Establishing a Framework to Evaluate Real-World Endpoints

July 10, 2018
Washington, DC

Introduction

Advances in data analytics and data capture through electronic health records (EHRs) and medical/pharmacy claims have brought the opportunities and challenges associated with using real-world evidence (RWE) to the forefront of the US healthcare industry. Increasingly, the promise of RWE to contribute to a more complete picture of the benefits and risks associated with therapies, when paired with results from randomized, controlled clinical trials, is being realized. RWE provides an opportunity to collect data rapidly on a broader patient population outside of a strict clinical trial protocol to help identify new indications or rare safety events, provide more generalizability of clinical trial results, and confirm clinical benefit in the post-market setting. Further, integration of the various sources of real-world data (RWD), including EHRs, clinical decision and support and hospital-based systems, administrative billing and claims databases, patient registries, longitudinal cohort studies, and patient reported outcomes tools, will yield a more robust dataset of RWE. However, the methods to aggregate data and the implications of integrating these multiple data sets as they evolve (especially in often dynamic post-approval settings) needs to be validated.

Applications for RWE extend the spectrum of therapeutics development from regulatory decision-making, to clinical use, to coverage and payment decisions. In the regulatory space, RWE has been utilized most frequently to evaluate drug safety through pharmacovigilance and adverse event monitoring in pre- and post-approval settings. However, RWE has increasingly been used to support effectiveness studies, in the form of historical data, as a surrogate for control arms in clinical trials (in the rare disease setting, for instance). Beyond regulatory decisions, RWE is frequently used to support clinical trial design, development of clinical practice guidelines, confirmation of population/subgroup size, and payment decisions including formulary placement.

These current applications of RWE in healthcare are quite limited with respect to the potential uses once appropriate standards and guardrails are implemented. Indeed, the pharmaceutical industry, FDA, and Congress recognize the importance of further developing this resource as evidenced by numerous recent publications by the FDA, passage of the 21st Century Cures Act (Cures Act), and the Prescription Drug User Fee Act (PDUFA) VI reauthorization. The Cures Act, passed in December of 2016, requires the FDA to develop a framework and issue guidance regarding the use of RWE to support a new indication for an already approved drug or post-market studies as a requirement for regulatory approval. Interestingly, FDA has already issued similar guidance regarding use of RWE for medical devices which includes supporting new indications and in post-approval studies. PDUFA VI builds upon the requirements of the Cures Act by instructing the FDA to consider stakeholder input through hosting of

public workshops as it develops its guidance for use of RWE. Other uses of RWE that could be imagined for future pharmaceutical approvals include expanded labels, pragmatic clinical trial design, and confirming benefit in the case of converting an Accelerated Approval to full approval status. In addition to potential regulatory uses, RWE could provide helpful information about the long-term value of a product and could inform future value assessments. For example, long-term efficacy endpoints that may not have been incorporated in pre-market clinical trial might be able to be captured using RWE, which requires increased understanding of how time-on-treatment or treatment discontinuation rates correlate to overall survival.

Significant progress has been made in data collection efforts to support use of RWE in regulatory settings, however challenges remain, chiefly with combining, organizing, and analyzing data from various information sources. Friends of Cancer Research proposes a pilot project, comprised of six leading healthcare data organizations, to develop a dataset curation process and validation framework to operationalize RWD collection and explore potential real-world endpoints that may be fit for regulatory purposes as well as assessing long-term benefits of a product.

Pilot Project Overview

Immunotherapies are being used to treat patients with cancers that have historically had few treatment options, which has generated high level of interest in their use and development. While immunotherapies have resulted in significant improvements in some patients, many other patients do not respond or only respond for a limited time. This has raised questions about the value of these new drugs. Applying current value frameworks to immune checkpoint inhibitors has proved difficult as they tend to underestimate the benefits of long-term survival and treatment-free survival.[1] This is likely due to the reliance on pivotal trial data, and in the setting of expedited approvals, assessments of the full clinical endpoints have not been completed. Thus, conclusions are often based on surrogate efficacy endpoints. At the initiation of this pilot project, three immune checkpoint inhibitors were approved for use in non-small cell lung cancer (NSCLC), which presented an opportunity to collect a robust amount of data for analysis from the post-market setting.

This pilot project was initiated to help determine whether RWD can be used to develop an early perspective on real-world outcomes, as defined by real-world endpoints from EHR and claims data, and whether these data correlate to overall survival (OS) in the context of randomized control trials (RCTs) for patients treated with novel therapies. The pilot project evaluates the performance of real-world endpoints across multiple data sets by focusing on a common question: ***What outcomes can be evaluated for aNSCLC patients treated with immune checkpoint inhibitors?***

To answer this question, a framework of necessary data elements, characteristics, and internal validation processes were proposed along with a set of definitions for real-world endpoints in the

[1] Ben-Aharon O, Magnezi R, Leshno M, Goldstein DA. Association of Immunotherapy with Durable Survival as Defined by Value Frameworks for Cancer Care. JAMA Oncol. 2018, 4(3):326–332.

context of their use in RCTs, FDA's regulatory framework, and data availability in EHR and claims systems. The pilot project will help evaluate whether the various data sets included in this study can achieve a similar level of correlation and statistical significance using a common framework.

