# Project: Clinical Effectiveness of Treatment Options for Advanced Non-Small Cell Lung Cancer


## Research Protocol

July 19, 2016
Version 1.0

# TABLE OF CONTENTS

## BACKGROUND, OBJECTIVES, AND RESEARCH QUESTIONS

### Background

Lung cancer is the number one cause of cancer death in the United States, expected to cause 158,000 deaths in 2016, and accounting for 26.5% of all cancer deaths.[1] It is the second most common cancer in both men (after prostate cancer) and women (after breast cancer), with 118,000 and 106,000 new cases expected in 2016, respectively.[2] The median age at death is 72.[1] Lung cancer rates reflect smoking behavior, and the incidence of lung cancer peaked in men in 1992 with 69.5 cases per 100,000 and in women in 2005 with 53.8 cases per 100,000; those rates declined by 2013 to 52.2 and 47.7 cases per 100,000, respectively, reflecting earlier declines in the prevalence of smoking.[3,4]

Lung cancer includes different pathological types, broadly divided into small-cell lung cancer and non-small-cell lung cancer (NSCLC).[5] NSCLC makes up about 85% of lung cancers and comprises mainly squamous cell carcinoma, adenocarcinoma, and large-cell lung cancer.[6] Stage at diagnosis is a primary factor in patient survival, and patients with NSCLC commonly present with advanced disease (i.e., distant spread, malignant effusion, or bilateral lung disease); 24% have regional spread at presentation, and 55% have distant spread.[3] Prognosis is generally poor at diagnosis; five-year survival from 2006-2012 was 31.4% in patients with regional spread and 4.9% in patients with distant spread. In recent years, the treatment of some advanced NSCLCs has changed based on the determination of driver mutations in tumors.

Mutations affecting the kinase region of the epidermal growth factor receptor (EGFR) are found in approximately 10% of patients with adenocarcinoma in the United States, but in up to 50% of patients from Asia.[5,7] EGFR mutations are more common in NSCLC among non-smokers and less common in squamous cell carcinoma (approximately 2.7% with EGFR mutations).[8]

In NSCLC patients with EGFR mutations (referred to as EGFR+ NSCLC), tyrosine kinase inhibitors (TKIs) have become first-line therapy.[8] The main TKIs used as first-line therapy for advanced NSCLC include afatinib (Gilotrif®, Boehringer Ingelheim), erlotinib (Tarceva®, Genentech), and gefitinib (Iressa®, AstraZeneca). There is some evidence that the type of EGFR mutation may influence response to TKI therapy.[7,9,10] A course of treatment with first-line TKI therapy typically costs approximately $90,000 per year.[11]

Patients with NSCLC without a driver mutation are typically treated with a platinum-based chemotherapy doublet (e.g., cisplatin + paclitaxel, carboplatin + gemcitabine, etc.) as first-line therapy,[8], which have been the comparators for most trials of TKIs.[10] Conclusions about the relative efficacy and toxicity among the TKIs are somewhat limited by the predominance of indirect evidence from trials comparing them each with chemotherapy.

Despite treatment with a TKI, nearly all patients with advanced NSCLC will eventually progress.[5] A common mechanism of TKI resistance is a T790M mutation in EGFR. Commonly, patients who have this mutation are treated second-line with osimertinib (Tagrisso®, AstraZeneca), a TKI that is effective in EGFR+ tumors with a T790M mutation.[8] For patients who progress on osimertinib, guidelines suggest proceeding with chemotherapy doublet treatment as in patients without a driver mutation.[8]

Other newer agents are also being used for advanced NSCLC. Immunotherapy aimed at altering checkpoint inhibition through the programmed death 1 (PD-1) receptor or its ligand (PD-L1) shows promise in at least some patients with NSCLC.[12] Agents focused on this pathway include nivolumab (Opdivo®, Bristol-Myers Squibb) and pembrolizumab (Keytruda®, Merck), which are antibodies to PD-1, as well as atezolizumab (Tecentriq®, Genentech), an antibody to PD-L1. PD-1 immunotherapy is recommended as second-line treatment in patients with advanced NSCLC without a driver mutation who progress on a chemotherapy doublet.[8]

Most patients studied in trials of PD-1 immunotherapy have received prior treatment with a chemotherapy doublet, whether or not they were EGFR+ and/or had received prior TKI therapy.[13-15] The alternative treatment in this setting would typically be single-agent chemotherapy with an agent that was not used in the original doublet, such as docetaxel.

