

# Comparing Findings From a Friends of Cancer Research Exploratory Analysis of Real-World End Points With the Cancer Analysis System in England

Pia Horvat, PhD<sup>1</sup>; Christen M. Gray, PhD<sup>1</sup>; Alexandrina Lambova, BSc<sup>2</sup>; Jennifer B. Christian, PharmD, MPH, PhD<sup>3</sup>; Laura Lasiter, PhD<sup>4</sup>; Mark Stewart, PhD<sup>4</sup>; Jeff Allen, PhD<sup>4</sup>; Paul Clarke, BA<sup>5,6</sup>; Cong Chen, PhD<sup>5,6</sup>; and Adam Reich, BA<sup>1</sup>

**PURPOSE** This study compared real-world end points extracted from the Cancer Analysis System (CAS), a national cancer registry with linkage to national mortality and other health care databases in England, with those from diverse US oncology data sources, including electronic health care records, insurance claims, unstructured medical charts, or a combination, that participated in the Friends of Cancer Research Real-World Evidence Pilot Project 1.0. Consistency between data sets and between real-world overall survival (rwOS) was assessed in patients with immunotherapy-treated advanced non–small-cell lung cancer (aNSCLC).

**PATIENTS AND METHODS** Patients with aNSCLC, diagnosed between January 2013 and December 2017, who initiated treatment with approved programmed death ligand-1 (PD-[L]1) inhibitors until March 2018 were included. Real-world end points, including rwOS and real-world time to treatment discontinuation (rwTTD), were assessed using Kaplan-Meier analysis. A synthetic data set, Simulacrum, on the basis of conditional random sampling of the CAS data was used to develop and refine analysis scripts while protecting patient privacy.

**RESULTS** Characteristics (age, sex, and histology) of the 2,035 patients with immunotherapy-treated aNSCLC included in the CAS study were broadly comparable with US data sets. In CAS, a higher proportion (46.7%) of patients received a PD-(L)1 inhibitor in the first line than in US data sets (18%-30%). Median rwOS (11.4 months; 95% CI, 10.4 to 12.7) and rwTTD (4.9 months; 95% CI, 4.7 to 5.1) were within the range of US-based data sets (rwOS, 8.6-13.5 months; rwTTD, 3.2-7.0 months).

**CONCLUSION** The CAS findings were consistent with those from US-based oncology data sets. Such consistency is important for regulatory decision making. Differences observed between data sets may be explained by variation in health care settings, such as the timing of PD-(L)1 approval and reimbursement, and data capture.

JCO Clin Cancer Inform 5:1155-1168. © 2021 by American Society of Clinical Oncology

Creative Commons Attribution Non-Commercial No Derivatives 4.0 License 

## INTRODUCTION

Randomized controlled trials (RCTs) are the gold standard for assessing safety and efficacy of drugs; however, RCTs have limitations such as controlled settings, selective patient populations, and short-term study outcomes that might not be generalizable to heterogeneous real-world populations and settings.<sup>1-3</sup>

There is a need to understand the effectiveness of medical interventions in representative real-world populations and for accelerated clinical evidence, which can be met by improved data analytics and unprecedented availability of real-world data (RWD).<sup>4,5</sup> Regulatory bodies including the US Food and Drug Administration, the European Medicines Agency, and health technology assessment agencies, such as the National Institute for Health and Care Excellence, are increasingly acknowledging the importance of RWD.<sup>6-8</sup> However, for RWD to be used widely to supplement or augment clinical trials, the validity of

real-world clinical end points in specific RWD must be established.

A pilot RWD project conducted by Friends of Cancer Research (*Friends*) evaluated the reliability of real-world end points among programmed death-1 or programmed death ligand-1 (PD-[L]1) inhibitor-treated advanced non–small-cell lung cancer (aNSCLC; patients with advanced-stage IIIB-IV or early-stage recurrent or progressed non–small-cell lung cancer [NSCLC]) patients by convening six oncology-focused US health care data sets from participating organizations, including Cancer Research Network, Cota Healthcare, Flatiron Health, IQVIA, OptumLabs Data Warehouse, and PCORnet.<sup>5</sup> The findings of the pilot project demonstrated that worthwhile data can be aggregated from diverse research-ready RWD and that real-world end points measured from these data are consistent with each other and are directionally similar to those observed in RCTs,<sup>9</sup> especially in relation to

## ASSOCIATED CONTENT

### Appendix

Author affiliations and support information (if applicable) appear at the end of this article.

Accepted on October 8, 2021 and published at [ascopubs.org/journal/cci](https://ascopubs.org/journal/cci) on December 3, 2021; DOI <https://doi.org/10.1200/CCI.21.00013>

## CONTEXT

### Key Objective

The Cancer Analysis System (CAS) is a national cancer registry with linkage to national mortality and other health care databases in England. Real-world end points, including overall survival and time to treatment discontinuation, were analyzed from the CAS database and compared with diverse US oncology data sources used in the Friends of Cancer Research Real-World Evidence Pilot Project 1.0.

### Knowledge Generated

The CAS analysis demonstrates the consistent performance and validity of real-world end points across geographically and structurally diverse settings. It also validates the findings from the CAS database by way of comparison with multiple US data sets.

### Relevance

Evaluation and comparison of the strengths, limitations, and validity of specific real-world data for addressing defined clinical questions can facilitate development of guidelines on fit-for-purpose real-world data and defining standards for real-world evidence studies intended to inform regulatory decisions.

overall survival. That said, none of the US-based data sets that participated in the pilot captured mortality systematically in all patients and the approaches to measure mortality varied across the US health care data sources.

The present study expands on the work of the pilot study by including the Cancer Analysis System (CAS) in England. CAS is a cancer registry that covers more than 99% of all patients with cancer in England and contains data on patient and tumor characteristics, treatments, hospitalizations, and mortality. It combines data from the Cancer Outcomes and Services Dataset (COSD), collected by Public Health England's National Disease Registration Service, with linkage to other national data sets, including the Systemic Anticancer Therapy (SACT), Hospital Episode Statistics, the National Radiotherapy Dataset, and mortality data from the Office of National Statistics (ONS).<sup>10</sup> The CAS database is a vital resource for understanding patient care and outcomes in England and for supporting drug research and development globally. The present study aims to validate and compare real-world end points extracted from CAS with those from the six US oncology data sets from *Friends'* original pilot. Considering the population and health care system differences between the US data sets and CAS, along with the complete capture of mortality data in CAS, this study will help in the validation of real-world end points across diverse populations and health care settings.