Pilot Project Study Design and Objectives

This is a retrospective observational analysis of data derived from EHR and claims data. The data sets generated for the study include all relevant, retrospective patient-level data available for eligible individuals up to the data cutoff date, pending approval by a third-party de-identification.

Objective 1: Describe the demographic and clinical characteristics of aNSCLC patients treated with immune checkpoint inhibitors (Table 1)

Objective 2: Assess ability to generate real-world endpoints (OS, rwPFS, rwTTP, TTNT, TTD) in aNSCLC patients treated with immune checkpoint inhibitors, and segmented by clinical and demographic characteristics (Tables 2, 3, and 4)

Objective 3: Assess performance of real-world endpoints (rwPFS, rwTTP, TTNT, TTD) as surrogate endpoints for OS (Table 5)

Methods

Project Details	
Cohort and inclusion / exclusion criteria	aNSCLC patients treated with an immune checkpoint inhibitor (i.e., nivolumab, pembrolizumab, atezolizumab) Inclusion: • At least two documented clinical visits on or after January 1, 2011 until data cutoff date • Pathology consistent with NSCLC[2] • Has evidence of IIIB or IV NSCLC or has early stage NSCLC with a recurrence or progression described/documented in the EHR or claims • Treatment with immune checkpoint inhibitor, as documented by a structured medication order or claim as evidence of having received the treatment Exclusion: • Incomplete historical treatment data available within the database (i.e., patients whose advanced diagnosis date is more than 90 days before first activity date)

[2] For claims data, to minimize misclassification of aNSCLC, treatment with an IO agent following diagnosis of lung cancer was required. During the timeframe of this project, coverage for IO agents required evidence of advanced disease defined as either stage IIIB or IV NSCLC at initial diagnosis or early stage (stages I, II, and IIIA) NSCLC with a recurrence or progression.

EHR and Claims-derived endpoints definition and analytical guidance	**Overall survival (OS)** • *Data definition / computation*: length of time from the date the patient initiates the study treatment to the date of death or proxied by time to disenrollment. Patients without a date of death will be censored at their last known activity or date of disenrollment from the health plan identified and categorized as "due to death" if the date of death captured by SSA DMF was within 30 days prior or 60 days following. **Time to Next Treatment (TTNT)** • *Data definition / computation:* length of time from the date the patient initiates the study treatment to the date the patient initiates their next systemic treatment. When subsequent treatment is not received (e.g., continuing current treatment or disenrollment not due to confirmed death), patients will be censored at their last known activity. • Start date of regimen immediately after PD-(L)1 line (i.e., the subsequent systemic therapy after the initial PD-(L)1-containing regimen) **Time to Treatment Discontinuation (TTD)** • *Data definition / computation*: length of time from the date the patient initiates the PD-(L)1 regimen to the date the patient discontinues treatment. Patients still on treatment will be censored at their last known activity. • Event Date: Date of PD-(L)1 regimen discontinuation defined as last administration or non-cancelled order of a drug contained within the PD-(L)1 line regimen (between the line's start and end date) among patients that discontinued their immune checkpoint inhibitor therapy. Permanent discontinuation is defined as meeting one of the following conditions: o Having a subsequent systemic therapy after the initial PD-(L)1-containing regimen o Having a date of death while on the PD-(L)1-containing regimen Having a gap of more than 120 days between the last administration or non-cancelled order of the PD-(L)1 line and the patient's last visit or medication administration if there is no other systemic therapy after the PD-(L)1-containing regimen • Censor date: Patients without a discontinuation will be censored at their last known PD-(L)1 usage defined as the last administration or non-cancelled order of a drug contained within the PD-(L)1 regimen **Progression Event** • *Data definition / computation*: distinct episode in which the treating clinician concludes that there has been growth or worsening in the aNSCLC. The progression event (and date) is based on review of the patient chart. **Real-world Progression Free Survival (rwPFS)** • *Data definition / computation*: length of time from the date the patient initiates the PD-(L)1 regimen to the date that a progression event as evident in the EHR is documented in the patient's chart or the patient passes away.