Recently, clinicians have begun exploring the use of PD-1 agents in patients who have not received a chemotherapy doublet.[16] This includes using PD-1 immunotherapy as first-line treatment in patients with NSCLC without a driver mutation or as third-line therapy (after osimertinib) in patients with EGFR+ NSCLC. As with the TKIs, conclusions about the relative efficacy and toxicity among the PD-1 agents are limited by the predominance of indirect evidence from trials comparing them each with chemotherapy.

A course of PD-1 immunotherapy has been estimated to cost approximately $150,000 per year.[11] In addition to questions of the comparative effectiveness of these agents, both among the agents and compared with alternative therapies, it is uncertain whether tumor expression of PD-L1 is helpful in selecting appropriate patients for PD-1 based therapies.[12]

## Overview

This project will evaluate the health and economic outcomes of treatment regimens for advanced non-small-cell lung cancer with no driver mutation or with an EGFR mutation. Evidence will be collected from available randomized controlled trials as well as high-quality systematic reviews; higher-quality comparative cohort studies will also be evaluated as necessary. We will not restrict studies according to study duration or study setting; however, we will limit our review to those that capture the outcomes of interest. We will supplement our review of published studies with data from conference proceedings, regulatory documents, information submitted by manufacturers, and other grey literature when the evidence meets ICER standards (for more information, see http://icer-review.org/methodology/icers-methods/icer-value-assessment-framework/grey-literature-policy/).

**Quality Assessment Criteria**

We will use criteria published by the US Preventive Services Task Force (USPSTF) to assess the quality of clinical trials and cohort studies, using the categories "good," "fair," or "poor."[17]

**Good**: *Meets all criteria: Comparable groups are assembled initially and maintained throughout the study; reliable and valid measurement instruments are used and applied equally to the groups; interventions are spelled out clearly; all important outcomes are considered; and appropriate attention paid to confounders in analysis. In addition, intention to treat analysis is used for RCTs.*

**Fair***: Any or all of the following problems occur, without the fatal flaws noted in the "poor" category below: Generally comparable groups are assembled initially but some question remains whether some (although not major) differences occurred with follow-up; measurement instruments are acceptable (although not the best) and generally applied equally; some but not all important outcomes are considered; and some but not all potential confounders are addressed. Intention to treat analysis is done for RCTs. Specifically, for this review, differences in baseline characteristics and/or duration of follow-up were allowed only if appropriate statistical methods were used to control for these differences (e.g., multiple regression, survival analysis).*

**Poor***: Any of the following fatal flaws exists: groups assembled initially are not close to being comparable or maintained throughout the study; unreliable or invalid measurement instruments are used or not applied equally among groups (including not masking outcome assessment); and key confounders are given little or no attention. For RCTs, intention to treat or modified intention to treat (e.g., randomized and received at least one dose of study drug) analysis is lacking.*

**PICOTS Inclusion Criteria**

All search algorithms for the systematic literature review will be generated utilizing PICOTS related elements: Patient, Interventions, Comparisons, Outcomes, Timing, and Setting.

***Population***

The four populations of focus for the review will be adults with advanced NSCLC who:

P1) Have an EGFR+ tumor and have not previously been treated for advanced disease
P2) Have a tumor without a driver mutation and have not previously been treated for advanced disease
P3) Have a tumor without a driver mutation that has progressed after first-line treatment with a platinum-based chemotherapy doublet (e.g., cisplatin + paclitaxel, carboplatin + gemcitabine, etc.)
P4) Have an EGFR+ tumor that has progressed after first-line or first- and second-line TKI therapy (patients who do not develop a T790M mutation will only receive first-line TKI therapy)

***Interventions***

P1) The interventions will be the TKIs erlotinib, gefitinib, and afatinib
P2) The intervention will be a treatment sequence of PD-1 immunotherapy (i.e., nivolumab, pembrolizumab, or atezolizumab), followed by a platinum-based chemotherapy doublet at the time of progression
P3) The intervention will be PD-1 immunotherapy (after progression on a platinum-based chemotherapy doublet)
P4) The intervention will be PD-1 immunotherapy (after progression on first-line or first- and second-line TKI therapy)