## PATIENTS AND METHODS

### Study Design and Objectives

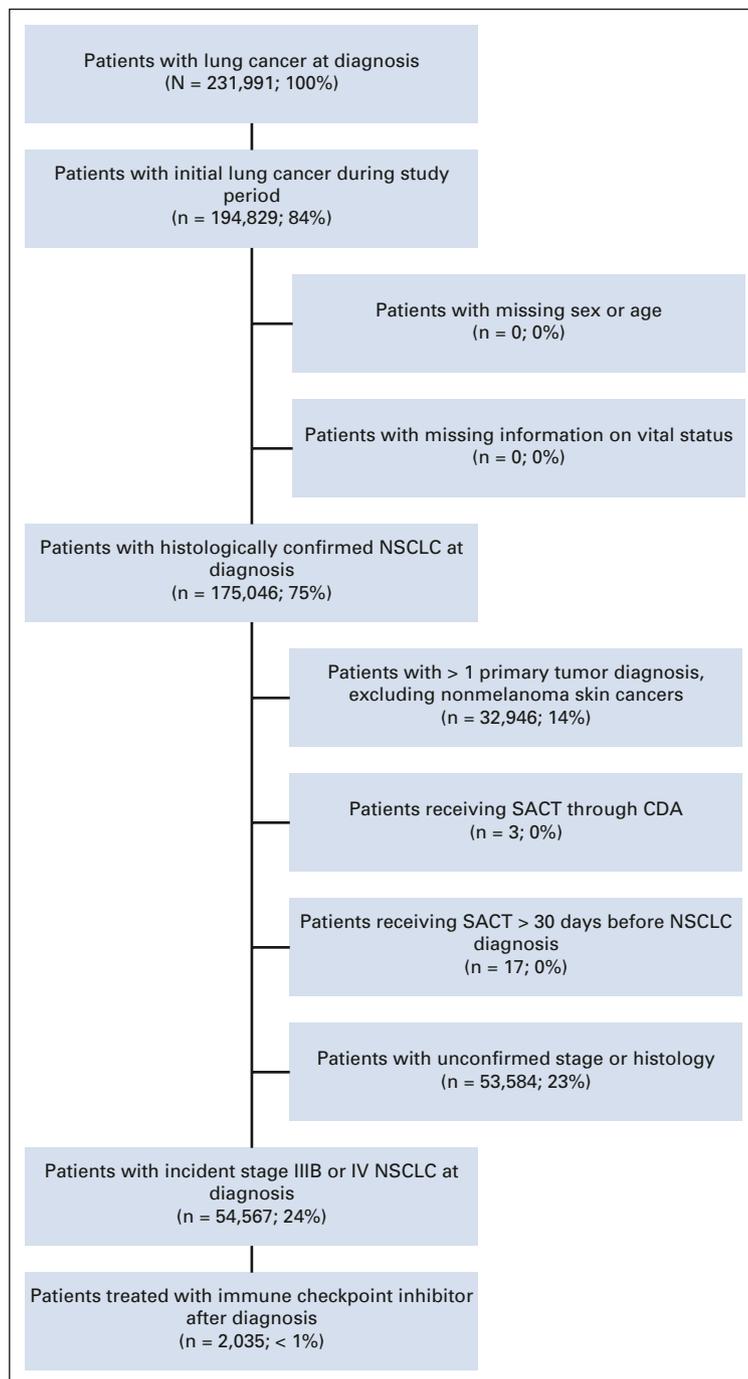
The present study followed the protocol of the US-based pilot project.<sup>5</sup> It used a retrospective observational cohort design in patients with aNSCLC (stage IIIB-IV) leveraging the CAS database. However, because of differences in content and structure of the CAS data and US data sets, some distinctions in the study design should be noted. For example, the CAS study cohort included only incident stage IIIB-IV

histologically confirmed NSCLC patients, unlike several of the US-based data sets, which also included early-stage patients with documented evidence of progression. The overarching objective of the present study was to apply the original *Friends'* pilot study protocol to an ex-US setting and a different type of RWD, namely, a national cancer registry linked to national mortality and other health care databases. The study also had the following secondary objectives:

1. To describe and compare the demographic and clinical characteristics of patients with aNSCLC treated with PD-(L)1 immune checkpoint inhibitors.
2. To assess real-world end points (real-world overall survival [rwOS] and real-world time to treatment discontinuation [rwTTD]) in patients with aNSCLC treated with PD-(L)1 immune checkpoint inhibitors, overall and by clinical and demographic characteristics.
3. To highlight the performance and consistency of real-world end points, particularly rwOS, in CAS (with complete ascertainment of mortality) with the US data sets.

### Patient Cohort

The study cohort flow diagram is shown in [Figure 1](#). Histologically confirmed incident stage IIIB-IV NSCLC (referred to as aNSCLC) patients diagnosed between January 1, 2013, and December 31, 2017, who initiated treatment with approved PD-(L)1 immune checkpoint inhibitors (ie, nivolumab, pembrolizumab, or atezolizumab) until March 2018 (most recent available SACT data<sup>11</sup>) were eligible for analysis. Patients in CAS were excluded if they had unconfirmed stage at diagnosis and a concomitant non-melanoma malignancy or had systemic cancer treatment recorded more than 30 days before the aNSCLC diagnosis. Patients retrieved from CAS had a shorter minimum potential follow-up of 3 months (until March 2018) compared with a minimum potential follow-up of 6 months for US-based data sets. Index date was defined as the initiation of therapy containing any PD-(L)1 inhibitor.



**FIG 1.** Study cohort flow diagram. CDA, Cancer Drugs Fund; NSCLC, non–small-cell lung cancer; SACT, Systemic Anticancer Therapy.

### Data Sources

Data were collected retrospectively from the CAS database. The CAS database comprises several linked databases. For the present study, COSD, SACT, and ONS were used. COSD contains patient demographics (eg, age, sex, ethnicity, and geographic region) and tumor characteristics (eg, staging, morphology, and performance status). The SACT dataset<sup>11</sup> is the national mandatory collection of systemic anticancer therapy from all National Health Service England chemotherapy providers. The CAS database contains patient-level data, which is subject to

strict data protection rules. The process to access the CAS data has historically been long and complex. To simplify access to the CAS database for research, Health Data Insight CIC has developed a programming tool called Simulacrum, which contains artificial patient data and allows analytical programs to be developed before running queries on the CAS database. Simulacrum mimics the structure and types of CAS data, enabling faster analyses by enabling debugging and validation of programming code before running analyses on the CAS database. The details of the data sets or health care data organizations,

which participated in the US-based pilot study, are described in the original publication.<sup>5</sup>

### End Point Definition

End point definitions used in this analysis were closely aligned with the US pilot study protocol and are given in Appendix Table A1 where end points had to be adapted.<sup>5</sup>

### Statistical Analysis

Descriptive statistics (counts [%] or means [with standard deviations]) were performed for demographic and clinical characteristics. The lines of therapy were derived using an algorithm for NSCLC developed with input from clinicians.<sup>12</sup> Continuous variables were summarized using medians and interquartile ranges, and frequencies were calculated for categorical variables. Kaplan-Meier analysis was performed for time-to-event end points (rwOS and rwTTD); these end points were subsequently summarized using median time (in months) with the associated 95% CIs. Spearman's rank-order correlation was used to estimate the correlations between rwOS and rwTTD.

