



# Transportability of Overall Survival Estimates From US to Canadian Patients With Advanced Non–Small Cell Lung Cancer With Implications for Regulatory and Health Technology Assessment

Sreeram V. Ramagopalan, PhD; Sanjay Popat, MD; Alind Gupta, PhD; Devon J. Boyne, PhD; Alexandre Lockhart, MSc; Grace Hsu, MSc; Dylan E. O'Sullivan, PhD; Jessica Inskip, PhD; Joshua Ray, MSc; Winson Y. Cheung, MD, MPH, FRCPC; Frank Griesinger, MD, PhD; Vivek Subbiah, MD

## Abstract

**IMPORTANCE** The external validity of survival outcomes derived from clinical practice data from US patients with advanced non–small cell lung cancer (NSCLC) is not known and is of potential importance because it may be used to support regulatory decision-making and health technology assessment outside of the US.

**OBJECTIVE** To evaluate whether overall survival (OS) estimates for a selected group of patients with advanced NSCLC from a large US clinical practice database are transportable to Canadian patients receiving the same systemic therapies.

**DESIGN, SETTING, AND PARTICIPANTS** This retrospective multicenter cohort study used transportability analysis to assess whether adjustment for pretreatment characteristics of eligible patient cohorts could reliably approximate OS estimated from US-based samples to Canadian populations. A total of 17 432 eligible adult patients who were diagnosed de novo with advanced NSCLC on or after January 1, 2011, were included in the analysis and followed up until September 30, 2020. Because data on race and ethnicity were available in the US database but not the Canadian database and because racial and ethnic distribution was likely to be similar between US and Canadian patients, these characteristics were not analyzed.

**EXPOSURES** Initiation of platinum-doublet chemotherapy or pembrolizumab monotherapy as first-line systemic treatment for advanced NSCLC.

**MAIN OUTCOMES AND MEASURES** OS measured from the time of initiation of the respective treatment regimen.

**RESULTS** Among 17 432 eligible patients, 15 669 patients from the US and 1763 patients from Canada were included in the analysis. Of those, 11 863 patients (sample size–weighted estimates of mean [SD] age, 68.0 [9.3] years; 6606 [55.7%] male; 10 100 from the US and 1763 from Canada) were included in the subset of patients with complete data for baseline covariates. A total of 13 532 US patients received first-line chemotherapy, and 2137 received first-line pembrolizumab monotherapy. Of those, 8447 patients (62.4%) in the first-line chemotherapy group and 1653 patients (77.3%) in the first-line pembrolizumab group had complete data on baseline covariates for outcome model estimation. A total of 1476 Canadian patients who received first-line chemotherapy and 287 patients who received first-line pembrolizumab monotherapy were identified from the target population. After standardization to baseline patient covariates in the Canadian cohorts, transported OS estimates revealed a less than 5% mean absolute difference from the observed OS in the target population (0.56% over 60 months of follow-up in the first-line chemotherapy group and

(continued)

## Key Points

**Question** Is overall survival of patients with advanced non–small cell lung cancer (NSCLC) that is estimated from US-based clinical practice data externally valid among Canadian patients who receive the same systemic therapies?

**Findings** In this cohort study of 17 432 patients with advanced NSCLC in the US and Alberta, Canada, among those who initiated platinum-doublet chemotherapy or pembrolizumab monotherapy as first-line systemic treatment, overall survival in the US and Canada had high concordance over 60 months of follow-up in the US and 30 months of follow-up in Canada after adjustment for differences in baseline patient characteristics.

**Meaning** These findings suggest that patient survival estimated from US-based clinical practice data may be valid for decision-making in Canada and that transportability analyses can provide evidence of external validity.

## + Supplemental content

Author affiliations and article information are listed at the end of this article.

Abstract (continued)

4.54% over 30 months of follow-up in the first-line pembrolizumab group). Negative control analysis using a mismatched outcome model revealed a 6.64% discrepancy and an incompatible survival curve shape. The results were robust to assumptions of random missingness for baseline covariates, to unadjusted differences in baseline metastases and comorbidities, and to differences in the standard of care between the US and Canada related to administration of second-line anti-programmed cell death 1 ligand 1 immunotherapy for patients who initiated first-line chemotherapy.

**CONCLUSIONS AND RELEVANCE** The results of this cohort study suggest that, under specific circumstances, OS estimates from US clinical practice data can be adjusted using baseline clinical characteristics to closely approximate OS in selected groups of Canadian patients with advanced NSCLC. These results may have implications for regulatory decision-making and health technology assessment in target populations outside of the US.

JAMA Network Open. 2022;5(11):e2239874. doi:10.1001/jamanetworkopen.2022.39874

## Introduction

Although randomized clinical trials (RCTs) remain the gold standard for evaluating the efficacy of therapeutic interventions, their results may not generalize well if systematic differences exist between patients enrolled in an RCT and those in the target population of patients in clinical practice who eventually receive these therapies.<sup>1</sup> Furthermore, when an RCT is infeasible due to lack of timeliness or ethical barriers, data from clinical practice can, in certain cases, be used to fill important evidence gaps. For these reasons, evidence derived from clinical practice data is increasingly being considered as a part of regulatory and reimbursement evidence packages. Examples include the selection of external control arms from clinical practice data for single-arm clinical trials and evaluation of postmarketing comparative effectiveness.<sup>2,3</sup>

Despite the increasing importance of clinical practice data, decision-makers are often concerned about bias. Considerable research efforts have focused on issues of internal validity, such as study design principles and/or statistical techniques to minimize confounding and immortal time bias.<sup>4-7</sup> However, relatively few studies have evaluated the external validity of evidence derived from clinical practice data beyond the study sample to assess whether results were consistent across settings, institutions, and geographic locations after adjustment for systematic differences across settings using transportability analysis. Evidence of external validity of the use of clinical practice data may support decision-making and policy in a different population when direct estimation of treatment effects in the target population is not yet feasible. For example, if transportability can be supported by evidence, clinical practice data from a US-based study could be useful for informing reimbursement considerations for novel therapies in other countries where sufficiently mature follow-up data or sample sizes for direct estimation are not available. Transportability, in the context of the current study, refers to the ability to generalize inferences from a study sample in the US to a target population in another country in which the US-based study sample is not a subset of the target population.<sup>8</sup>

