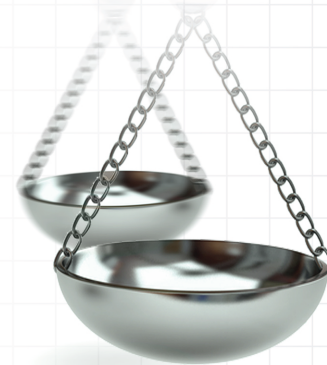




Comparative effectiveness from a single-arm trial and real-world data: alectinib versus ceritinib

Jessica Davies^{*,1}, Michael Martinec², Paul Delmar², Mathieu Coudert³, Walter Bordogna², Sophie Golding², Reynaldo Martina⁴ & Gracy Crane¹

¹Roche Products Ltd, 6 Flacon Way, Shire Park, Welwyn Garden City, AL7 1TW, UK

²F. Hoffmann-La Roche Ltd, Basel, Switzerland

³F. Hoffmann-La Roche Ltd, Boulogne-Billancourt Cedex, France

⁴Department of Biostatistics, University of Liverpool, Liverpool, UK

*Author for correspondence: Tel.: +44 7584 587 922; Jessica.davies@roche.com

Aim: To compare the overall survival of *anaplastic lymphoma kinase*-positive non-small-cell lung cancer patients who received alectinib with those who received ceritinib. **Materials & methods:** Two treatment arms (alectinib [n = 183] and ceritinib [n = 67]) were extracted from clinical trials and an electronic health record database, respectively. Propensity scores were applied to balance baseline characteristics. Kaplan–Meier and multivariate Cox regression were conducted. **Results:** After propensity score adjustment, baseline characteristics were balanced. Alectinib had a prolonged median overall survival (alectinib = 24.3 months and ceritinib = 15.6 months) and lower risk of death (hazard ratio: 0.65; 95% CI: 0.48–0.88). **Conclusion:** Alectinib was associated with prolonged overall survival versus ceritinib, which is consistent with efficacy evidence from clinical trials.

First draft submitted: 12 April 2018; Accepted for publication: 6 June 2018; Published online: 26 June 2018

Keywords: comparative effectiveness • non-small-cell lung cancer • real-world evidence

In health technology assessments, it has become increasingly important to demonstrate comparative effectiveness for new therapies, in addition to the traditional licensing requirements of quality of manufacture, safety and efficacy [1]. However, the race to license new therapies, especially for rare biomarker-defined populations such as *anaplastic lymphoma kinase*-positive non-small-cell lung cancer (*ALK*+ NSCLC), has meant that regulatory approval using single-arm Phase II studies can precede the availability of traditional, comparative evidence using randomized controlled trials (RCTs), which can take time to recruit rare biomarker-defined populations [2,3]. In addition, the highly segmented treatment landscape for *ALK* inhibitors has meant that it is not always possible to perform head-to-head comparison trials using the standard of care that can provide the best comparative evidence.

In the absence of comparative evidence from RCTs, external controls are often used to bridge the gap of providing comparative evidence using direct adjusted comparisons [4–7]. However, as these types of studies typically do not use individual patient-level data (IPD), they are susceptible to systematic variation or bias. Therefore, innovation in methods of gathering and assessing comparative evidence is needed in order to accelerate reimbursement decisions for new therapies in rarer cancers.

To overcome this problem, our study used US IPD from the Flatiron Health electronic health record (EHR) database [8] to create an external real-world data (RWD) control arm that provides comparative evidence. The external RWD control and two single-arm Phase II alectinib trials (NP28673 and NP28761) were retrospectively analyzed to compare overall survival (OS) in patients with *ALK*+ advanced NSCLC treated with alectinib versus ceritinib following crizotinib failure.

Materials & methods

Study population & data sources

The population was derived from IPD of adult patients (≥ 18 years) diagnosed with locally advanced (American Joint Committee on Cancer stage IIIB, not amenable to curative therapy) or metastatic (American Joint Committee on Cancer Stage IV) NSCLC, with a documented *ALK* rearrangement, and who had experienced disease progression on crizotinib, two treatment arms were constructed: alectinib and ceritinib. The alectinib treatment arm was derived by pooling data from patients enrolled in two Phase II studies (Global study NP28673 [NCT01801111] [9]; and North American study NP28761 [NCT01871805] [10]). To derive the ceritinib arm, IPD were extracted by applying inclusion and exclusion criteria from the NP28673 and NP28761 clinical trials to the electronic health record-derived database (Flatiron Health). After blending the two treatment cohorts, additional exclusion criteria were applied to address the imbalances between the ceritinib and alectinib cohorts in stage of diagnosis and treatment with crizotinib postprogression (further details can be found in the Supplementary Materials).

Statistical analysis

The primary end point was OS, defined as the time from the date of initiation of alectinib or ceritinib until death from any cause. Patients were censored on their date of last visit, if they were still on their previous treatment at the end of the study period or if their death could not be confirmed.

To reduce indication bias, a logit propensity score was estimated by regressing treatment assignment (alectinib vs ceritinib) on key prognostic factors that were previously identified and similarly defined across populations (age, gender, race, stage at initial diagnosis and prior lines of therapy [11,12]). The propensity score was applied via inverse probability treatment weighting with stabilized weights, the most general approach for the primary analysis. Trimming was applied preweighting to include only patients with an overlapping distribution of propensity scores [13,14]. Graphical display of the propensity score was used to assess the distribution pre- and post-weighting. Standardized mean differences were computed for all covariates in the unadjusted (no propensity scores) and adjusted sample to assess the balance of the cohort. Any difference observed greater than 0.1 pooled standard deviation units indicated an imbalance [11,13].

Median OS and corresponding 95% CIs were provided for each of the treatment arms and compared with a log-rank test. A doubly robust approach was taken with a multivariate Cox proportional hazard model to estimate the average treatment effect of alectinib on the reweighted cohort.