## RESULTS

### Patient Identification and Demographic Characteristics

Table 1 shows the demographic and clinical characteristics of patients with aNSCLC treated with PD-(L)1 inhibitors in the CAS and six US oncology data sets.<sup>5</sup> Overall, 2,035 patients with aNSCLC were included in the analysis. The median age at advanced NSCLC diagnosis was 67 (interquartile range [IQR]: 60.0-73.0) years and 68.3 (IQR: 61.6-73.7) years at PD-(L)1 inhibitor initiation (Appendix Table A2). Median age was comparable with the US pilot study; however, the proportion of PD-(L)1-treated patients in the 65-74 years age group was higher in CAS (47.7%), compared with 30%-40% in the US data sets. Gender distribution in CAS (54.3% male) was within the range of the US data sets. In terms of ethnicity, the patient population in CAS was more homogeneous (91.6% White) than in the US data sets (65%-87%). The majority of CAS patients (81.7%) were diagnosed with stage IV disease, and 18.3% with stage IIIB. In the four US data sets with data on disease stage at initial diagnosis, 62%-91% of patients were diagnosed with stage IV disease. Distribution of histology was broadly similar to the US data sets (71.2 v range 66%-74% in US data sets), with a higher proportion of nonsquamous histology at diagnosis (Appendix Table A2). In the CAS, a considerably higher proportion (46.7%) of PD-(L)1 inhibitors were given as first-line treatment than in US data sets (18%-30%). Almost similar proportion (44.1%) of PD-(L)1 inhibitors in CAS were given in second line and relatively few (7.9% of patients) in third line or beyond. The majority (93.4%) of CAS patients treated with a PD-(L)1 inhibitor (in any therapy line) did not receive a subsequent therapy line during the study observation period. A subgroup analysis of patients below 50 years showed that the majority of patients received PD-(L)1

inhibitor as their second-line therapy (51.7%; n = 45) and 91% (n = 79) did not receive a subsequent line after their PD-(L)1 inhibitor treatment (Appendix Table A4). The median time from advanced diagnosis to PD-(L)1 inhibitor initiation of 4.9 (IQR: 1.6-11.7) months in CAS was shorter than in US data sets (ranging from 6 to 8 months), likely because of higher first-line PD-(L)1 usage in CAS. At 10.6 (IQR: 3.9-16.0) months, the median follow-up time from PD-(L)1 inhibitor initiation was slightly longer in CAS than in US data sets (6-9 months).

### Real-World End Points

Table 2 shows the median time-to-event estimates for real-world end points in CAS. The median rwOS (Fig 2) was 11.4 months (95% CI, 10.4 to 12.7), consistent with the range observed in US data sets (from 8.6 to 13.5 months). The median rwTTD for CAS was 4.9 months (95% CI, 4.7 to 5.1), within the range of US data sets (3.2-7.0 months).

Median rwOS and 95% CI segmented by treatment setting and patient demographic characteristics are given in Table 3. Median rwOS in CAS was higher for the age groups 65-74 and  $\geq 75$  years (both around 12.2 months). Unexpectedly, survival improved with age among patients in CAS and a similar pattern was reported only in US data set E. In CAS, median rwOS at 13.5 months was higher in female than in male patients (9.95 months), consistent with the pattern seen in most of the US data sets. Within the CAS database, stage III patients had higher median rwOS (15.8 months) than many US data sets (B, C, and D). For CAS patients, median rwOS was higher for nonsquamous than squamous aNSCLC (11.8 months) within the range of US data sets (from 8.7 to 14.2 months). Median rwOS decreased almost linearly with increasing therapy line in CAS, whereas in US data sets, there was greater variability in survival across therapy lines with gains in survival also observed in higher lines. In the first-line setting, the median rwOS in CAS was 13.4 months (95% CI, 11.9 to 15.6). In the second-line setting, it was 10.2 months (95% CI, 8.9 to 11.8). In the US-based data sets, there was a wide range in first-line median rwOS from 9.4 to 20.8 months. In the second-line setting, where the US range for median rwOS was narrower 7.9-11.7 (95% CI, 6.5 to 10.5-10.9 to 12.8), the CAS results were also consistent with the observed range (Appendix Table A3). Unlike the US data sets, mortality information is systematically completed for all patients in CAS through linkage with ONS. Correlations between rwOS and rwTTD for CAS and US data sets are shown in Appendix Table A5 (0.7 in CAS compared with the US range of 0.6-0.9).

## DISCUSSION

This study compared the findings from the US-based Friends' pilot project with a UK Cancer Registry (CAS database) to characterize the validity of real-world end points for addressing clinically relevant questions on treatment effectiveness, determination of unmet need, and use of PD-(L)1 inhibitors across diverse health care settings. This

**TABLE 1.** Description of Demographic and Clinical Characteristics of Patients With aNSCLC Treated With PD-(L)1 Checkpoint Inhibitor

Patient Characteristic	CAS (N = 2,035)	Data Sets A-F (N = 269,6,924)
Demographic characteristics		
Median age at advanced diagnosis, years (IQR)	67.00 (60.00-73.00)	64-70 (14-15)
Median age at PD-(L)1 inhibitor initiations, years (IQR)	68.29 (61.55-73.68)	65-71 (14)
Age categories at PD-(L)1 inhibitor initiation, years, No. (%)		
≤ 49	87 (4.28)	8-219 (3-5)
50-64	567 (27.86)	65-2,048 (24-45)
65-74	971 (47.71)	94-2,504 (33-39)
≥ 75	410 (20.15)	86-2,153 (15-38)
Age categories at PD-(L)1 inhibitor initiation, years, No. (%)		
< 75	1,625 (79.85)	167-4,771 (62-85)
≥ 75	410 (20.15)	86-2,153 (15-38)
Sex, No. (%)		
Female	931 (45.75)	125-3,172 (44-49)
Male	1,104 (54.25)	143-3,752 (51-56)
Unknown or missing	0	0-5
Race or ethnicity, No. (%)		
White	1,865 (91.65)	160-4,969 (65-87)
Black or African American	—	14-594 (8-13)
Asian	27 (1.33)	6-155 (1-19)
Chinese	1,662 (81.67)	
Others	42 (2.06)	1-580 (1-9)
Mixed	—	
Unknown or missing	67 (3.29)	0-2,085
Clinical characteristics		
Stage at initial diagnosis, No. (%)		
0 or occult	—	0-2 (0)
I	—	18-496 (6-7)
II	—	17-426 (6-7)
III <sup>a</sup>	373 (18.33)	17-1,494 (7-23)
IV	1,662 (81.67)	161-4,335 (62-91)
Not reported		10-175
Histology, No. (%)		
Nonsquamous NSCLC	1,448 (71.15)	194-4,679 (66-74)
Squamous cell carcinoma	521 (25.60)	61-1,983 (17-30)
NSCLC NOS	24 (1.18)	10-262 (3-10)
Others specified	42 (2.06)	—
Unknown or missing		4
Line of first PD-(L)1 inhibitor, No. (%)		
1 (no prior therapy received)	950 (46.68)	77-2,074 (18-30)
2	898 (44.13)	87-3,357 (32-56)
3	129 (6.34)	51-1,012 (15-20)
≥ 4	33 (1.62)	44-481 (3-20)
Missing or unknown	25 (1.23)	
Follow-up time, months, median (Q1, Q3)		
From advanced diagnosis to PD-(L)1 inhibitor initiation	4.86 (1.61, 11.70)	6-8 (2.13-4.15)
From advanced diagnosis	16.79 (10.61, 26.05)	14-18 (8.25-10.31)
From PD-(L)1 inhibitor initiation (any LOT)	10.61 (3.91, 16.00)	6-9 (2.12-4.13)

Abbreviations: aNSCLC, advanced non–small-cell lung cancer; CAS, Cancer Analysis System; IQR, interquartile range; LOT, line of therapy; NA, not applicable; NOS, not otherwise specified; NSCLC, non–small-cell lung cancer; PD-(L)1, programmed death-1 or programmed death ligand-1; Q1, quartile 1; Q3, quartile 3.

