35.26% vs. 25.03%, p<0.0001, q <0.001), seen in both females (HR+: 38.80% vs. HR- 20.47%, p<0.0001, q<0.0001) and males (HR+: 29.19% vs. HR-: 21.29%, p<0.0001, q <0.0001) (Table 5, Figure 2B). Interestingly, when looking at *KRAS* and *TP53* comutations, there were higher percentage of *KRAS* and *TP53* comutations in HR- NSCLC, particularly in females (HR+: 40.65% vs. HR-: 54.10%, p<0.0001, q =<0.0001) (Table 5).

In samples tested for ALK IHC expression (n=1052), there was an increased proportion of ALK IHC expression in HR+ NSCLC compared with HR- (10/163, 6.13% vs. 21/889, 2.36%, p = 0.0190, q =0.0434). When stratified by gender, we saw a trend towards increased prevalence of ALK+ IHC in HR+ NSCLC versus HR-NSCLC in females (9/97, 9.28% vs. 14/405, 3.46%, p = 0.0259, q=0.0518) but not in males (1/66, 1.52% vs. 7/484, 1.45%, p>0.9999, q >0.9999) (Table 5, Figure 2C).

We did not see a significant difference in HR+ and HR- NSCLC in receptor tyrosine-protein kinase erbB-2 (*ERBB2*) and v-raf murine sarcoma oncogene homolog B1 (*BRAF*) mutations in our population (Table 5, Figures 2D, E).

TABLE 2 (A) Hormone receptor status total and percentage in NSCLC by gender and age. (B) Estrogen receptor/progesterone receptor status total and percentage by gender.

| A)	HR+	%	HR-	%	Total	p-value (female vs. male)	q-value (female vs. male)		
Total	504	18.31	2752	84.52	3256	<0.0001	<0.0001		
Female	318	19.55	1309	80.45	1627				
Male	186	11.42	1443	88.58	1629				
Age ≤45						p-value (female vs. male)	q-value (female vs. male)		
Female	25	25.25	74	74.75	99	0.0565	0.0942		
Male	6	11.32	47	88.68	53				
Age ≥65						p-value (female vs. male)	q-value (female vs. male)		
Female	160	19.63	655	80.37	815	<0.0001	<0.0001		
Male	95	10.54	806	89.46	901				
B)	ER +/PR+	%	ER +/PR-	%	ER-/PR+	%	Total	p-value (ER+ in female vs. male)	q-value (ER+ in female vs. male)
Total									
Female	38	11.95	217	68.24	63	19.81	318	0.0216	0.0432
Male	17	9.14	115	61.83	54	29.03	186		
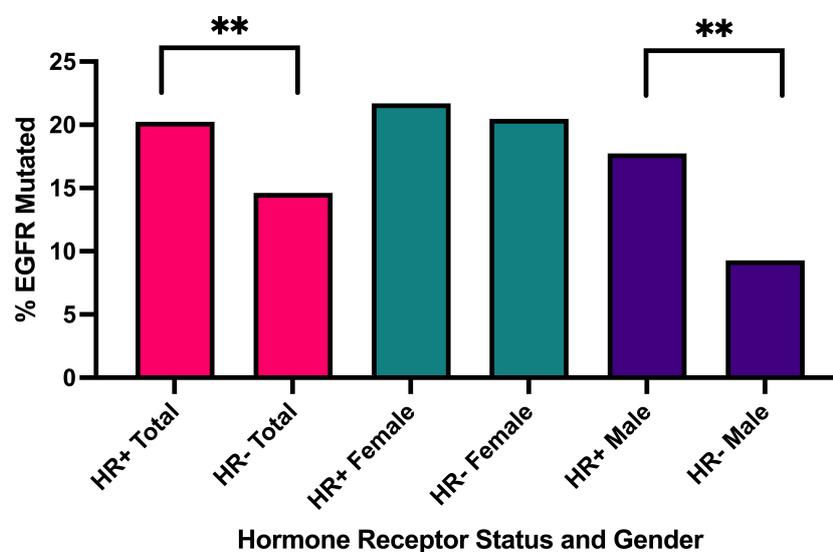
FIGURE 1
Percentage of *EGFR* mutated cases in all HR+ cases versus HR- cases, in all HR+ female cases versus HR- male cases, in all HR+ male cases versus HR- male cases. **q<0.001.

TABLE 3 (A) Hormone receptor status and *EGFR* mutations in NSCLC total and by gender (B) Estrogen receptor/progesterone receptor status total and percentage by gender.

A)	HR+		HR-		Total		p-value (HR + vs. HR-)	q-value (HR + vs. HR-)		
EGFR	HR+		HR-		Total					
EGFR mutated	102		402		504					
EGFR wildtype	402		2350		2752					
%	20.24		14.61		15.48		0.0020	0.0070		
Female	HR+		HR-		Total					
EGFR mutated	69		268		337					
EGFR wildtype	249		1041		1290					
%	21.70		20.47		20.71		0.6436	0.7509		
Female Age ≤ 45	HR+		HR-		Total					
EGFR mutated	10		22		32					
EGFR wildtype	15		52		67					
%	40.00		29.72		32.32		0.4584	0.6418		
Female Age ≥ 65	HR+		HR-		Total					
EGFR mutated	39		132		171					
EGFR wildtype	121		523		644					
	24.38		20.15		20.98		0.2357	0.4125		
Male	HR+		HR-		Total					
EGFR mutated	33		134		167					
EGFR wildtype	153		1309		1462					
%	17.74		9.29		10.25		0.0008	0.0056		
Male Age ≤ 45	HR+		HR-		Total					
EGFR mutated	1		7		8					
EGFR wildtype	5		40		45					
%	16.67		14.89		11.86		>0.9999	>0.9999		
Male Age ≥ 65	HR+		HR-		Total					
EGFR mutated	14		55		69					
EGFR wildtype	81		751		832					

(Continued)

TABLE 3 Continued

A)	HR+		HR-		Total	p-value (HR + vs. HR-)	q-value (HR + vs. HR-)		
%	14.74		6.82		7.66	0.0124	0.0289		
B)	ER +/PR +	%	ER +/PR-	%	ER-/ PR+	%	Total	p-value (ER+ in female vs. male)	q-value (ER+ in female vs. male)
EGFR Mutated									
Female	10	14.49	45	65.22	14	20.29	69	>0.9999	>0.9999
Male	6	18.18	20	60.61	7	21.21	33		
EGFR wildtype									
Female	28	11.24	172	69.08	49	19.68	249	0.0157	0.0314
Male	11	7.19	95	62.09	47	30.72	153		

TABLE 4 Hormone receptor status and *EGFR* subtypes by total and by gender.