Sensitivity & secondary analyses

Multiple sensitivity analyses were conducted to assess the robustness of the primary analysis. The following cohorts were used: patients with follow-up time in the top tenth percentile excluded; missing covariates imputed (race and stage at diagnosis); CNS metastases as a prognostic factor; patients with one to three lines of prior treatment (patients with more than three lines of treatment were excluded); and age as a continuous variable. A secondary matched analysis was performed through Genetic Matching (GenMatch), by applying an evolutionary search algorithm that maximized the balance of observed confounders across the alectinib and ceritinib treatment arms [15]. A naïve comparison with an independent ceritinib clinical trial cohort (ASCEND-2, NCT01685060 [16]) was also conducted to evaluate the generalizability of the RWD outcomes to a clinical trial cohort [17]. Further description can be found in the Supplementary Materials.

Results

The primary analysis included a pooled cohort of 183 patients who received alectinib from the two Phase II studies, and 67 patients who received ceritinib from the Flatiron database (Figure 1).

Demographic & clinical characteristics

Prior to adjustment, patients in the alectinib arm were younger, more heavily pretreated and had a higher prevalence of CNS metastases than the ceritinib arm. Both treatment cohorts were more likely to be female, Caucasian, have adenocarcinoma histology, have an initial diagnosis of stage IV and no history of smoking (Table 1). After reweighting, all covariates were balanced with standardized mean differences below the 10% threshold. The highest difference that remained was prior lines of treatment (6.4%) (Table 2). The balance was further confirmed upon visual graphical inspection; the distribution of the propensity score by treatment arm was comparable after reweighting (Figure 2).

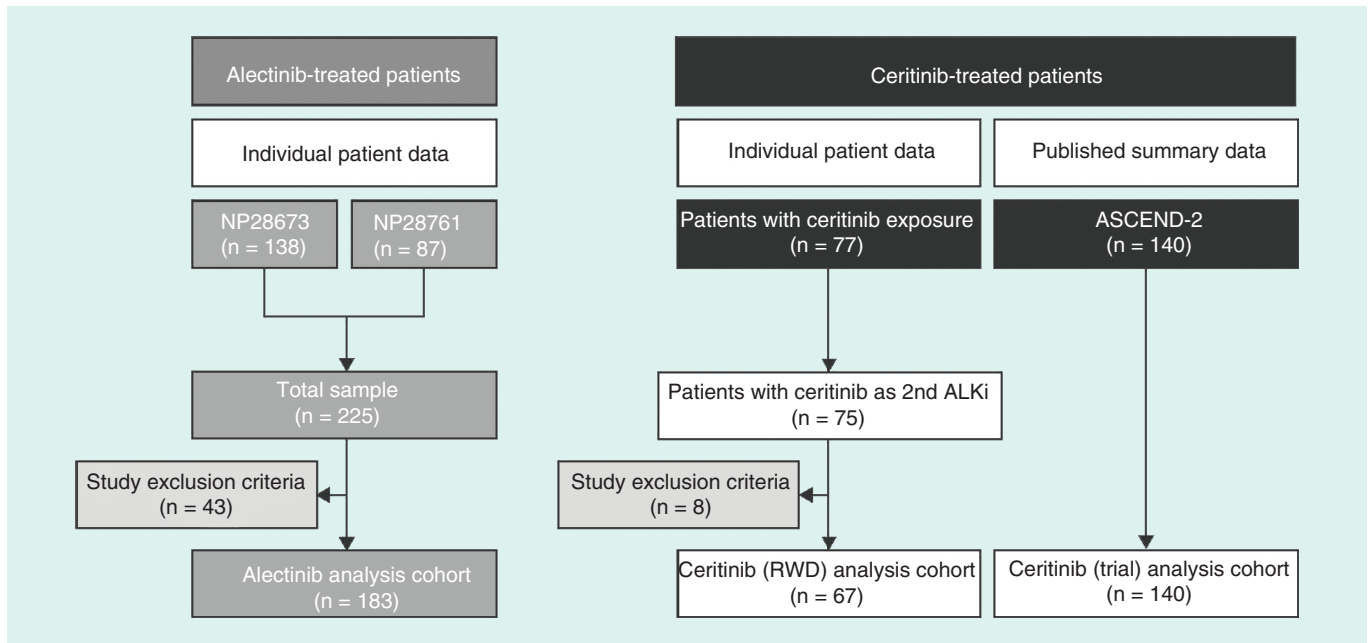


Figure 1. Cohort attrition.

ALKi: ALK inhibitor; RWD: Real-world data.

Treatment effectiveness

The median follow-up was 23 months for the alectinib arm and 12 months for the ceritinib arm. The crude median OS of the unweighted alectinib arm (n = 183) was 26 months (95% CI: 21–not reached [NR]) and 16 months (n = 67; 95% CI: 12–NR) for the unweighted ceritinib arm. Median OS of the reweighted cohorts was 24 months (95% CI: 21–NR) for alectinib (n = 183) and 16 months (n = 64; 95% CI: 16–19) for ceritinib (Figure 3). The multivariate Cox model showed that alectinib was associated with a significantly lower risk of death compared with ceritinib (hazard ratio [HR]: 0.65; 95% CI: 0.48–0.88; p = 0.006) (Figure 4). Female, nonwhite race, younger age and one prior treatment were also associated with longer survival (Figure 4). The estimated comparative effectiveness of alectinib versus ceritinib was consistent across the sensitivity analyses with an adjusted HR ranging from 0.53 (95% CI: 0.40–0.70) to 0.67 (95% CI: 0.49–0.92) (Table 3).

Naïve comparison with ceritinib trial

Similar imbalances observed between the alectinib arm and ceritinib arm were also observed between patients enrolled in a ceritinib trial group (ASCEND-2) and the ceritinib RWD arm (Supplementary Table 1). Median OS was consistent between ceritinib (RWD) (16 months, 95% CI: 12–NR) and ceritinib (trial) (15 months, 95% CI: 14–NR; Figure 5A). Prolonged OS compared with ceritinib (trial) was observed in the unweighted alectinib cohort (median 26 months, 95% CI: 21–NR; Figure 5B).