<sup>a</sup>Only stage IIIB patients were included in the analysis of CAS data.

**TABLE 2.** Median Time and 95% CI for Real-World Extracted End Points Comparing CAS With US Data Sets

Data Set	rwOS	rwTTD
CAS	11.40 (10.38 to 12.65)	4.86 (4.67 to 5.09)
A	13.50 (12.80 to 14.50)	7.03 (6.27 to 9.97)
B	15.78 (12.2 to 24.59)	3.25 (2.76 to 3.75)
	8.58 (7.56 to 10.36) <sup>a</sup>	
C	8.67 (6.83 to 10.02)	4.70 (3.68 to 5.52)
D	9.15 (8.82 to 9.51)	3.21 (3.21 to 3.44)
E	12.69 (11.70 to 13.87)	3.63 (3.40 to 3.87)
F	12.30 (9.61 to 16.94)	4.60 (3.71 to 6.32)

Abbreviations: CAS, Cancer Analysis System; rwOS, real-world overall survival; rwTTD, real-world time to treatment discontinuation.

<sup>a</sup>Sites with social security or state death data, censored at the estimated earliest date such that data should be available if no death was observed.

study is based on the UK National Health Service, which operates on a fundamentally different health care model (ie, free for all at the point of use). Furthermore, this study helps to assess the performance of real-world end points because of the completeness of patient and mortality information and validate the CAS database by way of comparison with diverse US data sets.

A comparison of the present study with the US-based pilot study highlights a few key differences between CAS and the US data sets in terms of population and database characteristics. CAS captures clinical characteristics related to TNM staging, whereas this information is generally not available in US-based claims data. Thus, unlike most US-based data sets, which also included patients early-stage progressed NSCLC, the present study included only incident stage IIIB-IV NSCLC, with more than 80% of patients diagnosed in stage IV. In CAS, the median follow-up time from PD-(L)1 inhibitor initiation was longer than that in the US data sets, mainly attributed to the large proportion of PD-(L)1 inhibitors in patients within first-line setting. The estimated overall median rwOS in CAS was within the range of estimates observed in the US data sets although the observed US range was relatively wide. Variations in rwOS in the US data sets are likely due to challenges in accessing mortality data, as deaths are not routinely recorded in most electronic health care records or insurance claims. In contrast to CAS where mortality is ascertained through regularly updated linkage with official national mortality statistics, some US-based RWD rely on published obituary or insurance data to supplement the gaps in mortality information.<sup>13</sup> Combining data from multiple research-ready real-world sources under a common framework can enhance the reliability of real-world end points. Therefore, combining different data types and sources allows timely availability of data, addresses missing data, and improves completeness to create robust data sources.<sup>13</sup>

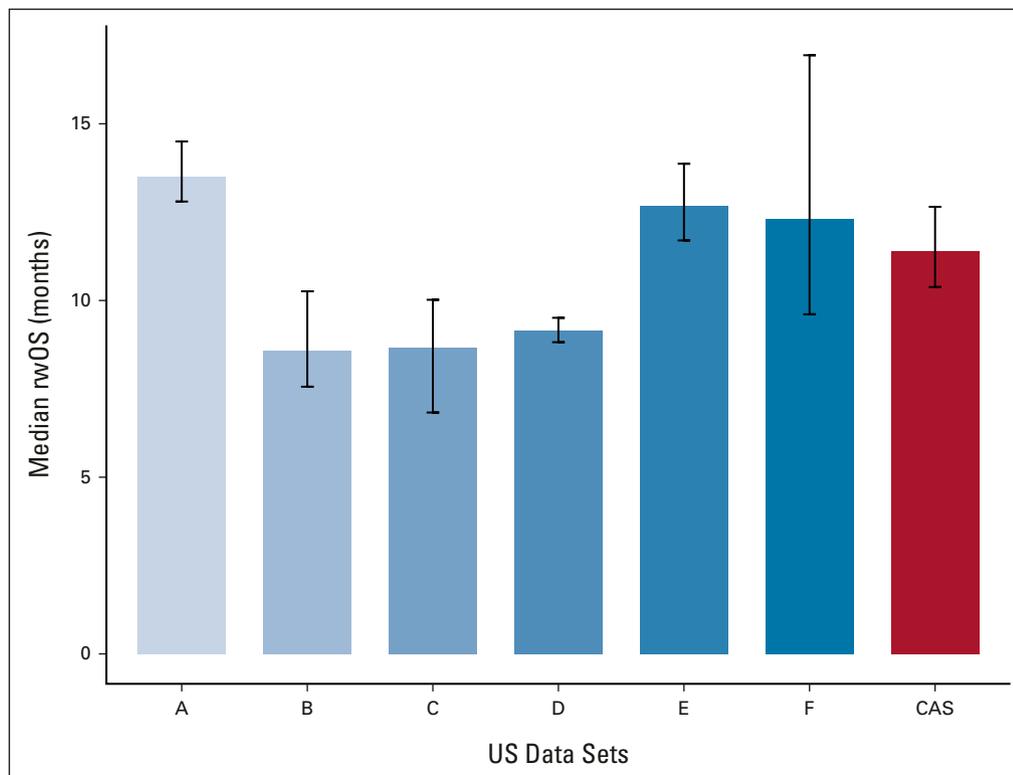
RCTs demonstrate efficacy of PD-(L)1 inhibitors under optimal settings, whereas real-world evidence (RWE)

provides realistic estimates of effectiveness in routine clinical practice. There can be differences in the results of RWD and RCTs; however, considering the complexities in health care systems and high degree of variation in treatment response in the real world, RWD may differ from RCTs but still be valid.<sup>14</sup>

The CAS database has several key strengths including the use of synthetic data to facilitate data access while maintaining patient confidentiality, linkage between a national population-based cancer registry of England with a national systemic therapy database (SACT) and other national databases, and most importantly, the inclusion of mortality data obtained via linkage with ONS. Mortality data are reliably captured through regular updates as structured fields and undergo regular evaluations of validity. This includes a tracing process annually, which ensures the completeness of the annual mortality update. Completeness of survival or mortality data emphasizes and validates the overall findings of the study.