The goal of this study was to evaluate whether risk estimates for overall survival (OS) from a large nationally representative clinical practice database of US patients with advanced non-small cell lung cancer (NSCLC) were transportable to Canadian patients. Specifically, we used transportability analysis to assess whether adjustment for pretreatment characteristics of eligible patient cohorts could reliably and robustly approximate OS estimated from US-based samples to the respective Canadian populations who initiated either first-line platinum-doublet chemotherapy or first-line pembrolizumab monotherapy. An exploratory analysis of patient cohorts who received second-line docetaxel monotherapy after experiencing disease progression while receiving first-line chemotherapy was performed separately due to the inability to harmonize important baseline

variables between US and Canadian databases beyond the first-line treatment setting. In addition, we quantified bias due to potential violations of the strong ignorability assumption for transportability<sup>8</sup> (ie, the assumption that there were no unmeasured pretreatment patient characteristics, effect modifiers, or bias from differences in the standard of care between the US and Canada related to administration of immunotherapy as a subsequent treatment line after experiencing disease progression while receiving first-line platinum-doublet chemotherapy).

---

## Methods

This cohort study was approved by the Health Research Ethics Board of Alberta, Canada, and the WIRB-Copernicus Group Institutional Review Board. A waiver of informed consent for all analyses was granted because the data from both the US and Canadian databases were deidentified. This study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline for cohort studies.<sup>5</sup>

## Data Sources

Data for US patients with advanced NSCLC were selected from the Flatiron Health (FH) database, which is a deidentified longitudinal database derived from electronic health records from more than 280 cancer clinics.<sup>9</sup> For the target population, patients were identified using population-based data from the Canadian province of Alberta; the database comprises linked demographic information, health care encounters, and electronic medical records for the province. Patients with advanced NSCLC were further identified through the Alberta Cancer Registry, which captures information on more than 99% of cancers diagnosed in the province.<sup>10</sup>

## Eligibility Criteria

Eligibility criteria for patient selection were harmonized between the US and Canadian data sets. At the time of treatment initiation (index date), eligible patients were 18 years or older and previously diagnosed de novo with advanced NSCLC (stage IIIb, IIIc, or IV) on or after January 1, 2011. This approach was used because data for patients diagnosed with early-stage cancer who progressed to advanced disease were not reliable in the Alberta data set. Patients were followed up until September 30, 2020. To account for potentially incomplete historical treatment data,<sup>11</sup> US patients in the FH database with a greater than 90-day gap between diagnosis of advanced NSCLC and first recorded visit or medication administration were excluded in accordance with best practices, and Canadian patients in the Alberta database were excluded if they did not initiate therapy within 180 days of diagnosis. Patients with tumor histological characteristics that were categorized as *not otherwise specified* in the US and Canadian databases were excluded.

Patients with missing data for 1 or more baseline covariates of interest were excluded from the US cohort for outcome model fitting. As a sensitivity analysis, model coefficients were also estimated using multiple imputation with chained equations under the missing at random assumption.<sup>12</sup> For imputation, all variables used in the transportability models were included along with mean Eastern Cooperative Oncology Group (ECOG) performance status as an auxiliary variable. For the target Canadian cohort used to evaluate the transportability of OS, missing data for baseline ECOG performance status and smoking status were imputed using stochastic single imputation with logistic regression; no other covariates had any missingness.

## Treatment Regimens

Two treatment groups were assessed as a part of the transportability analysis: (1) patients who initiated first-line platinum-doublet chemotherapy comprising cisplatin or carboplatin plus 1 other agent (paclitaxel, pemetrexed, gemcitabine, vinorelbine, or etoposide) as first-line systemic treatment after diagnosis of advanced NSCLC (first-line chemotherapy group) and (2) patients who initiated first-line pembrolizumab monotherapy (first-line pembrolizumab group). As an exploratory

analysis, we assessed a third treatment group comprising patients who initiated docetaxel monotherapy as second-line treatment after previous exposure to chemotherapy but no exposure to anti-programmed cell death 1 ligand 1 (PD-L1) immunotherapy or anti-cytotoxic T-lymphocyte-associated antigen-4 immunotherapy (second-line docetaxel group).

These 3 treatment regimens were defined such that outcomes could be expected to be relatively homogeneous within each treatment group based on clinical input. Any dose was permitted at the discretion of the treating clinician. Because information on ECOG performance status and metastases that developed after diagnosis was not readily available in the Canadian data set, transportability analysis was not performed for this patient subset, and only baseline characteristics and unadjusted OS were reported.

### Baseline Covariates

The following baseline covariates were included in the outcome model for adjustment: age (in years), sex (male vs female), cancer stage at advanced NSCLC diagnosis (IIIb or IIIc vs IV), ECOG performance status (0-1 vs  $\geq 2$ ), tumor histological characteristics (squamous vs nonsquamous), smoking history (ever vs never), time since advanced NSCLC diagnosis (in months), and time since January 1, 2011 (in years), as a parametric representation of the year of treatment initiation. The baseline covariates of comorbidities and metastases were expected to be recorded and measured differently between the US samples and the Canadian target groups and were therefore assessed as part of the quantitative bias analysis. Data on self-reported race and ethnicity were available in the US database, but these characteristics were not measured in the Canadian database; because racial and ethnic distribution was likely to be similar between US and Canadian patients, these characteristics were not analyzed.

### Outcome

Overall survival was defined as the time (in months) from index date to all-cause death. For the FH data set, the 15th of each month was imputed as the date of death. Patients with estimated survival time of less than 0 months due to this imputation had survival set to 0. Patients with missing information on date of death were censored at their last recorded date of structured activity or the administrative cutoff date of September 30, 2020, whichever was earlier.