	Patients without a progression event or date of death will be censored at the end of the patient's chart. **Real-world Time to Progression (rwTTP)** • *Data definition / computation*: length of time from the date the patient initiated the PD-(L)1 regimen to the date that a progression event is documented in the patient's EHR (excludes death as an event). Patients without a progression event will be censored at the end of the patient's chart. • Event date: Patient's first progression date more than 14 days after PD-(L)1 initiation as described in the index date definition. Death will not be considered a progression event in TTP • Censor date: Patients without a progression date more than 14 days after the index date or date of death (for PFS) will be censored at the last date the patient could have been assessed for progression (e.g., last clinic note date) **Index Date** • *Data definition / computation*: the earliest PD-(L)1 inhibitor initiation in the advanced setting anchored to start (e.g., first administration or non-cancelled order) of the immune-checkpoint inhibitor-containing regimen (nivolumab, pembrolizumab, atezolizumab).
Analyses	<u>Table 1:</u> • Assess ability to identify aNSCLC patients treated with immune checkpoint inhibitors • Description of demographic and clinical characteristics of aNSCLC patients treated with immune checkpoint inhibitors, example characteristics include: ○ Demographic: gender, age, SES, region ○ Clinical: histology, smoking status, group stage at time of initial diagnosis, follow up, biomarker status (e.g., ALK, EGFR, PD-L1), hepatic and renal function • Description of population characteristics for overall population and by treatment setting / line of therapy (e.g., 1st line metastatic, 2nd line, 3rd line plus) <u>Table 2, Table 3, and Table 4:</u> • Assess ability to generate real-world endpoints (OS, TTNT, TTD) for aNSCLC patients treated with immune checkpoint inhibitors within the advanced treatment setting (range and median figures) • Evaluate these endpoints when patient cohort is segmented by treatment setting and demographic /clinical characteristics <u>Table 5</u>: • Assess correlation of real-world endpoints (TTNT, TTD) to overall survival (OS)

Summary of Data Sources for Pilot Project Study

Cancer Research Network

The Cancer Research Network originated as an NCI-funded consortium of research groups affiliated with a dozen integrated health care systems across the US, and among whom the Health Care Systems Research Network was formed. In the early 2000's, the CRN created the Virtual Data Warehouse (VDW), a common data model to facilitate collaborative research. Data in the VDW are extracted from multiple source databases and maintained by each research group with the possibility of pooling data under specific IRB-approved research protocols. For most participating institutions, the VDW has essentially complete information on care dating back to 1996 or earlier for most data domains. Domains include health plan enrollment periods, cancer registries, encounters including diagnoses and procedures, prescription and infusion medications, laboratory results, and other areas. The data provided are results from one of the participating CRN organizations.

Cota

The Cota Real-World Evidence (RWE) database is a HIPAA-compliant, de-identified data source drawn from the electronic health records (EHR) of contributing academic, for-profit, and community oncologist provider sites and hospital systems. The database includes detailed demographic, diagnostic, molecular and genomic testing, treatment, and outcome data. As of 2018, Cota's RWE is comprised of rich longitudinal patient records collected from over 40 unique locations across North America. For the purposes of this pilot study, patient data was sourced from a predominantly community setting (98%).

Flatiron Health

Flatiron Health is a longitudinal, demographically and geographically diverse database derived from electronic health record (EHR) data from over 265 cancer clinics (~800 sites of care) including more than 2 million active US cancer patients available for analysis. The patient-level data in the EHRs includes structured data (e.g., laboratory values, and prescribed drugs) in addition to unstructured data collected via technology-enabled chart abstraction from physician's notes and other unstructured documents (e.g., biomarker reports).

IQVIA™

IQVIA™ is a leading global provider of information, innovative technology solutions and contract research services focused on using data and science to help healthcare clients find better solutions for their patients. For this engagement, IQVIA provided data sourced through Oncology Electronic Medical Records (EMR) from multiple partners, including TransMed. The data are comprised predominately of community practices (90%+). The integrated EMR platform includes activity from all payer types and all practice sizes across the United States. Results for this analysis were calculated primarily based on structured EMR fields.

Mayo Clinic Analysis using OptumLabs® Data Warehouse

OptumLabs® is an open, collaborative research and innovation center founded in 2013 as a partnership between Optum and Mayo Clinic. Its core linked data assets include de-identified claims data for privately insured and Medicare Advantage enrollees and de-identified electronic health record (EHR)

data from a nationwide network of provider groups. This pilot project was a retrospective analysis of claims data from the OptumLabs® Data Warehouse (OLDW), which includes de-identified claims data for privately insured and Medicare Advantage enrollees in a large, private, U.S. health plan. The database contains longitudinal health information on enrollees, representing a diverse mixture of ages, ethnicities and geographical regions across the United States. The health plan provides comprehensive full insurance coverage for physician, hospital, and prescription drug services.

PCORnet Sites

This pilot project included 11 PCORnet partner sites who had previously participated in a PCORnet Rapid Cycle Project. The 11 sites are based in healthcare systems within three PCORnet networks across 10 US states and include 10 academic medical centers. These sites were selected from 80 PCORnet partner sites because they could rapidly provide tumor registry data and linked electronic health records in PCORnet Common Data Model (CDM) format. The pooled database contributed to the RWE Endpoints Pilot Project consisted of tumor registry data from each site and linked CDM diagnosis, procedures, prescribing, dispensing, medication administration, and death data tables. Data sources for the CDM include institutional billing and electronic health record data. The study cohort includes patients with a single primary advanced stage non-small cell lung cancer (NSCLC) diagnosis who were either diagnosed at stage 3b or 4 or who had an ICD9/10 diagnosis code for secondary metastasis.

