




Analysis of a Real-World Progression Variable and Related Endpoints for Patients with Five Different Cancer Types

Aracelis Z. Torres · Nathan C. Nussbaum · Christina M. Parrinello ·
Ariel B. Bourla · Bryan E. Bowser · Samuel Wagner · David C. Tabano ·
Daniel George · Rebecca A. Miksad 

Received: November 23, 2021 / Accepted: February 14, 2022 / Published online: April 17, 2022
© The Author(s) 2022

ABSTRACT

Introduction: We previously demonstrated that real-world progression (rwP) can be ascertained from unstructured electronic health record (EHR)-derived documents using a novel abstraction approach for patients with advanced non-small cell lung cancer (base case). The objective of this methodological study was to assess the reliability, clinical relevance, and the need for disease-specific adjustments of this abstraction approach in five additional solid tumor types.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12325-022-02091-8>.

A. Z. Torres · N. C. Nussbaum · C. M. Parrinello ·
A. B. Bourla · B. E. Bowser · R. A. Miksad (✉)
Flatiron Health, Inc., 233 Spring Street, Fifth Floor,
New York, NY 10025, USA
e-mail: ramiksad@flatiron.com

N. C. Nussbaum
New York University School of Medicine, New York,
NY, USA

C. M. Parrinello
Pine Mountain Consulting, LLC., Redding, CT, USA

S. Wagner · D. C. Tabano
Bristol-Myers Squibb, Princeton, NJ, USA

D. George
Duke University, Durham, NC, USA

Methods: Patients with metastatic breast cancer (mBC), advanced melanoma (aMel), small cell lung cancer (SCLC), metastatic renal cell carcinoma (mRCC), and advanced gastric/esophageal cancer (aGEC) were selected from a real-world database. Disease-specific additions to the base case were implemented as needed. The resulting abstraction approach was applied to each disease cohort to capture rwP events and dates. To provide comprehensive clinical context, real-world progression-free survival (rwPFS) and time to progression (rwTTP) were compared to real-world overall survival (rwOS), time to next treatment (rwTTNT), and time to treatment discontinuation (rwTTD). Endpoint estimates were assessed using the Kaplan–Meier method. Correlations between real-world endpoints and rwOS were calculated using Spearman's ρ .

Results: Additions to the base-case rwP abstraction approach were required for mBC, aMel, and SCLC. Inter-abstractor agreement for rwP occurrence, irrespective of date, ranged from 88% to 97%. Occurrence of clinically relevant downstream events (new antineoplastic systemic therapy start, antineoplastic systemic therapy end, or death relative to the rwP event) ranged from 59% (aMel) to 72% (mBC). Median rwPFS ranged from 3.7 (aMel) to 7.7 (mBC) months, and median rwTTP ranged from 4.6 (aMel) to 8.3 (mRCC) months. Correlations between rwOS and rwPFS ranged from 0.52 (aMel) to 0.82 (SCLC). The correlation between

rwOS and rwTTD was often lower relative to other comparisons (range 0.40–0.62).

Conclusion: Derivation of a rwP variable from EHR documentation is feasible and reliable across the five solid tumors. Endpoint analyses show that rwP produces clinically meaningful information.

Keywords: Abstraction; Electronic health record; Oncology; Real-world data; Survival

Key Summary Points

Why carry out this study?

The identification of cancer progression events (i.e., disease worsening) is important for assessing therapeutic benefit. Measuring cancer progression using real-world data (RWD) requires distinct methodology from measuring progression in a clinical trial setting

We previously developed a novel method to reliably ascertain real-world cancer progression (rwP) in a cohort of patients with advanced non-small cell lung cancer (aNSCLC) using de-identified data from an electronic health record (EHR)-derived database (base case). We conducted this methodological study to determine whether the same method could identify cancer progression in five additional solid tumor types with a range of disease characteristics: metastatic breast cancer (mBC), advanced melanoma (aMel), small cell lung cancer (SCLC), metastatic renal cell carcinoma (mRCC), and advanced gastric/esophageal cancer (aGEC)

What was learned from this study?

Our results show that, with disease-specific additions to the base case for mBC, aMel, and SCLC, derivation of rwP from EHR documentation is feasible across the five additional cancers, despite differences in tumor biology

In addition, rwP can be used in endpoint analyses to produce clinically meaningful information that may be valuable in research

We believe our approach for reliably measuring rwP in multiple tumor types is a key metric to assess interventions to improve outcomes and enhance survival and quality of life for patients with cancer

INTRODUCTION

Real-world data (RWD) are data collected as part of routine health assessment and care delivery from sources outside of a typical clinical trial setting, such as electronic health records (EHRs), medical claims and billing data, or disease registries [1, 2]. While not gathered for research purposes, RWD can generate real-world evidence (RWE) that may be leveraged to help address issues of external validity and generalizability for drug development programs [1–4]. RWE has also been used in regulatory and clinical decision-making, outcomes research, and safety surveillance [4–9]. Insights gained from RWE can be especially useful in scenarios where conducting a randomized controlled trial (RCT) is difficult or unethical, such as in rare, genomically defined cancers [4, 6, 10]. In order to support research efforts that use RWE, the US Food and Drug Administration (FDA) has developed a framework that outlines current use cases for RWE as well as opportunities in which RWE may have the potential to supplement or complement clinical trial data [5, 11]. In these contexts, robust and well-characterized outcome measures are crucial.

While improved overall survival (OS) remains a standard threshold to demonstrate the therapeutic benefit of an anticancer intervention, the identification of cancer progression events (and the evaluation of associated endpoints) is also important, especially to understand the durability of treatment effect [12–15]. In clinical trials of solid tumors, tumor burden changes are typically measured by applying Response Evaluation Criteria in Solid

Tumors (RECIST) to imaging in order to determine response or progression events based on changes in tumor size [16, 17]. However, in a previous study, we found that it was not feasible to routinely use RECIST to retrospectively assess tumor size changes based on community oncology EHR-derived documents because of the lack of consistent documentation in imaging reports [18]. Accordingly, the analysis of non-randomized RWD to ascertain real-world progression (rwP) requires specific methodological considerations and abstraction approaches that account for issues such as data sourcing, quality, and completeness in order to generate meaningful and actionable insights [10, 19, 20].

Previously, we demonstrated that rwP can be identified from unstructured EHR-derived documents (e.g., clinician notes) using a novel abstraction approach that was applied to a cohort of patients with advanced non-small cell lung cancer (aNSCLC) [18]. This approach was found to be reliable, clinically meaningful, and scalable to a cohort of more than 30,000 patients; rwP-based endpoints (e.g., rwPFS, real-world time to progression [rwTTP]) were associated with downstream clinical events and were correlated with treatment-based endpoints (e.g., real-world time to next treatment [rwTTNT]) and with rwOS [21, 22].

Curating rwP from EHR documents for other tumor types would enable new uses of RWD, such as evaluating new therapies in a rapidly changing treatment landscape. Given that different cancer types have different biologies as well as progression patterns and dynamics, the overall objective of this methodological study was to assess the performance of the previously developed rwP abstraction approach in aNSCLC (the base case) for five additional solid tumor types. First, we evaluated feasibility of the base-case approach for each tumor type and if disease-specific additions to the base case were needed. We hypothesized that different source evidence may be referenced in routine clinical care in various disease settings. For the resulting abstraction approach for each tumor type, we repeated the characterization analysis performed for the aNSCLC base case; we assessed the feasibility, reliability, and performance (e.g.,

likelihood of a downstream clinical event relative to rwP and association of rwP-based endpoints with rwOS). Comparisons to treatment-based endpoints provided additional clinical context. Finally, we determined the suitability of implementing each abstraction approach in very large cohorts through a holistic, qualitative assessment of the rwP variable's characteristics. We hypothesize that ascertaining rwP from EHR documentation using a previously developed approach with disease-specific additions as needed is feasible for five additional tumor types that demonstrate a range of disease characteristics.