This study also highlights some limitations of the CAS database. First, the CAS database currently does not contain biomarker data although this information is expected in the future. Information on patients' biomarker status would be relevant in interpreting the results of this study, given that response to treatment with PD-(L)1 inhibitors has been shown to correlate with the NSCLC molecular profile. Previous studies have failed to demonstrate unequivocal survival benefits of PD-(L)1 inhibitor monotherapy in patients with epidermal growth factor receptor (EGFR) or anaplastic lymphoma kinase (ALK) genetic alterations compared with standard chemotherapy.<sup>15</sup> Second, disease progression and treatment response are also not directly captured. Treatment-related intermediate end points, such as time to next treatment or death, can be derived from the available data and used to make inferences about possible disease progression. However, such intermediate treatment-related end points are generally limited to patients who have received a subsequent therapy. Compared with most electronic health care records, CAS has a longer lag time as it combines data from national cancer registry and other national databases, which contributed to a shorter observed follow-up time in our study. Because of the high completeness of mortality information in CAS, stronger insights can be gleaned into the value of nontraditional end points, such as time to treatment discontinuation (TTD). The strong correlation between overall survival and TTD in these data further elucidates the role of TTD and its potential use as an earlier indicator of clinical benefit. This could prove to be useful in evaluating the effectiveness of therapies in a broader range of RWD sets, particularly in instances where mortality information may be less accurate or incomplete.

Compared with the US-based study, a higher proportion of patients in CAS received treatment with PD-(L)1 inhibitor in the first-line setting. This was largely due to the time period studied, which reflects the later approval and



**FIG 2.** Comparison of median rwOS (with 95% CI) in CAS with six US data sets in patients with aNSCLC treated with PD-(L)1 inhibitors. aNSCLC, advanced non-small-cell lung cancer; CAS, Cancer Analysis System; OS, overall survival; PD-(L)1, programmed death-1 or programmed death ligand-1; rwOS, real-world overall survival.

reimbursement—and in the case of nivolumab also, more restricted use through the Cancer Drugs Fund—of PD-(L)1 inhibitors in England. The earlier approval and reimbursement of pembrolizumab in England, relative to the later approval and conditional reimbursement (through CDF) of nivolumab, has also led to a disproportionate number of patients treated with pembrolizumab in our study. Patients who received other PD-(L)1 inhibitors (ie, nivolumab) through the CDF are under-represented in our analysis because clinical outcomes data were not accessible for these patients.

RCTs are widely used by regulators because of their strong internal validity, despite having limited generalizability to real-world settings.<sup>16</sup> The CAS database represents more than 99% of the population of England and is nationally representative for patients with cancer. Thus, the results of this study are broadly representative of patients with aNSCLC treated with PD-(L)1 inhibitors in England over the observed time period. National coverage makes CAS a valuable resource for assessing real-world end points, and it may also be used for research in rare cancers, where RCTs may be less feasible. Furthermore, studies similar to ours can establish reference ranges for direct comparison of RWD with clinical trial data in oncology paving the way for alternative study designs.<sup>17</sup>

RWD collection practices for treatment and clinical parameters have improved over time in terms of data

completeness and quality, and this trend is expected to continue.<sup>18</sup> In 2021, CAS is expected to incorporate biomarker data, which can provide additional insights into clinical characteristics of patients. Extension of initiatives such as *Friends* to multicountry settings will further support the development of best practices for the generation and evaluation of RWE to supplement RCTs in regulatory decision making and inform the development of future regulatory guidance.<sup>14</sup> RWE can support regulatory decision making about new or expanded medication indications in a number of ways, from the more established use of external controls for single-arm trials, which have also been used as the basis of US Food and Drug Administration-accelerated approvals, to the growing interest in the use of pragmatic trials and nonrandomized RWE from health care databases, particularly in situations where real-world outcomes and clinical practice patterns differ significantly from the tightly controlled RCT settings.<sup>6,19</sup> Studies similar to ours that establish research processes and evaluate the strengths, limitations, and validity of specific RWD for addressing defined clinical questions can support development of guidelines on fit-for-purpose RWD and defining standards for RWE studies intended to inform regulatory decisions.<sup>19</sup>

In conclusion, this study corroborates and extends the conclusions of the original pilot study that RWD can

**TABLE 3.** Median rwOS and 95% CI, Segmented by Treatment Setting and Patient Characteristics Comparing CAS With US Data Sets

Variable	CAS			Data Sets A-F	
	No. of Patients (N = 2,035)	No. of Events (n = 1,244)	Median OS (months) (95% CI)	No. of Patients (N = 269-6,924)	Median OS (months) (95% CI)
Age, years					
≤ 49	87	57	6.54 (4.47 to 9.03)	8-219	9.1-18.1 (2.7 to NA-11.9 to 21.6)
50-64	567	347	10.51 (9.3 to 13.67)	65-2,047	8.3-16.9 (6.8 to 11.4-9.4 to NA)
65-74	971	590	12.19 (10.71 to 13.6)	94-2,504	8.6-13.4 (6.7 to 12.3-12.1 to 14.9)
≥ 75	410	250	12.22 (9.86 to 14.09)	51-2,153	7-13.2 (5.0 to 15.7-11.8 to 14.6)
Sex					
Female	931	532	13.5 (11.79 to 15.51)	187-950	8.4-14.8 (7.3 to 8.9-13.2 to 16.9)
Male	1,104	712	9.95 (8.9 to 11.4)	143-3,172	7.9-13.2 (6.8 to 10.1-12.1 to 14.5)
Stage					
0 or I				9-498	5.7-12.1 (0.5 to 9.7-10.7 to 14.0)
II				13-426	6.4-13.9 (1.5 to NA-4.0 to NA)
III <sup>a</sup>	373	196	15.84 (13.54 to 18.4)	39-1,494	9.6-14.9 (6.7 to NA-9.4 to 22.1)
IV	1,662	1,048	10.51 (9.59 to 11.93)	161-4,334	8.3-12.1 (7.8 to 8.8-8.3 to 20.7)
Unknown				10-91	NA-7.9 (6.2 to 10.6)
Histology					
NSQ	1,448	867	11.83 (10.38 to 13.21)	194-4,678	9.6-14.2 (9.1 to 10.3-12.7 to 15.8)
SQ	521	343	10.51 (8.48 to 12.32)	61-1,983	6.8-13.9 (4.9 to 8.8-8.6 to NA)
NOS	24	14	10.18 (4.99 to NA)	10-220	10.3-20.4 (5.1 to 13.3-6.1 to NA)
Others specified	42	20	21.29 (13.83 to NA)		
LOT number of first PD-(L)1 inhibitor					
1	950	526	13.37 (11.86 to 15.57)	74-2,074	9.4-20.8 (6.4 to 12.2-14.8 to 25.1)
2	898	590	10.15 (8.94 to 11.83)	87-3,357	7.9-11.7 (6.5 to 10.5-10.9 to 12.8)
3	129	95	7.46 (6.01 to 10.55)	51-1,011	9-15.3 (7.8 to 9.9-9.6 to NA)
4+	33	25	4.8 (3.25 to NA)	34-481	5.1-14.2 (2.1 to 11.9-10.1 to 17.2)
Missing	25	8	NA (14.59 to NA)		

Abbreviations: CAS, Cancer Analysis System; LOT, line of therapy; NA, not applicable; NOS, not otherwise specified; NSQ, nonsquamous; OS, overall survival; PD-(L)1, programmed death-1 or programmed death ligand-1; rwOS, real-world overall survival; SQ, squamous.

<sup>a</sup>Only stage IIIB patients were included in the analysis of CAS data.

generate clinically meaningful and timely evidence on the efficacy of new cancer treatments used across diverse real-world settings. It further describes the usefulness of readily extractable end points, such as TTD, to assess clinical benefit.