### Statistical Analysis

#### Outcome Model

Due to the inability to pool US and Canadian data sets because of data-sharing limitations, an outcome regression approach for transportability was used. For this approach, a prespecified outcome regression model for survival was fit as a function of patient-level covariates in the study sample and standardized using the covariate distributions in the target population to obtain marginal survival probabilities.<sup>8</sup>

For the transportability analysis, a pooled logistic regression model was fitted on a maximum of 60 months of follow-up data for the US cohorts to model the probability of survival as a function of baseline covariates in the US study sample. The specification of this regression model (Q model<sup>13</sup>) included no interaction terms, but quadratic terms were included for continuous variables. Time (in months) was modeled as a cubic spline with manually specified knot locations that maximized overlap of parametric estimates with Kaplan-Meier estimates for the US cohort. Coefficients from a pooled logistic regression analysis are equivalent to those from a Cox regression analysis under assumptions described elsewhere<sup>14</sup>; we verified that coefficients were similar, in our case, to a Cox proportional hazards model using standard time-to-event data format.

To estimate standardized survival curves, the fitted models were used to estimate individual-level survival probabilities up to a maximum of 60 months using baseline covariates for either US or Canadian patient cohorts (depending on the analysis of interest), from which the cumulative mean survival probability by month was calculated. To assess transportability of survival estimates under the specified Q model fitted on US data, standardized parametric estimates of OS in the target

Canadian cohorts (estimated) were compared with Kaplan-Meier estimates (observed) for visual overlap. For the purposes of this study, we chose a threshold of 5% mean absolute difference between estimated and observed OS estimates in the target population to represent sufficient similarity. Percentile-based 95% CIs were generated using 1000 iterations of nonparametric bootstrapping, in which resampling with replacement was conducted by patient rather than observation (with patient-month as the unit of measure). Monthly survival probabilities from this discrete time parametric model were plotted as a smooth function of time.

### Quantitative Bias

To assess the consequences of potential underrecorded metastases and comorbidities in the FH database for the transportability results, we performed a tipping point analysis by (singly) imputing values for mismeasured metastases and comorbidities. First, logistic regression models were estimated for metastases and comorbidities as a function of survival time (in months), event indicator at the end of follow-up, and baseline covariates. Next, these models were used to impute metastases and comorbidities for patients lacking recorded information on these conditions. This overimputation scenario reflected a previous distribution in which a positive recording status (ie, status recorded in the FH database) corresponded to the presence of metastases or comorbidities, but a nonpositive recording status could correspond to either the presence or absence of metastases and comorbidities in the FH data. For the transportability analysis,  $\delta$  adjustment<sup>15</sup> was applied to simulate increase in prevalence from the observed data until the mean absolute difference was 5% or greater (tipping point), with  $\delta$  representing a multiplicative shift to the intercept term in the logistic regression imputation models for metastases and comorbidities.

To assess the sensitivity of results to unadjusted differences (US vs Canada) in the prevalence of PD-L1 immunotherapy as a subsequent treatment line after experiencing disease progression while receiving first-line platinum-doublet chemotherapy, marginal risks under 2 hypothetical dynamic treatment regimens were estimated using G computation.<sup>16</sup> For this analysis, a time-varying indicator of recorded cancer progression was also included in the model along with a 3-way interaction between time (in months), immunotherapy initiation, and chemotherapy initiation to model time-varying hazards. Two hypothetical interventions were modeled: chemotherapy to immunotherapy and immunotherapy to chemotherapy. These interventions were conditional on covariate and treatment history (disregarding drug costs). For the chemotherapy to immunotherapy intervention, if a patient was estimated to initiate chemotherapy (either alone or in combination with immunotherapy) as second-line treatment at some time (denoted by  $t$ ) after the index date (ie, the start of first-line chemotherapy), then they were instead assigned to immunotherapy at that time ( $t$ ). For the immunotherapy to chemotherapy intervention, the converse of this process was performed. These interventions represented hypothetical worst-case scenarios in which all patients who discontinued first-line platinum-doublet chemotherapy could receive either immunotherapy only or chemotherapy only (targeted, with other agents permitted as normal). Maximum risk differences and 95% CIs were estimated using nonparametric bootstrapping with the `gfoRmula` package for R software, version 0.3.2 (R Foundation for Statistical Computing).<sup>17</sup>

---

## Results

### Baseline Characteristics

Among patients in the US and Canada who initiated either first-line platinum-doublet chemotherapy or first-line pembrolizumab monotherapy, 17 432 (15 669 from the US and 1763 from Canada) met eligibility criteria and were included in the analysis. Among eligible patients in the US database, ECOG performance status was missing for 5063 of 13 537 patients (37.4%) who initiated first-line platinum-doublet chemotherapy and 482 of 2133 patients (22.6%) who initiated first-line pembrolizumab monotherapy. A total of 11 863 patients (sample size–weighted estimates of mean [SD] age, 68.0 [9.3] years; 6606 [55.7%] male; 10 100 from the US and 1763 from Canada) were

included in the subset of patients with complete data on baseline covariates for outcome model estimation.

Among patients who initiated first-line platinum-doublet chemotherapy, a total of 13 532 US patients and 1476 Canadian patients met eligibility criteria and were included in the analysis. Within the subset of this treatment group including US patients with complete data for baseline covariates, 8447 were in the US cohort, and 1476 were in the Canadian cohort. The largest differences between the US vs Canadian cohorts were observed for tumor histological characteristics (nonsquamous: 5168 patients [61.2%] vs 1228 patients [83.2%]; squamous: 3279 patients [38.8%] vs 248 patients [16.8%]; standardized mean difference [SMD] for nonsquamous vs squamous, 0.507), cancer stage at diagnosis (stage IIIb or IIIc: 2679 patients [31.7%] vs 264 patients [17.9%]; stage IV: 5768 patients [68.3%] vs 1212 patients [82.1%]; SMD for stage IIIb or IIIc vs stage IV, 0.324), and age at index date (mean [SD], 67.34 [9.25] years vs 65.07 [9.53] years; SMD, 0.242) (Table).