METHODS

Data Source

This retrospective study used data from the Flatiron Health database, a nationwide longitudinal, de-identified database derived from EHR data containing patient-level structured and unstructured data curated via technology-enabled abstraction [23, 24]. The analyses utilized data available at the time each experiment was conducted; by the end of the observation period used for the last set of analyses (December 31, 2018), the database originated from approximately 280 cancer clinics (ca. 800 sites of care) in the USA and Puerto Rico. The majority of patients in the database were drawn from community oncology settings and the rest from academic medical centers. All data were drawn from EHR documentation in these care settings (i.e., care received elsewhere not documented in the source EHR is not available).

The previously published rwP abstraction approach studied in aNSCLC was defined as the "base case". Five additional tumor types were evaluated in this study: metastatic breast cancer (mBC), advanced melanoma (aMel), small cell lung cancer (SCLC), metastatic renal cell carcinoma (mRCC), and advanced gastric/esophageal cancer (aGEC). These tumor types were selected on the basis of the availability of a curated dataset and because assessment of disease burden was typically based on imaging obtained at a regular cadence, similar to

aNSCLC in the base case. The specific inclusion and exclusion criteria for the individually analyzed cohorts varied by tumor type owing to disease-specific diagnostic and treatment nuances; however, all patients were required to have confirmation of diagnosis with the relevant tumor type through review of unstructured EHR documents and structured visit data. All diagnoses were confirmed by trained abstractors following standard procedures and protocols [23]. In addition, at the time of study entry all patients had initiated systemic therapy as indicated by oncologist-defined, rule-based line of therapy comprised of structured antineoplastic medication orders, structured antineoplastic medication administrations, and/or oral therapy data from unstructured EHR documents, depending on the tumor type (Table 1).

This observational study was performed in accordance with the Helsinki Declaration of 1964 and its later amendments. Institutional review board (IRB) approval of the study protocol, with a waiver of informed consent, was obtained prior to study conduct, and covers the data from all sites represented. Approval was granted by the WCG IRB. (“The Flatiron Health Real-World Evidence Parent Protocol”; IRB registration number IRB00000533; Protocol approval ID tracking number 420180044.)

Feasibility and Assessment for Disease-Specific Additions to the Base Case

The base-case approach for aNSCLC leverages documentation in the EHR of the clinician’s interpretation of source evidence (i.e., radiographic, pathologic, or clinical assessment only) [18]. For initial feasibility assessment of the base-case approach, rwP events were abstracted for 20–30 randomly selected patients for each of the five additional solid tumor types, starting at the cohort entry date (diagnosis date, advanced diagnosis date, or metastatic diagnosis date, depending on the tumor type). This feasibility step also explored whether additions to the base case to incorporate disease-specific clinical and documentation factors were necessary. Research scientists with relevant oncology expertise performed a qualitative assessment of the evidence

sources referenced in EHR documentation to support determination of worsening disease that are not captured by the base-case approach (e.g., physical exam, tumor markers, etc.). The disease-specific approaches were modified as relevant in an iterative fashion (Fig. 1).

Inter-Abstractor Agreement for and Availability of rwP Variable

The resulting disease-specific rwP abstraction approaches were applied to a larger cohort of randomly selected patients who had received treatment. Descriptive statistics were used to characterize baseline demographic and clinical characteristics for each cohort. We calculated median time to the patient’s first, second, and third imaging assessment, indexed to the initiation of first-line therapy unless otherwise specified, by identifying radiology document titles for individual unstructured notes. We also assessed median time to the patient’s first, second, and third clinician note to understand if patients were being followed as well as evaluated at anticipated time points consistent with expert clinical opinion of real-world practice patterns. For these median time-to-event analyses (e.g., time to first, second, third (1) imaging assessment or (2) clinician note), patients were also required to have started first-line systemic therapy for the tumor type at least 1 year prior to the end of the observation period, with the exception of aMel which was anchored to second-line systemic therapy. In addition, the imaging assessment/clinician note date was considered the event date and the censor date was the date of death (or end of the observation period for patients without a date of death).

In order to evaluate whether the capture of rwP aligned with clinical expectations, we calculated the proportion of patients by cancer type for the mBC, SCLC, mRCC, aGEC cohorts with at least one rwP event occurring at any point during the observation period and at least 14 days after the start of the patient’s index systemic therapy. We also assessed whether a rwP event was further corroborated by the occurrence of a clinically relevant downstream event in the curated dataset (i.e., new

Table 1 Definition of rwP and inclusion/exclusion criteria by disease

	mBC	aMel	SCLC	mRCC	aGEC
Base case (previously published rwP definition ^a)	Documentation in clinician’s EHR note of cancer progression based on the clinician’s interpretation of source evidence, specifically radiographic, pathologic, or clinical assessment only				
Disease-specific addition ^b	Inclusion of clinician’s interpretation of (1) physical exam; or (2) tumor markers only as source evidence	Inclusion of clinician’s interpretation of physical exam as source evidence	Instructions to abstractors included reminder that patients might be more frequently seen in the inpatient setting	No additions	No additions
Broad cohort inclusion criteria					
ICD codes	ICD-9 174.x or 175.x or ICD-10 C50x	ICD-9 172.x or ICD-10 C43x or D03x	ICD-9 162.x or ICD-10 C34x or C39.9	ICD-9 189.x or ICD-10 C64x or C65x	ICD-9 150.x or 151.x or ICD-10 C15.x or C16.x
Required clinical visits	≥ 2 on or after January 1, 2011	≥ 2 on or after January 1, 2011	≥ 2 on or after January 1, 2013	≥ 2 on or after January 1, 2011	≥ 2 on or after January 1, 2011
Diagnosis date as confirmed in EHR	Metastatic diagnosis date on or after January 1, 2011	Advanced diagnosis date on or after January 1, 2011	Diagnosis date on or after January 1, 2013	Metastatic diagnosis date on or after January 1, 2011	Advanced diagnosis date on or after January 1, 2011
Analytic cohort criteria					
Included	≥ 1 line of therapy Metastatic diagnosis date between January 1, 2011 and December 31, 2013	≥ 2 lines of therapy	≥ 1 line of therapy	≥ 1 line of therapy	≥ 1 line of therapy
Excluded					
Death within 14 days of:	1L start	2L start	1L start	1L start	1L start
OR					

Table 1 continued

	mBC	aMel	SCLC	mRCC	aGEC
Interval \geq 90 days prior to first structured activity ^c from:	Metastatic diagnosis date	Advanced diagnosis date	Diagnosis date	Metastatic diagnosis date	Advanced diagnosis date
Total patients, <i>n</i>	884	152	359	323	370
End date of observation period	August 31, 2017	October 31, 2017	March 31, 2018	July 31, 2018	December 31, 2018

IL first line, *aGEC* advanced gastric/esophageal cancer, *aMel* advanced melanoma, *EHR* electronic health record, *ICD* International Classification of Diseases, *mBC* metastatic breast cancer, *mRCC* metastatic renal cell carcinoma, *aNSCLC* advanced non-small cell lung cancer, *SCLC* small cell lung cancer

^aPreviously developed rwP abstraction approach in aNSCLC used as the “base case” (see “Methods” section)

^bRefers to the need for disease-specific additions to the “base-case” rwP abstraction approach (see “Methods” section)

^cStructured activity refers to the first visit date, administration date, or order date on or after the diagnosis date

antineoplastic systemic therapy start, antineoplastic systemic therapy end, or death) within 15 days prior to or up to 60 days after the rwP event. We hypothesized that patients who started second-line therapy are likely to have progressed on prior therapy; therefore, we evaluated the percentage of patients with at least one rwP event in a subgroup of patients with evidence of second-line therapy. In the case of aMel, this evaluation was conducted on the full analytic cohort since all patients were required to have evidence of second-line systemic therapy.

Duplicate abstraction by two independent abstractors was conducted on a random subsample of at least 100 patients (ranging from 20.8% to 66.4% of the analyzed cohorts) for each tumor type to assess inter-abstractor agreement. Final sample sizes reflected study-specific factors at the time of each study. In the subset of patients abstracted in duplicate, the reliability of the capture of the first rwP event observed was evaluated by calculating (1) event agreement (i.e., the proportion of instances where both abstractors agreed on whether a rwP event did or did not occur, regardless of date documented for the event) and (2) date agreement (i.e., among patients for whom both

abstractors agreed that at least one rwP event occurred, the proportion of instances where there was agreement on the date documented for the first event: exact date, within a \pm 15-day window, and within a \pm 30-day window).

Time-to-Event Analyses and Correlation with rwOS

Consistent with the base-case abstraction approach evaluation, the use of rwP as an outcome in time-to-event analyses (rwPFS and rwTTP) was compared to rwOS and to treatment-based real-world endpoints (rwTTNT and real-world time to treatment discontinuation [rwTTD]) (Supplementary Material Table 1) in order to pressure test performance of the rwP-based endpoints (i.e., do they behave as would be expected?). Dates of death for rwPFS and rwOS calculations were based on a previously described composite mortality variable [25]. Endpoint estimates were assessed using the Kaplan–Meier method and indexed to the start of first-line systemic therapy for mBC, SCLC, mRCC, and aGEC. In contrast, patients with aMel were required to have at least two lines of therapy to be included in the analysis, and

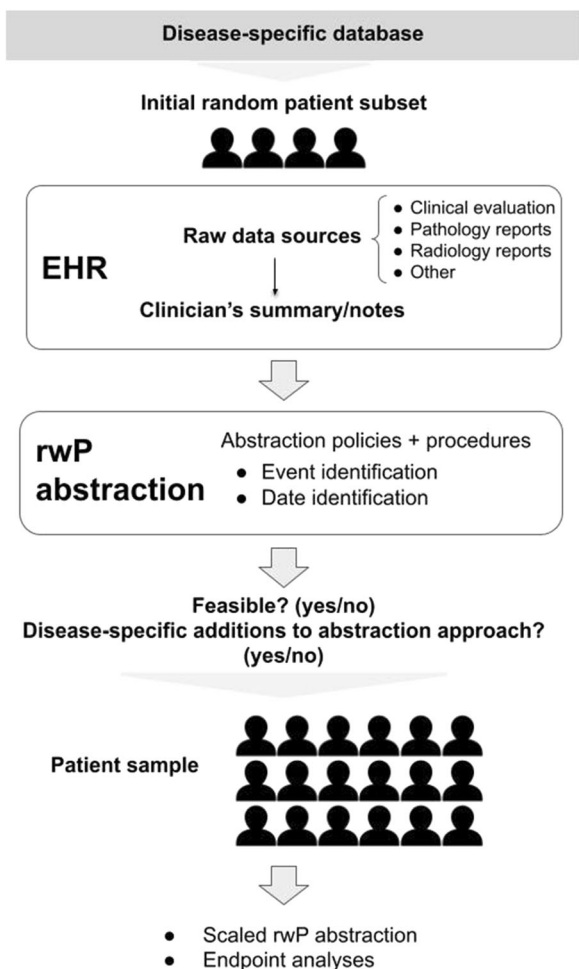


Fig. 1 Patient capture and rwP abstraction approach. EHR, electronic health record; rwP, real-world progression

analyses were indexed to the start of second-line systemic therapy. Kaplan–Meier curves and median time-to-event estimates were reported. The correlation between each real-world endpoint and rwOS was calculated using Spearman’s ρ and restricted to patients with a death and the real-world endpoint event of interest.

Assessment of Suitability for Implementation in Large Cohorts

For the final step of the evaluation, we determined the suitability of the rwP abstraction approach for each tumor type to be implemented in large cohorts (i.e., thousands or tens of thousands of patients) through a qualitative,

holistic assessment. Research scientists with clinical and quantitative expertise considered the variable’s feasibility, inter-abstractor agreement, availability, correlation results, and performance in time-to-event analyses along with correlation results. Researchers with the relevant expertise were assembled for each disease evaluated for a cross-functional scientific team approach. First, the source evidence for progression (e.g., imaging, clinical exam) had to be sufficiently documented in the EHR to allow abstraction of rwP events consistent with the disease-specific expectations. Second, the capture of rwP had to be sufficiently reliable based on the duplicate abstraction output. Third, the presence of clinically relevant downstream events had to corroborate that the rwP event aligned with related measures (antineoplastic systemic therapy stop, antineoplastic systemic therapy start, or death) as expected. Supplemental in-depth chart reviews were conducted on a random subset of patients that did not appear to have a downstream event in curated data for confirmation and hypothesis generation. Finally, the observed correlations and time-to-event estimates needed to align with clinical expectations as well as with the published literature as applicable given the difference in populations of the largely clinical trial-based literature.

RESULTS

Across the five tumor types evaluated, the cohort sizes ranged from 152 (aMel) to 884 (mBC) and the observation period end dates spanned from August 31, 2017 (mBC) to December 31, 2018 (aGEC; Table 1). Cohorts (see key patient characteristics, Supplementary Material Table 2a–e) predominantly comprised patients from community sites (range 80.9–100.0%) and proportions of patients with later-stage disease (stage III/IV) at initial diagnosis ranged from 46.7% to 71.6%.

Assessment of Disease-Specific Additions

Disease-specific additions to the base-case abstraction approach were necessary for

Table 2 Assessments of feasibility and performance by disease for a random sample of patients meeting inclusion/exclusion criteria

	mBC (Total N = 884)	aMcl ^a (Total N = 152)	SCIC (Total N = 359)	mRCC (Total N = 323)	aGEC (Total N = 370)
Evaluation cadence (months), median (95% CI)					
NCCN guideline recommendation	Baseline: cross-sectional imaging and bone scan	Baseline: cross-sectional and brain imaging	Baseline: cross-sectional and brain imaging	Baseline: cross-sectional imaging	Baseline: cross-sectional imaging
evaluation cadence for tumor type/stage corresponding to Flatiron Health dataset ^b	CT chest/abdomen/pelvis with contrast: baseline prior to new therapy and every 2–4 cycles (chemotherapy) or every 2–6 months (ET)	Treatment response assessment for active treatment other than complete surgical resection: clinical examination and/or imaging (cross-sectional ± brain)	Treatment response assessment during systemic therapy: CT chest/abdomen/pelvis with contrast after every 2–3 cycles of systemic therapy and at completion of therapy	Treatment response assessment: chest, abdominal, and pelvic CT or MRI imaging every 6–16 weeks as per physician discretion, patient clinical status, and therapeutic schedule. Imaging interval adjusted according to rate of disease change and sites of active disease	Treatment response assessment: chest, abdominal, and pelvic CT or MRI imaging every 6–16 weeks as per physician discretion, patient clinical status, and therapeutic schedule. Imaging interval adjusted according to rate of disease change and sites of active disease
	Bone scan: baseline prior to new therapy and every 4–6 cycles (chemotherapy) or every 4–6 months (ET)	Imaging during treatment (at clinically appropriate intervals) recommended for:	For patients with asymptomatic brain metastases receiving systemic therapy before brain RT; brain MRI (preferred) or CT with contrast every 2 cycles of systemic therapy and at completion of therapy	Consider MRI (preferred) or CT of head at baseline and as clinically indicated	
	PET/CT: As clinically indicated	Stage III (clinical satellite or in-transit primary or local, satellite, and/or in-transit recurrence)		Annual surveillance scans at physician discretion	
		Nodal recurrence in previously dissected nodal bed that is unresectable or incompletely resected		MRI of spine as clinically indicated	
		Limited (resectable) distant metastatic disease		Bone scan as clinically indicated	
		Disseminated (unresectable) distant metastatic disease			

Table 2 continued

	mBC(Total N = 884)	aMeI ^a (Total N = 152)	SCLC(Total N = 359)	mRCC(Total N = 323)	aGEC(Total N = 370)
1st image ^c	2.0 (1.8, 2.1)	1.4 (1.2, 1.8)	1.1 (1.0, 1.3)	1.6 (1.5, 1.9)	1.4 (1.3, 1.5)
2nd image ^c	4.7 (4.4, 5.0)	3.2 (2.9, 4.0)	3.0 (2.6, 3.3)	3.8 (3.4, 4.4)	3.3 (3.0, 3.6)
3rd image ^c	7.3 (6.9, 8.0)	5.2 (4.7, 6.8)	4.6 (4.1, 5.1)	6.2 (5.7, 7.0)	5.4 (5.1, 5.8)
1st note	0.2 (0.2, 0.3)	0.2 (0.2, 0.3)	0.2 (0.2, 0.2)	0.3 (0.2, 0.4)	0.1 (0.1, 0.2)
2nd note	1.0 (0.9, 1.1)	0.9 (0.9, 1.0)	0.8 (0.7, 0.9)	1.0 (0.9, 1.2)	0.6 (0.5, 0.7)
3rd note	1.8 (1.6, 1.8)	1.5 (1.4, 1.6)	1.3 (1.2, 1.4)	1.9 (1.7, 2.1)	1.2 (1.0, 1.2)
Occurrence of rwP, % (95% CI)					
≥ 1 rwP event	72 (69, 75)	–	58 (53, 64)	60 (54, 65)	55 (50, 60)
Patients with a 2L treatment start date	630	152	129	167	147
≥ 1 rwP event in patients who started 2L treatment	89 (87, 92)	94 (89, 97)	94 (88, 97)	90 (85, 94)	90 (85, 95)
Inter-abstractor agreement, % (95% CI)					
Patients abstracted in duplicate (i.e., reviewed by two independent abstractors)	n = 184	n = 101	n = 112	n = 101	n = 105
Event agreement ^d	92 (89, 96)	97 (94, 100)	92 (87, 97)	88 (82, 94)	95 (91, 99)
Date agreement, n ^e	127	93	60	63	54
Date, exact	65 (57, 74)	60 (50, 70)	68 (57, 80)	73 (62, 84)	65 (52, 78)
Date, ± 15 days	75 (67, 82)	80 (71, 88)	80 (70, 90)	81 (71, 91)	76 (65, 87)
Date, ± 30 days	77 (70, 84)	84 (76, 91)	85 (76, 94)	86 (77, 94)	78 (67, 89)
Patients with rwP event	n = 640	n = 89	n = 210	n = 193	n = 202

Table 2 continued

	mBC(Total N = 884)	aMel ^a (Total N = 152)	SCLC(Total N = 359)	mRCC(Total N = 323)	aGEC(Total N = 370)
Downstream events, % (95% CI)	72 (69, 75)	56 (46, 66)	59 (53, 65)	64 (57, 71)	68 (62, 74)

LL first line, 2L second line, aGEC advanced gastric/esophageal cancer, aMel^a advanced melanoma, CT computed tomography, ET endocrine therapy, mBC metastatic breast cancer, mRCC metastatic renal cell carcinoma, MRI magnetic resonance imaging, NA not applicable, PET, positron emission tomography, RT radiation therapy, rwP real-world progression, SCLC small cell lung cancer

^aAll patients with aMel were required to have at least two lines of therapy; as a result, the aMel analyses are indexed to the start of 2L

^bNCCN recommendations reported in this table are for the following stages for each tumor type: mBC, *metastatic disease^c; aMel, *stage IV or recurrence with distant metastatic disease^c; SCLC, *extensive stage^c; mRCC, *relapsed or stage IV and surgically unresectable disease^c; aGEC, NA

^cAn image is defined as EHR document titles that included mention of *radiology^d

^dEvent agreement was calculated on all patients abstracted in duplicate as the proportion of instances where both abstractors agreed on whether a rwP event did or did not occur, regardless of date documented for the event

^eDate agreement was calculated on a subset of patients abstracted in duplicate for whom both abstractors agreed that at least one rwP event occurred

evaluating rwP in mBC, aMel, and SCLC, but not in RCC or aGEC (Table 1).

Inter-Abstractor Agreement for and Availability of rwP Variable

The first imaging assessment was observed within a median of 2.0 months after the start of first-line systemic therapy (range of medians across tumor types 1.1–2.0 months). For the analyses indexed to first-line systemic therapy (i.e., excluding aMel), at least one rwP event following the patient's first-line systemic therapy, start date was identified in the majority of patients (range 55–72%). When the analyses were restricted to patients with evidence of having started second-line systemic therapy, the proportion with a rwP event identified at some point following the patient's first-line systemic therapy start increased (range 89–94%) (Table 2).

For all tumor types, inter-abstractor agreement as to whether the patient had at least one rwP event or not, irrespective of date, ranged from 88% to 97% (Table 2). When both abstractors identified at least one rwP event, date agreement was lower for the exact date (agreement range 60–73%) compared with expanded time windows of ± 15 days (agreement range 75–81%) and ± 30 days (agreement range 77–86%).

Time-to-Event Analyses and Correlations with rwOS

We observed variability by tumor type in the occurrence of clinically relevant downstream events corresponding with a rwP event in the curated database. For example, 72% of patients with mBC who had a rwP event had a new antineoplastic systemic therapy start, antineoplastic systemic therapy end, or death occur within the specified time window relative to the rwP event, while this was observed in only 59% of patients with SCLC who had a rwP event (Table 2).

Median rwPFS ranged from 3.7 (aMel) to 7.7 (mBC) months while median rwTTP ranged from 4.6 (aMel) to 8.3 (mRCC) months

(Fig. 2a–e). Correlations between rwOS and rwPFS ranged from a moderate association of 0.52 (aMel) to a strong association of 0.82 (SCLC) (Fig. 3a–e) [26]. In contrast, the range of correlations between rwTTD and rwOS (0.40–0.62) consisted, overall, of lower values.

Assessment of Suitability for Implementation in Large Cohorts

The final rwP abstraction approach for each of the five solid tumor types was determined to be suitable to implement across a large cohort of patients (i.e., at scale) on the basis of the researchers' qualitative, holistic assessment of the variable. Additional chart reviews provided supplemental information. For example, independent hand-review of randomly selected charts of patients who had an abstracted rwP event without a downstream event captured in the curated database identified four primary reasons: (1) subsequent treatment was not a systemic therapy captured in the curated dataset (e.g., radiation therapy for brain metastases); (2) clinical decision was made to continue on the same treatment; (3) patient choice to not change therapy; and (4) loss to follow-up (e.g., hospice referral).

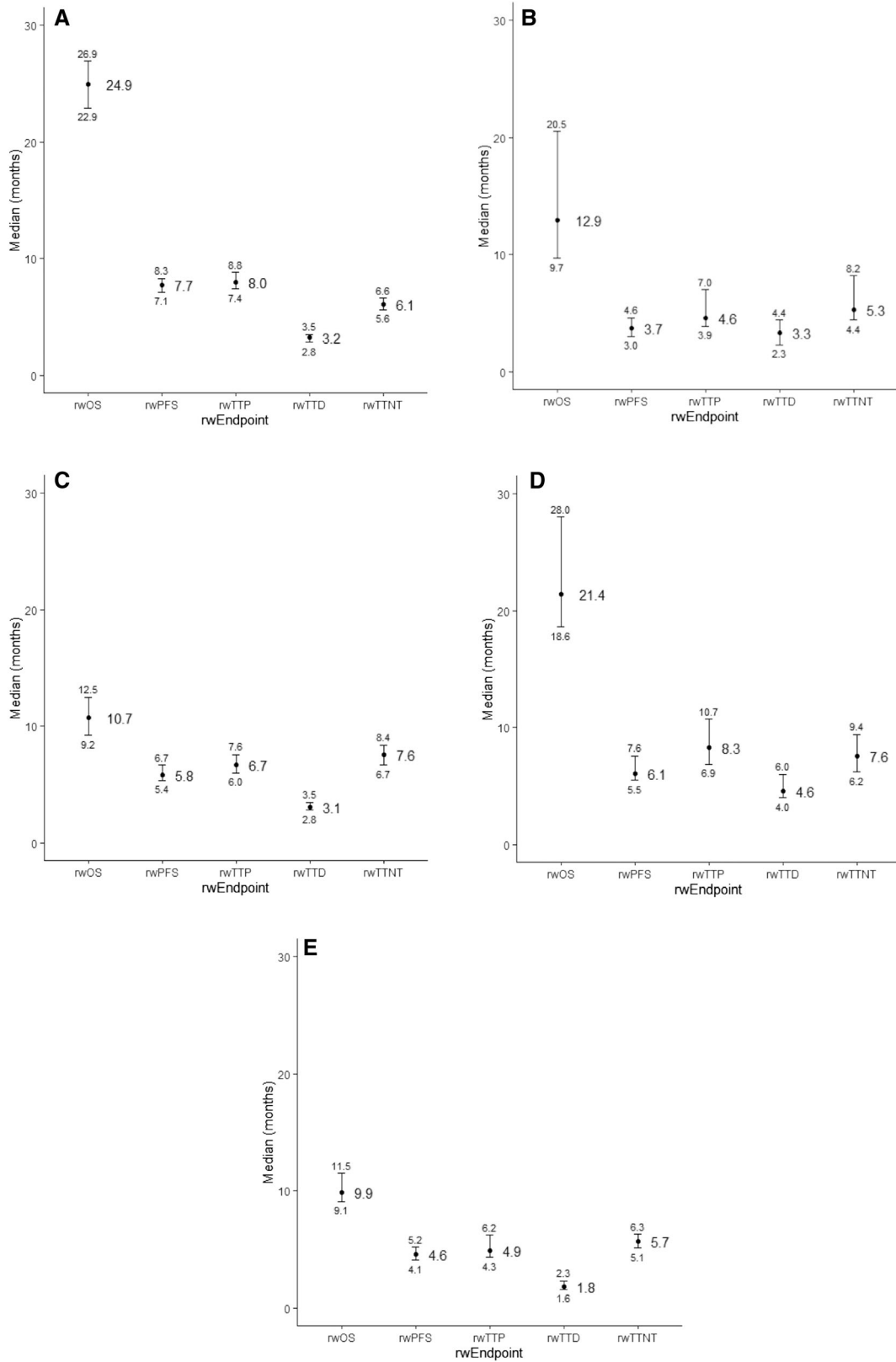
DISCUSSION

The current methodological study demonstrates that deriving clinically meaningful rwP events from EHR documentation is feasible and reliable across multiple tumor types, despite differing underlying cancer biology [22]. Consistent with prior evaluation of variable characteristics using the base-case approach previously in aNSCLC, results from the rwP-based time-to-event analyses (using the same approach with some disease-specific additions) were aligned with clinical expectations for each tumor type when considered in the context of rwOS. Although direct comparisons were not possible because of differences in patient populations, the association observed between rwOS and rwPFS by disease was directionally similar to the clinical trial literature for RECIST-based PFS [27–40].

Contextualization of rwP-based endpoint characteristics relative to treatment-based endpoints demonstrated that correlations with rwOS for rwTTD and rwPFS differed by tumor types as generally expected on the basis of the literature [27, 29–41]. For example, in SCLC the correlation between rwOS and rwTTD was lower compared to rwP and rwOS because at the time of the analysis, first-line therapy for SCLC consisted of a fixed number of therapy cycles [42]. In these types of situations, interpretation of the relationship between TTD and OS is most interpretable in the subset of patients who progress before the planned finish of the fixed treatment time course. In other settings, patients may switch therapies for reasons other than worsening of disease (e.g., treatment side effects); in contrast, patients may appear to continue on the same therapy despite evidence of disease progression, such as in cases of suspected immunotherapy-related pseudoprogression. These observations of nuanced clinical treatment decisions suggest potential challenges in using changes in treatment as direct indicators of the presence or absence of progression and, therefore, as an efficacy endpoint (e.g., time to next treatment).

Overall, rwPFS and rwTTP results were directionally similar to reported clinical trial results; the largest gap was observed in mBC, with shorter PFS for real-world patients. Potential hypotheses include selection bias (greater presence of aggressive disease subtypes [visceral disease or triple-negative] in the real-world cohort compared to clinical trials), and documentation of progression based on clinical assessment (e.g., worsening pain from bone metastasis) at an earlier time point than images, demonstrating changes in target lesion sizes consistent with RECIST-based progression are available [43–55].

As with all RWD, several considerations may apply to this work. First, disease-specific rwOS estimates observed in these real-world patient cohorts were generally shorter than in relevant clinical trials. This finding is unsurprising given that trial eligibility criteria tend to select for healthier patients, such as exclusion of patients with organ dysfunction [56]. Second, rwPFS estimates can be subject to heterogeneity in the



◀**Fig. 2** Forest plot of medians (95% CI) for rwOS, rwPFS, rwTTP, rwTTD, rwTTNT by disease. **A** mBC. **B** aMel^a. **C** SCLC. **D** mRCC. **E** aGEC. ^aIndexed to second-line start date. aGEC, advanced gastric/esophageal cancer; aMel, advanced melanoma; mBC, metastatic breast cancer; mRCC, metastatic renal cell carcinoma; rwOS, real-world overall survival; rwP, real-world progression; rwPFS, real-world progression-free survival; rwTTD, real-world time to discontinuation; rwTTP, real-world time to progression; rwTTNT, real-world time to next treatment; SCLC, small cell lung cancer

frequency and type of outcome assessments in routine practice as compared with protocol-dictated clinical trial assessments. Third, while inclusion and exclusion criteria for this study were minimal, those implemented may have introduced bias. For example, in order to identify progression in the context of treatment, the analytic cohort was limited to patients who lived at least 14 days after starting index treatment, which may have resulted in survival bias. Fourth, results in this study provide researchers critical context for knowing whether the data are sufficient for potential future use cases. For example, inter-abstractor date agreement for a rwP event was lower for exact date agreement than for the ± 15 -day and 30-day windows. Accordingly, researchers should consider what date agreement thresholds are necessary for answering a particular research question. The level of acceptable uncertainty (and differential between comparator groups) for the date of the event and the potential effects of censoring should be considered for time-to-event analysis. Fifth, as with all EHR-derived RWD, miscategorization (e.g., of advanced/metastatic diagnosis) is possible as a result of inaccurate documentation or abstractor error.

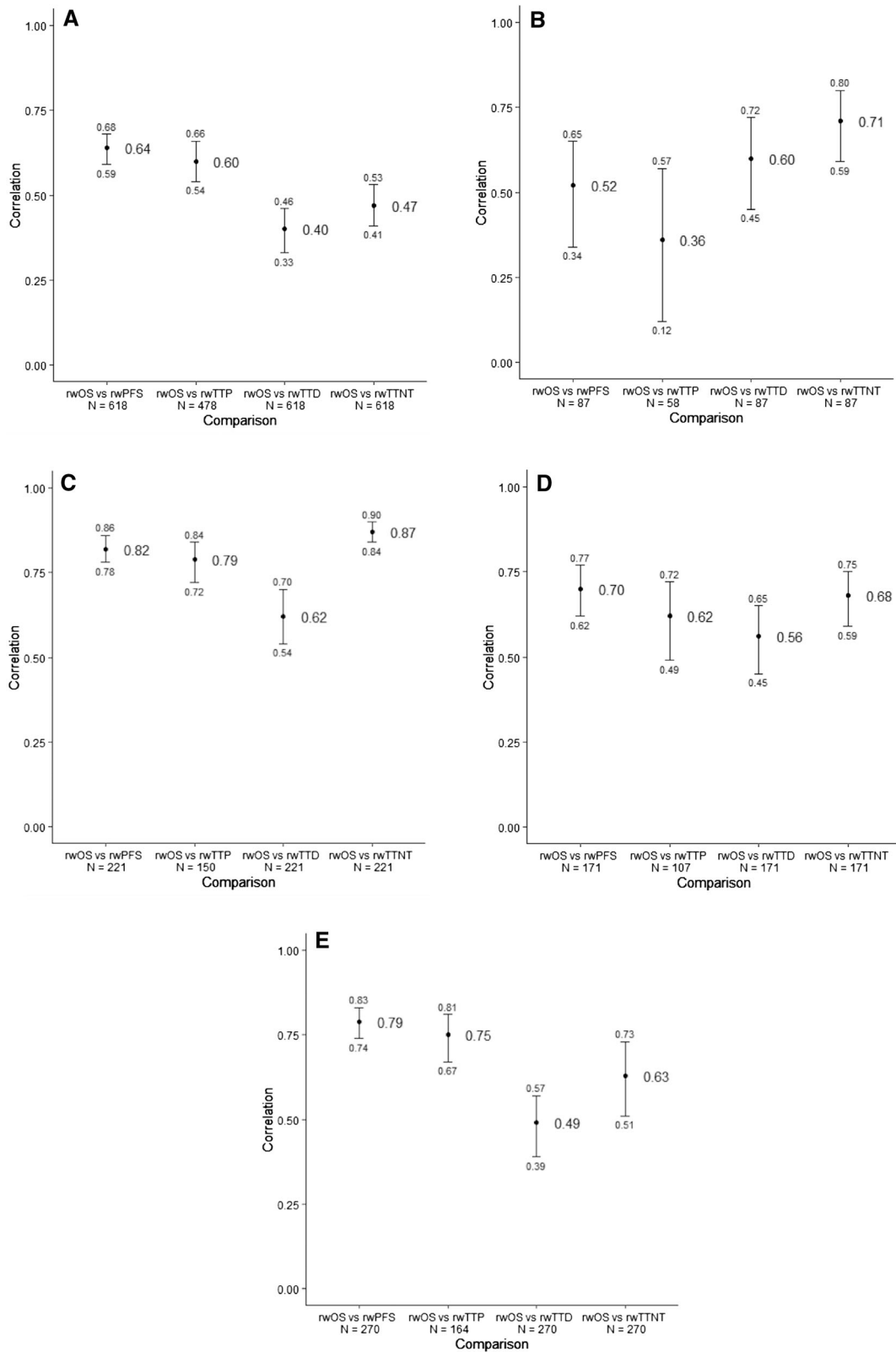
One of the strengths of our approach to ascertain rwP is the ability to revisit the source EHR evidence to identify gaps. This access to source data unlocks opportunities for quality check procedures and may facilitate enhancement of explanatory portions of abstraction policies. For example, abstractor disagreements emerging from differential interpretation of unstructured documents (or because of a lack of explicit documentation) triggered

improvements in quality assurance/control procedures in how documents were reviewed. This standard quality approach across disease databases results in clearer, more streamlined and robust workflow practices. In addition, chart review provided insights into how the concept of rwP fits into individual patient journeys, including reasons why a downstream event was not observed in curated data around the time of a rwP event.

Clinicians routinely assess and document progression of cancer on the basis of multiple factors, which may or may not include imaging results. As previously shown [18], limitations in EHR data (e.g., radiology report documentation patterns lack the level of tumor measurement detail required for RECIST) constrain the ability to consistently derive RECIST-defined progression from routine EHR written documentations, regardless of tumor type. Access to raw files of imaging captured during routine care would enable concordance analyses at the patient level between the application of RECIST to radiographic images and the derivation of rwP from EHR documentation. The comparability of endpoints captured in the real-world versus the trial setting is a key scientific question, particularly in the context of using RWD in external comparator cohorts.

This study focuses on solid tumors; derivation of progression events for liquid/hematologic malignancies may require distinct processes, given differences in how disease burden is ascertained (i.e., often by labs or pathology, rather than by radiographic imaging) [57–60]. Even within the realm of solid tumors, this study highlights the need for tumor-specific considerations regarding cited source evidence, which in turn can inform the methodology behind the approach for ascertaining rwP for specific cancers. Therefore, cohort characteristics and idiosyncrasies in disease-specific standards of care warrant particular attention when the intention is to produce generalizable results.

Future steps to support development of robust and reliable variables and endpoints for RWE include standardization of approaches across various RWD sources. Appropriate deployment of novel endpoints also requires



◀**Fig. 3** Forest plot of Spearman’s correlations (95% CI) with rwOS by disease. **A** mBC. **B** aMel^a. **C** SCLC. **D** mRCC. **E** aGEC. ^aIndexed to second-line start date. Note: Only patients who had both an event for death and the endpoint of interest were included in the correlation analysis. aGEC, advanced gastric/esophageal cancer; aMel, advanced melanoma; mBC, metastatic breast cancer; mRCC, metastatic renal cell carcinoma; rwOS, real-world overall survival; rwP, real-world progression; rwPFS, real-world progression-free survival; rwTTD, real-world time to discontinuation; rwTTP, real-world time to progression; rwTTNT, real-world time to next treatment; SCLC, small cell lung cancer

alignment in the field on the minimum set of metrics needed for characterization frameworks and on the criteria and standards for quality evaluation (e.g., completeness). Additionally, a system that enables continuous variable and endpoint performance reappraisal over time is needed to account for the evolution in documentation or data capture in real-world settings, changes in data sources, and changes in patterns of care. Finally, future work could define the dimensions for a checklist, which may include pre-specified thresholds, to determine “fitness for use” of a given variable (e.g., minimum date agreement at the cohort level) for specific research questions. Defining thresholds for real-world endpoint assessment requires further study because published literature may not always be relevant or available.

CONCLUSION

Derivation of a rwP variable from EHR documentation is feasible and reliable across the five solid tumors evaluated. Endpoint analyses show that rwP reflects clinically meaningful information, as demonstrated by the correlations of rwP with downstream events and other endpoints. On the basis of these findings, we determined that our rwP abstraction approach is suitable for implementation in large, real-world EHR datasets and may enable outcome analyses.

ACKNOWLEDGEMENTS

Funding. This study, including the journal’s Rapid Service and Open Access fees, was sponsored by Flatiron Health, Inc., which is an independent subsidiary of the Roche Group.

Medical Writing, Editorial and Other Assistance. The authors would like to thank Raina Mathur (Flatiron Health, Inc) and Kelly Magee (Flatiron Health, Inc.) for their contributions to this study, and Cody Patton (Flatiron Health, Inc.) for medical writing and publication management support.

Authorship. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

Author Contributions. Study design (literature, background search): Aracelis Z. Torres, Nathan C. Nussbaum, Christina M. Parrinello, Ariel B. Bourla, Bryan E. Bowser, Rebecca A. Miksad. Data collection: Flatiron Health, Inc. Data analysis (figures as applicable): Aracelis Z. Torres, Christina M. Parrinello. Data interpretation: Aracelis Z. Torres, Nathan C. Nussbaum, Christina M. Parrinello, Ariel B. Bourla, Bryan E. Bowser, Samuel Wagner, David C. Tabano, Daniel George, Rebecca A. Miksad. Writing and review of manuscript: Aracelis Z. Torres, Nathan C. Nussbaum, Christina M. Parrinello, Ariel B. Bourla, Bryan E. Bowser, Samuel Wagner, David C. Tabano, Daniel George, Rebecca A. Miksad.

Disclosures. The study sponsor, Flatiron Health Inc, is funding the Open Access Fees for publication of this manuscript. At the time the study was conducted, Aracelis Z. Torres, Nathan C. Nussbaum, Christina M. Parrinello, Ariel B. Bourla, Bryan E. Bowser, and Rebecca A. Miksad report employment at Flatiron Health, Inc., which is an independent subsidiary of the Roche Group. Aracelis Z. Torres, Nathan C. Nussbaum, Christina M. Parrinello, Ariel B. Bourla, Bryan E. Bowser, and Rebecca A. Miksad

report equity ownership in Flatiron Health, Inc. Aracelis Z. Torres, Nathan C. Nussbaum, Christina M. Parrinello, Ariel B. Bourla, Bryan E. Bowser, and Rebecca A. Miksad report stock ownership in Roche. Aracelis Z. Torres reports current employment at Verana Health. Nathan C. Nussbaum reports current employment at Verily Life Sciences. Samuel Wagner reports BMS employment and stock ownership in BMS. At the time of the study, David C. Tabano reports BMS employment; David C. Tabano's current affiliation is Genentech. Daniel George reports the following: American Association for Cancer Research – Sr Editor – Personal/Financial; Astellas – Consultant, Research, Advisory Board – Institutional and Personal/Financial; AstraZeneca – Research, Consultant, Advisory Board, CAPI-281 Steering Committee member – Institutional and Personal/Financial; Axess Oncology – Independent Contractor – Personal/Financial; Bayer H/C Pharmaceuticals – Consultant, Speaker, Honorarium, Travel accommodations, SC – Personal/Financial; BMS – Consultant, Research, Steering Committee – Institutional and Personal/Financial; Calithera – Research – Institutional; Capio Biosciences – Scientific Advisory Board – Personal/Financial; Constellation Pharmaceuticals – Consultant (09/2020) – Personal/Financial; EMD Serono – Honorarium – Personal/Financial; Exelixis, Inc – Research, Consultant, Speaker, Honorarium, Travel accommodations – Institutional and Personal/Financial; Flatiron – Consultant – Personal/Financial; Ipsen – Honorarium – Personal/Financial; Janssen Pharmaceuticals – Research, Consultant, Independent Data Monitoring Committee (IDMC) – Institutional and Personal/Financial; Merck Sharp & Dohme – Consultant – Personal/Financial; Michael J Hennessey Associates – Honorarium, Consultant – Personal/Financial; Millennium Medical Publishing, *Clinical Advances in Hematology & Oncology* – Co-Editor-in-Chief – Personal/Financial; Modra Pharmaceuticals B.V. – Advisory Board – Personal/Financial; Myovant Sciences, Inc – Consultant – Personal/Financial; NCI Genitourinary Steering Committee member (Leidos Biomedical Research Inc) – Personal/Financial; Nektar Therapeutics – Steering Committee – Personal/Financial; Novartis – Research – Institutional;

Physician Education Resource LLC – Consultant – Personal/Financial; Pfizer – Research, Consultant, Steering Committee, Honorarium – Institutional and Personal/Financial; Propella TX – Consultant (formerly Vizuri) – Personal/Financial; RevHealth, LLC – Consultant – 01/2021 – Personal/Financial; Sanofi – Research, Consultant, Speaker, Honorarium, Travel accommodations – Institutional and Personal/Financial; UroGPO – Honorarium – Personal/Financial; UroToday – Honorarium, Travel accommodations – Personal/Financial.

Compliance with Ethics Guidelines. This observational study was performed in accordance with the Helsinki Declaration of 1964 and its later amendments. Institutional review board (IRB) approval of the study protocol, with a waiver of informed consent, was obtained prior to study conduct, and covers the data from all sites represented. Approval was granted by the WCG IRB. (“The Flatiron Health Real-World Evidence Parent Protocol”; IRB registration number IRB00000533; Protocol approval ID tracking number 420180044.)

Data Availability. The data that support the findings of this study have been originated by Flatiron Health, Inc. These de-identified data may be made available upon request, and are subject to a license agreement with Flatiron Health; interested researchers should contact <DataAccess@flatiron.com> to determine licensing terms.

Open Access. This article is licensed under a Creative Commons Attribution-Non-Commercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not

permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

- 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):2293–7.
- Khozin S, Blumenthal GM, Pazdur R. Real-world data for clinical evidence generation in oncology. *J Natl Cancer Inst*. 2017;109:11. <https://doi.org/10.1093/jnci/djx187>.
- Rothwell PM. External validity of randomised controlled trials: to whom do the results of this trial apply? *Lancet*. 2005;365(9453):82–93.
- Frieden TR. Evidence for health decision making—beyond randomized, controlled trials. *N Engl J Med*. 2017;377(5):465–75.
- US Food & Drug Administration. Framework for FDA's real-world evidence program. 2018. <https://www.fda.gov/media/120060/download>. Accessed 23 Mar 2021.
- Franklin JM, Schneeweiss S. When and how can real world data analyses substitute for randomized controlled trials? *Clin Pharmacol Ther*. 2017;102(6):924–33.
- Cave A, Kurz X, Arlett P. Real-world data for regulatory decision making: challenges and possible solutions for Europe. *Clin Pharmacol Ther*. 2019;106(1):36–9.
- Klonoff DC. The expanding role of real-world evidence trials in health care decision making. *J Diabetes Sci Technol*. 2020;14(1):174–9.
- Ramagopalan SV, Simpson A, Sammon C. Can real-world data really replace randomised clinical trials? *BMC Med*. 2020;18(1):13–8.
- Miksad RA, Abernethy AP. Harnessing the power of real-world evidence (RWE): a checklist to ensure regulatory-grade data quality. *Clin Pharmacol Ther*. 2018;103(2):202–5.
- Bonamici S. H.R.34—21st century cures act. 114th Congress. 2016. <https://www.congress.gov/bill/114thcongress/house-bill/34>.
- Blumenthal GM, Kluetz PG, Schneider J, Goldberg KB, McKee AE, Pazdur R. Oncology drug approvals: evaluating endpoints and evidence in an era of breakthrough therapies. *Oncologist*. 2017;22(7):762–7.
- Pazdur R. Endpoints for assessing drug activity in clinical trials. *Oncologist*. 2008;13(Suppl 2):19–21.
- US Food & Drug Administration. Clinical trial endpoints for the approval of cancer drugs and biologics guidance for industry. 2018. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-trial-endpoints-approval-cancer-drugs-and-biologics>. Accessed 23 Mar 2021.
- Wilson MK, Karakasis K, Oza AM. Outcomes and endpoints in trials of cancer treatment: the past, present, and future. *Lancet Oncol*. 2015;16(1):32.
- Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45(2):228–47.
- Seymour L, Bogaerts J, Perrone A, et al. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. *Lancet Oncol*. 2017;18(3):e143–52.
- Griffith SD, Tucker M, Bowser B, et al. Generating real-world tumor burden endpoints from electronic health record data: comparison of RECIST, radiology-anchored, and clinician-anchored approaches for abstracting real-world progression in non-small cell lung cancer. *Adv Ther*. 2019;36(8):2122–36.
- Eichler HG, Koenig F, Arlett P, et al. Are novel, nonrandomized analytic methods fit for decision making? The need for prospective, controlled, and transparent validation. *Clin Pharmacol Ther*. 2020;107(4):773–9.
- Stewart M, Norden AD, Dreyer N, et al. 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*. 2019;3:1–15.
- Khozin S, Miksad RA, Adami J, et al. Real-world progression, treatment, and survival outcomes during rapid adoption of immunotherapy for advanced non-small cell lung cancer. *Cancer*. 2019;125(22):4019–32.
- Griffith SD, Miksad RA, Calkins G, et al. Characterizing the feasibility and performance of real-world tumor progression end points and their association with overall survival in a large advanced non-small-cell lung cancer data set. *JCO Clin Cancer Inform*. 2019;3:1–13.

23. Ma X, Long L, Moon S, Adamson BJS, Baxi SS. Comparison of population characteristics in real-world clinical oncology databases in the US: Flatiron health, SEER, and NPCR. 2020. <https://www.medrxiv.org/content/10.1101/2020.03.16.20037143v2>. Accessed 23 Mar 2021.
24. 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. 2020. <https://arxiv.org/abs/2001.09765>. Accessed 23 Mar 2021.
25. Curtis MD, Griffith SD, Tucker M, et al. Development and validation of a high-quality composite real-world mortality endpoint. *Health Serv Res.* 2018;53(6):4460–76.
26. Mindrila D, Balentyne P. Scatterplots and correlation. In: Moore DS, Notz WI, Flinger MA, editors. *The basic practice of statistics* (6th ed). New York: Freeman; 2013.
27. Paoletti X, Oba K, Bang YJ, et al. Progression-free survival as a surrogate for overall survival in advanced/recurrent gastric cancer trials: a meta-analysis. *J Natl Cancer Inst.* 2013;105(21):1667–70.
28. Shitara K, Ikeda J, Mizota A, et al. Progression-free survival and time to progression as surrogate markers of overall survival in patients with advanced gastric cancer: literature-based analysis of 36 randomized trials. *J Clin Oncol.* 2011;29(4):103.
29. Shitara K, Matsuo K, Muro K, Doi T, Ohtsu A. Correlation between overall survival and other endpoints in clinical trials of second-line chemotherapy for patients with advanced gastric cancer. *Gastric Cancer.* 2014;17(2):362–70.
30. Petrelli F, Barni S. Surrogate end points and post-progression survival in renal cell carcinoma: an analysis of first-line trials with targeted therapies. *Clin Genitourin Cancer.* 2013;11(4):385–9.
31. Négrier S, Bushmakina AG, Cappelleri JC, et al. Assessment of progression-free survival as a surrogate end-point for overall survival in patients with metastatic renal cell carcinoma. *Eur J Cancer.* 2014;50(10):1766–71.
32. Halabi S, Rini B, Escudier B, Stadler WM, Small EJ. Progression-free survival as a surrogate endpoint of overall survival in patients with metastatic renal cell carcinoma. *Cancer.* 2014;120(1):52–60.
33. Heng DY, Xie W, Bjarnason GA, et al. Progression-free survival as a predictor of overall survival in metastatic renal cell carcinoma treated with contemporary targeted therapy. *Cancer.* 2011;117(12):2637–42.
34. Foster NR, Qi Y, Shi Q, et al. Tumor response and progression-free survival as potential surrogate endpoints for overall survival in extensive stage small-cell lung cancer: findings on the basis of North Central Cancer Treatment Group trials. *Cancer.* 2011;117(6):1262–71.
35. Foster NR, Renfro LA, Schild SE, et al. Multitrial evaluation of progression-free survival as a surrogate end point for overall survival in first-line extensive-stage small-cell lung cancer. *J Thorac Oncol.* 2015;10(7):1099–106.
36. Imai H, Mori K, Wakuda K, et al. Progression-free survival, post-progression survival, and tumor response as surrogate markers for overall survival in patients with extensive small cell lung cancer. *Ann Thorac Med.* 2015;10(1):61–6.
37. Imai H, Mori K, Watase N, et al. Clinical significance of the relationship between progression-free survival or postprogression survival and overall survival in patients with extensive disease-small-cell lung cancer treated with carboplatin plus etoposide. *Can Respir J.* 2016;2016:5405810.
38. Flaherty KT, Hennig M, Lee SJ, et al. Surrogate endpoints for overall survival in metastatic melanoma: a meta-analysis of randomised controlled trials. *Lancet Oncol.* 2014;15(3):297–304.
39. Miksad RA, Zietemann V, Gothe R, et al. Progression-free survival as a surrogate endpoint in advanced breast cancer. *Int J Technol Assess Health Care.* 2008;24(4):371–83.
40. Burzykowski T, Buyse M, Piccart-Gebhart M, et al. Evaluation of tumor response, disease control, progression-free survival, and time to progression as potential surrogate end points in metastatic breast cancer. *J Clin Oncol.* 2008;26(12):1987–92.
41. Shitara K, Ikeda J, Yokota T, et al. Progression-free survival and time to progression as surrogate markers of overall survival in patients with advanced gastric cancer: analysis of 36 randomized trials. *Invest New Drugs.* 2012;30(3):1224–31.
42. Kalemkerian GP, Loo BW, Akerley W, et al. NCCN guidelines insights: small cell lung cancer, version 2.2018. *J Natl Compr Canc Netw.* 2018;16(10):1171–82.
43. Finn RS, Martin M, Rugo HS, et al. Palbociclib and letrozole in advanced breast cancer. *N Engl J Med.* 2016;375(20):1925–36.
44. Hortobagyi GN, Stemmer SM, Burris HA, et al. Ribociclib as first-line therapy for HR-positive, advanced breast cancer. *N Engl J Med.* 2016;375(18):1738–48.

45. Swain SM, Kim S, Cortés J, et al. Pertuzumab, trastuzumab, and docetaxel for HER2-positive metastatic breast cancer (CLEOPATRA study): overall survival results from a randomised, double-blind, placebo-controlled, phase 3 study. *Lancet Oncol.* 2013;14(6):461–71.
46. Robert C, Schachter J, Long GV, et al. Pembrolizumab versus ipilimumab in advanced melanoma. *N Engl J Med.* 2015;372(26):2521–32.
47. Horn L, Mansfield AS, Szczerba A, et al. First-line atezolizumab plus chemotherapy in extensive-stage small-cell lung cancer. *N Engl J Med.* 2018;379(23):2220–9.
48. Motzer RJ, Hutson TE, Cella D, et al. Pazopanib versus sunitinib in metastatic renal-cell carcinoma. *N Engl J Med.* 2013;369(8):722–31.
49. Motzer RJ, Hutson TE, Tomczak P, et al. Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. *N Engl J Med.* 2007;356(2):115–24.
50. Motzer RJ, Tannir NM, McDermott DF, et al. Nivolumab plus ipilimumab versus sunitinib in advanced renal-cell carcinoma. *N Engl J Med.* 2018;378(14):1277–90.
51. Ohtsu A, Shah MA, Van Cutsem E, et al. Bevacizumab in combination with chemotherapy as first-line therapy in advanced gastric cancer: a randomized, double-blind, placebo-controlled phase III study. *J Clin Oncol.* 2011;29(30):3968–76.
52. Fuchs CS, Shitara K, Di Bartolomeo M, et al. Ramucirumab with cisplatin and fluoropyrimidine as first-line therapy in patients with metastatic gastric or junctional adenocarcinoma (RAINFALL): a double-blind, randomised, placebo-controlled, phase 3 trial. *Lancet Oncol.* 2019;20(3):420–35.
53. Hecht JR, Bang Y, Qin SK, et al. Lapatinib in combination with capecitabine plus oxaliplatin in human epidermal growth factor receptor 2-positive advanced or metastatic gastric, esophageal, or gastroesophageal adenocarcinoma: TRIO-013/LOGiC—a randomized phase III trial. *J Clin Oncol.* 2016;34(5):443–51.
54. Im S, Lu Y, Bardia A, et al. Overall survival with ribociclib plus endocrine therapy in breast cancer. *N Engl J Med.* 2019;381(4):307–16.
55. Taberero J, Hoff PM, Shen L, et al. Pertuzumab plus trastuzumab and chemotherapy for HER2-positive metastatic gastric or gastro-oesophageal junction cancer (JACOB): final analysis of a double-blind, randomised, placebo-controlled phase 3 study. *Lancet Oncol.* 2018;19(10):1372–84.
56. Jin S, Pazdur R, Sridhara R. Re-evaluating eligibility criteria for oncology clinical trials: analysis of investigational new drug applications in 2015. *J Clin Oncol.* 2017;35(33):3745–52.
57. Cheson BD, Bennett JM, Kopecky KJ, et al. Revised recommendations of the international working group for diagnosis, standardization of response criteria, treatment outcomes, and reporting standards for therapeutic trials in acute myeloid leukemia. *J Clin Oncol.* 2003;21(24):4642–9.
58. Hallek M, Cheson BD, Catovsky D, et al. iwCLL guidelines for diagnosis, indications for treatment, response assessment, and supportive management of CLL. *Blood.* 2018;131(25):2745–60.
59. Younes A, Hilden P, Coiffier B, et al. International Working Group consensus response evaluation criteria in lymphoma (RECIL 2017). *Ann Oncol.* 2017;28(7):1436–47.
60. Kumar S, Paiva B, Anderson KC, et al. International Myeloma Working Group consensus criteria for response and minimal residual disease assessment in multiple myeloma. *Lancet Oncol.* 2016;17(8):e328–46.