EGFR subtypes	HR+	%	HR-	%	Total	p-value (HR+ vs. HR-)	q-value (HR+ vs. HR-)
Exon 19 del	42	41.18	154	38.31	196	0.6494	>0.9999
L858R	27	26.47	112	27.86	139	0.8056	>0.9999
Exon 20 ins	9	8.82	19	4.73	28	0.1423	>0.9999
G719X|S768I|L861Q	4	3.92	21	5.22	25	0.7991	>0.9999
Uncommon nonclassical mutations	12	11.76	69	17.16	81	0.2273	>0.9999
T790M	8	7.84	27	6.72	35	0.6655	>0.9999
Total	102		402		504		
EGFR subtypes Female	HR+	%	HR-	%	Total		
Exon 19 del	27	39.13	99	36.94	126	0.7808	>0.9999
L858R	19	27.54	76	28.36	95	>0.9999	>0.9999
Exon 20 ins	8	11.59	13	4.85	21	0.3599	>0.9999
G719X|S768I|L861Q	3	4.35	15	5.60	18	>0.9999	>0.9999
Uncommon nonclassical mutations	8	11.59	45	16.79	53	0.2273	>0.9999
T790M	4	5.80	20	7.46	24	0.7957	>0.9999
Total	69		268				
EGFR subtypes Male	HR+	%	HR-	%	Total		
Exon 19 del	15	45.45	55	41.04	70	0.6959	>0.9999
L858R	8	24.24	36	26.87	44	0.8288	>0.9999
Exon 20 ins	1	3.03	6	4.48	7	>0.9999	>0.9999
G719X|S768I|L861Q	1	3.03	6	4.48	7	>0.9999	>0.9999
Uncommon nonclassical mutations	4	12.12	24	17.91	28	0.6037	>0.9999
T790M	4	12.12	7	5.22	11	0.2304	>0.9999
Total	33		134		167		
EGFR TP53 comutation	HR+	%	HR-	%	Total		
Female	42	60.87	149	55.60	191	0.4963	0.5004
Male	15	45.45	91	67.91	106	0.0254	0.0762
Total	57	55.88	240	59.70	297	0.5004	0.5004

TABLE 5 Hormone receptor status and *KRAS, TP53*, ALK IHC+, *BRAF*, and *ERBB2* mutations by total and by gender.

	HR+		HR-		Total	p-value (HR+ vs. HR-)	q-value (HR+ vs. HR-)
TP53	HR+		HR-		Total		
TP53 mutated	211		1461		1672		
TP53 wildtype	285		1250		1535		
%	42.54		53.89		52.14	<0.0001	<0.0001
Indeterminate	8		41		49		
Female	HR+		HR-		Total		
TP53 mutated	137		642		779		
TP53 wildtype	175		650		825		
%	43.91		49.69		48.57	0.0676	0.1082
Indeterminate	6		17		23		
Male	HR+		HR-		Total		
TP53 mutated	74		819		893		
TP53 wildtype	110		600		710		
%	40.22		57.72		55.71	<0.0001	<0.0001
Indeterminate	2		24		26		
KRAS	HR+		HR-		Total		
KRAS mutated	177		684		861		
KRAS wildtype	325		2049		2374		
%	35.26		25.03		26.62	<0.0001	<0.0001
Indeterminate	2		19		21		
Female	HR+		HR-		Total		
KRAS mutated	123		268		391		
KRAS wildtype	194		1041		1235		
%	38.80		20.47		24.05	<0.0001	<0.0001
Indeterminate	1				1		
Male	HR+		HR-		Total		
KRAS mutated	54		304		358		
KRAS wildtype	131		1124		1255		
%	29.19		21.29		22.19	<0.0001	<0.0001
Indeterminate	1		15		16		
KRAS TP53 comutations		%		%			
Female	50	40.65	145	54.10	195		
Male	19	35.19	122	40.13	141	<0.0001	0.0165
ALK	HR+		HR-		Total		
ALK IHC+	10		21		31		
ALK IHC-	153		868		1021		
%	6.13		2.36		2.95	0.0190	0.0434
Unknown	341		1863		2204		

(Continued)

TABLE 5 Continued

	HR+		HR-		Total	p-value (HR+ vs. HR-)	q-value (HR+ vs. HR-)
Female	HR+		HR-		Total		
ALK IHC+	9		14		23		
ALK IHC-	88		391		479		
%	9.28		3.46		4.58	0.0259	0.0518
Unknown	221		904		1125		
Male	HR+		HR-		Total		
ALK IHC+	1		7		358		
ALK IHC-	65		477		1255		
%	1.52		1.45		22.19		
Indeterminate	120		959		1079	>0.9999	>0.9999
	HR+		**HR-**		**Total**		
ERBB2	HR+		HR-		Total		
ERBB2 mutated	7		18		25		
ERBB2 wildtype	491		2701		3192		
%	1.41		0.66		0.78	0.0937	0.1363
Indeterminate	6		33		39		
Female	HR+		HR-		Total		
ERBB2 mutated	5		7		12		
ERBB2 wildtype	310		1290		1600		
%	1.59		0.54		0.74	0.0659	0.1082
Indeterminate	3		12		15		
Male	HR+		HR-		Total		
ERBB2 mutated	2		11		13		
ERBB2 wildtype	181		1411		1592		
%	1.09		0.77		0.81	0.6525	0.7457
Indeterminate	3		21		24		
	HR+		**HR-**		**Total**		
BRAF	HR+		HR-		Total		
BRAF mutated	19		115		134		
BRAF wildtype	483		2627		3110		
%	3.78		4.19		4.13	0.8070	0.8608
Indeterminate	2		10		12		
Female	HR+		HR-		Total		
BRAF mutated	12		72		84		
BRAF wildtype	304		1309		1613		
%	3.80		5.21		4.95	0.3875	0.5167
Indeterminate	2		6		8		

(Continued)

TABLE 5 Continued

	HR+		HR-		Total	p-value (HR+ vs. HR-)	q-value (HR+ vs. HR-)
Male	HR+		HR-		Total		
BRAF mutated	7		43		50		
BRAF wildtype	179		1396		1575		
%	3.76		2.99		3.07	0.5022	0.6181
Indeterminate	0		4		4		