Subsequent treatments

Subsequent therapies were allowed for both treatment groups. At data cut-off, 56 patients in the alectinib arm were still receiving alectinib treatment and four patients in the ceritinib arm were still receiving ceritinib treatment. In total, 53 patients (42%) in the alectinib arm received at least one subsequent therapy; 32 patients (25%) received ceritinib and three patients (2%) were retreated with crizotinib. In the ceritinib arm, 34 patients (54%) received at least one subsequent therapy; eight patients (12%) received alectinib and two patients (3%) were retreated with crizotinib.

Discussion

Although not a substitute for RCTs, comparisons with RWD can bridge evidence gaps when the standard of care has shifted quickly, or when RCT evidence is not available, as in the case of the crizotinib failure (CF) setting in

Table 1. Patient characteristics by treatment cohort.

Characteristic	Alectinib (n = 183)	Ceritinib RWD (n = 67)	Total (n = 250)	p-value
Mean age (SD) ^{†,‡}	52.5 (± 11.2)	59.8 (± 11.4)	54.5 (± 11.7)	< 0.0001
Age group, n (%) ^{†,§} :				
– <65 years	160 (87)	41 (61)	201 (80)	< 0.0001
– ≥65 years	23 (13)	26 (39)	49 (20)	
Gender, n (%) [†] :				
– Male	85 (47)	30 (45)	115 (46)	0.89
– Female	98 (54)	37 (55)	135 (54)	
Race, n (%) [†] :				
– White	133 (73)	49 (73)	182 (73)	1
– Other	50 (27)	18 (27)	68 (27)	
ACA histology, n (%) [§]	175 (96)	61 (91)	236 (94)	0.029
ECOG PS, n (%) [†] :				
– 0	64 (35)	8 (12)	72 (29)	0.11
– 1	101 (55)	8 (12)	109 (44)	
– 2	18 (10)	5 (7)	23 (9)	
Missing	0	46 (69)	46 (18)	
Stage at diagnosis, n (%) ^{†,§} :				
– IIIB	13 (7)	8 (12)	21 (8)	0.3
– IV	170 (93)	59 (88)	229 (92)	
CNS metastasis, n (%) ^{†,§} :				
– Yes	112 (61)	23 (34)	135 (54)	0.0002
– No	71 (39)	44 (66)	115 (46)	
History of smoking, n (%) [§] :				
– Yes	62 (34)	30 (45)	92 (37)	0.14
– No	121 (66)	37 (55)	158 (63)	
Prior chemotherapy, n (%)	136 (74)	19 (28)	155 (62)	< 0.0001
Prior radiation, n (%)	84 (46)	21 (31)	105 (42)	0.043
Prior lines of treatment, n (%)				
– 1	52 (28)	38 (57)	90 (36)	< 0.0001
– 2	66 (36)	20 (30)	86 (34)	
– ≥3	65 (36)	9 (13)	74 (30)	
Range of prior lines	(1–8)	(1–5)	(1–8)	

[†]SD between treatment cohorts (p < 0.05).
[‡]Wilcoxon rank-sum test.
[§]χ² test.
ACA: Adenocarcinoma; ECOG PS: Eastern Cooperative Oncology Group performance status; RWD: Real-world data; SD: Standard deviation.

patients with *ALK*+ NSCLC. Our analysis provides insight into the treatment effect among second generation *ALK*-inhibitors in the CF setting.

Prior to reweighting, patient characteristics were imbalanced. The ceritinib arm had a higher median age compared with the ceritinib trial arm, most likely due to the difference in patient selection for clinical trials versus treatment in clinical practice. Patients in the ceritinib arm were similar to a previous cohort of CF patients that had a median age of 67 years [18]. Patients in the alectinib arm had a higher proportion of pre-treatment, mainly comprising chemotherapy (platinum-doublet) regimens. This may be due to the fact that many patients in the alectinib arm were diagnosed before crizotinib was approved for first-line treatment. Patients receiving alectinib also reported a higher prevalence of CNS metastases at baseline than in the ceritinib arm, possibly due to under-reporting in clinical practice compared with the more rigorous clinical trial assessment. The prevalence of CNS metastases observed in the ceritinib arm was similar to a CF population from a global real-world cohort, which reported a range of CNS prevalence after CF of 23% in the USA to 54% in Europe [19].

As this study is a comparison of a nonrandomized sample, similarity of baseline characteristics was essential. To address any potential bias, the study cohort was reweighted using inverse probability treatment weighting, based

Table 2. Baseline characteristics pre- and post-weighting.

Characteristic	Pre-weighting			Post-weighting		
	Alectinib	Ceritinib	SMD, %	Alectinib	Ceritinib	SMD, %
Patients, n	183	64 [†]	–	251 [‡]	236 [‡]	–
≥65 years, n (%)	23 (13)	23 (36)	56.7	51 (20)	48 (21)	0.9
Female, n (%)	98 (54)	34 (53)	0.9	134 (53)	121 (51)	3.7
White, n (%)	133 (73)	48 (75)	5.3	184 (73)	171 (72)	1.8
Stage IV at diagnosis, n (%)	170 (93)	57 (89)	13.4	231 (92)	218 (92)	2.2
Prior lines of treatment, n (%):						
– 1	52 (28)	35 (55)		91 (36)	89 (38)	
– 2	66 (36)	20 (31)	62.8	86 (34)	85 (36)	6.4
– ≥3	65 (36)	9 (14)		74 (29)	63 (27)	

[†]Three patients removed after applying trimming.

[‡]Population sizes are larger due to weighting. Stabilized weights were used in order to have an average weight close to one (and pseudo sample size close to original sample size). SMD: Standardized mean difference.

Table 3. Sensitivity analyses of overall survival.