Table 1. Description of demographic and clinical characteristics of aNSCLC patients treated with PD-(L)1 checkpoint inhibitors

Demographics	Data Set A PD-(L)1-treated N=2595	Data Set B PD-(L)1-treated N=557	Data Set C PD-(L)1-treated N=435	Data Set D PD-(L)1-treated N=6924	Data Set E PD-(L)1-treated N=2860	Data Set F PD-(L)1-treated N=269
Age at advanced diagnosis (years), median [IQR]	68 [15]	64 [14]	66 [14]	69 [14]	68 [14]	70 [14]
Age at PD-(L)1 inhibitor initiation (years), median [IQR]	69 [14]	65 [14]	68 [14]	69 [14]	69 [14]	71 [14]
Age categories at PD-(L)1 inhibitor initiation (categorical):						
≤49 years	120 (5%)	24 (4%)	21 (5%)	219 (3%)	80 (3%)	8 (3%)
50-64 years	888 (34%)	251 (45%)	129 (30%)	2048 (30%)	863 (30%)	65 (24%)
65-74 years	866 (33%)	198 (36%)	169 (39%)	2504 (36%)	1047 (37%)	94 (35%)
75+ years	721 (28%)	84 (15%)	116 (27%)	2153 (31%)	870 (30%)	102 (38%)
Age categories at PD-(L)1 inhibitor initiation (binary):						
<75 years	1874 (72%)	473 (85%)	319 (73%)	4771 (69%)	1990 (70%)	167 (62%)
75+ years	721 (28%)	84 (15%)	116 (27%)	2153 (31%)	870 (30%)	102 (38)
Gender:						
Female	1147 (44%)	276 (50%)	212 (49%)	3172 (46%)	1351 (47%)	125 (46%)
Male	1448 (56%)	281 (50%)	222 (51%)	3752 (54%)	1509 (53%)	143 (53%)
Unknown/Missing	0	0	≤5	0	0	1
Race/ethnicity:						
White	1704 (78%)	478 (86%)	284 (65%)	4969 (79%)	676 (87%)	160 (87%)
Black or African American	282 (13%)	67 (12%)	37 (9%)	594 (9%)	44 (6%)	14 (8%)
Asian	52 (2%)	6 (1%)	83 (19%)	155 (3%)	13 (2%)	9 (5%)
Other Race	142 (7%)	6 (1%)	31 (7%)	580 (9%)	42 (5%)	1 (1%)
Unknown/Missing	415	0	0	626	2085	85
Median household income (zip-level):						
1 (lowest median household income)			103 (24%)	1003 (15%)		
2			105 (24%)	1539 (22%)		
3			114 (26%)	1833 (27%)		
4 (highest median household income)			112 (26%)	2525 (37%)		
Unknown			≤5	24		
CLINICAL CHARACTERISTICS						
Group stage at initial diagnosis:						
Stage 0 / Occult		0		2 (0%)		
Stage I		23 (6%)		496 (7%)		18 (7%)
Stage II		22 (6%)		426 (6%)		17 (7%)
Stage III		88 (23%)	39 (9%)	1494 (22%)		17 (7%)
Stage IV		248 (65%)	396 (91%)	4335 (64%)		161 (62%)
Group stage is not reported		176		171		10
Histology:						
Non-squamous cell carcinoma		370 (66%)	320 (74%)	4679 (70%)	1981 (69%)	194 (73%)
Squamous cell carcinoma		147 (26%)	73 (17%)	1983 (30%)	659 (23%)	61 (23%)