Prevalence of mutations in NSCLC cases ages ≤ 45

There were 152 NSCLC cases in our study population in which the age of collection was 45 years old or younger. Given the small sample size, we were not able to detect statistically significant differences, but we saw that females had a higher frequency of *EGFR* mutations overall (32/99, 32.32% vs. 8/53, 15.09%) and seen in both HR+ NSCLC (10/25, 40.00% vs. 1/6, 16.67%) and in HR- NSCLC (22/74, 29.73% vs. 7/47, 14.89%). Furthermore, there appeared to be a greater percentage of ALK+ IHC cases in females (4/22, 18.18% vs. 1/19, 5.26%), but very small sample sizes of ALK+ tested in HR+ (*n*= 6). (Table 6)
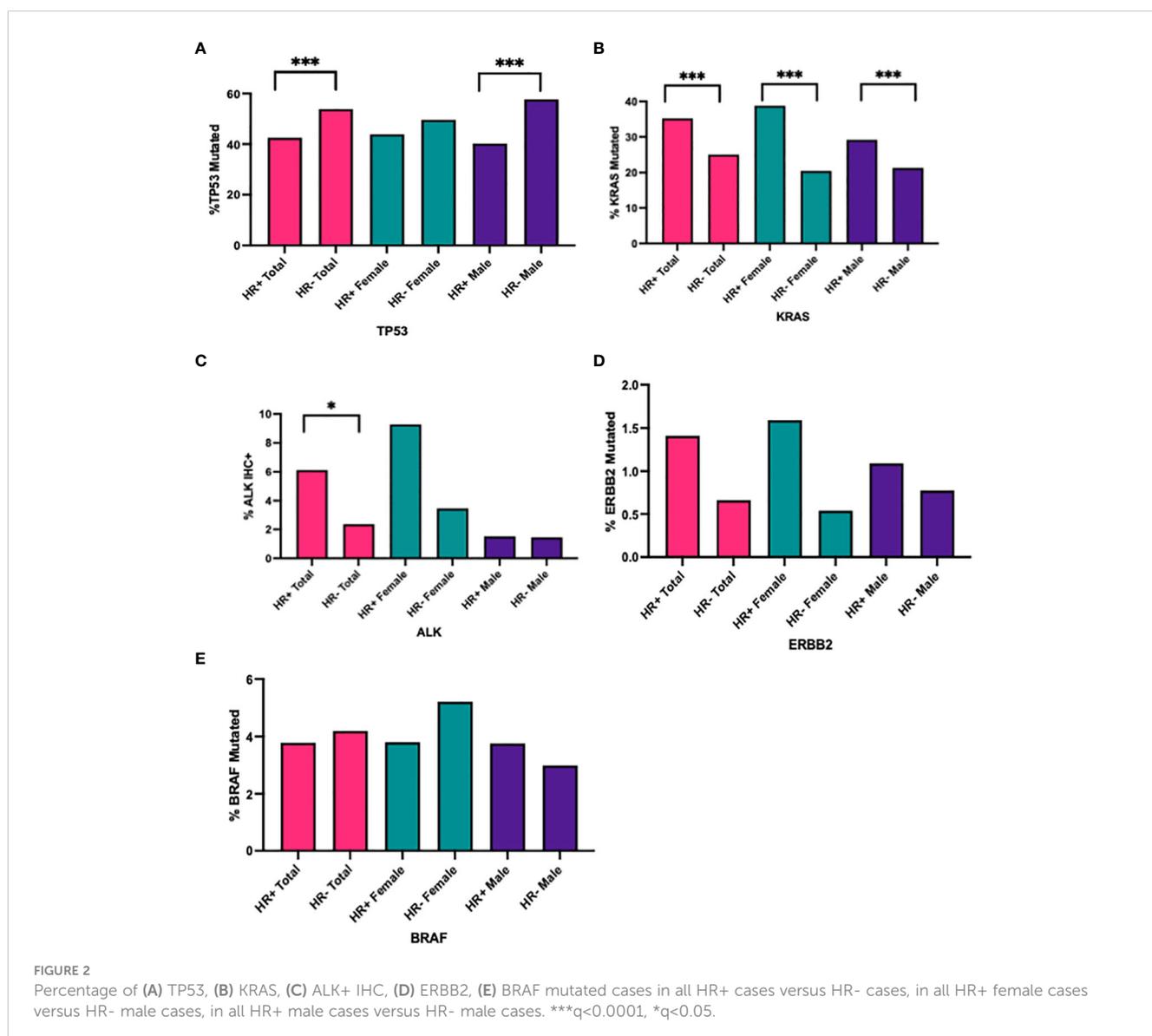

FIGURE 2
Percentage of **(A)** TP53, **(B)** KRAS, **(C)** ALK+ IHC, **(D)** ERBB2, **(E)** BRAF mutated cases in all HR+ cases versus HR- cases, in all HR+ female cases versus HR- male cases, in all HR+ male cases versus HR- male cases. ***q<0.0001, *q<0.05.

TABLE 6   Prevalence of mutations in patients age ≤45 by gender and by hormone receptor status.

Prevalence of Mutations in age ≤45	Overall, *n*=152	Female, *n*=99	Male, *n*=53	HR+ Female, *n*=25	HR+ Male, *n*=6	HR- Female, *n*= 74	HR- Male, *n* = 47
EGFR n (%)	40 (26.32)	32 (32.32)	8 (15.09)	10 (40.00)	1 (16.67)	22 (29.73)	7 (14.89)
KRAS n (%)	23 (15.13)	15 (15.15)	8 (15.09)	3 (12.00)	2 (33.33)	12 (16.21)	6 (12.77)
TP53 n (%)	80 (52.63)	48 (48.48)	32 (60.38)	8 (32.00)	2 (33.33)	40 (54.05)	30 (63.83)
ERBB2 n (%)	4 (2.63)	2 (2.02)	2 (3.77)	0 (0.00)	0 (0.00)	2 (2.7)	2 (4.26)
BRAF n (%)	3 (1.97)	1 (1.01)	2 (3.77)	0 (0.00)	0 (0.00)	1 (1.35)	2 (4.26)
ALK+ IHC *n*/total tested (%)	5/41 (12.20)	4/22 (18.18)	1/19 (5.26)	1/6 (16.67)	None tested	3/16 (18.75)	1/19 (5.26)

TMB analysis

The overall mean TMB was 11.01 mutations/Mb among the 503 patients with TMB tested; the mean TMB was 10.58 mutations/Mb in females (*n*= 243) and 11.43 mutations/Mb in males (*n*=260). In HR+ NSCLC, the mean TMB was 9.88 mutations/Mb while in HR- NSCLC the mean TMB was 11.20 mutations/Mb (p=0.2550, q=0.3385). By gender, females trended towards a higher TMB in HR

+ NSCLC (10.88 mutations/Mb vs. 8.27 mutations/Mb, p = 0.0891, q= 0.2163). (Table 7)

In our *KRAS* mutant NSCLC cases, the mean TMB was 11.17 mutations/Mb in HR+ cases versus 10.21 mutations/Mb in HR- cases (p=0.4478, q = 0.7165). Specifically by gender, females had a statistically significantly higher TMB (12.88 mutations/Mb vs. 7.28 mutations/Mb, p = 0.0056, q = 0.0448). There were no significant differences in TMB in *KRAS TP*53 co-mutated cases

TABLE 7   Tumor Mutational Burden by gender and in *KRAS* mutant cases.

TMB	HR+		p-value/q-value (female vs. male)	HR-		p-value/q-value (female vs. male)	Overall		p-value/q-value (female vs. male)	p-value/q-value (HR+ vs. HR-)
Overall Mean (SD)	9.88 (5.82)	*n*=68		11.20 (9.23)	*n*=435		11.01 (8.85)	*n*= 503	0.2821/ 0.3385	0.2550/ 0.3385
Male	8.27 (4.65)	*n*=26	0.0721/ 0.2163	11.78 (10.36)	*n*=234	0.1545/ 0.3090	11.43 (9.99)	*n*=260		0.0891/ 0.2163
Female	10.88 (6.29)	*n*=42		10.52 (7.67)	*n*=201		10.58 (7.44)	*n*=243		0.7740/ 0.7740
TMB in KRAS mutants										
KRAS mutants overall	11.17 (4.70)	*n*=23		10.21 (5.66)	*n*=109		10.37 (5.50)	*n*=132	0.1971/ 0.3942	0.4478/ 0.7165
Female	12.88 (4.5)	*n*=16	0.0056/ 0.0448	10.38 (6.02)	*n*=63	0.7156/ 0.7964	10.88 (5.81)	*n*=79		0.1259/ 0.2957
Male	7.28 (2.29)	*n*=7		9.98 (5.16)	*n*=46		9.62 (4.97)	*n*=53		0.1843/ 0.3942
*KRAS TP*53 comutated	12.75 (5.42)	*n*=12		12.20 (6.70)	*n*=44		12.32 (6.40)	*n*=56		0.7964/ 0.7964
KRAS mutant with no *TP*53 mutation	9.45 (3.17)	*n*=11		8.86 (4.40)	*n*=65		8.94 (4.26)	*n*=76		0.6704/ 0.7964

in HR+ and HR- nor in *KRAS* mutant cases with no *TP53* co-mutation. (Table 7)