Cohort	Alectinib patients		Ceritinib patients		Adjusted HR (95% CI)	p-value
	n	Median OS (95% CI)	n	Median OS (95% CI)		
Exclude patients with follow-up time in top tenth percentile	164	24 (21–NR)	57	16 (16–19)	0.65 (0.47–0.89)	0.007
Exclude patients initiating treatment <12 weeks prior to data cut-off	183	24 (20–NR)	63	16 (16–19)	0.65 (0.48–0.88)	0.006
Impute missing covariates (race and stage at diagnosis)	208	24 (20–NR)	70	16 (12–19)	0.58 (0.44–0.76)	< 0.001
Include CNS metastases as prognostic factor	152	–	64	–	0.67 (0.49–0.92)	0.01
Not including stage at diagnosis as a prognostic factor	225	26 (21–NR)	68	16 (12–19)	0.53 (0.40–0.70)	< 0.001
Patients with 1–3 lines of treatment [†]	148	24 (20–NR)	63	16 (14–19)	0.66 (0.48–0.90)	0.009
Age as continuous [‡]	170	23 (20–NR)	64	18 (16–19)	0.66 (0.49–0.89)	0.006
Include Asian race category	174	24 (20–NR)	64	17 (15–NR)	0.67 (0.49–0.91)	0.01
GenMatch	183	26 (24–NR)	67	16 (16–16)	0.57 (0.50–0.65)	< 0.001

[†]The sensitivity analysis was conducted including patients that had one to three lines of prior treatment. Patients with >3 lines of treatment were excluded.

[‡]Adjusted for gender, race, age, stage at diagnosis and prior treatment.

GenMatch: Genetic Matching; HR: Hazard ratio; NR: Not reached; OS: Overall survival.

on a calculated propensity score to harmonize the two treatment arms. This is a robust method to reduce bias and estimate average treatment effects across groups after balancing observed confounders [13,20–23]. After reweighting, balance in baseline covariates between the two treatment arms was achieved based on standardized mean differences <10%, which is considered negligible [11].

Alectinib was associated with a lower risk of death in the primary analysis and sensitivity analyses compared with ceritinib. The lower risk of death associated with alectinib is aligned when examining the efficacy evidence from the Phase II and III trials for alectinib. In the two Phase II studies of alectinib in the CF setting, median duration of response with alectinib was 11.2 and 13.5 months [9,10]. A lower duration of response to ceritinib of 9.7 months was reported in a similar Phase II trial (ASCEND-2) [16]. A randomized Phase III study (ALUR) of CF patients treated with alectinib versus chemotherapy confirmed alectinib's superiority over chemotherapy with a median progression-free survival (PFS) of 9.6 months compared with 1.4 months in patients that received chemotherapy (HR: 0.15; 95% CI: 0.08–0.29) [24]. The differentiated profile of alectinib compared with other ALK inhibitors is further supported by data from the Phase III randomized study (ALEX), which compared the treatment of alectinib to crizotinib in previously untreated patients with ALK rearrangement. This trial showed the positive treatment effect of alectinib over crizotinib in terms of PFS (HR for disease progression or death, 0.47; 95% CI: 0.34–0.65; median PFS, not estimable [95% CI, 17.7 months – not estimable] with alectinib vs 11.1 months [95% CI: 9.1–13.1] with crizotinib) [25].

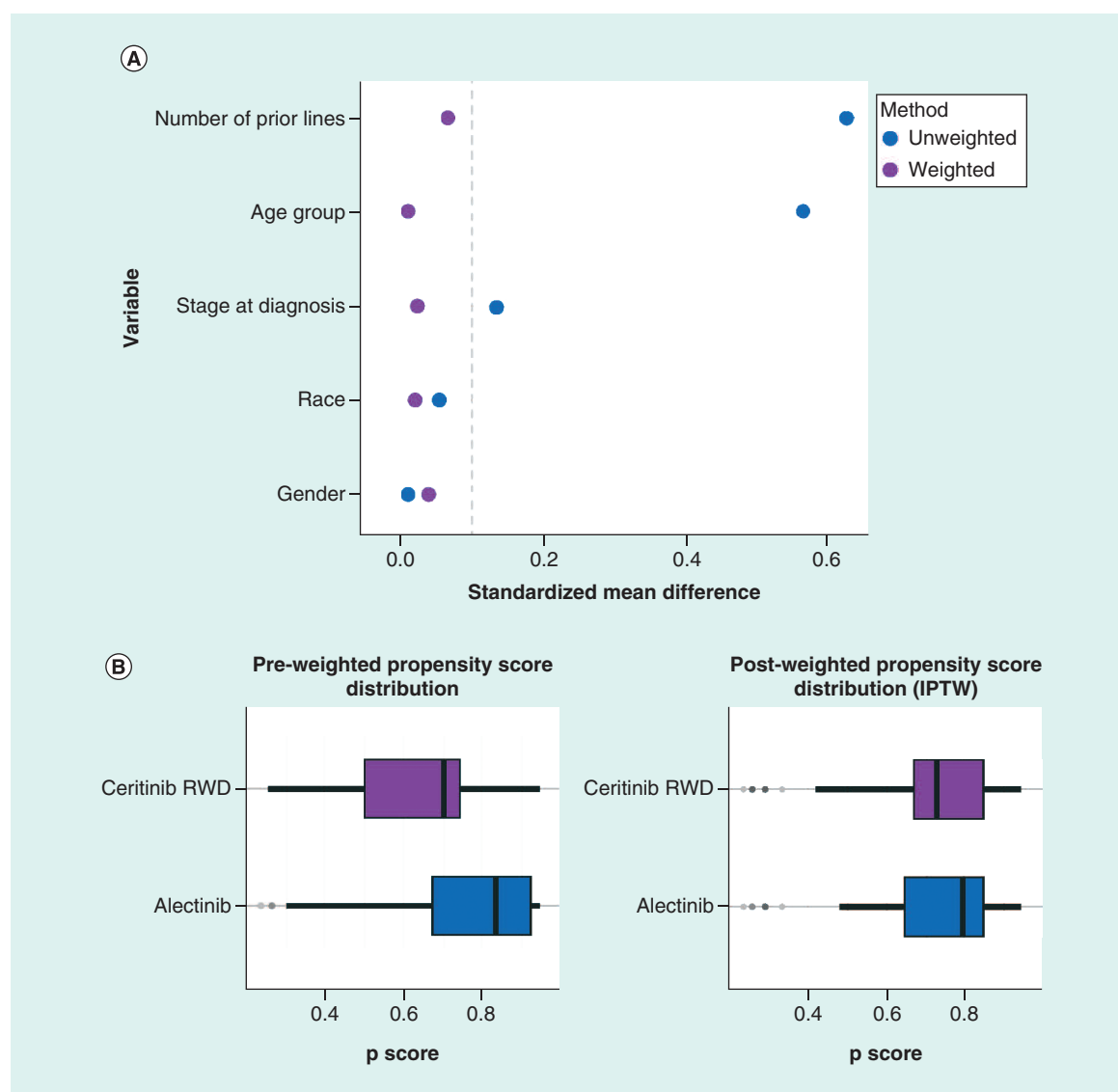


Figure 2. Pre- and post-matched weighted propensity score diagnostics. (A) Standardized mean differences in pre- and post-matched weighted samples, **(B)** Pre- and post-matched weighted distribution of propensity score (p score). IPTW: Inverse probability treatment weighting; RWD: Real-world data.