Among patients who initiated first-line pembrolizumab monotherapy, a total of 2137 US patients and 287 Canadian patients met eligibility criteria and were included in the analysis. In the subset of this treatment group including US patients who had complete data for baseline covariates, 1653 were in the US cohort, and 287 were in the Canadian cohort (Table). Smaller differences between the US vs Canadian cohorts were observed for this treatment group. For example, the differences in tumor histological characteristics (nonsquamous: 1256 patients [76.0%] vs 244 patients [85.0%]; squamous: 397 patients [24.0%] vs 43 patients [15.0%]; SMD for nonsquamous vs squamous, 0.229) and cancer stage at diagnosis (stage IIIb or IIIc: 94 patients [5.7%] vs 27 patients [9.4%]; stage

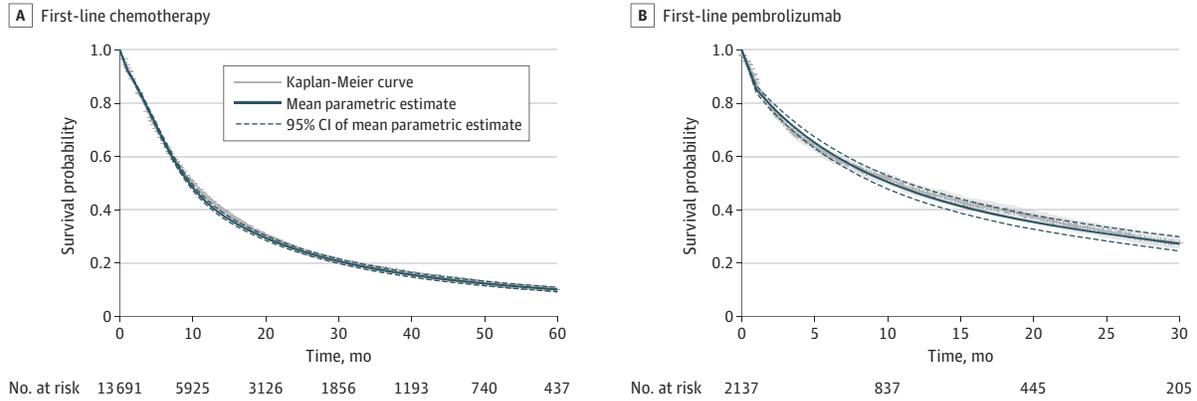
**Table. Baseline Characteristics of US and Canadian Patients With Complete Data for Covariates**

Characteristic	Patients, No./total No. (%)					
	First-line chemotherapy			First-line pembrolizumab		
	US (n = 8447)	Canada (n = 1476)	SMD	US (n = 1653)	Canada (n = 287)	SMD
Age at index date, mean (SD), y	67.34 (9.25)	65.07 (9.53)	0.242	71.64 (9.81)	69.01 (8.95)	0.280
Sex						
Female	3602/8447 (42.6)	703/1476 (47.6)	0.111	803/1653 (48.6)	149/287 (51.9)	0.066
Male	4845/8447 (57.4)	773/1476 (52.4)		850/1653 (51.4)	138/287 (48.1)	
Cancer stage at diagnosis						
IIIb or IIIc	2679/8447 (31.7)	264/1476 (17.9)	0.324	94/1653 (5.7)	27/287 (9.4)	0.140
IV	5768/8447 (68.3)	1212/1476 (82.1)		1559/1653 (94.3)	260/287 (90.6)	
ECOG performance status						
0-1	6625/8447 (78.4)	1091/1476 (73.9)	0.106	1107/1653 (67.0)	209/287 (72.8)	0.127
≥2	1822/8447 (21.6)	385/1476 (26.1)		546/1653 (33.0)	78/287 (27.2)	
Tumor histological characteristics						
Nonsquamous	5168/8447 (61.2)	1228/1476 (83.2)	0.507	1256/1653 (76.0)	244/287 (85.0)	0.229
Squamous	3279/8447 (38.8)	248/1476 (16.8)		397/1653 (24.0)	43/287 (15.0)	
Smoking history						
Ever	7808/8447 (92.4)	1343/1476 (91.0)	0.051	1521/1653 (92.0)	255/287 (88.9)	0.106
Never	639/8447 (7.6)	133/1476 (9.0)		132/1653 (8.0)	32/287 (11.1)	
Time from diagnosis to index date, median (IQR), mo	1.12 (0.72-1.63)	1.84 (1.25-2.76)	0.330	1.25 (0.89-1.81)	1.81 (1.30-2.52)	0.148
Time since January 1, 2011, median (IQR), y <sup>a</sup>	5.28 (3.53-7.02)	4.58 (2.50-6.44)	0.297	7.68 (6.78-8.68)	7.89 (7.28-8.52)	0.192
No. of comorbidities						
0	6188/8447 (73.3)	837/1476 (56.7)	0.362	1062/1653 (64.2)	169/287 (58.9)	0.109
≥1	2259/8447 (26.7)	639/1476 (43.3)		591/1653 (35.8)	118/287 (41.1)	
No. of sites of metastases						
0-1	7304/8447 (86.5)	877/1473 (59.5)	0.638	1367/1653 (82.7)	170/285 (59.6)	0.527
≥2	1143/8447 (13.5)	596/1473 (40.5)		286/1653 (17.3)	115/285 (40.4)	

Abbreviations: ECOG, Eastern Cooperative Oncology Group; SMD, standardized mean difference.