Discussion

In our study, we found that a higher percentage of women have lung cancers that are hormone receptor positive and that among hormone receptor positive NSCLC patients, women had a significantly greater prevalence of ER-a positivity. Preclinical studies have examined ER and *EGFR* simultaneously and have found that estrogen through its receptor can stimulate lung cancer cell proliferation, resistance to cell death, angiogenesis, and metastasis (24). Epidemiological evidence is lacking. In a study in the Women's Health Initiative (WHI) there was no statistically significant association between HRT and the incidence of NSCLC (25). However, no similar investigation has been conducted with hormone based oral contraceptives. Clearly more observational research is needed to conclusively address this question.

We found that HR+ was associated with increased prevalence of *EGFR* mutations in NSCLC patients age≥65 and in males overall, which is interesting as patients with *EGFR* mutant lung cancer typically have a lower median age than the average age of U.S. lung cancer patients and seen more in females (26). Both estrogen signaling and *EGFR* signaling can promote proliferation by inducing tumor angiogenesis through vascular endothelial growth factor (*VEGF*) secretion and other growth factors (27). *EGFR* signaling activation increases the expression and activity of aromatase in NSCLC cells and estrogen can induce epidermal growth factor (*EGF*) production and activate *EGFR* signaling (24). Studies have shown a correlation between both ER-a and ER-b expression and the presence of *EGFR* mutations (19, 28). Further studies are needed to better understand the role of estrogen in older men, but it has been shown that among male patients with advanced NSCLC, those with high serum levels of free β-estradiol had significantly worse survival than those with lower β-estradiol so hormone therapy targeting β-estradiol may have benefit in older men (29).

Meanwhile in women, we did not see significance difference in prevalence of *EGFR* mutations in HR+ versus HR- cases and in women age ≥ 65 years. However, we saw a noticeably higher percentage of females ≤ 45 years with *EGFR* mutations with an even higher percentage (10/25, 40.0%) seen in HR+ NSCLC. In addition, females ≤ 45 years trended towards having higher prevalence of HR+ NSCLC compared to males. Young lung cancer patients have a different profile, as many young lung cancer patients are never smokers, have actionable mutations (most common being *ALK* and *EGFR*), and have predominantly adenocarcinoma histology (30, 31). Comparing between young women and men, the lung cancer incidence in young women has been more rapid than the incidence in young men with much of this driven by increases in adenocarcinoma incidence rates in women (32). Much of the reasoning for this remains unclear but strong family genetics may play a role in lifetime nonsmoking women being more suspectable to lung cancer (33). Plus, research has shown that female sex, age of diagnosis ≤ 60, and those with a family

history of cancer had lower DNA repair capacity so further understanding of DNA repair genes beyond *BRCA* may identify targets driving these increases (34).

Our study reflects these patterns, but also shows that HR+ NSCLC and HR+/*EGFR* mutated NSCLC are more common in young women. Comparisons between premenopausal and postmenopausal NSCLC women have shown that that adenocarcinoma is more prevalent in premenopausal women (35). In premenopausal women, estrogens are produced by their ovaries through ER-a and thus targeting ER-a may help aid in the treatment of lung cancer in young women (28).

Various hormonal markers and their association with NSCLC clinical outcomes have been previously investigated. High levels of circulating estrogen have been associated with worse survival both in women and men (36). Overexpression of aromatase leads to poor survival in postmenopausal women with NSCLC (37). ER-b overexpression has been shown to be a predictive factor of poor survival in women particularly when co-expressed with aromatase (38, 39). As our data shows higher prevalence of HR+ with *EGFR* mutations in older age NSCLC patients, future studies directed towards response to TKIs based on aromatase levels and specific ER receptor expression is warranted. Also, since the time period of our study, there have been new novel treatments in patients with *EGFR* Exon 20 insertions and given that 8 of our 9 HR+ *EGFR* Exon 20 insertion cases were female, it may be worth investigating the role of ER+ specifically with *EGFR* Exon 20 insertions (40, 41).

Our study also showed that *TP53* mutations were negatively associated with the presence of hormone receptors. This could be in part because the estrogen receptor positive tumors were more likely to be *EGFR* mutated and this subtype is less commonly associated with *TP53* mutations. Smoking has been associated with *TP53* mutations and not with *EGFR* associated cancers (42). However, *EGFR* and *TP53* co-mutations were seen at a similar prevalence in HR+ and HR- NSCLC, but they were more prevalent in females in our HR+ NSCLC population. *TP53* co-mutation with *EGFR* has conferred worse overall survival to first line *EGFR* TKI use in real world settings and our gender disparity findings in HR+ NSCLC suggest that further investigation is needed (43).

Multiple driver mechanisms and the impact of co-mutations has become increasingly recognized in NSCLC. We showed a significant prevalence in ALK IHC positive NSCLC in HR+ NSCLC compared to HR- NSCLC; this combination has not been studied much in lung cancer and may be worth further investigation first by evaluating HR+ in *ALK* fusion NSCLC. Our study also showed a significant increase in *KRAS* mutations in HR+ NSCLC yet a significant decrease in *KRAS TP53* co-mutations in HR+ NSCLC. Almotlak et al. showed in ER-b/*KRAS* mutant mice models that the combination of an ER-b blocker, fulvestrant, with a pan-HER tyrosine kinase inhibitor dacomitinib had a synergistic anti-tumor effect in treating ER-b positive lung cancer. Furthermore, they showed that sequential immunotherapy improved treatment response, suggesting that this combination may provide a novel approach for HR+ *KRAS* mutated NSCLC (44, 45). On further analysis incorporating TMB analysis, we saw a trend towards lower TMB in HR+ NSCLC but we saw a trend towards higher TMB in females in HR+ NSCLC and that HR+ *KRAS* mutant females

specifically had a significantly higher TMB in comparison to males. *KRAS* G12C mutations, which have therapeutic implications, are more seen in women with a younger median age and less of a smoking history (5, 46). As we see that *KRAS* mutant women in HR + NSCLC have significantly higher TMB but not in HR- NSCLC compared to men, there may be additional benefit incorporating hormone therapy in this subset. Future studies also evaluating *KRAS* mutant subtype, as never smokers are more likely to have G>A transition mutations, along with PD-L1 scores and *STK11/KEAP1* mutations maybe beneficial in better understanding higher incidence of *KRAS* mutations in HR+ NSCLC.