NSCLC histology not otherwise specified (NOS)		40 (7%)	42 (10%)	262 (3%)	220 (8%)	10 (4%)
Missing						4
Smoking status:						
History of smoking			340 (78%)	6185 (90%)	448 (92%)	182 (87%)
No history of smoking			94 (22%)	717 (10%)	38 (8%)	28 (13%)
Unknown/Not documented			≤5	22	2374	210
PD-L1 tested on or prior to PD-(L)1 inhibitor start	326 (13%)			2384 (34%)	96	80/96 (83%)
PD-L1 expression status (among those tested):						
PD-L1 positive				512 (22%)	45 (50%)	65 (68%)
PD-L1 negative/not detected				691 (29%)	45 (50%)	29 (30%)
Unsuccessful/indeterminate test				1012 (42%)	0	2 (2%)
Results pending/unknown				169 (7%)	6	173
ALK tested on or prior to PD-(L)1 inhibitor start	258 (10%)			4513 (65%)	582	143/173 (83%)
ALK status (among those tested):						
Rearrangement present				57 (1%)	8 (1%)	1 (1%)
Rearrangement not present				4145 (92%)	570 (99%)	170 (98%)
Results pending/unknown				68 (2%)	0	2 (1%)
Unsuccessful/indeterminate test				243 (5%)	4	96
EGFR tested on or prior to PD-(L)1 inhibitor start	543 (21%)		171(39%)	4684 (68%)	953	115/142 (81%)
EGFR status (among those tested)[2,3]:						
Mutation positive				305 (7%)	68 (11%)	6/142 (4%)
Mutation negative				4161 (89%)	525 (89%)	135/142 (95%)
Results pending/unknown				60 (1%)	358	1/142 (1%)
Unsuccessful/indeterminate test				158 (3%)	2	127
No prior therapy received	690 (27%)		80 (18%)	2074 (30%)	777 (27%)	77 (29%)
Line number of first PD-(L)1 inhibitor in advanced setting:						
1	690 (27%)		80 (18%)	2074 (30%)	777 (27%)	77 (29%)
2	1440 (56%)		205 (47%)	3357 (49%)	1414 (49%)	87 (32%)
3	380 (15%)		85 (20%)	1012 (15%)	448 (16%)	51 (19%)
4+	85 (3%)		65 (15%)	481 (7%)	221 (8%)	54 (20%)
Patients receiving a second PD-(L)1 inhibitor in a subsequent line:						
No			402 (92%)	1740 (25%)		
No subsequent therapy received				4879 (71%)		
Yes	93 (4%)		33 (8%)	305 (4%)	112	14
Line number of second PD-(L)1 inhibitor in advanced setting:						
2	28 (30%)		11 (33%)	99 (33%)	9 (8%)	5 (36%)
3	45 (48%)		10 (30%)	134 (44%)	51 (46%)	4 (29%)
4+	20 (22%)		12 (36%)	72 (24%)	52 (46%)	5 (36%)
N/A			402			
Time from advanced diagnosis to first PD-(L)1 inhibitor initiation (months), median [IQR]		7 [11]	8 [11]	6 [11]	8 [14]	7 [12]

Structured follow up time[3]					
Structured follow-up time from advanced diagnosis (months), median [IQR]	18 [18]	18 [21]	14 [17]	18 [20]	18 [18]
Structured follow-up time from PD-(L)1 inhibitor initiation (months), median [IQR]	8 [13]	9 [13]	6 [10]	8 [11]	8 [9]

[3] Structured follow-up time is calculated from the relevant time-point for each patient until their last structured activity (i.e., most recent visit or administration)

Table 2. Median time and 95% confidence interval for real-world extracted endpoints

Data Set	rwOS	rwTTNT	rwTTD	rwTTP	rwPFS
Data Set A	13.50 [12.80, 14.50][4]	22.50 [NA]	7.03 [6.27, 9.97]		
Data Set B	15.78 [12.2, 24.59]; 8.58 [7.56, 10.26][5]		3.25[2.76, 3.75]		
Data Set C	8.67 [6.83, 10.02]	11.60 [8.80, 16.10]	4.70 [3.68, 5.52]		
Data Set D	9.15 [8.82, 9.51]	14.03 [12.89, 15.15]	3.21 [3.21, 3.44]	5.41 [5.18, 5.67]	3.28 [3.18, 3.41]
Data Set E	12.69 [11.7, 13.87]	12.07 [11.24, 13.48]	3.63 [3.40, 3.87]		
Data Set F	12.30 [9.61, 16.94]	12.50 [9.29, NA]	4.60 [3.71, 6.32]	9.37 [7.42, 11.93]	9.37 [7.42, 11.93]

Data sets measured median time for real-world extracted endpoints utilizing a common definition as described in the pilot project methods

Table 3. One-year real-world overall survival landmark analysis post PD-(L)1 initiation

Data Set	One-year rwOS Landmark Analysis
Data Set A	0.57 [0.52, 0.57][4]
Data Set B	0.54 [0.48, 0.57]; 0.41 [0.34, 0.47][5]
Data Set C	0.40 [0.35, 0.46]
Data Set D	0.42 [0.41, 0.43]
Data Set E	0.51 [0.49, 0.53]
Data Set F	0.40 [0.34, 0.48]

[4] OS was calculated as days between I/O initiation and disenrollment.