***Comparators***

P1) The comparator will be a platinum-based chemotherapy doublet
P2) The comparator will be a treatment sequence of a platinum-based chemotherapy doublet, followed by PD-1 immunotherapy at the time of progression
P3) The comparator will be single-agent chemotherapy (e.g., docetaxel)
P4) The comparator will be a platinum-based chemotherapy doublet

***Outcomes***

This review will examine key clinical outcomes that occur in all four populations of patients being treated for advanced NSCLC, including surrogate outcomes common to cancer trials. We will also engage with patient groups and clinical experts to ascertain which outcomes are of greatest importance to patients and determine whether patient-reported outcomes or other evidence sources can be found to enrich the available data. Initial discussions with patient groups indicate that patients with NSCLC have particularly high levels of distress, even when compared with other cancer patients, that NSCLC and its treatments cause substantial disruption of work and family life over time, and that burdens of therapy to the patient and family can accumulate over time as patients live longer and remain on therapy longer; these outcomes may not have been captured well in clinical trials.

Outcomes of interest will include:

- Overall survival
- Disease progression-related measures (progression-free survival, time to progression)
- Objective response rate
- Symptom control
- Health-related quality of life
- Treatment-related adverse events
    - Rates of key adverse events by type (e.g., systemic, gastrointestinal, dermatologic, etc.)
    - Rates of Grade 3 or 4 adverse events
    - Discontinuation due to adverse events
    - Treatment-related deaths

Evidence tables will be developed for each outcome and meta-analysis will be used to quantitatively summarize outcomes for therapies. To assess comparisons within classes of TKIs and PD-1 immunotherapies, network meta-analysis will be used to combine direct and indirect evidence.

***Timing***

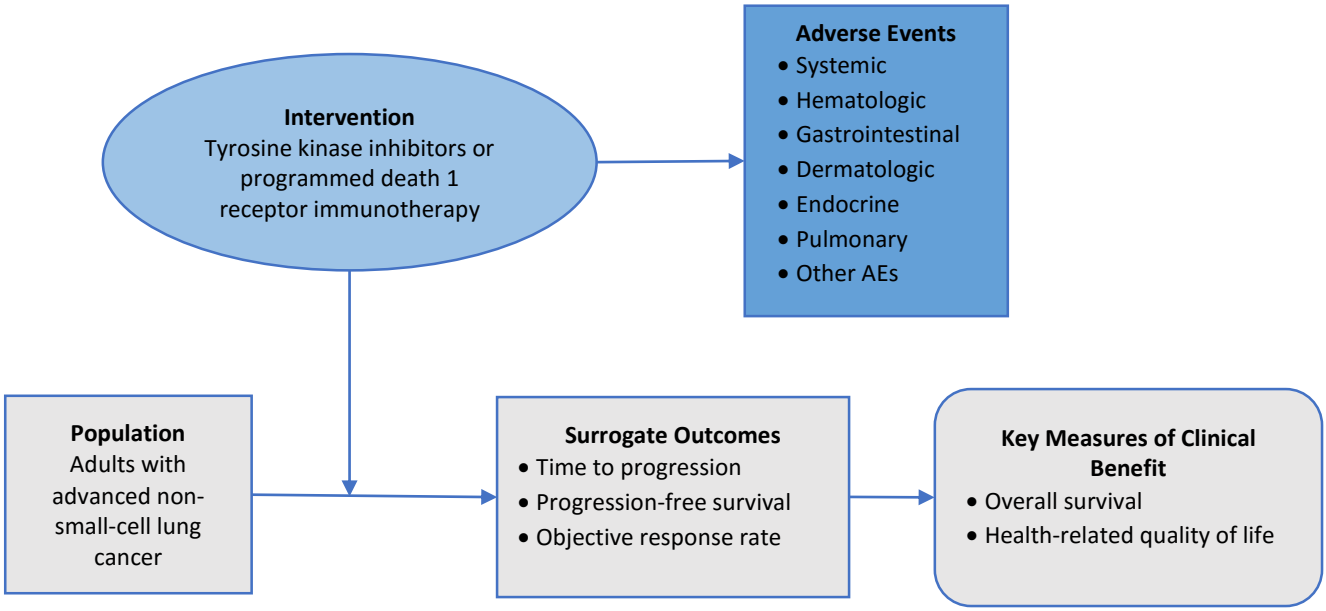
Evidence on intervention effectiveness and harms will be derived from studies of any duration and time period.