Despite the variation in local clinical practice and data collection, there was considerable consistency in the

findings between CAS and the US data sets. The observed differences in results could be largely explained by underlying differences in health care settings, including the timing of and variation in approvals and reimbursement, and data structure. This supports the premise that RWE can be informative for clinical, payer, policy, and regulatory decision making.

#### AFFILIATIONS

<sup>1</sup>IQVIA, London, United Kingdom

<sup>2</sup>IQVIA, Sofia, Bulgaria

<sup>3</sup>IQVIA, Durham, NC

<sup>4</sup>Friends of Cancer Research, Washington, DC

<sup>5</sup>Health Data Insight CIC, Cambridge, United Kingdom

<sup>6</sup>National Cancer Registration and Analysis Service, London, United Kingdom

#### CORRESPONDING AUTHOR

Pia Horvat, PhD, IQVIA, Real World Solutions, 210 Pentonville Rd, London N1 9JY, United Kingdom; e-mail: pia.horvat@iqvia.com.

#### PRIOR PRESENTATION

Presented at the Friends of Cancer Research Virtual Meeting: An International Framework for Real-World Evidence, September 21, 2020.

## AUTHOR CONTRIBUTIONS

**Conception and design:** Pia Horvat, Jennifer B. Christian, Laura Lasiter, Mark Stewart, Jeff Allen, Adam Reich

**Provision of study materials or patients:** Paul Clarke, Cong Chen

**Collection and assembly of data:** Paul Clarke, Cong Chen

**Data analysis and interpretation:** Pia Horvat, Christen M. Gray, Alexandrina Lambova, Jennifer B. Christian, Mark Stewart, Adam Reich

**Manuscript writing:** All authors

**Final approval of manuscript:** All authors

**Accountable for all aspects of the work:** All authors

**Pia Horvat**

**Employment:** IQVIA

**Christen M. Gray**

**Employment:** AstraZeneca, IQVIA

**Alexandrina Lambova**

**Employment:** IQVIA

**Travel, Accommodations, Expenses:** IQVIA

**Jennifer B. Christian**

**Employment:** IQVIA

**Stock and Other Ownership Interests:** GlaxoSmithKline, IQVIA

**Laura Lasiter**

**Employment:** Regulatory and Quality Solutions (I), AgNovos (I), Qiagen (I), Smith & Nephew (I)

No other potential conflicts of interest were reported.

## AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to [www.asco.org/rwc](http://www.asco.org/rwc) or [ascopubs.org/cci/author-center](http://ascopubs.org/cci/author-center).

Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians ([Open Payments](http://OpenPayments)).

## ACKNOWLEDGMENT

This work uses data that have been provided by patients and collected by the NHS as part of their care and support. The data are collated, maintained, and quality-assured by the National Cancer Registration and Analysis Service, Public Health England (PHE).

## REFERENCES

- Kim H-S, Lee S, Kim JH: Real-world evidence versus randomized controlled trial: Clinical research based on electronic medical records. *J Korean Med Sci* 33:e213, 2018
- Makady A, Boer A, Hillege H, et al: What is real-world data? A review of definitions based on literature and stakeholder interviews. *Value Health* 20:858-865, 2017
- Ramagopalan SV, Simpson A, Sammon C: Can real-world data really replace randomised clinical trials? *BMC Med* 18:13, 2020
- Blonde L, Khunti K, Harris SB, et al: Interpretation and impact of real-world clinical data for the practicing clinician. *Adv Ther* 35:1763-1774, 2018
- Stewart M, Norden AD, Dreyer N: An exploratory analysis of real-world end points for assessing outcomes among immunotherapy-treated patients with advanced non-small-cell lung cancer. *JCO Clin Cancer Inform* 3:1-15, 2019
- Baumfeld Andre E, Reynolds R, Caubel P, et al: Trial designs using real-world data: The changing landscape of the regulatory approval process. *Pharmacoepidemiol Drug Saf* 29:1201-1212, 2020
- Cave A, Kurz X, Arlett P: Real-world data for regulatory decision making: Challenges and possible solutions for Europe. *Clin Pharmacol Ther* 106:36-39, 2019
- US Food and Drug Administration: Framework for FDA's Real-World Evidence Program. US Food and Drug Administration, 2018. <https://www.fda.gov/media/120060/download>.
- Huang G, Sun X, Liu D: The efficacy and safety of anti-PD-1/PD-L1 antibody therapy versus docetaxel for pretreated advanced NSCLC: A meta-analysis. *Oncotarget* 9:4239-4248, 2017
- Henson KE, Elliss-Brookes L, Coupland VH: Data resource profile: National Cancer Registration dataset in England. *Int J Epidemiol* 49:16-16h, 2019
- Bright CJ, Lawton S, Benson S, et al: Data resource profile: The Systemic Anti-Cancer Therapy (SACT) dataset. *Int J Epidemiol* 49:15-15l, 2020
- Soares M, Antunes L, Redondo P, et al: Real-world treatment patterns and survival outcomes for advanced non-small cell lung cancer in the pre-immunotherapy era in Portugal: A retrospective analysis from the IO optimise initiative. *BMC Pulm Med* 20:240, 2020
- Curtis MD, Griffith SD, Tucker M: Development and validation of a high-quality composite real-world mortality endpoint. *Health Serv Res* 53:4460-4476, 2018
- Dreyer NA, Hall M, Christian JB: Modernizing regulatory evidence with trials and real-world studies. *Ther Innov Regul Sci* 54:1112-1115, 2020
- Velez MA, Burns TF: Is the game over for PD-1 inhibitors in *EGFR* mutant non-small cell lung cancer? *Transl Lung Cancer Res* 8:S339-S342, 2019 (suppl 4)
- Dreyer NA: Advancing a framework for regulatory use of real-world evidence: When real is reliable. *Ther Innov Regul Sci* 52:362-368, 2018
- Gray CM, Grimson F, Layton D, et al: A framework for methodological choice and evidence assessment for studies using external comparators from real-world data. *Drug Saf* 43:623-633, 2020
- New quality and completeness checks on submissions of Systemic Anti-Cancer Therapies (SACT) data-in place June 2017. Public Health England, 2017. [www.chemodataset.nhs.uk](http://www.chemodataset.nhs.uk)
- Franklin JM, Liaw K-L, Iyasu S, et al: Real-world evidence to support regulatory decision making: New or expanded medical product indications. *Pharmacoepidemiol Drug Saf* 30:685-693, 2021