[5] Sites with social security or state death data, censored at estimated earliest date such data should be available if no death was observed

Table 4. Median times and 95% confidence interval (indexed to initial PD-(L)1 inhibitor line start in advanced setting) segmented by treatment setting and demographic characteristics as described in Table 1

Demographics		Data Set A		Data Set B		Data Set C		Data Set D		Data Set E		Data Set F
	N	rwTTD (Months) Median [95% CI]	N	rwTTD (Months) Median [95% CI]	N	rwTTD (Months) Median [95% CI]	N	rwTTD (Months) Median [95% CI]	N	rwTTD (Months) Median [95% CI]	N	rwTTD (Months) Median [95% CI]
		rwOS (Months) Median [95% CI][6]	N	rwOS (Months) Median [95% CI] [7]		rwOS (Months) Median [95% CI]		rwOS (Months) Median [95% CI]		rwOS (Months) Median [95% CI]		rwOS (Months) Median [95% CI]
Age categories at PD-(L)1 inhibitor initiation:												
≤49 years	100	5.87 [3.97, 9.80]	24	3.97 [1.84, 17.03]	21	6.44 [1.28, 16.29]	219	2.89 [2.30, 3.64]	80	2.33 [1.43, 4.40]	8	8.52 [2.77, NA)
		18.1 [11.87, 21.63]	16	9.07 [2.66, NA]		12.02 [4.27, NA)		9.28 [7.77, 12.07]		10.20 [7.96, 17.36]		NA
50-64 years	723	6.23 [5.20, 7.27]	251	3.62 [2.96, 4.14]	129	5.35 [3.35, 9.23]	2047	3.21 [2.92, 3.44]	863	3.53 [3.17, 4.03]	65	6.16 [3.71, 12.26]
		13.60 [11.83,14.6]	163	8.78 [6.81, 11.41]		9.33 [6.73, 13.27]		9.34 [8.43, 10.26]		13.84 [11.80, 15.32]		16.94 [9.37, NA]
65-74 years	728	7.40 [6.10, 9.60]	198	2.76 [2.27, 3.75]	169	4.63 [3.45, 6.44]	2504	3.41 [3.21, 3.67]	1047	3.77 [3.30, 4.23]	94	5.38 [3.42, 6.90]
		13.40[12.07,14.93]	135	8.52 [6.54, 9.70]		8.97 [5.78, 11.14]		9.34 [8.79, 10.26]		12.16 [10.49, 14.33]		12.30 [7.55, 21.58]
75+ years	593	7.70 [6.37, 9.37]	84	2.47 [1.45, 4.64]	116	3.57 [1.84, 5.26]	2153	3.25 [3.18, 3.61]	870	3.77 [3.30, 4.23]	102	3.67 [2.82, 5.74]
		13.22 [11.83,14.61]	49	10.65 [5.03, NA]		6.83 [4.24, 9.13]		8.79 [8.23, 9.28]		13.02 [10.62, 14.79]		10.00 [8.71, 15.55]
Gender:												
Female	950	6.80 [5.90, 8.23]	276	3.53 [2.76, 3.98]	212	4.76 [3.45, 7.72]	3751	2.98 [2.75, 3.21]	1351	3.83 [3.53, 4.27]	125	5.03 [3.77,7.03]
		13.70 [12.8, 15.23]	188	8.94 [6.67, 12.33]		9.33 [7.42, 13.44]		8.43 [7.93, 8.98]		14.79 [13.15, 16.87]		13.23 [8.77, 23.33]
Male	1194	7.23 [6.10, 8.43]	281	3.16 [2.37, 3.81]	222	4.62 [3.35, 5.52]	3172	3.51 [3.21, 3.70]	1509	3.30 [2.87, 3.77]	143	3.87 [2.93, 6.40]
		13.20 [12.13, 14.5]	175	7.92 [6.77, 10.06]		7.49 [6.34, 10.02]		9.84 [9.38, 10.72]		11.08 [10.16, 12.39]		11.87 [8.40, 20.40]
CLINICAL CHARACTERISTICS												
Group stage at initial diagnosis:												
Stage 0 / 1			23	3.45 [0.92, 5.69]			498	4.36 [3.67, 5.28]			18	5.18 [3.43, 15.65]
			9	5.69 [0.49, 9.70]				12.07 [10.69, 14.03]				7.03 [4.87, NA]

[6] rwOS was calculated as time between I/O initiation and disenrollment

[7] rwOS estimates include sites with social security or state death data available; excluded are sites with only local/EHR death data available