***Setting***

All relevant settings will be considered, including inpatient, clinic, and outpatient settings.

**Analytic Framework**

The proposed analytic framework for this project is depicted below.




## EVIDENCE REVIEW METHODS

### Search Methods and Data Sources

Procedures for the systematic literature review assessing the evidence on advanced NSCLC will follow established best methods.[18,19] The review will be conducted in accordance with the Preferred Reporting

Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[20] The PRISMA guidelines include a list of 27 checklist items, which are described further in Appendix A.

We will search MEDLINE, EMBASE, Cochrane Database of Systematic Reviews, and Cochrane Central Register of Controlled Trials for relevant studies. Each search will be limited to English-language studies of human subjects and will exclude articles indexed as guidelines, letters, editorials, narrative reviews, case reports, or news items. The search strategies include a combination of indexing terms (MeSH terms in MEDLINE and EMTREE terms in EMBASE), as well as free-text terms, and are presented in Tables 1-2 on the following pages. We will also include abstracts from conference proceedings in the literature search. In order to supplement the above searches and ensure optimal and complete literature retrieval, we will perform a manual check of the references of recent relevant reviews and meta-analyses.

**Table 1: Search Strategy of Medline 1996 to Present with Daily Update**

| | | |
|---|---|---|
| 1 | Erlotinib.ti,ab | 4110 |
| 2 | Gefitinib.ti,ab | 4313 |
| 3 | Afatinib.ti,ab | 417 |
| 4 | Nivolumab.ti,ab | 325 |
| 5 | Pembrolizumab.ti,ab | 216 |
| 6 | Atezolizumab.ti,ab | 16 |
| 7 | 1 or 2 or 3 or 4 or 5 or 6 | 7633 |
| 8 | randomized controlled trial.pt. | 412234 |
| 9 | controlled clinical trial.pt. | 90416 |
| 10 | randomized.ab. | 343460 |
| 11 | placebo.ab. | 168630 |
| 12 | drug therapy.fs. | 1844178 |
| 13 | randomly.ab. | 247596 |
| 14 | trial.ab. | 354302 |
| 15 | groups.ab. | 1545446 |
| 16 | observational study.pt. | 20770 |
| 17 | exp case-control studies/ | 772861 |
| 18 | exp cohort studies/ | 1521358 |
| 19 | exp cross-over studies/ | 37845 |
| 20 | exp matched-pair analysis/ | 4379 |
| 21 | multicenter study.pt. | 197912 |
| 22 | 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 | 3717029 |
| 23 | 16 or 17 or 18 or 19 or 20 or 21 | 1832443 |

| | | |
|---|---|---|
| 24 | comparative study.pt. or compare.ab,ti. or compares.ab,ti. or compared.ab,ti. or comparing.ab,ti. or comparison.ab,ti. or comparison.ab,ti. or comparative.ab,ti. or effective.ab,ti. or effectiveness.ab,ti. or versus.ab,ti. or vs.ab,ti. | 5822578 |
| 25 | 23 and 24 | 923801 |
| 26 | 22 or 25 | 4216211 |
| 27 | exp carcinoma, non-small-cell lung/ | 37351 |
| 28 | (lung and (cancer$ or carcin$ or neoplasm$ or tumour$ or tumor$) and ((non-small or nonsmall) and cell)).ti,ab. | 42181 |
| 29 | nsclc.ti,ab. | 25173 |
| 30 | 27 or 28 or 29 | 50626 |
| 31 | 7 and 26 and 30 | 2953 |
| 32 | exp animals/ | 20053975 |
| 33 | humans.sh. | 15834501 |
| 34 | 32 not 33 | 4219474 |
| 35 | 31 not 34 | 2937 |
| 36 | limit 35 to english language | 2695 |
| 37 | (case reports or comment or congresses or editorial or letter or review).pt | 5093706 |
| 38 | 36 not 37 | **1719** |