There are potential limitations to this analysis. Unobserved confounding may have impacted the results which propensity score-based approaches cannot address. Eastern Cooperative Oncology Group performance status (ECOG PS) is a key prognostic factor and criterion of the alectinib trials, but it was not available consistently in the Flatiron database as it is not generally collected in clinical practice. When ECOG PS was available, any patient in the ceritinib cohort with an ECOG PS >2 was excluded. However, this information was not available for more than 60% of patients. Given the high rate of missing data, imputation methods were not used as they would not be meaningful. A recent systematic review of cohorts initiating second-line and third-line treatment in clinical practice reported that most patients had a Karnofsky performance status ≥ 70 or an ECOG PS <2 [26]. Thus, it is plausible that many of the patients in the ceritinib arm had an ECOG PS <2 , given that the treatment setting was second and third line, aligning to patients in the alectinib arm.

Propensity-based approaches can also be at risk of model misspecification. To evaluate the potential impact of misspecification on the treatment effect, we conducted a sensitivity analysis with a nonbiased propensity score approach, GenMatch [27]. The treatment effect stayed consistent in the GenMatch analysis, which suggests little impact of this potential bias on the observed treatment effect.

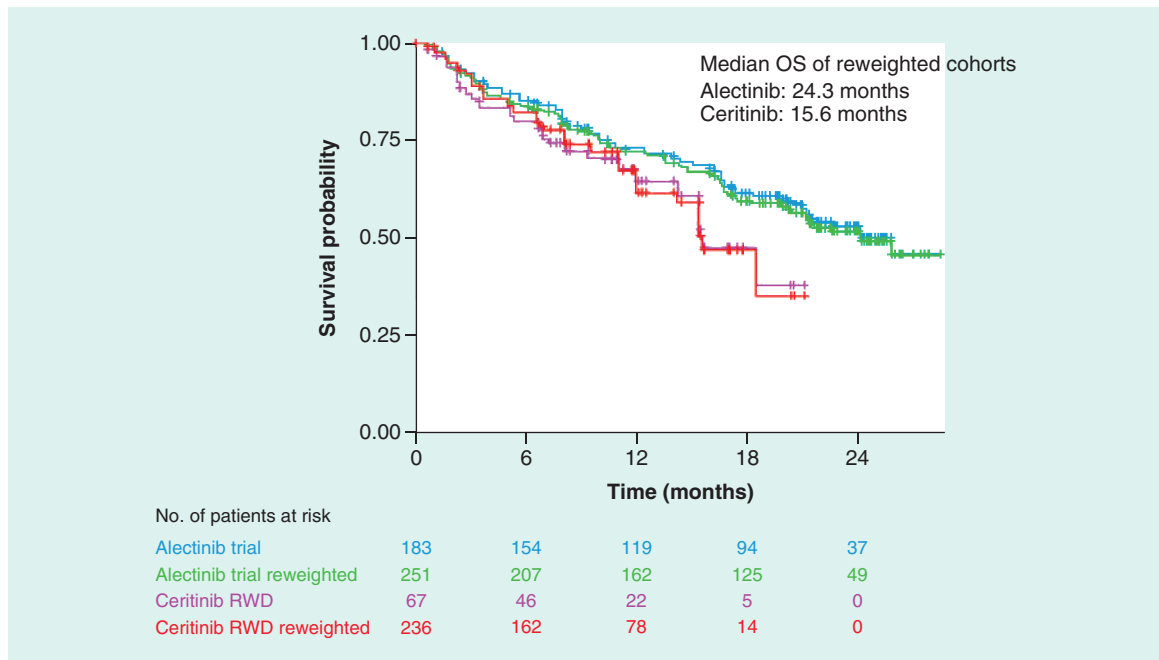


Figure 3. Kaplan–Meier analysis of overall survival in the unweighted and weighted alectinib and ceritinib populations.

OS: Overall survival; RWD: Real-world data.

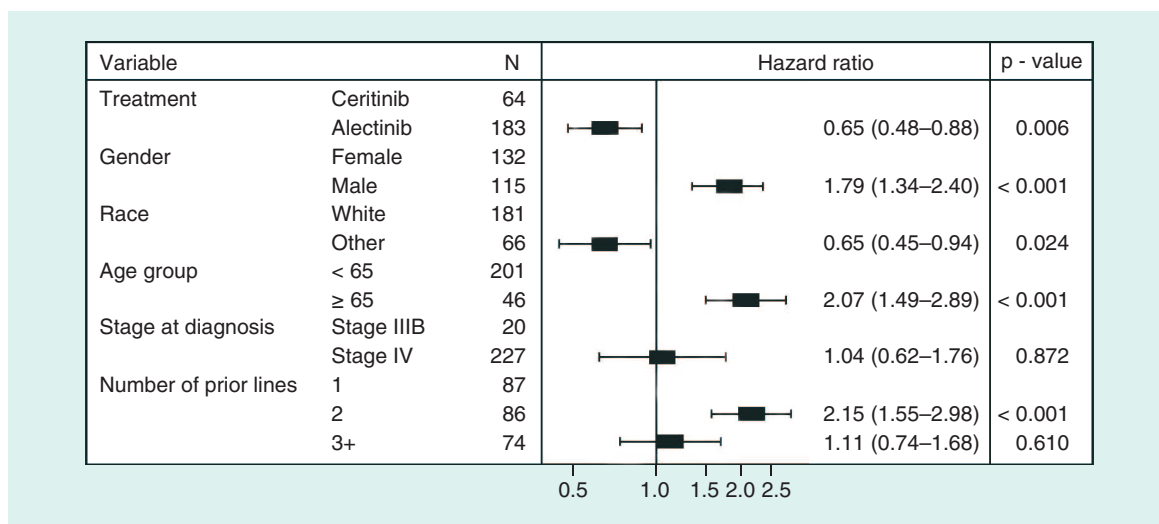


Figure 4. Multivariate Cox model analysis of overall survival. The first group listed for each variable is the reference group.

OS: Overall survival.

OS estimates in the ceritinib RWD arm may have been immature given the limited follow-up time, and the small sample size which may have biased results. However, the median OS estimates from the ASCEND-2 trial, which had a longer follow-up and a larger sample size, were similar to the ceritinib RWD cohort, providing evidence that the OS estimates from the ceritinib RWD arm are generalizable. In a sensitivity analysis excluding patients with >10% of follow-up to evaluate if the longer follow-up in the alectinib was driving the HR, the HR of alectinib versus ceritinib of 0.65 (95% CI: 0.47–0.89) stayed consistent with the primary result of 0.65 (95% CI: 0.48–0.88), which suggests the higher proportion of patients with longer follow-up time in the alectinib arm had a minimal impact on the estimated treatment effect.

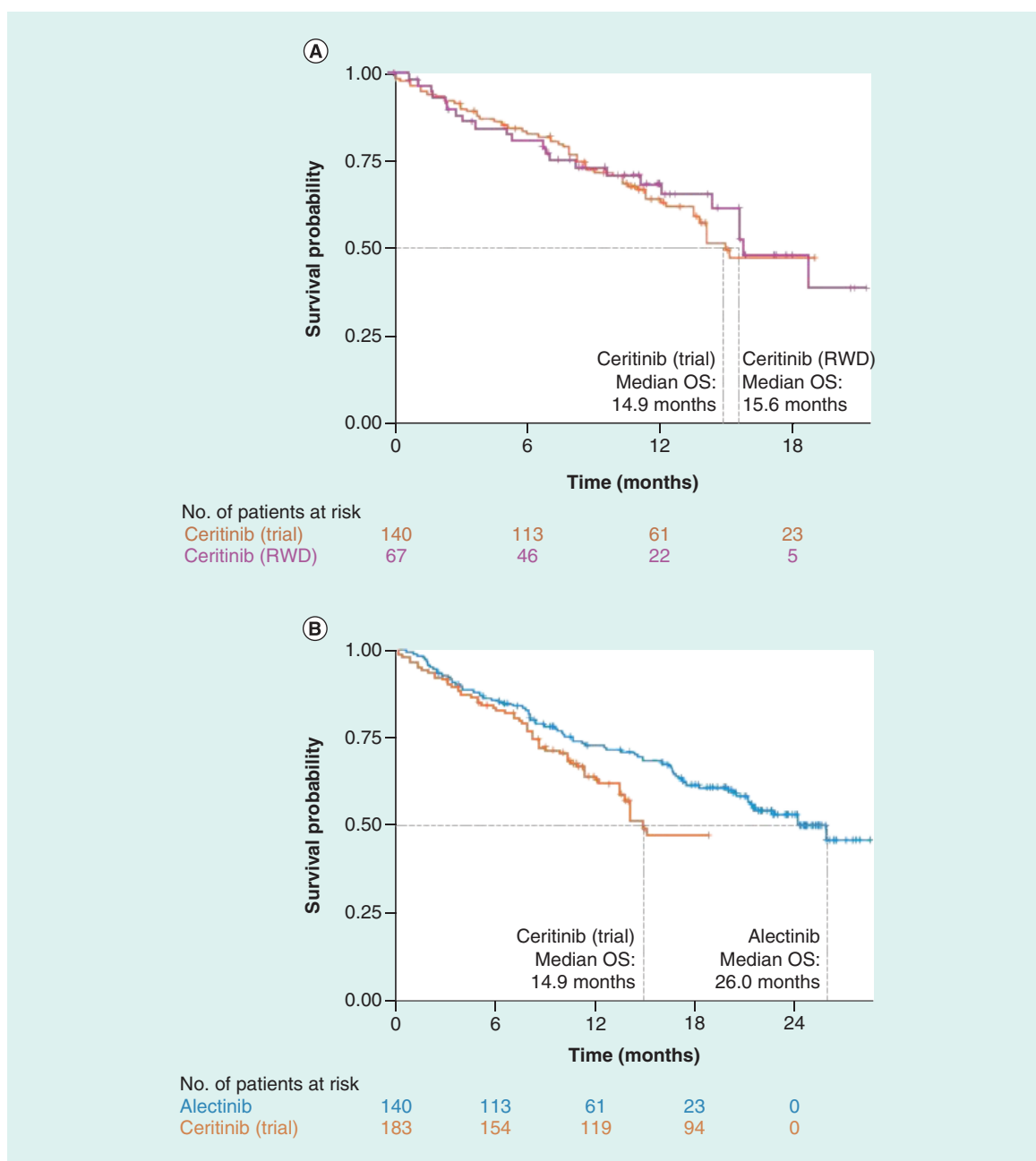


Figure 5. Kaplan–Meier analysis of (A) real-world and clinical trial OS data for ceritinib and (B) alectinib pooled data versus redigitized ASCEND-2. All populations are unweighted. OS: Overall survival; RWD: Real-world data.

Finally, there are inherent differences in the data collection between the two treatment arms by nature of design of the clinical trial and RWD, which could potentially lead to surveillance bias. We believe if this did exist, it would be biased towards the null, because data from the alectinib arm were collected for the purpose of the clinical trial. Moreover, collection of the outcome (date of death) from the Flatiron Health database we believe matches that of a clinical trial, with a sensitivity of 91% and specificity of 96% of capturing death events of the advanced NSCLC electronic health record derived database [28].

To our knowledge, this is the first comparative effectiveness analysis to evaluate OS among patients with *ALK*+ advanced NSCLC who received treatment with alectinib or ceritinib following CF. Alectinib has a clearly differentiated profile compared with other *ALK* inhibitors and this has been demonstrated in Phase II and III

randomized studies in both the CF setting as well as in the treatment naïve setting [24,25,29]. Therefore, our results from this analysis are consistent with the observations from the two randomized Phase III studies that have been conducted until now (ALUR and ALEX). In addition, a similar median OS was observed among ceritinib treated patients in the real world (median OS: 15.6 months) and in a clinical trial setting (median OS: 14.9 months) thereby increasing the credibility of the results observed with the RWD ceritinib. It is therefore, the totality of evidence which in this case includes real world data as well as trial data that is being used to demonstrate the clinical effectiveness of alectinib.

Conclusion

In this adjusted comparison, alectinib was associated with a lower risk of death compared with ceritinib, which aligns with the totality of clinical trial evidence that has been used to license and reimburse alectinib. Median survival estimates were similar between the ceritinib and ceritinib (trial) cohorts, supporting the generalizability of the ceritinib RWD and the plausibility of the results reported in this analysis. In conclusion, we propose that the utility of RWD in providing external controls should be evaluated further with more case studies and validated with IPD from a trial control arm. This will be especially useful in cases where novel medicines are being licensed in rare populations, where it is more difficult to perform large RCTs to generate comparative evidence.

Summary points

- Due to the small proportion of patients with *anaplastic lymphoma kinase*-positive non-small-cell lung cancer (ALK+ NSCLC), development of new tyrosine kinase inhibitors has been based on single-arm trials, and therefore, generating comparative evidence in this setting has been challenging.
- This study specifically looked at how newly emerging areas, such as electronic health records (EHR) systems, can help identify a rare molecularly defined population (such as ALK+ NSCLC), and how this can be used to generate comparative evidence within the setting of single-arm studies.
- EHR data collected from 1 January 2011 to 28 February 2016 (ceritinib) was combined with individual patient data from two single-arm Phase II studies (alectinib) to estimate treatment effectiveness.
- Data were harmonized by applying inclusion and exclusion criteria from the trials to the EHR database. Propensity score and GenMatch modelling were used to address imbalances of measured confounders. Overall survival (OS) was then compared with a multivariate Cox proportional hazards model.
- A sensitivity analysis evaluating OS of the control cohort from EHR was conducted with a naïve comparison between the EHR cohort and digitized trial data from the control treatment.
- Alectinib was associated with a significantly lower risk of death compared with ceritinib (hazard ratio: 0.65, 95% CI: 0.48–0.88, $p = 0.006$), in line with efficacy data from the single-arm trials. Alectinib also achieved a longer median OS compared with ceritinib (24.3 vs 15.6 months, respectively).
- Median OS estimates were similar between the ceritinib ASCEND-2 trial and RWD populations (ASCEND-2 15 months, 95% CI: 14–not reached, RWD 16 months, 95% CI: 12–not reached). In our comparison, alectinib was associated with prolonged OS versus ceritinib.
- This study is a proof of concept study to test whether a comparator arm could be generated from RWD and average treatment effect could be estimated in a rare molecularly defined population, where a Phase III study may not be feasible. More case studies are needed to validate the approach.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/suppl/10.2217/cer-2018-0032

Acknowledgements

The authors would like to thank C Pasina for her support in study setup and interpretation of results.

Author's contributions

Study concepts/design: J Davies, G Crane, M Martinec, P Dleamar, M Coudert and R Martina. Data acquisition: J Davies, G Crane, M Martinec and S Golding. Data analysis/interpretation and statistical analysis: J Davies, M Martinec, P Delmar, M Coudert, W Bordogna, S Golding, R Martina and G Crane.

Financial & competing interests disclosure

This study was funded by F Hoffmann-La Roche Ltd, including the study design, data collection and analysis and the writing of the manuscript. No grant number is applicable. J Davies, M Martinec, P Delmar, M Coudert, W Bordogna, S Golding and G Crane are all employed by F Hoffmann-La Roche Ltd. R Martina has received consulting fees from F Hoffmann-La Roche Ltd. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Third-party medical writing assistance, under the direction of the authors, was provided by R Hubbard of Gardiner-Caldwell Communications and was funded by F Hoffmann-La Roche Ltd.

Ethical conduct of research

Institutional Review Board approval of the study protocol was obtained prior to study conduct and included a waiver of informed consent. Data provided were de-identified and provisions were in place to prevent re-identification in order to protect patients' confidentiality.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

- Hall P. Real-world data for efficient health technology assessment. *Eur. J. Cancer* 79, 235–237 (2017).
- Vivot A, Boutron I, Béraud-Chaulet G *et al.* Evidence for treatment-by-biomarker interaction for FDA-approved oncology drugs with required pharmacogenomic biomarker testing. *Sci. Rep.* 7(1), 6882 (2017).
- Hatswell AJ, Baio G, Berlin *et al.* Regulatory approval of pharmaceuticals without a randomised controlled study: analysis of EMA and FDA approvals 1999–2014. *BMJ Open* 6 (6), e011666 (2016).
- Shaw AT, Yeap BY, Solomon BJ *et al.* Effect of crizotinib on overall survival in patients with advanced non-small-cell lung cancer harbouring *ALK* gene rearrangement: a retrospective analysis. *Lancet Oncol.* 12(11), 1004–1012 (2011).
- Mateos MV, Hernandez JM, Hernandez MT *et al.* Bortezomib plus melphalan and prednisone in elderly untreated patients with multiple myeloma: results of a multicenter Phase 1/2 study. *Blood* 108(7), 2165–2172 (2006).
- Di Lorenzo G, Casciano R, Malangone E *et al.* An adjusted indirect comparison of everolimus and sorafenib therapy in sunitinib-refractory metastatic renal cell carcinoma patients using repeated matched samples. *Expert Opin. Pharmacother.* 12(10), 1491–1497 (2011).
- Tan DS, Araújo A, Zhang J *et al.* Comparative efficacy of ceritinib and crizotinib as initial *ALK*-targeted therapies in previously treated advanced NSCLC: an adjusted comparison with external controls. *J. Thorac Oncol.* 11(9), 1550–1557 (2016).
- Flatiron Health Data, NY, USA (February 2016). <https://flatiron.com/real-world-evidence/>
- Ou SHI, Ahn JS, De Petris L *et al.* Alectinib in crizotinib-refractory *ALK*-rearranged non-small-cell lung cancer: a Phase II global study. *J. Clin. Oncol.* 34(7), 661–668 (2016).
- **Global Phase II study of alectinib in crizotinib-refractory anaplastic lymphoma kinase-positive non-small-cell lung cancer (*ALK*+ NSCLC). Alectinib was active and well tolerated, including in patients with central nervous system metastases.**
- Shaw AT, Gandhi L, Gadgeel S *et al.* Alectinib in *ALK*-positive, crizotinib-resistant, non-small-cell lung cancer: a single-group, multicentre, Phase II trial. *Lancet Oncol.* 17(2), 234–242 (2016).
- **Phase II study of alectinib in patients with *ALK*+ NSCLC who progressed on crizotinib in North America. Alectinib showed clinical activity and was well tolerated, providing evidence for alectinib as a suitable treatment in this patient population.**
- Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav. Res.* 46(3), 399–424 (2011).
- **Describes different propensity score methods and balance diagnostics for examining whether the propensity score model has been adequately specified.**
- Hadley JH, Yabroff R, Barrett MJ, Penson DF, Saigal CS, Potosky AL. Comparative effectiveness of prostate cancer treatments: evaluating statistical adjustments for confounding in observational data. *J. Natl Cancer Inst.* 102(23), 1780–1793 (2010).
- Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat. Med.* 34(28), 3661–3679 (2015).
- **Describes methods to allow assessment of whether measured baseline covariates are balanced between treatment groups in a weighted sample in inverse probability treatment weighting.**
- Lee BK, Lessler J, Stuart EA. Weight trimming and propensity score weighting. *PLoS ONE* 6(3), e18174 (2011).

15. Diamond A, Sekhon JS. Genetic matching for estimating causal effects: a general multivariate matching method for achieving balance in observational studies. *Rev. Econ. Stat.* 95(3), 932–945 (2013).
- **Describes genetic matching, a type of multivariate matching that uses an evolutionary search algorithm to determine the weight of each covariate.**
16. Crino L, Ahn M, De Marinis F *et al.* Multicenter Phase II study of whole-body and intracranial activity with ceritinib in patients with *ALK*-rearranged non-small-cell lung cancer previously treated with chemotherapy and crizotinib: results from ASCEND-2. *J. Clin. Oncol.* 34(24), 2866–2873 (2016).
- **The Phase II ASCEND-2 trial evaluated efficacy and safety of certinib in patients with *ALK*+ NSCLC previously treated with chemotherapy and who had progressed on crizotinib. Certinib treatment showed clinically meaningful responses, including in patients with central nervous system metastases.**
17. Guyot P, Ades AE, Ouwens MJ *et al.* Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan–Meier survival curves. *BMC Med. Res. Methodol.* 2(12), 9 (2012).
- **Describes a method that derives a close approximation to the original individual patient time-to-event data from the published Kaplan–Meier survival curves from which they were generated.**
18. Guerin A, Sasane M, Zhang J *et al.* *ALK* rearrangement testing and treatment patterns for patients with *ALK*-positive non-small-cell lung cancer. *Cancer Epidemiol.* 39(3), 307–312 (2015).
19. Cadranet J, Park K, Arrieta O *et al.* Characteristics, treatment patterns, and survival among *ALK*+ non-small-cell lung cancer (NSCLC) patients treated with crizotinib: a chart review study. *Lung Cancer* 98, 9–14 (2016).
20. Al-Gilani M, Skillington SA, Kallogjeri D, Haughey B, Piccirillo JF. Surgical vs nonsurgical treatment modalities for T3 glottic squamous cell carcinoma. *JAMA Otolaryngol. Head Neck Surg.* 142(10), 940–946 (2016).
21. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1), 41–55 (1983).
22. Eulenburg C, Suling A, Neuser P *et al.* Propensity scoring after multiple imputation in a retrospective study on adjuvant radiation therapy in lymph-node positive vulvar cancer. *PLoS ONE* 11(11), e0165705 (2016).
23. Galsky MD, Stensland KD, Moshier E *et al.* Effectiveness of adjuvant chemotherapy for locally advanced bladder cancer. *J. Clin. Oncol.* 34(8), 825–832 (2016).
24. Novello S, Mazières J, Oh IJ *et al.* Alectinib versus chemotherapy in crizotinib-pretreated anaplastic lymphoma kinase (*ALK*)-positive non-small-cell lung cancer: results from the Phase III ALUR study. *Ann. Oncol.* 29(6), 1409–1416 (2018).
25. Peters S, Camidge DR, Shaw AT *et al.* Alectinib versus crizotinib in untreated *ALK*-positive non-small-cell lung cancer. *N. Engl. J. Med.* 377(9), 829–838 (2017).
26. Davies J, Patel M, Gridelli C, de Marinis F, Waterkamp D, McCusker ME. Real-world treatment patterns for patients receiving second-line and third-line treatment for advanced non-small-cell lung cancer: a systematic review of recently published studies. *PLoS ONE* 12(4), e0175679 (2017).
27. Imai K, King G, Stuart EA. Misunderstandings between experimentalists and observationalists about causal inference. *J. R. Stat. Soc. Ser. A Stat. Soc.* 171(2), 481–502 (2018).
28. Curtis MD, Griffith SD, Tucker M *et al.* Development and validation of a high-quality composite real-world mortality endpoint. *Health Serv. Res.* doi:10.1111/1475-6773.12872 (2018) (Epub ahead of print).
29. Yang JC, Ou SH, De Petris L *et al.* Pooled systemic efficacy and safety data from the pivotal Phase II studies (NP28673 and NP28761) of alectinib in *ALK*-positive non-small-cell lung cancer. *J. Thorac. Oncol.* 12(10), 1552–1560 (2017).