With regards to the therapeutic implications of our findings, there have been several studies of anti-estrogen therapy in lung cancer particularly looking at *EGFR* mutated NSCLC. Garon et al. conducted a phase II study looking at erlotinib with fulvestrant in advanced stage NSCLC and did not find a significant difference in overall response rate (ORR), progression free survival (PFS), or overall survival (OS) (47). Meanwhile another randomized phase II trial investigating *EGFR*-TKI naïve postmenopausal women with advanced lung cancer combining gefitinib with fulvestrant showed tolerability but did not show PFS benefit (48). However, it should be noted that these two studies did not limit enrollment to patients who were HR+ nor limit enrollment to patients with *EGFR* mutations.

Our study showed that men had a greater prevalence of PR+ overall. Little has been studied regarding anti-progesterone therapy in NSCLC, however, a recent preclinical study showed that PR contains a polyproline domain (PPD) that inhibits NSCLC cell proliferation and has a synergistic effect when given in combination with *EGFR* TKIs while another preclinical study demonstrated that progesterone can inhibit lung adenocarcinoma cell growth via membrane progesterone receptor alpha (49, 50). Further work targeting progesterone receptors should be considered particularly given that 23% of our HR+ NSCLC cases that were ER-a-/PR+.

The strength of our study was the large cohort of 3,256 NSCLC patients available for testing, compared to most other studies with much smaller sample sizes (17, 19, 38, 51). A limitation was that our markers were not directly comparable to other studies. For IHC of ER-a and PR in our study, we used Sp1 transcription factor (SP1) and Calnexin antibody (IE2) respectively and looked at nuclear staining. Other studies have used mouse monoclonal PAI-1 antibody (1D5), anti-estrogen alpha receptor (6F11), or rabbit polyclonal estrogen alpha receptor (HC-20) antibody clone or have not specified when checking for ER-a positivity and mouse anti-progesterone receptor (MAB429) has also been used to evaluate for PR+ (17, 19, 38, 51). Consequently, there have been large variation in detection rates; for example, a review of studies looking at ER-a positivity in NSCLC showed detection rates ranging from 0-97% (51). Another limitation was that our IHC panel only examined ER-a but not ER-b. Like ER-a, studies looking at ER-b have used different antibody clones (Tau antibody (H-150), anti-estrogen receptor beta antibody (14C8), and estrogen receptor beta 1 antibody (PPG5/10)) with varying percentages of detection from 19-98% looking at expression both in the nucleus and cytoplasm (51). Finally, our dataset did not have information on *KRAS*

mutation subtypes, PD-L1, or information on *STK11/KEAP1* which may be useful in better understanding the differences in the higher prevalence of *KRAS* mutations in HR+ NSCLC in both genders and HR+ *KRAS* mutant females having a significantly higher TMB than males. Thus, gauging absolute percentages of HR+ between studies should be cautioned and future studies should standardize the IHC being used to evaluate HR+ in NSCLC.

Further clinical trials in the future evaluating HR+ NSCLC with specific mutations should have more specific inclusion criteria regarding hormone positivity. For example, a Phase I trial recently investigating a combination treatment with aromatase inhibitor exemestane and a carboplatin-based therapy for postmenopausal women with advanced NSCLC showed a significant correlation between overall response rate with level of positive aromatase IHC expression (52). Also, in both NSCLC and breast cancer, there have been new novel agents since the previous phase II studies were completed. There are new TKIs not only in *EGFR* but for *ALK* rearrangements and in *KRAS* G12C (46, 53, 54). A new class of selective estrogen receptor degraders has shown promise in ER+/HER-2- breast cancer; a recent phase 3 trial on elacestrant showed significant benefit in patients with *ESR1* mutation versus standard of care and another phase 2 on camizestrant demonstrated superior PFS when compared to fulvestrant (55, 56). As a greater majority of NSCLC has ER-a expression, these new class of endocrine therapies in breast cancer focusing on *ESR1* mutations may hold promise in future studies in HR+ NSCLC (57).Thus, additional clinical trials with more selective inclusion parameters and investigation of new TKIs and estrogen modulator combinations should be investigated in HR+ NSCLC.

Data availability statement

The aggregate summarized Caris datasets generated during and/or analyzed during the current study can be requested from corresponding author on reasonable request. The deidentified sequencing data are owned by Caris Life Sciences. Qualified researchers can apply for access to these summarized data by contacting Joanne Xiu, PhD (jxiu@carisls.com) and signing a data usage agreement.

Ethics statement

Ethical approval was not required for the studies involving humans because In keeping with 45 CFR 46.101(b) (4), this study was performed utilizing retrospective, deidentified clinical data. Therefore, this study was considered IRB exempt and patient consent was not required. The studies were conducted in accordance with the local legislation and institutional requirements. The human samples used in this study were acquired from a by-product of routine care or industry. Written informed consent to participate in this study was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and the institutional requirements.

Author contributions

RH: Conceptualization, investigation, visualization, supervision, project administration, data curation, methodology, formal analysis, writing–original draft, writing–review and editing. DC: Conceptualization, investigation, data curation, writing–original draft, writing–review and editing. BX: Conceptualization, investigation, data curation, writing–original draft, writing–review and editing. RF: Conceptualization, investigation, data curation, software, formal analysis, writing–original draft,writing–review and editing. WC: Conceptualization, investigation, writing–original draft,writing–review and editing. LR: Investigation, writing—review and editing. HB: Investigation, writing—review and editing. CK: Investigation, writing—review and editing. MN: Investigation, writing—review and editing. HM: Investigation, writing—review and editing. AMV: Investigation, writing—review and editing. GL: Investigation, writing—review and editing. MS: Investigation, writing—review and editing. AJW: Investigation, writing—review and editing. AIS: Investigation, writing—review and editing. SVL: Investigation, writing—review and editing. JJN: Conceptualization, investigation, visualization, supervision, project administration, writing–original draft, writing–review and editing. All authors contributed to the article and approved the submitted version.