**Table 2: Search strategy of Cochrane Central Register of Controlled Trials June 2016**

| | | |
|---|---|---|
| 1 | Erlotinib.ti,ab | 491 |
| 2 | Gefitinib.ti,ab | 265 |
| 3 | Afatinib.ti,ab | 85 |
| 4 | Nivolumab.ti,ab | 40 |
| 5 | Pembrolizumab.ti,ab | 22 |

| | | |
|---|---|---|
| 6 | Atezolizumab.ti,ab | 2 |
| 7 | 1 or 2 or 3 or 4 or 5 or 6 | 839 |
| 8 | exp carcinoma, non-small-cell lung/ | 2110 |
| 9 | (lung and (cancer$ or carcin$ or neoplasm$ or tumour$ or tumor$) and ((non-small or nonsmall) and cell)).ti,ab. | 4620 |
| 10 | nsclc.ti,ab. | 3406 |
| 11 | 8 or 9 or 10 | 5147 |
| 12 | exp animals/ | 477411 |
| 13 | humans.sh. | 477407 |
| 14 | 12 not 13 | 4 |
| 15 | 7 and 11 | 511 |
| 16 | 15 not 14 | 511 |
| 17 | Limit 16 to English language | **458** |

**Table 3: Search Strategy of Embase on June 8, 2016**

| | | |
|---|---|---|
| 49 | #48 NOT [medline]/lim | 804 |
| 48 | 4 AND 46 AND 47 | 1426 |
| 47 | 5 OR 6 OR 7 OR 8 OR 9 OR 10 | 14,176 |
| 46 | 23 OR 45 | 2,249,347 |
| 45 | 35 AND 44 | 1,182,711 |
| 44 | 36 OR 37 OR 38 OR 39 OR 40 OR 41 OR 42 OR 43 | 7,313,846 |
| 43 | 'vs.':ab | 858,224 |
| 42 | 'vs.':ti | 54,249 |
| 41 | 'versus':ab | 518,090 |
| 40 | 'versus':ti | 174,320 |
| 39 | 'effective*':ab | 1,799,105 |
| 38 | 'effective*':ti | 169,404 |
| 37 | 'compar*':ab | 5,145,166 |
| 36 | 'compar*':ti | 668,020 |
| 35 | 24 OR 25 OR 26 OR 27 OR 28 OR 29 OR 30 OR 31 OR 32 OR 33 OR 34 | 2,495,810 |
| 34 | 'cross-over study'/de OR 'cross-over study' | 50,555 |

| | | |
|---|---|---|
| 33 | 'matched-pair analysis'/de OR 'matched-pair analysis' | 227,404 |
| 32 | 'case* and control*':ab | 20,539 |
| 31 | 'case* and control*':ti | 222 |
| 30 | 'cohort*':ab | 537,509 |
| 29 | 'cohort*':ti | 79,595 |
| 28 | 'case control study'/de OR 'case control study' | 136,746 |
| 27 | 'follow-up'/de OR 'follow-up' | 1,399,521 |
| 26 | 'prospective study'/de OR 'prospective study' | 383,879 |
| 25 | 'longitudinal study'/de OR 'longitudinal study' | 105,387 |
| 24 | 'cohort analysis'/de OR 'cohort analysis' | 243,267 |
| 23 | 22 AND 13 | 1,189,208 |
| 22 | 21 NOT (18 OR 20) | 2,671,928 |
| 21 | 11 OR 12 | 19,427,629 |
| 20 | 19 NOT 14 | 3,377,708 |
| 19 | 16 OR 17 | 11,184,305 |
| 18 | 15 NOT 14 | 4,509,109 |
| 17 | 'random sampl*':ti OR 'random digit*':ti OR 'random effect*':ti OR 'random survey':ti OR 'random regression':ab | 1,966 |
| 16 | 'random sampl*':ti OR 'random digit*':ti OR 'random effect*':ti OR 'random survey':ti OR 'random regression':ti | 1,558 |
| 15 | book:pt OR editorial:pt OR letter:pt OR review:pt | 116,080 |
| 14 | 'randomized controlled trial'/de OR 'randomized controlled trial' | 506,725 |
| 13 | [humans]/lim | 17,112,198 |
| 12 | random*:ti OR placebo*:ti OR 'single blind*':ti OR 'double blind*':ti OR 'triple blind*':ab | 218,834 |