Stage II	22	3.68 [1.41, 6.38]			426	3.90 [3.28, 4.95]			17	4.10 [2.80, NA]
	13	6.38 [1.48, NA]				11.84 [10.59, 13.28]				13.90 [4.03, NA]
Stage III	88	4.14 [2.76, 3.88]	39	4.43 [1.38, 10.35]	1494	3.67 [3.44, 4.13]			63	5.73 [2.90, 9.67]
	62	9.60 [6.67, NA]		8.97 [2.73, 13.44]		9.84 [9.18, 10.79]				14.87 [9.37, 22.10]
Stage IV	248	3.35 [2.76, 3.88]	396	4.70 [3.68, 5.52]	4334	2.89 [2.75, 3.18]			161	4.32 [3.19, 6.47]
	187	8.77 [6.77, 10.85]		8.67 [6.83, 10.02]		8.26 [7.80, 8.79]				12.10 [8.30, 20.73]
Unknown	176	5.79 [1.87, 3.85]							10	8.83 [3.74, NA]
	92	7.92 [6.15, 10.55]								NA
Histology:										
Non-squamous cell carcinoma	370	3.16 [2.53, 3.75]	320	4.76 [3.81, 5.81]	4678	3.34 [3.21, 3.51]	1981	3.60 [3.30, 3.90]	194	4.15 [3.23, 5.97]
	240	9.69 [7.56, 12.33]		8.67 [6.7, 10.18]		9.61 [9.11, 10.30]		14.14 [12.69, 15.81]		11.87 [8.77, 21.80]
Squamous cell carcinoma	147	3.25 [2.47, 4.01]	73	4.14 [1.68, 7.33]	1983	3.21 [2.98, 3.54]	659	3.77 [3.30, 4.30]	61	7.55 [3.67, 12.97]
	93	6.80 [4.87, 8.78]		8.38 [4.8, 12.06]		8.66 [7.84, 9.25]		10.36 [9.24, 11.77]		13.90 [8.57, NA]
NSCLC histology not otherwise specified (NOS)	40	3.88 [1.94, 5.10]	42	3.80 [2.07, 9.13]			220	3.57 [2.83, 5.17]	10	5.07 [2.29, NA]
	30	10.26 [5.13, 13.32]		7.92 [3.58, 17.15]				11.84 [7.89, 16.08]		20.40 [6.13, NA]
Smoking status:										
History of smoking			340	4.53 [3.45, 5.35]	6185	3.28 [3.21, 3.48]	448	5.17 [3.90, 6.53]	182	4.27 [3.43, 6.19]
				8.67 [6.7, 10.18]		9.21 [8.85. 9.64]		19.17 [14.30, 24.46]		11.43 [8.40, 20.73]
No history of smoking			94	5.52 [3.22, 9.20]	716	2.75 [2.49, 3.11]	38	3.30 [2.37, 9.07]	28	6.08 [3.23, 8.71]
				8.34 [6.04, 12.88]		8.69 [8.03, 9.77]		14.50 [4.57, NA]		12.10 [8.71, NA]
Unknown/Not documented							2374	3.53 [3.30, 3.77]	59	5.97 [2.77, 10.00]
								12.00 [10.82, 13.08]		13.21 [9.37, NA]
PD-L1 expression status (among those tested):										
PD-L1 positive					512	4.10 [3.38, 4.82]	45	5.63 [2.83, 18.23]	65	3.68 [2.82, 5.97]
						10.79 [9.05, 13.28]		NA		9.63 [7.55, NA]
PD-L1 negative/not detected					690	2.75 [2.49, 3.11]	45	9.63 [NA]	29	6.32 [4.19, NA]
						8.69 [7.48, 9.84]		NA		20.80 [11.43, NA]
Line number of first PD-(L)1 inhibitor in advanced setting:										
1	592	9.10 [7.97, 12.40]	80	5.26 [2.63, 6.64]	2074	3.90 [3.67, 4.23]	777	5.90 [4.93, 6.80]	77	5.03 [3.67, 8.83]
		19.83 [17.23, 22.23]		9.17 [5.68, 17.05]		10.36 [9.48, 11.18]		20.78 [14.79, 25.12]		15.87 [9.87, NA]

2	1174	6.57 [6.03, 7.50] 11.7 [10.97, 12.83]	205	4.53 [3.22, 5.81] 7.75 [5.75, 10.28]	3357	3.21 [2.82, 3.21] 8.66 [8.13, 9.15]	1414	3.00 [2.83, 3.30] 10.68 [9.83, 11.93]	87	4.81 [2.74, 7.00] 9.07 [7.39, 20.40]
3	304	4.47 [3.8, 6.1] 12.8 [10.7, 14.67]	85	4.63 [2.30, 8.05] 9.33 [6.04, 12.02]	1011	2.98 [2.75, 3.44] 9.02 [7.80, 9.97]	448	3.53 [3.00, 4.23] 10.72 [9.04, 13.87]	51	5.67 [3.43, 8.71] 15.29 [9.63, NA]
4+	74	3.83 [2.83, 5.47] 14.2 [10.1, 17.17]	65	4.76 [3.94, 11.10] 8.67 [5.26, 13.27]	481	2.59 [2.30, 3.18] 8.52 [6.89, 10.46]	221	3.30 [2.60, 4.23] 12.00 [8.25, 15.58]	54	3.47 [2.77, 7.16] 10.43 [6.90, 21.80]

Table 5. Correlation between real-world overall survival and real-world extracted endpoints using Spearman's rank correlation coefficient