Funding

Acknowledgments

Conflict of interest

RH is a consultant for Targeted Oncology and received honorarium from DAVA Oncology and The Dedham Group outside of the submitted work. BX is an employee of Janssen Pharmaceuticals. RF is an employee of Caris Life Sciences. L.E.R. reports research funding from Genentech/Roche, Merck Serono, Boehringer Ingelheim, Novartis, Pfizer, Syndax, Loxo, Merck, Bristol Myers Squibb, Guardant Health, Heat Biologics, Amgen, Calithera Biosciences, Daiichi Sankyo/UCB Japan, NantHealth outside of the submitted work. HB consulting fees and advisory board involvement from BMS, Lilly, Genentech, Pfizer, Merck, EMD-Serono, Boehringer Ingelheim, AstraZeneca, Novartis, Genmab, Regeneron, BioNTech, Amgen, Axiom, PharmaMar, Takeda, Mirati, Daiichi, Guardant, Natera, Oncocyte, Beigene, iTEO, Jazz, Janssen, Da Volterra, Puma, BerGenBio, Bayer, and Iobiotech; reports being on the data and safety monitoring board of Takeda, Incyte, Novartis, Springworks, and University of Pennsylvania CAR-T program; reports stock options in Sonnetbio, Inspirna, and Nucleai; reports honoraria from Amgen, Pfizer, Daiichi, Regeneron; reports research support from BMS, Lilly, and Amgen; reports travel support from Amgen, BMS, Merck, Lilly, EMD-Serono, Genentech, and Regeneron outside of the submitted work. CK reports grants from AstraZeneca, BMS, Regeneron, Tesaro, Karyopharm, Debiopharm, Mirati, Genentech, Spectrum, and Merck; grants and personal fees from Novartis and Janssen; and personal fees from PierianDx, Sanofi, Diffusion, Mirati, Jazz Pharmaceuticals outside the submitted work. MN reports personal fees from AstraZeneca, Daiichi Sankyo, Takeda, Novartis, EMD Serono, Janssen, Pfizer, Caris Life Sciences, Blueprint Medicines, Regeneron, Mirati, and Lilly, and non-financial support from An Heart outside of the submitted work. H.M. reports personal fees from AstraZeneca and Genentech and research funding from AstraZeneca outside of the submitted work. AMV reports employment from Caris Life Sciences, personal fees from Bristol Myers Squibb, Caris Life Sciences, Compugen, ConcertoHealthAI, Elsevier, and Inivata; non-financial support from Roche/Genentech and AstraZeneca; and grants from Amgen, Merck, Replimune, EMD Serono, Immunomedics/Gilead outside the submitted work. GL reports stock in Lucence Diagnostics, Xilis, honoraria from Boehringer Ingelheim, Blueprint Medicines, AstraZeneca, Merck, consulting fees from Pfizer, AstraZeneca, research funding from Merck, EMD Serono, AstraZeneca, Blueprint Medicines, Tesaro, Bavarian Nordic, Novartis, G1 Therapeutics, Adaptimmune, BMS, GlaxoSmithKline, AbbVie, Rgenix, Pfizer, Roche, Genentech, Lilly, Janssen, Lucence, Silis, E.R. Squibb Sons, LLC, and non-financial support from Boehringer Ingelheim, Pfizer, E.R. Squibb Sons, LLC, Janssen, Seattle Genetics, Celgene, Ipsen, Pharmacyclics, Merck, AstraZeneca outside of the submitted work. M.A.S reports grants and personal fees from Genentech, AstraZeneca, Merck, and Novartis; personal fees from Jazz, Lilly, Mirati, Regeneron, Guardant, AbbVie, Blueprint, and Janssen; grants from Spectrum, BeiGene, and Daiichi Sankyo during the conduct of the study; grants and personal fees from Genentech, AstraZeneca, Merck, and Novartis outside of the submitted work. AJW reports personal fees from Premier, Beigene, Incyte, Novocure, Janssen, GlaxoSmithKline, Regeneron, and AstraZeneca; other support from BeyondSpring and Odronate outside the submitted work. AIS reports grants and personal fees from Amgen and Mirati; personal fees from Novartis, Pfizer, Merck, and AstraZeneca outside the submitted work. SVL reports personal fees from AstraZeneca, Beigene, Daiichi Sankyo, G1 Therapeutics, Guardant Health, Inivata, Janssen, Jazz Pharmaceuticals, PharmaMar, Regeneron, Takeda, Novartis, Eisai, Sanofi, Catalyst Pharmaceuticals, Candel Therapeutics, MSD Oncology, and Amgen; grants and personal fees from Blueprint, Bristol Myers Squibb, Genentech, Lilly, Merck, Pfizer, Elevation Oncology, Gilead Sciences, and Turning Point Therapeutics; grants from Alkermes, Bayer, Merus, Rain Therapeutics, and RAPT

outside of the submitted work. JJN reports personal fees from AstraZeneca, Naveris, AADi, Bioatla, Mindmed, stock in Epic Sciences, Cansera, Quantgene, Indee P/L, and research funding from Merck and Genentech outside of the submitted work.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Freedman ND, Leitzmann MF, Hollenbeck AR, Schatzkin A, Abnet CC. Cigarette smoking and subsequent risk of lung cancer in men and women: analysis of a prospective cohort study. *Lancet Oncol* (2008) 9:649–56. doi: 10.1016/S1470-2045 (08)70154-2

2. Jemal A, Miller KD, Ma J, Siegel RL, Fedewa SA, Islami F, et al. Higher lung cancer incidence in young women than young men in the United States. *New Engl J Med* (2018) 378:1999–2009. doi: 10.1056/NEJMoa1715907

3. Patel JD. Lung cancer in women. *J Clin Oncol* (2005) 23:3212–8. doi: 10.1200/ JCO.2005.11.486

4. Jemal A, Travis WD, Tarone RE, Travis L, Devesa SS. Lung cancer rates convergence in young men and women in the United States: Analysis by birth cohort and histologic type. *Int J Cancer* (2003) 105:101–7. doi: 10.1002/ijc.11020

5. Dogan S, Shen R, Ang DC, Johnson ML, D'Angelo SP, Paik PK, et al. Molecular epidemiology of EGFR and KRAS mutations in 3,026 lung adenocarcinomas: higher susceptibility of women to smoking-related KRAS-mutant cancers. *Clin Cancer Res* (2012) 18:6169–77. doi: 10.1158/1078-0432.CCR-11-3265

6. Luo K, Luo X, Cao W, Hochalter JB, Paiano V, Sipe CJ, et al. Cigarette smoking enhances the metabolic activation of the polycyclic aromatic hydrocarbon phenanthrene in humans. *Carcinogenesis* (2021) 42:570–7. doi: 10.1093/carcin/bgaa137

7. Bastian LA, Gray KE, DeRycke E, Mirza S, Gierisch JM, Haskell SG, et al. Differences in active and passive smoking exposures and lung cancer incidence between veterans and non-veterans in the women's health initiative. *Gerontologist* (2016) 56 Suppl 1:S102–11. doi: 10.1093/geront/gnv664

8. North CM, Christiani DC. Women and lung cancer: what is new? *Semin Thorac Cardiovasc Surg* (2013) 25:87–94. doi: 10.1053/j.semtcvs.2013.05.002

9. Bell DW, Brannigan BW, Matsuo K, Finkelstein DM, Sordella R, Settleman J, et al. Increased prevalence of EGFR-mutant lung cancer in women and in East Asian populations: analysis of estrogen-related polymorphisms. *Clin Cancer Res* (2008) 14:4079–84. doi: 10.1158/1078-0432.CCR-07-5030

10. Lee YJ, Cho BC, Jee SH, Moon JW, Kim SK, Chang J, et al. Impact of environmental tobacco smoke on the incidence of mutations in epidermal growth factor receptor gene in never-smoker patients with non–small-cell lung cancer. *J Clin Oncol* (2010) 28:487–92. doi: 10.1200/JCO.2009.24.5480

11. Toh C-K, Gao F, Lim W-T, Leong S-S, Fong K-W, Yap S-P, et al. Never-smokers with lung cancer: epidemiologic evidence of a distinct disease entity. *J Clin Oncol* (2006) 24:2245–51. doi: 10.1200/JCO.2005.04.8033

12. Zhu H, Zhang S. Body mass index and lung cancer risk in never smokers: a meta-analysis. *BMC Cancer* (2018) 18:635. doi: 10.1186/s12885-018-4543-y