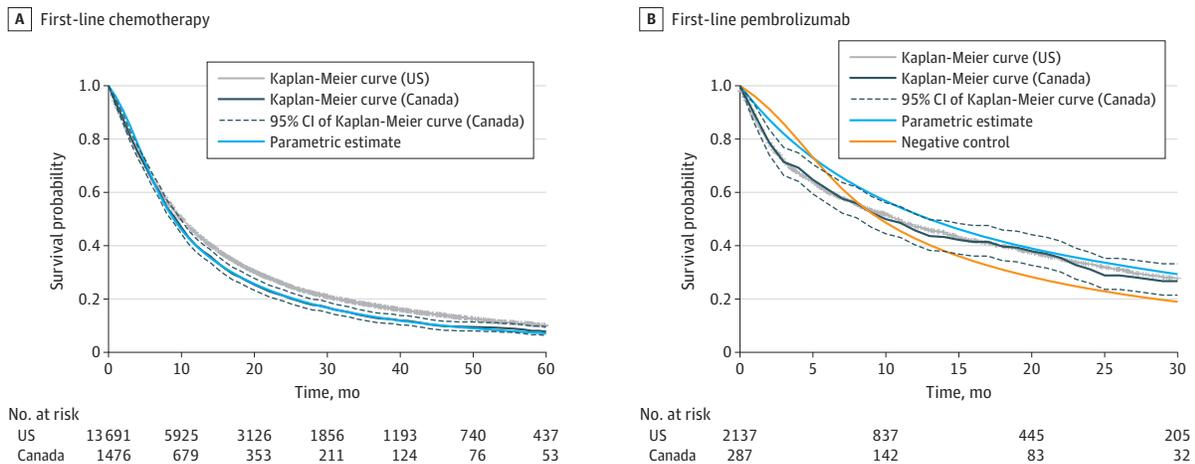
<sup>a</sup> All eligible patients were previously diagnosed de novo with advanced NSCLC on or after January 1, 2011.

Figure 1. Survival Curves for US Patients



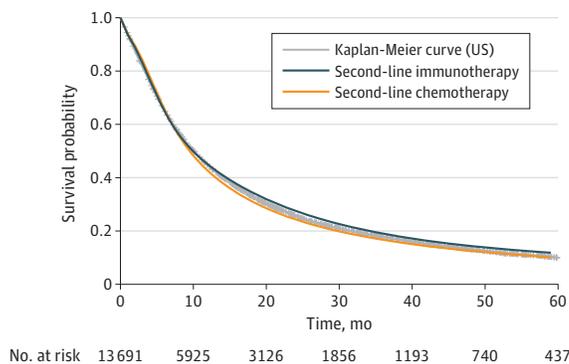
Comparison of unadjusted Kaplan-Meier curves vs standardized parametric estimates. Kaplan-Meier curves and parametric estimates for the sample population were expected to overlap by design (positive control).

Figure 2. Transportability Results



Overall survival was estimated from US data standardized to covariates in the Canadian cohorts compared with the observed overall survival in Canada and the US. In panel B, the negative control corresponded to the overall survival estimated from US data for first-line chemotherapy standardized to first-line pembrolizumab among Canadian patients.

Figure 3. Bias Analysis



Overall survival curves under hypothetical scenarios in which patients who received first-line platinum-doublet chemotherapy could only receive second-line immunotherapy or second-line chemotherapy, regardless of drug costs. The index date (time zero) corresponds to the time of initiation of first-line treatment. The gray Kaplan-Meier curve (US) represents observed risks. Numbers at risk pertain to US patients.

IV: 1559 patients [94.3%] vs 260 patients [90.6%]; SMD for stage IIIb or IIIc vs stage IV, 0.140) were smaller in this treatment group compared with the first-line chemotherapy group.

Baseline summary statistics for the second-line docetaxel cohorts are shown in the eTable in the [Supplement](#). A total of 541 US patients and 138 Canadian patients were identified; of those, 301 US patients and 138 Canadian patients were included in the subset of patients with complete data for baseline covariates. In general, differences in patient characteristics between the US vs Canadian cohorts were relatively minor. For example, differences in age at index date (mean [SD] 65.00 [9.52] years vs 63.53 [8.67] years; SMD, 0.161), tumor histological characteristics (nonsquamous: 227 patients [75.4%] vs 92 patients [66.7%]; squamous: 74 patients [24.6%] vs 46 patients [33.3%]; SMD for nonsquamous vs squamous, 0.193), and cancer stage at diagnosis (stage IIIb or IIIc: 30 patients [10.0%] vs 23 patients [16.7%]; stage IV: 271 patients [90.0%] vs 115 patients [83.3%], SMD for stage IIIb or IIIc vs stage IV, 0.198) were small.

### Transportability Analysis

After estimating parameters of the outcome regression models using complete case data for US patient cohorts, we first assessed their goodness of fit on total US data as a positive control. Parametric model-based estimates of OS standardized to individual-level covariates in the US closely approximated the respective Kaplan-Meier curves for all eligible patients who initiated first-line chemotherapy and first-line pembrolizumab (**Figure 1**). Regression coefficients were similar to pooled estimates after multiple imputation of missing baseline covariates in the US data (maximum difference, 0.05 for the first-line chemotherapy model and 0.13 for the first-line pembrolizumab model), suggesting that the modeling was robust to the assumption of random missingness for baseline data.

Using these complete-case fitted models, we estimated OS in the Canadian population after adjusting for differences in baseline covariates. For the first-line chemotherapy cohort, OS standardized to baseline covariates in the Canadian cohort had high concordance with Kaplan-Meier estimates for the Canadian cohort (**Figure 2A**), with a mean absolute difference of 0.56% over 60 months of follow-up. For the first-line pembrolizumab cohort, Canada-standardized OS initially revealed slightly overestimated survival probability but progressively shifted closer to the Kaplan-Meier curve at later time points (**Figure 2B**), yielding a mean absolute error of 4.54% over 30 months of follow-up. Notably, Kaplan-Meier estimates for the respective US and Canadian cohorts were similar without any adjustment for baseline covariates; this similarity was particularly noticeable in the first-line pembrolizumab cohort, which had relatively small differences in patient characteristics at baseline (**Table**). This pattern was also observed for the second-line docetaxel group (eFigure in the [Supplement](#)).

As a negative control, we estimated OS standardized to the Canadian first-line pembrolizumab cohort using parameters from the US first-line chemotherapy model (ie, a mismatched outcome model). The estimated survival had a sigmoid shape that did not resemble the shape of Kaplan-Meier curves for the pembrolizumab cohorts and consistently underestimated survival after month 8, with a mean absolute difference of 6.64% from the Canadian Kaplan-Meier estimates (**Figure 2**).