Data Set	Comparison	N	Correlation (95% CI)
Data Set A	rwOS vs rwTTNT	83	0.36
	rwOS vs rwTTD	254	0.63
Data Set B	rwOS vs rwTTNT		
	rwOS vs rwTTD	225	0.62 (0.54, 0.69)
Data Set C	rwOS vs rwTTNT	96	0.70 (0.58, 0.79)
	rwOS vs rwTTD	295	0.89 (0.86, 0.91)
Data Set D	rwOS vs rwTTNT	1203	0.61 (0.57, 0.64)
	rwOS vs rwTTD	4337	0.80 (0.79, 0.81)
	rwOS vs rwPFS	4337	0.75 (0.74, 0.76)
	rwOS vs rwTTP	2286	0.60 (0.57, 0.63)
Data Set E	rwOS vs rwTTNT	358	0.62 (0.54, 0.68)
	rwOS vs rwTTD	1456	0.77 (0.75, 0.79)
Data Set F	rwOS vs rwTTNT	39	0.46 (0.33, 0.81)
	rwOS vs rwTTD	142	0.80 (0.66, 0.85)
	rwOS vs rwPFS	142	0.84 (0.62, 0.86)
	rwOS vs rwTTP	55	0.56 (0.21, 0.71)

The correlation analysis is restricted to patients with a death date and documented event as described in the definitions and algorithms

Conclusions from Pilot Project Study

1. There is a high level of shared characteristics among the varying data sets despite varying sample sizes, data capture processes, and data sources demonstrating the feasibility of identifying aNSCLC patients treated with immune checkpoint inhibitors from diverse RWD sources.
2. The pilot project demonstrated that several extractable endpoints from EHR and claims data correlate with OS. Further validation is required to determine whether these endpoints are reliable surrogates for OS outside of a traditional clinical trial and whether they can support regulatory and payer decision-making.
3. Survival among patients as assessed through EHR and claims data fall within the range of median OS values observed in several immune checkpoint inhibitor trials.[8]
4. Assessment of extracted endpoints from EHR and claims data demonstrate that efficacy of immune checkpoint inhibitors is relatively consistent across a variety of patient characteristics, such as age and sex.

Assumptions and Limitations of Pilot Project Data Sets

- Ability to collect reliable data will vary across data providers
- Approaches to analysis may vary even when using a common protocol; A careful review and collaboration is needed to align on a consistent and reliable approach
- Verified diagnosis and diagnosis date, clinical stage and cell type, planned chemotherapy regimen (dose and schedule) and other clinical and socioeconomic factors cannot always be determined from the available EHR and claims data
- Verifying and determining date of death may also prove challenging. Although discharge status and some diagnosis codes may be a source of mortality information, but some data partners rely on linkage to the public SSA death master file (DMF). The public DMF has been shown to under identify deaths[9]
- For claims-based data, some patients with advanced disease may enroll in clinical trials and some or all the care received in a clinical trial setting may not generate insurance claims, thus, data for these patients may not be fully captured or captured at all
- Approaches to the analyses may vary even when using a common protocol and careful review and collaboration is needed to align on a consistent and reliable approach
- Some biomarkers may not routinely be assessed in the real-world setting, but more would have been included in this analysis if a chart review had been conducted or the use of natural language processing (NLP)
- Provider data (EHR) may not identify all chemotherapy as patients may seek care inside and outside a provider group that contributes to the EHR data (e.g., chemotherapy at an academic center then move to a community setting). This may or may not be a source of missing information in the advance NSCLC setting

[8] Huang G, Sun X, Liu D, et al. The efficacy and safety of anti-PD-1/PD-L1 antibody therapy versus docetaxel for pretreated advanced NSCLC: a meta-analysis. Oncotarget, 4239-4248

[9] Jones B, V. D. (2015, March). Measuring Mortality Information in Clinical Data Warehouses. AMIA Jt Summits Transl Sci Proc, 450-5

Discussion Questions

These questions may help guide the discussion during the meeting:

1. Are there processes to handle challenges associated with the availability and consistency of data across provider types and settings?

2. How to overcome difficulties associated with determining events like death?

3. What opportunities or incentives exist to help improve the format, quality, and validity of RWE?

4. Are there lessons from clinical trials, or registration trials, that need to be considered for RW data?

5. What opportunities exist for FDA decision-making to be supported by RWE?

6. What opportunities exist to expand to other endpoints such as patient reported outcomes (PROs) and patient-generated health data?

7. Are there other extractable endpoints for EHR- or claims-based algorithms that should be validated?

8. What is the role and use of real-world endpoints, such as TTD, TTNT, or PFS, for payer decision-making, particularly in the context of accelerated approval or breakthrough therapy designation?

9. How important is RWE in the development of new payment designs, such as value-based payment, risk-sharing arrangements, and outcomes-based agreements?

10. How timely does the data have to be for regulatory or reimbursement? How quickly must the data be analyzed/reported?

11. For reimbursement/value-based payment/risk sharing, are data from all data sets (A-F) available to payers? Manufacturers?