## APPENDIX

TABLE A1. Operational Definitions

Term	Definitions
End point	
rwOS	Length of time from the index date that the patient initiates treatment with a PD-(L)1 inhibitor to the date of death or end of follow-up, whichever occurred earliest. For claims data, health plan disenrollment dates are incorporated if deaths are not captured among those who leave health plan coverage. For CAS data, patients were censored at their last vital status date
rwTTD	Length of time from the index date that the patient initiates treatment with a PD-(L)1 inhibitor to the date that the patient discontinues the treatment. The study treatment discontinuation date was defined as the last administration or noncancelled order of a drug contained within the PD-(L)1 regimen. Discontinuation was defined as having a subsequent systemic therapy after the initial PD-(L)1-containing regimen, having a gap of more than 120 days with no systemic therapy after the last administration or having a date of death while on the PD-(L)1-containing regimen. Patients without a discontinuation were censored at their last known PD-(L)1 use. For CAS data, patients were censored at their last vital status date
Other elements	
Structured follow-up time	Length of time from the date that the patient initiates PD-(L)1 therapy or advanced diagnosis date for each patient until the last structured activity (ie, most recent visit or administration), unenrollment when relevant, death, or end of the follow-up period (ie, last structured activity)
LOT	LOT may be available from review of structured medication data, text fields, or other unstructured data from chart review. The first LOT was identified on the basis of the first date of receipt of any anticancer medication for treatment of aNSCLC. A treatment regimen was defined as the combination of anticancer medications that were received within the first 30 days of treatment with the first anticancer drug (in CAS, the first LOT includes all systemic anticancer medications given during the first 28 days). The second LOT was identified after a gap of 120 days or more in infusion or oral anticancer drug therapy or if the combination of drugs being received was changed (in CAS, second LOT was identified after a treatment gap of $\geq 70$ days or after a change in treatment). Subsequent LOTs were defined similarly

Abbreviations: aNSCLC, advanced non–small-cell lung cancer; CAS, Cancer Analysis System; LOT, line of therapy; PD-(L)1, programmed death-1 or programmed death ligand-1; rwOS, real-world overall survival; rwTTD, real-world time to treatment discontinuation.

**TABLE A2.** Description of Demographic and Clinical Characteristics of Patients With aNSCLC Treated With PD-(L)1 Checkpoint Inhibitor

Patient Characteristic	CAS (N = 2,035)	Data Set A (n = 2,595)	Data Set B (n = 556)	Data Set C (n = 435)	Data Set D (n = 6,924)	Data Set E (n = 2,860)	Data Set F (n = 269)
Demographic characteristics							
Median age at advanced diagnosis, years (IQR)	67.00 (60.00-73.00)	68 (15)	64 (14)	66 (14)	69 (14)	68 (14)	70 (14)
Median age at PD-(L)1 inhibitor initiation, years (IQR)	68.29 (61.55-73.68)	69 (14)	65 (14)	68 (14)	69 (14)	69 (14)	71 (14)
Age categories at PD-(L)1 inhibitor initiation, years, No. (%)							
≤ 49	87 (4.28)	120 (5)	24 (4)	21 (5)	219 (3)	80 (3)	8 (3)
50-64	567 (27.86)	888 (34)	252 (45)	129 (30)	2,048 (30)	863 (30)	65 (24)
65-74	971 (47.71)	866 (33)	194 (35)	169 (39)	2,504 (36)	1,047 (37)	94 (35)
≥ 75	410 (20.15)	721 (28)	86 (15)	116 (27)	2,153 (31)	870 (30)	102 (38)
Age categories at PD-(L)1 inhibitor initiation, years, No. (%)							
< 75	1,625 (79.85)	1,874 (72)	470 (85)	319 (73)	4,771 (69)	1,990 (70)	167 (62)
≥ 75	410 (20.15)	721 (28)	86 (15)	116 (27)	2,153 (31)	870 (30)	102 (38)
Sex, No. (%)							
Female	931 (45.75)	1,147 (44)	275 (49)	212 (49)	3,172 (46)	1,351 (47)	125 (46)
Male	1,104 (54.25)	1,448 (56)	281 (51)	222 (51)	3,752 (54)	1,509 (53)	143 (53)
Unknown or missing	0	0	0	5	0	0	1
Race or ethnicity, No. (%)							
White	1,865 (91.65)	1,704 (78)	477 (86)	284 (65)	4,969 (79)	676 (87)	160 (87)
Black or African American	—	282 (13)	67 (12)	37 (9)	594 (9)	44 (6)	14 (8)
Asian	27 (1.33)	52 (2)	6 (1)	83 (19)	155 (3)	13 (2)	9 (5)
Others	42 (2.06)	142 (7)	6 (1)	31 (7)	580 (9)	42 (5)	1 (1)
Unknown or missing	67 (3.29)	415	0	0	626	2,085	85
Clinical characteristics							
Stage at initial diagnosis, No. (%)							
0 or occult			0		2 (0)		
I			23 (6)		496 (7)		18 (7)
II			22 (6)		426 (6)		17 (7)
III <sup>a</sup>	373 (18.33)		88 (23)	39 (9)	1,494 (22)		17 (7)
IV	1,662 (81.67)		248 (65)	396 (91)	4,335 (64)		161 (62)
Not reported			175		171		10
Histology, No. (%)							
Nonsquamous NSCLC	1,448 (71.15)		369 (66)	320 (74)	4,679 (70)	1,981 (69)	194 (73)
Squamous cell carcinoma	521 (25.06)		147 (26)	73 (17)	1,983 (30)	659 (23)	61 (23)
NSCLC NOS	24 (1.18)		40 (7)	42 (10)	262 (3)	220 (8)	10 (4)
Others specified	42 (2.06)		—	—	—	—	—
Unknown or missing							4
Line of first PD-(L)1 inhibitor, No. (%)							
1 (no prior therapy received)	950 (46.68)	690 (27)	144 (26)	80 (18)	2,074 (30)	777 (27)	77 (29)
2	898 (44.13)	1,440 (56)	272 (49)	205 (47)	3,357 (49)	1,414 (49)	87 (32)
3	129 (6.34)	380 (15)	96 (17)	85 (20)	1,012 (15)	448 (16)	51 (19)

(continued on following page)

**TABLE A2.** Description of Demographic and Clinical Characteristics of Patients With aNSCLC Treated With PD-(L)1 Checkpoint Inhibitor (continued)

Patient Characteristic	CAS (N = 2,035)	Data Set A (n = 2,595)	Data Set B (n = 556)	Data Set C (n = 435)	Data Set D (n = 6,924)	Data Set E (n = 2,860)	Data Set F (n = 269)
≥ 4	33 (1.62)	85 (3)	44 (8)	65 (15)	481 (7)	221 (8)	54 (20)
Missing or unknown	25 (1.23)						
Follow-up time, months, median (Q1, Q3)							
From advanced diagnosis to PD-(L)1 inhibitor initiation	4.86 (1.61, 11.70)		7 (3, 14)	8 (4, 15)	6 (2, 13)	8 (3, 17)	7 (2, 14)
From advanced diagnosis	16.79 (10.61, 26.05)		18 (10, 28)	18 (10, 31)	14 (8, 25)	18 (10, 30)	18 (10, 28)
From PD-(L)1 inhibitor initiation (any LOT)	10.61 (3.91, 16.00)		8 (3, 16)	9 (3, 16)	6 (2, 12)	8 (3, 14)	8 (4, 13)

Abbreviations: aNSCLC, advanced non–small-cell lung cancer; CAS, Cancer Analysis System; IQR, interquartile range; LOT, line of therapy; NA, not applicable; NOS, not otherwise specified; NSCLC, non–small-cell lung cancer; PD-(L)1, programmed death-1 or programmed death ligand-1; Q1, quartile 1; Q3, quartile 3.