13. Yamane Y, Iwasaki M, Kawase A, Tsuchihara K, Ishii G, Ohmatsu H, et al. Impact of dietary habits on epidermal growth factor receptor (EGFR) mutation status of Japanese patients with lung adenocarcinomas. *J Clin Oncol* (2011) 29:1567–7. doi: 10.1200/jco.2011.29.15_suppl.1567

14. DeRouen MC, Hu L, McKinley M, Gali K, Patel M, Clarke C, et al. Incidence of lung cancer histologic cell-types according to neighborhood factors: A population based study in California. *PloS One* (2018) 13:e0197146. doi: 10.1371/ journal.pone.0197146

15. Hubaux R, Becker-Santos DD, Enfield KS, Lam S, Lam WL, Martinez VD. Arsenic, asbestos and radon: emerging players in lung tumorigenesis. *Environ Health* (2012) 11:89. doi: 10.1186/1476-069X-11-89

16. Schwartz AG, Wenzlaff AS, Prysak GM, Murphy V, Cote ML, Brooks SC, et al. Reproductive factors, hormone use, estrogen receptor expression and risk of non–small-cell lung cancer in women. *J Clin Oncol* (2007) 25:5785–92. doi: 10.1200/ JCO.2007.13.3975

17. Ishibashi H, Suzuki T, Suzuki S, Niikawa H, Lu L, Miki Y, et al. Progesterone receptor in non–small cell lung cancer—A potent prognostic factor and possible target for endocrine therapy. *Cancer Res* (2005) 65:6450–8. doi: 10.1158/0008-5472.CAN-04-3087

18. Skjefstad K, Richardsen E, Donnem T, Andersen S, Kiselev Y, Grindstad T, et al. The prognostic role of progesterone receptor expression in non-small cell lung cancer patients: Gender-related impacts and correlation with disease-specific survival. *Steroids* (2015) 98:29–36. doi: 10.1016/j.steroids.2015.01.020

19. Raso MG, Behrens C, Herynk MH, Liu S, Prudkin L, Ozburn NC, et al. Immunohistochemical expression of estrogen and progesterone receptors identifies a subset of NSCLCs and correlates with EGFR mutation. *Clin Cancer Res* (2009) 15:5359–68. doi: 10.1158/1078-0432.CCR-09-0033

20. Thorne-Nuzzo T, Williams C, Catallini A, Clements J, Singh S, Amberson J, et al. A sensitive ALK immunohistochemistry companion diagnostic test identifies patients eligible for treatment with crizotinib. *J Thorac Oncol* (2017) 12:804–13. doi: 10.1016/ j.jtho.2017.01.020

21. Salem ME, Weinberg BA, Xiu J, El-Deiry WS, Hwang JJ, Gatalica Z, et al. Comparative molecular analyses of left-sided colon, right-sided colon, and rectal cancers. *Oncotarget* (2017) 8:86356–68. doi: 10.18632/oncotarget.21169

22. Illumina Inc. *TruSeq AmpliCon-Cancer Panel* . Available at: https://www.illumina.com/Documents/products/datasheets/datasheet_truseq_amplicon_cancer_panel.pdf (Accessed 18 July 2023).

23. Philip PA, Azar I, Xiu J, Hall MJ, Hendifar AE, Lou E, et al. Molecular characterization of *KRAS* wild-type tumors in patients with pancreatic adenocarcinoma. *Clin Cancer Res* (2022) 28:2704–14. doi: 10.1158/1078-0432.CCR-21-3581

24. Rodriguez-Lara V, Hernandez-Martinez J-M, Arrieta O. Influence of estrogen in non-small cell lung cancer and its clinical implications. *J Thorac Dis* (2018) 10:482–97. doi: 10.21037/jtd.2017.12.61

25. Chlebowski RT, Schwartz AG, Wakelee H, Anderson GL, Stefanick ML, Manson JE, 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* (2009) 374:1243–51. doi: 10.1016/S0140-6736(09)61526-9

26. Nieva J, Reckamp KL, Potter D, Taylor A, Sun P. Retrospective analysis of real-world management of EGFR-mutated advanced NSCLC, after first-line EGFR-TKI treatment: US treatment patterns, attrition, and survival data. *Drugs Real World Outcomes* (2022) 9:333–45. doi: 10.1007/s40801-022-00302-w

27. Le X, Nilsson M, Goldman J, Reck M, Nakagawa K, Kato T, et al. Dual EGFR-VEGF pathway inhibition: A promising strategy for patients with EGFR-mutant NSCLC. *J Thorac Oncol* (2021) 16:205–15. doi: 10.1016/j.jtho.2020.10.006

28. Hsu L-H, Chu N-M, Kao S-H. Estrogen, estrogen receptor and lung cancer. *Int J Mol Sci* (2017) 18:1–17. doi: 10.3390/ijms18081713

29. Ross H, Oldham FB, Bandstra B, Sandalic L, Bianco J, Bonomi P, et al. Serum-free estradiol (E2) levels are prognostic in men with chemotherapy-naive advanced non-small cell lung cancer (NSCLC) and performance status (PS) 2. *J Clin Oncol* (2007) 25:7683–3. doi: 10.1200/jco.2007.25.18_suppl.7683

30. Siegel DA, Fedewa SA, Henley SJ, Pollack LA, Jemal A. Proportion of never smokers among men and women with lung cancer in 7 US States. *JAMA Oncol* (2021) 7:302–4. doi: 10.1001/jamaoncol.2020.6362

31. Gitlitz BJ, Novello S, Vavalà T, Bittoni M, Sable-Hunt A, Pavlick D, et al. The genomics of young lung cancer: comprehensive tissue genomic analysis in patients under 40 with lung cancer. *JTO Clin Res Rep* (2021) 2:100194. doi: 10.1016/ j.jtocrr.2021.100194

32. Fidler-Benaoudia MM, Torre LA, Bray F, Ferlay J, Jemal A. Lung cancer incidence in young women *vs* . young men: A systematic analysis in 40 countries. *Int J Cancer* (2020) 147:811–9. doi: 10.1002/ijc.32809

33. Wu AH, Fontham ETH, Reynolds P, Greenberg RS, Buffler P, Liff J, et al. Family history of cancer and risk of lung cancer among lifetime nonsmoking women in the United States. *Am J Epidemiol* (1996) 143:535–42. doi: 10.1093/ oxfordjournals.aje.a008783

34. Wei Q, Cheng L, Amos CI, Wang LE, Guo Z, Hong WK, et al. Repair of tobacco carcinogen-induced DNA adducts and lung cancer risk: a molecular epidemiologic study. *J Natl Cancer Inst* (2000) 92:1764–72. doi: 10.1093/jnci/92.21.1764

35. Oton AB, Belani C, Cai C, Owonikoko T, Gooding W, Siegfried J, et al. Comparison of survival for non-small cell lung cancer (NSCLC) between premenopausal and postmenopausal women: An analysis of the National Surveillance, Epidemiology and End Results (SEER) Database. *J Clin Oncol* (2006) 24:7038–8. doi: 10.1200/jco.2006.24.18_suppl.7038