| | | |
|---|---|---|
| 11 | random*:ti OR placebo*:ti OR 'single blind*':ti OR 'double blind*':ti OR 'triple blind*':ti | 218,575 |
| 10 | 'atezolizumab':ti OR 'atezolizumab':ab | 39 |
| 9 | 'pembrolizumab':ti OR 'pembrolizumab':ab | 477 |
| 8 | 'nivolumab':ti OR 'nivolumab':ab | 680 |
| 7 | 'afatinib':ti OR 'afatinib':ab | 925 |
| 6 | 'gefitinib':ti OR 'gefitinib':ab | 7,089 |
| 5 | 'erlotinib':ti OR 'erlotinib':ab | 8,255 |
| 4 | 1 OR 2 OR 3 | 77,297 |
| 3 | lung:ab AND (cancer*:ab OR carcin*:ab OR neoplasm*:ab OR tumour*:ab OR tumor*:ab) AND ('non small':ab OR nonsmall:ab) AND cell:ab | 54,746 |
| 2 | lung:ti AND (cancer*:ti OR carcin*:ti OR neoplasm*:ti OR tumour*:ti OR tumor*:ti) AND ('non small':ti OR nonsmall:ti) AND cell:ti | 38,925 |
| 1 | 'non small cell lung carcinoma'/de | 75,476 |

## Selection of Eligible Studies

Subsequent to the literature search and removal of duplicate citations using both online and local software tools, study selection will be accomplished through two levels of screening, at the abstract and full-text level. Two reviewers will screen the titles and abstracts of all publications identified through electronic searches according to the inclusion and exclusion criteria defined by the PICOTS elements; a third reviewer will work with the initial two reviewers to resolve any issues of disagreement through consensus. No study will be excluded at abstract-level screening due to insufficient information. For example, an abstract that does not report an outcome of interest in the abstract would be accepted for further review in full text.

Citations accepted during abstract-level screening will be retrieved in full text for review. Reasons for exclusion will be categorized according to the PICOTS elements during both title/abstract and full-text review.

## Data Extraction Strategy

For the systematic literature review, the data extraction will be performed in the following steps:

1. Two reviewers will extract information from the full articles.
2. Extracted data will be reviewed for logic, and a random proportion of data will be validated by a third investigator for additional quality assurance.

Information from the accepted studies will be extracted into data extraction forms.

### Publication Bias Assessment

Given the emerging nature of the evidence base for these newer treatments, multiple assessments of publication bias will be performed. We will first scan the ClinicalTrials.gov site to identify studies completed more than two years ago which would have met our inclusion criteria, and for which no findings have been published. We will provide qualitative analysis of the objectives and methods of these studies, in order to ascertain whether there may be a biased representation of study results in the published literature.

### Evidence Synthesis

Data on relevant outcomes will be summarized in evidence tables, and synthesized qualitatively in the text of the report. An evidence table shell is presented in Appendix B.

In addition, quantitative indirect comparisons using Bayesian network meta-analysis (NMA) will be considered where possible.[21] Review of the deviance information criterion (DIC) statistics as well as comparison of the residual deviance (resdev) to the number of unconstrained data points will be used to assess model fit under multiple alternative assumptions. Given the number of comparisons to be made among NSCLC regimens, and the expectation of at least some degree of heterogeneity in patient populations and/or study design, there is a general preference for a random-effects approach. However, the network is also expected to be constructed of primarily single-study connections, which may limit the feasible approach to use of a fixed-effects model (to preserve statistically-significant effects observed in trials) and selected subgroup analyses (to address heterogeneity).[22]

Quantitative analyses will focus attention on the effects of the regimens of interest on progression-free and/or overall survival, and will be conducted using the NetMetaXL tool (http://www.netmetaxl.com/), a publicly-available and validated Excel-based tool for specifying and analyzing Bayesian indirect comparisons in a WinBUGS environment. For these outcomes, adjusted hazard ratios from the randomized trials will be log-transformed and entered into the spreadsheet, and 95% confidence intervals will be used to specify variance estimates (i.e., standard errors). A total of 40,000 iterations each will be used for both “burn-in” (for model convergence) and model (for model results) simulations.