<sup>a</sup>Only stage IIIB patients were included in the analysis of CAS data.

**TABLE A3.** Median rwOS and 95% CI, Segmented by Treatment Setting and Patient Characteristics

Variable	CAS		Data Set A		Data Set B		Data Set C		Data Set D		Data Set E		Data Set F	
Age, years														
≤ 49	57	6.5 (4.5 to 9.0)	100	18.1 (11.9 to 21.6)	16	9.1 (2.7 to NA)	21	12.0 (4.3 to NA)	219	9.3 (7.8 to 12.1)	80	10.2 (7.9 to 17.4)	8	NA
50-64	347	10.5 (9.3 to 13.7)	723	13.6 (11.8 to 14.6)	164	8.3 (6.8 to 11.4)	129	9.3 (6.7 to 13.3)	2,047	9.3 (8.4 to 10.3)	863	13.8 (11.8 to 15.3)	65	16.9 (9.4 to NA)
65-74	590	12.2 (10.7 to 13.6)	728	13.4 (12.1 to 14.9)	131	8.6 (6.7 to 12.3)	169	8.9 (5.8 to 11.1)	2,504	9.3 (8.8 to 10.3)	1,047	12.2 (10.5 to 14.3)	94	12.3 (7.6 to 21.6)
≥ 75	250	12.2 (9.9 to 14.1)	593	13.2 (11.8 to 14.6)	51	7.0 (5.0 to 15.7)	116	6.8 (4.2 to 9.1)	2,153	8.8 (8.2 to 9.3)	870	13 (10.6 to 14.8)	102	10 (8.7 to 15.6)
Sex														
Female	532	13.5 (11.8 to 15.5)	950	13.7 (12.8 to 15.2)	187	8.9 (6.7 to 12.3)	212	9.3 (7.4 to 13.4)		8.4 (7.3 to 8.9)		14.8 (13.2 to 16.9)		13.2 (8.8 to 23.3)
Male	712	10.0 (8.9 to 11.4)	1,194	13.2 (12.1 to 14.5)	175	7.9 (6.8 to 10.1)	222	7.5 (6.4 to 10.0)	3,172	9.8 (9.4 to 10.7)	1,509	11.1 (10.2 to 12.4)	143	11.9 (8.4 to 20.4)
Stage														
0 or I					9	5.7 (0.5 to 9.7)			498	12.1 (10.7 to 14.0)			18	7 (4.9 to NA)
II					13	6.4 (1.5 to NA)			426	11.8 (10.6 to 13.3)			17	13.9 (4.0 to NA)
III <sup>a</sup>	196	15.8 (13.5 to 18.4)			62	9.6 (6.7 to NA)	39	8.9 (2.7 to 13.4)	1,494	9.8 (9.2 to 10.8)			63	14.9 (9.4 to 22.1)
IV	1,048	10.5 (9.6 to 11.9)			187	8.8 (6.8 to 10.6)	396	8.7 (6.8 to 10.0)	4,334	8.3 (7.8 to 8.8)			161	12.1 (8.3 to 20.7)
Unknown					91	7.9 (6.2 to 10.6)							10	NA
Histology														
NSQ	867	11.8 (10.4 to 13.2)			239	9.7 (7.7 to 12.6)	320	8.7 (6.7 to 10.2)	4,678	9.6 (9.1 to 10.3)	1,981	14.2 (12.7 to 15.8)	194	11.9 (8.8 to 21.8)
SQ	343	10.5 (8.5 to 12.3)			93	6.8 (4.9 to 8.8)	73	8.4 (4.8 to 12.1)	1,983	8.7 (7.8 to 9.3)	659	10.4 (9.2 to 11.8)	61	13.9 (8.6 to NA)
NOS	14	10.2 (5.0 to NA)			30	10.3 (5.1 to 13.3)	42	7.9 (3.6 to 17.2)			220	11.8 (7.9 to 16.1)	10	20.4 (6.1 to NA)
Others specified	20	21.3 (13.8 to NA)												
LOT number of first PD-(L)1 inhibitor														
1	526	13.4 (11.9 to 15.6)	592	19.8 (17.2 to 22.2)	74	9.4 (6.4 to 12.2)	80	9.2 (5.7 to 17.1)	2,074	10.4 (9.5 to 11.2)	777	20.8 (14.8 to 25.1)	77	15.9 (9.9 to NA)
2	590	10.2 (8.9 to 11.8)	1,174	11.7 (10.9 to 12.8)	191	7.9 (6.5 to 10.5)	205	7.8 (5.8 to 10.3)	3,357	8.7 (8.1 to 9.2)	1,414	10.7 (9.8 to 11.9)	87	9.1 (7.4 to 20.4)
3	95	7.5 (6.0 to 10.6)	304	12.8 (10.7 to 14.7)	63	10.7 (6.9 to 17.5)	85	9.3 (6.0 to 12.0)	1,011	9 (7.8 to 9.9)	448	10.7 (9.0 to 13.9)	51	15.3 (9.6 to NA)
4+	25	4.8 (3.3 to NA)	74	14.2 (10.1 to 17.2)	34	5.1 (2.1 to 11.9)	65	8.7 (5.3 to 13.3)	481	8.5 (6.9 to 10.5)	221	12 (8.3 to 15.6)	54	10.4 (6.9 to 21.8)
Missing	8	NA (14.6 to NA)												

Abbreviations: CAS, Cancer Analysis System; LOT, line of therapy; NA, not applicable; NOS, not otherwise specified; NSQ, nonsquamous; OS, overall survival; PD-(L)1, programmed death-1 or programmed death ligand-1; rwOS, real-world overall survival; SQ, squamous.

<sup>a</sup>Only stage IIIB patients were included in the analysis of CAS data.

**TABLE A4.** LOT for Age Group Below 50 Years

Variable	CAS, No. (%)
Overall	87 (100)
LOT number of first PD-(L)1 inhibitor (index LOT)	
1	30 (34.48)
2	45 (51.72)
3	—
4+	6 (6.90)
Missing or unknown	—

Abbreviations: CAS, Cancer Analysis System; LOT, line of therapy; PD-(L)1, programmed death-1 or programmed death ligand.

**TABLE A5.** Correlations Between rwOS and rwTTD in CAS Compared With US Data Sets Using Spearman's Rank Correlation Coefficient

Data Set	Comparison	No.	Correlation (95% CI)
CAS	rwOS v rwTTD	920	0.73 (0.69 to 0.76)
A	rwOS v rwTTD	254	0.63 (0.55 to 0.70)
B	rwOS v rwTTD	254	0.62 (0.54 to 0.69)
C	rwOS v rwTTD	295	0.89 (0.86 to 0.91)
D	rwOS v rwTTD	4,337	0.80 (0.79 to 0.81)
E	rwOS v rwTTD	1,456	0.77 (0.75 to 0.79)
F	rwOS v rwTTD	142	0.80 (0.66 to 0.85)

NOTE. Correlation analysis was restricted to those patients who had experienced both death and treatment discontinuation.

Abbreviations: CAS, Cancer Analysis System; rwOS, real-world overall survival; rwTTD, real-world time to treatment discontinuation.