36. Rodriguez-Lara V, Peña-Mirabal E, Baez-Saldaña R, Esparza-Silva AL, García-Zepeda E, Cerbon Cervantes MA, et al. Estrogen receptor beta and CXCR4/CXCL12 expression: differences by sex and hormonal status in lung adenocarcinoma. *Arch Med Res* (2014) 45:158–69. doi: 10.1016/j.arcmed.2014.01.001

37. Mah V, Marquez D, Alavi M, Maresh EL, Zhang L, Yoon N, et al. Expression levels of estrogen receptor beta in conjunction with aromatase predict survival in non-small cell lung cancer. *Lung Cancer* (2011) 74:318–25. doi: 10.1016/j.lungcan.2011.03.009

38. Stabile LP, Dacic S, Land SR, Lenzner DE, Dhir R, Acquafondata M, et al. Combined analysis of estrogen receptor beta-1 and progesterone receptor expression identifies lung cancer patients with poor outcome. *Clin Cancer Res* (2011) 17:154–64. doi: 10.1158/1078-0432.CCR-10-0992

39. Deng F, Li M, Shan W-L, Qian L-T, Meng S-P, Zhang X-L, et al. Correlation between epidermal growth factor receptor mutations and the expression of estrogen receptor-β in advanced non-small-cell lung cancer. *Oncol Lett* (2017) 13:2359–65. doi: 10.3892/ol.2017.5711

40. Park K, Haura EB, Leighl NB, Mitchell P, Shu CA, Girard N, et al. Amivantamab in EGFR exon 20 insertion-mutated non-small-cell lung cancer progressing on platinum chemotherapy: initial results from the CHRYSALIS phase I study. *J Clin Oncol* (2021) 39:3391–402. doi: 10.1200/JCO.21.00662

41. Zhou C, RaMalingam SS, Kim TM, Kim S-W, Yang JC-H, Riely GJ, et al. Treatment outcomes and safety of mobocertinib in platinum-pretreated patients with EGFR exon 20 insertion-positive metastatic non-small cell lung cancer: A phase 1/2 open-label nonrandomized clinical trial. *JAMA Oncol* (2021) 7:e214761. doi: 10.1001/jamaoncol.2021.4761

42. Gibbons DL, Byers LA, Kurie JM. Smoking, p53 mutation, and lung cancer. *Mol Cancer Res* (2014) 12:3–13. doi: 10.1158/1541-7786.MCR-13-0539

43. Le X, Molife C, Leusch MS, Rizzo MT, Peterson PM, Caria N, et al. TP53 co-mutation status association with clinical outcomes in patients with EGFR-mutant non-small cell lung cancer. *Cancers (Basel)* (2022) 14:1–17. doi: 10.3390/cancers14246127

44. Almotlak AA, Farooqui M, Siegfried JM. Inhibiting pathways predicted from a steroid hormone gene signature yields synergistic antitumor effects in NSCLC. *J Thorac Oncol* (2020) 15:62–79. doi: 10.1016/j.jtho.2019.09.195

45. Almotlak AA, Farooqui M, Soloff AC, Siegfried JM, Stabile LP. Targeting the ERβ/HER oncogenic network in KRAS mutant lung cancer modulates the tumor microenvironment and is synergistic with sequential immunotherapy. *Int J Mol Sci* (2021) 23:1–21. doi: 10.3390/ijms23010081

46. Skoulidis F, Li BT, Dy GK, Price TJ, Falchook GS, Wolf J, et al. Sotorasib for lung cancers with *KRAS* p.G12C mutation. *New Engl J Med* (2021) 384:2371–81. doi: 10.1056/NEJMoa2103695

47. Garon EB, Siegfried JM, Stabile LP, Young PA, Marquez-Garban DC, Park DJ, et al. Randomized phase II study of fulvestrant and erlotinib compared with erlotinib alone in patients with advanced or metastatic non-small cell lung cancer. *Lung Cancer* (2018) 123:91–8. doi: 10.1016/j.lungcan.2018.06.013

48. Mazieres J, Barlesi F, Rouquette I, Molinier O, Besse B, Monnet I, et al. Randomized phase II trial evaluating treatment with EGFR-TKI associated with antiestrogen in women with nonsquamous advanced-stage NSCLC: IFCT-1003 LADIE trial. *Clin Cancer Res* (2020) 26:3172–81. doi: 10.1158/1078-0432.CCR-19-3056

49. Xiao J, Chen X, Lu X, Xie M, He B, He S, et al. Progesterone/Org inhibits lung adenocarcinoma cell growth via membrane progesterone receptor alpha. *Thorac Cancer* (2020) 11:2209–23. doi: 10.1111/1759-7714.13528

50. Kaewjanthong P, Sooksai S, Sasano H, Hutvagner G, Bajan S, McGowan E, et al. Cell-penetrating peptides containing the progesterone receptor polyproline domain inhibits EGF signaling and cell proliferation in lung cancer cells. *PloS One* (2022) 17:e0264717. doi: 10.1371/journal.pone.0264717

51. Kawai H. Estrogen receptors as the novel therapeutic biomarker in non-small cell lung cancer. *World J Clin Oncol* (2014) 5:1020. doi: 10.5306/wjco.v5.i5.1020

52. Young PA, Márquez-Garbán DC, Noor ZS, Moatamed N, Elashoff D, Grogan T, et al. Investigation of combination treatment with an aromatase inhibitor exemestane and carboplatin-based therapy for postmenopausal women with advanced NSCLC. *JTO Clin Res Rep* (2021) 2:100150. doi: 10.1016/j.jtocrr.2021.100150

53. Solomon BJ, Besse B, Bauer TM, Felip E, Soo RA, Camidge DR, et al. Lorlatinib in patients with ALK-positive non-small-cell lung cancer: results from a global phase 2 study. *Lancet Oncol* (2018) 19:1654–67. doi: 10.1016/S1470-2045(18)30649-1

54. Hoy SM. Brigatinib: A review in ALK-inhibitor naïve advanced ALK-positive NSCLC. *Drugs* (2021) 81:267–75. doi: 10.1007/s40265-020-01449-y

55. Oliveira M, Pominchuk D, Nowecki Z, Hamilton E, Kulyaba Y, Andabekov T, et al. Camizestrant, a next-generation oral SERD vs. fulvestrant in post-menopausal women with advanced ER-positive HER2-negative breast cancer: Results of the randomized, multi-dose Phase 2 SERENA-2 trial. *San Antonio Breast Cancer Symposium San Antonio* (2022). doi: 10.1158/1538-7445.SABCS22-GS3-02

56. Bidard F-C, Kaklamani VG, Neven P, Streich G, Montero AJ, Forget F, et al. Elacestrant (oral selective estrogen receptor degrader) Versus Standard Endocrine Therapy for Estrogen Receptor-Positive, Human Epidermal Growth Factor Receptor 2-Negative Advanced Breast Cancer: Results From the Randomized Phase III EMERALD Trial. *J Clin Oncol* (2022) 40:3246–56. doi: 10.1200/JCO.22.00338

57. Kawai H, Ishii A, Washiya K, Konno T, Kon H, Yamaya C, et al. Estrogen receptor alpha and beta are prognostic factors in non-small cell lung cancer. *Clin Cancer Res* (2005) 11:5084–9. doi: 10.1158/1078-0432.CCR-05-0200