### Bias Analysis

No tipping points were identified for the prevalence of metastases at baseline. For comorbidities, the tipping point corresponding to a 5% or higher mean absolute error was identified at a prevalence of greater than 80%, which was approximately 4 times higher than recorded prevalence in the US and 2 times higher than that of the Canadian cohort.

Substantially more patients with advanced NSCLC received anti-PD-L1 immunotherapy (but not targeted agents) as second-line treatment in the US compared with Canada between 2011 and 2020. Because second-line PD-L1 immunotherapy has been associated with significant improvement in OS compared with second-line docetaxel in patients who experienced disease progression while receiving first-line chemotherapy,<sup>18-20</sup> we quantified the consequences of differences in the

prevalence of second-line immunotherapy between the US and Canada for OS in the first-line chemotherapy cohort. The US-standardized risks for the first-line chemotherapy cohort under these hypothetical treatment regimens are shown in **Figure 3**. Consistent with published results from RCTs,<sup>21,22</sup> initiation of second-line immunotherapy was associated with a superior OS profile compared with initiation of second-line chemotherapy under these hypothetical scenarios; the maximum risk difference of 3.50% (95% CI, 1.96%-4.97%) was observed at month 20 after first-line therapy initiation, and the mean absolute difference over 60 months was 2.13%.

---

## Discussion

In this cohort study of adult patients with de novo advanced NSCLC, we evaluated the validity of a transportability analysis for OS estimated from a large US clinical practice database of patients with cancer to clinical practice in Alberta, Canada, for selected treatment groups. Although patient outcomes have previously been compared without accounting for systematic differences between populations, this study is, to our knowledge, the first to formally evaluate transportability of OS estimates from US to Canadian patients. With increasing interest in the use of clinical practice data to supplement clinical trials and accelerate patient access to effective novel therapies, evidence of external validity beyond the study sample may help to inform regulatory decision-making and health technology assessment for target populations for whom direct estimation of an outcome is not feasible due to small sample sizes and/or immature outcome data, lack of timeliness, or low cost-effectiveness. This approach is of particular importance for advanced NSCLC, for which the therapeutic landscape has evolved rapidly over the last decade with the introduction of PD-L1 immunotherapy and targeted agents for patients with rare variants.<sup>23,24</sup> The results revealed that, conditional on having assumptions similar to those used in this study, OS estimates from US clinical practice data may be adjusted to attain valid approximations of OS in Canada.

This study adjusted for participant-level baseline characteristics to evaluate transportability. Beyond participant-level characteristics, there may be differences between the US and Canada with regard to health care delivery, patient adherence, staff experience, and standards for treatment dose and/or frequency that were not examined in this study. For example, due to the nature of health care delivery in these countries, most of the US data were derived from community centers, whereas the Canadian data were derived almost entirely from academic settings. Despite differences in health care systems, given that the US and Canada are largely comparable in terms of health care delivery, standards of care for cancer, and cancer outcomes,<sup>25</sup> we expect that individual patient characteristics are the most important factors when evaluating consequences for OS within specific treatment groups in an advanced cancer setting; thus, our conclusions are likely valid.

## Strengths and Limitations

This study has several strengths, including the formal and rigorous assessment of transportability of OS estimates derived from clinical practice data. We included a sensitivity analysis for partially missing data and a quantitative analysis of residual bias due to mismeasured clinical variables. In simulations using overimputation in which the prevalence of baseline metastases or comorbidities among those without a positive recording status was artificially increased, findings (assuming a 5% threshold for similarity) were robust to all scenarios tested for metastases. The results were also robust to scenarios in which the prevalence of comorbidities among all patients was greater than 80%, which is implausibly high when considering that their prevalence in the Canadian database, in which baseline comorbidities are measured with high accuracy, was approximately 40% in both first-line treatment groups. Using statistical methods to model time-varying treatments, we also quantified the impact of differences in standard of care between the US and Canada, specifically differences related to PD-L1 immunotherapy after disease progression during receipt of first-line platinum-doublet chemotherapy, which can be a point of criticism for claims of external validity.

Differences between these regimens, although significant, were small compared with the 5% threshold for similarity used in this study.

This study also has limitations. Transportability is a causal concept that relies on unverifiable assumptions, which have been described in detail elsewhere.<sup>8,26</sup> In particular, model misspecification can be a source of bias for parametric outcome models. These analyses were conducted among a selected group of patients. Therefore, consideration of whether the findings are generalizable to individuals with recurrent disease or to other disease sites, drug indications, or geographic regions should be given. Although we included a large set of potentially prognostic or effect-modifying variables in this study (either as a part of the main analysis or the sensitivity analyses), we cannot rule out the possibility of other unmeasured factors, such as PD-L1 expression for the first-line pembrolizumab cohort.<sup>27</sup> A potential limitation of this validation approach is the need for individual-level covariates in the target population, although we expect this limitation could be mitigated in some cases through the use of aggregate data and/or expert input regarding adjustment.

---

## Conclusions

This cohort study provides evidence that OS estimates among patients with advanced NSCLC from a US clinical practice database were transportable to clinical practice among patients in the province of Alberta, Canada. These findings suggest that, in specific scenarios in which sample size is insufficient, follow-up is immature, and OS estimates for the target population cannot be ascertained, the OS estimates of US patients with advanced NSCLC may be used as valid surrogates for their Canadian counterparts after adjustment for important baseline characteristics. The results may have implications for regulatory decision-making and health technology assessment of novel cancer therapies in Canada through the use of US-based clinical practice data.

---

## ARTICLE INFORMATION

**Accepted for Publication:** September 19, 2022.

**Published:** November 3, 2022. doi:[10.1001/jamanetworkopen.2022.39874](https://doi.org/10.1001/jamanetworkopen.2022.39874)

**Open Access:** This is an open access article distributed under the terms of the [CC-BY-NC-ND License](https://creativecommons.org/licenses/by-nc-nd/4.0/). © 2022 Ramagopalan SV et al. *JAMA Network Open*.