It is important to note that the above approach is appropriate when (a) the assumption of proportional hazards is valid; and (b) trials are not subject to high levels of crossover between treatment arms at the time of progression. Initial analysis suggests that this is the case for an assessment of progression-free survival among patients treated with TKIs, but assessment of overall survival may be diluted due to crossover rates. In such cases an alternative analytic approach such as rank-preserving structural failure

time models or inverse probability of censoring weighting should be considered.[23] We will seek such estimates from published analyses of relevant clinical trials or directly from manufacturers, as patient-level data are required to conduct these alternative analyses.

Preliminary analysis and other published commentary also suggests that the overall survival distribution in studies of PD-1 agents is subject to long tails and may in fact violate the proportional hazards assumption. We will test this by comparing the adjusted hazard ratio approach described above to an alternative in which digitized information from the overall survival curves for each regimen of interest will be used to develop estimates of hazard ratios at multiple timepoints, based on established WinBUGS-based methods.[24] The number of timepoints used will be dependent on data availability across the trials of the three agents. In this instance, 30,000 iterations will be used for both burn-in and model simulations.

## References


1. SEER. Stat fact sheets: lung and bronchus cancer. http://seer.cancer.gov/statfacts/html/lungb.html. Accessed May 13, 2016.

2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA: a cancer journal for clinicians.* 2016;66(1):7-30.

3. SEER. Cancer statistics review 1975-2013. http://seer.cancer.gov/csr/1975_2013/browse_csr.php. Accessed May 13, 2016.

4. Centers for Disease Control and Prevention. Smoking & tobacco use. http://www.cdc.gov/tobacco/data_statistics/tables/trends/cig_smoking/. Accessed June 9, 2016.

5. Herbst RS, Heymach JV, Lippman SM. Lung cancer. *The New England journal of medicine.* 2008;359(13):1367-1380.

6. Dela Cruz CS, Tanoue LT, Matthay RA. Lung cancer: epidemiology, etiology, and prevention. *Clinics in chest medicine.* 2011;32(4):605-644.

7. Sequist LV, Bell DW, Lynch TJ, Haber DA. Molecular predictors of response to epidermal growth factor receptor antagonists in non-small-cell lung cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* 2007;25(5):587-595.

8. NCCN Guidelines Version 4.2016. Non-small cell lung cancer. 2016; https://www.nccn.org/professionals/physician_gls/pdf/nscl.pdf. Accessed May 13, 2016.

9. Yang JC, Sequist LV, Geater SL, et al. Clinical activity of afatinib in patients with advanced non-small-cell lung cancer harbouring uncommon EGFR mutations: a combined post-hoc analysis of LUX-Lung 2, LUX-Lung 3, and LUX-Lung 6. *The Lancet. Oncology.* 2015;16(7):830-838.

10. Haspinger ER, Agustoni F, Torri V, et al. Is there evidence for different effects among EGFR-TKIs? Systematic review and meta-analysis of EGFR tyrosine kinase inhibitors (TKIs) versus chemotherapy as first-line treatment for patients harboring EGFR mutations. *Critical reviews in oncology/hematology.* 2015;94(2):213-227.

11. Truven Health Analytics. Red Book Online. https://www.micromedexsolutions.com. Accessed May 23, 2016.

12. Jing W, Li M, Zhang Y, et al. PD-1/PD-L1 blockades in non-small-cell lung cancer therapy. *OncoTargets and therapy.* 2016;9:489-502.

13. Borghaei H, Paz-Ares L, Horn L, et al. Nivolumab versus Docetaxel in Advanced Nonsquamous Non-Small-Cell Lung Cancer. *The New England journal of medicine.* 2015;373(17):1627-1639.

14. Herbst RS, Baas P, Kim DW, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. *Lancet (London, England).* 2016;387(10027):1540-1550.

15. Fehrenbacher L, Spira A, Ballinger M, et al. Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial. *Lancet (London, England).* 2016.