**Corresponding Author:** Sreeram V. Ramagopalan, PhD, Global Access, F. Hoffmann-La Roche, Grenzacherstrasse 124 CH-4070, Basel, Switzerland ([sreeram.ramagopalan@roche.com](mailto:sreeram.ramagopalan@roche.com)).

**Author Affiliations:** Global Access, F. Hoffmann-La Roche Ltd, Grenzacherstrasse, Basel, Switzerland (Ramagopalan, Ray); Royal Marsden Hospital and Institute of Cancer Research, London, United Kingdom (Popat); Real-World and Advanced Analytics, Cytel Inc, Cambridge, Massachusetts (Gupta, Lockhart, Hsu); Oncology Outcomes Research Initiative, University of Calgary, Calgary, Alberta, Canada (Boyne, O'Sullivan, Cheung); Department of Oncology, University of Calgary, Alberta, Canada (Boyne, O'Sullivan, Cheung); Roche Canada, Mississauga, Ontario, Canada (Inskip); Department of Hematology and Oncology, Pius-Hospital Oldenburg, Oldenburg, Germany (Griesinger); Department of Hematology and Oncology, University of Göttingen, Göttingen, Germany (Griesinger); Department of Investigational Cancer Therapeutics, Division of Cancer Medicine, The University of Texas MD Anderson Cancer Center, Houston, Texas (Subbiah).

**Author Contributions:** Drs Gupta and Boyne had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

**Concept and design:** Ramagopalan, Popat, Gupta, Hsu, Inskip, Ray, Cheung, Griesinger.

**Acquisition, analysis, or interpretation of data:** Ramagopalan, Popat, Gupta, Boyne, Lockhart, O'Sullivan, Cheung, Griesinger, Subbiah.

**Drafting of the manuscript:** Popat, Gupta, Lockhart, Hsu, Cheung.

**Critical revision of the manuscript for important intellectual content:** Ramagopalan, Popat, Gupta, Boyne, O'Sullivan, Inskip, Ray, Cheung, Griesinger, Subbiah.

*Statistical analysis:* Gupta, Boyne, Lockhart, O'Sullivan, Cheung.

*Obtained funding:* Ramagopalan, Cheung.

*Administrative, technical, or material support:* Ramagopalan, Hsu, Ray, Cheung, Subbiah.

*Supervision:* Ramagopalan, Popat, Gupta, Hsu, Cheung, Griesinger.

**Conflict of Interest Disclosures:** Dr Ramagopalan reported receiving personal fees from F. Hoffmann-La Roche during the conduct of the study. Dr Popat reported receiving personal fees from Amgen, AstraZeneca, Bayer, BeiGene, Boehringer Ingelheim, Bristol Myers Squibb, Daiichi Sankyo, Eli Lilly and Company, EQRx, F. Hoffmann-La Roche, GSK, Guardant Health, Janssen Pharmaceuticals, Merck Serono, Merck Sharp & Dohme, Novartis, Pfizer, Sanofi, Seattle Genetics (now Seagen), Takeda Pharmaceutical Company, and Turning Point Therapeutics outside the submitted work. Dr Gupta reported receiving grants from F. Hoffmann-La Roche during the conduct of the study. Dr Boyne reported receiving grants from F. Hoffmann-La Roche during the conduct of the study. Ms Hsu reported receiving grants from F. Hoffmann-La Roche during the conduct of the study. Dr O'Sullivan reported receiving grants from F. Hoffmann-La Roche during the conduct of the study. Dr Griesinger reported receiving grants from Amgen, AstraZeneca, Boehringer Ingelheim, Bristol Myers Squibb, Eli Lilly and Company, F. Hoffmann-La Roche, GSK, Janssen Pharmaceuticals, Merck Sharp & Dohme, Novartis, Pfizer, Sanofi, Siemens, and Takeda Pharmaceutical Company and personal fees from Abbvie, Amgen, AstraZeneca, Boehringer Ingelheim, Bristol Myers Squibb, Eli Lilly and Company, F. Hoffmann-La Roche, GSK, Janssen Pharmaceuticals, Merck & Co, Merck Sharp & Dohme, Novartis, Pfizer, Sanofi, and Takeda Pharmaceutical Company outside the submitted work. Dr Subbiah reported receiving grants from AbbVie, Agensys Corporation, Alfasigma, Altum, Amgen, Bayer, BERG Health, Blueprint Medicines Corporation, Boston Biomedical, Boston Pharmaceuticals, Celgene Corporation, D3 Bio, Dragonfly Therapeutics, Exelixis, FUJIFILM Pharmaceuticals, GSK, Idera Pharmaceuticals, Incyte Corporation, Inhibrx, Loxo Oncology, MedImmune, MultiVir, NanoCarrier Co, the National Cancer Institute Cancer Therapy Evaluation Program, the National Comprehensive Cancer Network, Northwest Biotherapeutics, Novartis, Pfizer, PharmaMar, Relay Therapeutics, Roche/Genentech, Takeda Pharmaceutical Company, Turning Point Therapeutics, the University of Texas MD Anderson Cancer Center, and Vegenics; travel support from the American Society of Clinical Oncology, the European Society for Medical Oncology, Helsinn Healthcare, Incyte Corporation, Novartis, and PharmaMar; consultancy or advisory board participation for Daiichi-Sankyo, Helsinn Healthcare, Incyte Corporation, Jazz Pharmaceuticals, Loxo Oncology/Eli Lilly and Company, MedImmune, Novartis, QED Therapeutics, Relay Therapeutics, and R-Pharm US; and a relationship with Medscape outside the submitted work. No other disclosures were reported.

**Funding/Support:** This study was supported by F. Hoffmann-La Roche.

**Role of the Funder/Sponsor:** The funder was involved in design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

**Additional Contributions:** Cameron Chernecki of Cytel Inc provided final edits and assistance with the submission of this article. He was not compensated for this work.

## REFERENCES

1. Malmivaara A. Generalizability of findings from randomized controlled trials is limited in the leading general medical journals. *J Clin Epidemiol*. 2019;107:36-41. doi:10.1016/j.jclinepi.2018.11.014
2. Sherman RE, Anderson SA, Dal Pan GJ, et al. Real-world evidence—what is it and what can it tell us? *N Engl J Med*. 2016;375(23):2293-2297. doi:10.1056/NEJMs1609216
3. Mishra-Kalyani PS, Kordestani LA, Rivera DR, et al. External control arms in oncology: current use and future directions. *Ann Oncol*. 2022;33(4):376-383. doi:10.1016/j.annonc.2021.12.015
4. Stuart EA, Cole SR, Bradshaw CP, Leaf PJ. The use of propensity scores to assess the generalizability of results from randomized trials. *J R Stat Soc Ser A Stat Soc*. 2001;174(2):369-386.
5. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ*. 2007;335(7624):806-808. doi:10.1136/bmj.39335.541782.AD
6. Dickerman BA, García-Albéniz X, Logan RW, Denaxas S, Hernán MA. Avoidable flaws in observational analyses: an application to statins and cancer. *Nat Med*. 2019;25(10):1601-1606. doi:10.1038/s41591-019-0597-x
7. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*. 2016;183(8):758-764. doi:10.1093/aje/kwv254
8. Degtiar I, Rose S. A review of generalizability and transportability. *ArXiv*. Preprint posted online February 23, 2021. doi:10.48550/arXiv.2102.11904

9. Birnbaum B, Nussbaum N, Seidl-Rathkopf K, et al. Model-assisted cohort selection with bias analysis for generating large-scale cohorts from the EHR for oncology research. *ArXiv*. Preprint posted online January 13, 2020. doi:10.48550/arXiv.2001.09765
10. North American Association of Central Cancer Registries. Certified registries. North American Association of Central Cancer Registries; 2021. Accessed October 21, 2021. <https://www.naacrr.org/certified-registries>
11. Khozin S, Carson KR, Zhi J, et al. Real-world outcomes of patients with metastatic non-small cell lung cancer treated with programmed cell death protein 1 inhibitors in the year following U.S. regulatory approval. *Oncologist*. 2019;24(5):648-656. doi:10.1634/theoncologist.2018-0307
12. Van Buuren S. *Flexible Imputation of Missing Data*. 2nd ed. Chapman and Hall/CRC; 2018. doi:10.1201/9780429492259
13. Snowden JM, Rose S, Mortimer KM. Implementation of G-computation on a simulated data set: demonstration of a causal inference technique. *Am J Epidemiol*. 2011;173(7):731-738. doi:10.1093/aje/kwq472
14. Green MS, Symons MJ. A comparison of the logistic risk function and the proportional hazards model in prospective epidemiologic studies. *J Chronic Dis*. 1983;36(10):715-723. doi:10.1016/0021-9681(83)90165-0
15. Ratitch B, O'Kelly M, Tosiello R. Missing data in clinical trials: from clinical assumptions to statistical analysis using pattern mixture models. *Pharm Stat*. 2013;12(6):337-347. doi:10.1002/pst.1549
16. Hernán MA, Robins JM. *Causal Inference: What If*. Chapman and Hall/CRC; 2020.
17. McGrath S, Lin V, Zhang Z, et al. gfoRmula: an R package for estimating the effects of sustained treatment strategies via the parametric g-formula. *Patterns (N Y)*. 2020;1(3):100008. doi:10.1016/j.patter.2020.100008
18. Rittmeyer A, Barlesi F, Waterkamp D, et al; OAK Study Group. Atezolizumab versus docetaxel in patients with previously treated non-small-cell lung cancer (OAK): a phase 3, open-label, multicentre randomised controlled trial. *Lancet*. 2017;389(10066):255-265. doi:10.1016/S0140-6736(16)32517-X
19. Borghaei H, Paz-Ares L, Horn L, et al. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N Engl J Med*. 2015;373(17):1627-1639. doi:10.1056/NEJMoa1507643
20. Ramagopalan S, Gupta A, Arora P, Thorlund K, Ray J, Subbiah V. Comparative effectiveness of atezolizumab, nivolumab, and docetaxel in patients with previously treated non-small cell lung cancer. *JAMA Netw Open*. 2021;4(11):e2134299. doi:10.1001/jamanetworkopen.2021.34299
21. Mazieres J, Rittmeyer A, Gadgeel S, et al. Atezolizumab versus docetaxel in pretreated patient with NSCLC: final results from the randomized phase 2 POPLAR and phase 3 OAK clinical trials. *J Thorac Oncol*. 2021;16(1):140-150. doi:10.1016/j.jtho.2020.09.022
22. Borghaei H, Paz-Ares L, Horn L, et al. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *New Engl J Med*. 2015;373(17):1627-1639. doi:10.1056/NEJMoa1507643
23. Melosky B, Wheatley-Price P, Juergens RA, et al. The rapidly evolving landscape of novel targeted therapies in advanced non-small cell lung cancer. *Lung Cancer*. 2021;160:136-151. doi:10.1016/j.lungcan.2021.06.002
24. Shields MD, Marin-Acevedo JA, Pellini B. Immunotherapy for advanced non-small cell lung cancer: a decade of progress. *Am Soc Clin Oncol Educ Book*. 2021;41:1-23. doi:10.1200/EDBK\_321483
25. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209-249. doi:10.3322/caac.21660
26. Dahabreh IJ, Robertson SE, Steingrimsson JA, Stuart EA, Hernán MA. Extending inferences from a randomized trial to a new target population. *Stat Med*. 2020;39(14):1999-2014. doi:10.1002/sim.8426
27. Xu Y, Wan B, Chen X, et al; AME Lung Cancer Collaborative Group. The association of PD-L1 expression with the efficacy of anti-PD-1/PD-L1 immunotherapy and survival of non-small cell lung cancer patients: a meta-analysis of randomized controlled trials. *Transl Lung Cancer Res*. 2019;8(4):413-428. doi:10.21037/tlcr.2019.08.09

#### SUPPLEMENT.

**eTable.** Baseline Patient Characteristics of the Second-Line Docetaxel Group

**eFigure.** Kaplan-Meier Curves for Second-Line Docetaxel Monotherapy