16. Schmid-Bindert G, Jiang T. First-line nivolumab (anti-PD-1) monotherapy in advanced NSCLC: the story of immune checkpoint inhibitors and "the sorcerers apprentice". *Translational lung cancer research.* 2015;4(3):215-216.

17. Appendix VII- Criteria for Internal Validity: Assessing Individual Studies. *Procedure Manual*: U.S. Preventive Services Task Force; 2011.

18. Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. *Annals of internal medicine.* 1997;126(5):376-380.

19. Higgins JPT, Green S. *Cochrane Collaboration Handbook for Systematic Reviews of Interventions.* The Cochrane Collaboration and John Wiley & Sons Ltd; 2008.

20. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine.* 2009;6(7):e1000097.

21. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *Bmj.* 2005;331(7521):897-900.

22. Cameron C, Coyle D, Richter T, et al. Systematic review and network meta-analysis comparing antithrombotic agents for the prevention of stroke and major bleeding in patients with atrial fibrillation. *BMJ open.* 2014;4(6):e004301.

23. Jonsson L, Sandin R, Ekman M, et al. Analyzing overall survival in randomized controlled trials with crossover and implications for economic evaluation. *Value in health* 2014;17:707-713.

24. Ouwens MJ, Philips Z, Jansen JP. Network meta-analysis of parametric survival curves. *Research synthesis methods*. 2010;1(3-4):258-271.

25. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *Bmj.* 2009;339:b2700.

## APPENDIX A. PRISMA CHECKLIST

The checklist below is drawn from Moher et al. 2009.[20] Additional explanation of each item can be found in Liberati et al. 2009.[25]

| Section/Topic | # | Checklist Item | Reported on Page # |
|---|---|---|---|
| **TITLE** | | | |
| Title | 1 | Identify the report as a systematic review, meta-analysis, or both. | |
| **ABSTRACT** | | | |
| Structured summary | 2 | Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number. | |
| **INTRODUCTION** | | | |
| Rationale | 3 | Describe the rationale for the review in the context of what is already known. | |
| Objectives | 4 | Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS). | |
| **METHODS** | | | |
| Protocol and registration | 5 | Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number. | |
| Eligibility criteria | 6 | Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale. | |
| Information sources | 7 | Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched. | |
| Search | 8 | Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated. | |
| Study selection | 9 | State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis). | |
| Data collection process | 10 | Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators. | |
| Data items | 11 | List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made. | |
| Risk of bias in individual studies | 12 | Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis. | |
| Summary measures | 13 | State the principal summary measures (e.g., risk ratio, difference in means). | |
| Synthesis of results | 14 | Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., $I^2$) for each meta-analysis. | |
| Risk of bias across studies | 15 | Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies). | |
| Additional analyses | 16 | Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified. | |
| **RESULTS** | | | |
| Study selection | 17 | Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram. | |
| Study characteristics | 18 | For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations. | |
| Risk of bias within studies | 19 | Present data on risk of bias of each study and, if available, any outcome-level assessment (see Item 12). | |
| Results of individual studies | 20 | For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group and (b) effect estimates and confidence intervals, ideally with a forest plot. | |
| Synthesis of results | 21 | Present results of each meta-analysis done, including confidence intervals and measures of consistency. | |
| Risk of bias across studies | 22 | Present results of any assessment of risk of bias across studies (see Item 15). | |
| Additional analysis | 23 | Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]). | |
| **DISCUSSION** | | | |
| Summary of evidence | 24 | Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., health care providers, users, and policy makers). | |
| Limitations | 25 | Discuss limitations at study and outcome level (e.g., risk of bias), and at review level (e.g., incomplete retrieval of identified research, reporting bias). | |
| Conclusions | 26 | Provide a general interpretation of the results in the context of other evidence, and implications for future research. | |
| **FUNDING** | | | |
| Funding | 27 | Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. | |

doi:10.1371/journal.pmed.1000097.t001

## APPENDIX B. DATA EXTRACTION SUMMARY TABLE SHELL

| Author & Year of Publication (Trial) | Study Design | Interventions (n) & Dosing Schedule | Inclusion and Exclusion Criteria | Patient Characteristics | Outcomes | Harms |
|---|---|---|---|---|---|---|
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |