

Considerations for Leveraging Real-World Endpoints in Oncology Drug Development

Use of real-world data (RWD) to generate real-world evidence (RWE) can support oncology drug development and regulatory decision-making. There is growing recognition that RWD, when analyzed appropriately, can generate RWE in broader patient populations than are able to be treated in clinical trials to inform medical product effectiveness, safety, and patient outcomes. Unlike traditional clinical trial settings where data are collected per protocol at pre-specified timepoints and reported uniformly for participants, there is significant heterogeneity in RWD within and across data sources. Inconsistent definitions and data missingness present challenges to using real-world (rw) endpoints for measuring treatment effectiveness. Strategies and methodologies for mitigating these challenges and alignment across stakeholders are needed to fully realize the potential of RWD. Friends of Cancer Research (*Friends*) initiated multiple research partnerships^{1,2,3,4} to develop and establish aligned methodologies for measuring rw-endpoints across RWD sources. Based on lessons learned from these research partnerships, a multi-stakeholder working group considered opportunities for using rw-endpoints and developed this resource to optimize use of rw-endpoints in oncology drug development (see table below).

There are multiple intended uses of RWD to support oncology development and may include generating RWE for signal detection to inform clinical development strategies, inform clinical trial design and patient access strategies, or directly be included as part of a regulatory submission. The intended use will impact the applicability of RWD and potential data quality considerations. For example, there should be justification for using RWD as part of a regulatory submission as well as evidence that the selected real-world dataset is fit-for-purpose. Further, caution should be taken when comparing rw-endpoints to clinical trial endpoints, given the inherent limitations of differing populations and measurements. Therefore, this work focuses on alignment across RWD sources, rather than comparison to clinical trial endpoints, through standardized methodologies for assessing rw-endpoints.

The table provides initial considerations for selecting rw-endpoints to measure treatment effectiveness. While rw-endpoints may be leveraged in many ways to support oncology drug development (e.g., rw-overall survival establishing natural history of a specific disease) that may be seen as more a benchmark, the definitions and minimum data elements listed are intended for comparative studies attributing an outcome to a specific treatment (e.g., causal inference). The definitions and data elements provided were jointly developed and implemented across collaborators participating in *Friends'* pilots evaluating rw-endpoints, which focused on patients with metastatic non-small cell lung cancer (mNSCLC) receiving systemic treatments (platinum doublet chemotherapy and/or immunotherapies). While the definitions and data elements listed herein are likely relevant to other solid tumor malignancies, additional data or validation may be needed to support use of these rw-endpoints in other tumor types and indications with disease specific requirements or endpoints. Furthermore, the strengths and limitations noted are informed by the mNSCLC rw-endpoint pilots conducted and may not be generalizable to other disease states.

1. [Establishing a Framework to Evaluate Real-World Endpoints, 2018 Friends of Cancer Research White Paper](#)
2. [The Friends of Cancer Research Real-World Data Collaboration Pilot 2.0: Methodological Recommendations from Oncology Case Studies, Rivera 2022, Clinical Pharmacology & Therapeutics](#)
3. [Real-world Overall Survival Using Oncology Electronic Health Record Data: Friends of Cancer Research Pilot, Lasiter 2022, Clinical Pharmacology & Therapeutics](#)
4. [rw-Response Endpoints in Patients with mNSCLC Treated with Chemotherapy Across rw-Datasets, 2023 ASCO Poster](#)

Working Group Collaborators

Thank you to our working group collaborators for informing the development of this table and considerations for using real-world endpoints in oncology drug development.

Amanda Bruno, Syneos Health, Formerly Bayer Pharmaceuticals

Gil Carrigan, Amgen Inc.

Victoria Chia, Amgen Inc.

Colleen Costello, Sanofi

Janet Espirito, Ontada

Laura Fernandes, COTA, Inc.

Elizabeth Garrett-Mayer, American Society Clinical Oncology

Shrividya Iyer, Eisai

Monika Izano, Syapse

Maria Karasarides, Bristol Myers Squibb Company

Sudeep Karve, Abbvie

Beata Korytowsky, Mirati

Mark Lanasa, BeiGene

Yanina Natanzon, ConcertAI

Irene Nunes, Flatiron Health

Vivek Pawar, EMD Serono

Danielle Potter, IQVIA

Lynn Sanders, Takeda Pharmaceutical Company

Regina Schwind, Tempus Labs, Inc.

Alessandria Strübing, Daiichi-Sankyo

We thank colleagues at the U.S. FDA for their input and collaboration. This document reflects discussions that occurred among stakeholder groups on various topics. This document should not be construed to represent FDA's views or policies.

rw-Endpoint and Definition	Minimum Data Elements Needed	Strengths & Limitations	Statistical Analysis Considerations
<p>rw-Overall Survival (rwOS) Length of time from the index treatment date to the date of death; for patients without a date of death, patients will be censored at the date of last structured recorded clinical activity, or end of follow-up period, whichever occurs earliest.</p>	<ul style="list-style-type: none"> • Date of index treatment initiation • Date of death or end of follow-up 	<p>Strengths</p> <ul style="list-style-type: none"> • Objectively defined. <p>Limitations</p> <ul style="list-style-type: none"> • Missingness of mortality information, which may not be random and could lead to biased estimates. Insufficient follow up time can also lead to a high proportion of censored patients which may overestimate survival. • Survival attributed to index treatment may be impacted by subsequent activities or therapies. These subsequent activities or therapies may be unavailable in EMR (e.g., start dates for oral medications may be difficult to obtain) due to incompleteness of data capture. • Real-world mortality information may not include cause of death to understand disease specific survival. 	<ul style="list-style-type: none"> • Capture median rwOS as well as landmark rwOS (e.g., 1 year and 5 year). • Additional data elements noting subsequent activity (subsequent therapies, etc.) or intercurrent events may be used to provide context to the rwOS endpoint. • Reduce immortal time bias (i.e., stratifying rwOS curves on factors that are determined after date of index treatment initiation).
<p>rw-Progression-Free Survival (rwPFS) Length of time from the index treatment date to the date of progression event or date of death. Patients without a progression event or date of death will be censored at the date of last structured recorded clinical activity reporting disease status, or end of follow-up period, whichever occurs earliest.</p>	<ul style="list-style-type: none"> • Date of index treatment initiation • Date of progression event through assessment of tumor response by clinician-based assessments • Date of death or end of follow-up • Date of index treatment discontinuation* • Date of next treatment initiation* <p>*Optional, to attribute progression event to index treatment</p>	<p>Strengths</p> <ul style="list-style-type: none"> • Less follow-up time is needed than rwOS, which may limit data missingness concerns. • Captures more direct effect of treatment activity on disease. <p>Limitations</p> <ul style="list-style-type: none"> • Subject to interval censoring bias, i.e., assessments may occur at different time intervals and using different methodologies or modalities. • Length of intervals may also be related to response to therapy. • Assessments may occur outside of available data source and lead to data missingness. • Capture of rwPFS based on clinician-based assessments is subjective (variable) and not based on RECIST criteria or have confirmation of progression. 	<ul style="list-style-type: none"> • Capture median rwPFS as well as landmark (e.g., 6 months and 1 year). • Account for interval censoring in data analysis. • Present breakdown of the type of PFS event: n(%) patients with progression event against n(%) patients with treatment discontinuation or n(%) patients with next line treatment start to understand the progression information captured. • Consider sensitivity analyses that assess rwPFS based on the type of progression data (e.g., tumor measurements from imaging, symptomatic progression). • Consider sensitivity analyses censoring patients who had no sign of progression but switched therapies.

rw-Endpoint and Definition	Minimum Data Elements Needed	Strengths & Limitations	Statistical Analysis Considerations
<p>rw-Response (rwR)</p> <ul style="list-style-type: none"> • Occurrence of a rwCR or rwPR after index treatment initiation during the study period among all patients. • This is often assessed as a rwR rate (rwRR), which is the proportion of patients with a rw-best overall response (rwBOR) of rwCR or rwPR. • $rwCR > rwPR > rwSD > rwPD$. 	<ul style="list-style-type: none"> • Assessment of tumor response by clinician-based assessments and date of assessment • Date of index treatment initiation • Date of index treatment discontinuation 	<p>Strengths</p> <ul style="list-style-type: none"> • Direct measurement of drug antitumor activity. • Less follow-up time is needed than rwOS and possibly rwPFS (depends on rwRR of drug). <p>Limitations</p> <ul style="list-style-type: none"> • For clinician-based assessments, subjective measure due to lack of standardized assessment framework in routine practice, including timing and frequency of assessment. • Absence of confirmatory scans on response in clinical practice. • Subject to observer or information bias. • Clinician-based assessments from one assessment to the next may use the last assessment as the new comparator, rather than the pre-treatment baseline. • Assessments may not be appropriately adjusted based on if the patient received any surgical resection or radiotherapy during index treatment. 	<ul style="list-style-type: none"> • Consider sensitivity analyses of patients with both images or image reports (to conduct a RECIST-like assessment) and clinician assessment to evaluate concordance of response. • Consider analysis of interval timing and frequency of clinician-based assessments to inform findings.
<p>rw-Duration of Response (rwDOR)</p> <p>The length of time from the date of the first documented assessment of rwCR or rwPR after the index date to the date of the first subsequent documented assessment of rwPD, rwMR or death, whichever comes first. For patients without rwPD, rwMR, or death, the patient will be censored at their last known response assessment of rwCR, rwPR, or rwSD, or the date of treatment discontinuation, whichever comes first.</p>	<ul style="list-style-type: none"> • Date of index treatment initiation • Date of first assessment of rwCR or rwPR • Date of first subsequent assessment of rwPD, rwMR, or death • Date of last assessment of rwCR or rwPR • Date of index treatment discontinuation • Date of next treatment initiation (Optional, if missing index treatment discontinuation) 	<p>Strengths</p> <ul style="list-style-type: none"> • Provides understanding of response durability. <p>Limitations</p> <ul style="list-style-type: none"> • Subject to interval censoring bias, i.e., assessments may occur at different time intervals and using different methodologies or modalities. • For clinician-based assessments, subjective measure due to lack of confirmatory scans and varying methodologies. • Requires various data points that may not be captured adequately for assessment. 	<ul style="list-style-type: none"> • Capture durable 6-month rwRR: The proportion of patients with at least one assessment of rwCR or rwPR who have not had an assessment of rwPD or rwMR or discontinuation of therapy within 6 months after the first documented assessment of rwCR or rwPR. • Capture median rwDOR and landmark (e.g., 3, 6, 9 months).

rw-Endpoint and Definition	Minimum Data Elements Needed	Strengths & Limitations	Statistical Analysis Considerations
<p>rw-Time to Treatment Discontinuation (rwTTD) Length of time from the index treatment date to the date the patient discontinues treatment. The treatment discontinuation date is defined as the last administration or non-cancelled order of the therapy. Discontinuation is defined as having a subsequent systemic therapy after the index therapy, having a gap of more than 120 days with no systemic therapy following the last administration prior to treatment completion, or having a date of death while on the therapy. Patients without a discontinuation will be censored at their last known usage of study treatment.</p>	<ul style="list-style-type: none"> • Date of index treatment initiation • Date of index treatment discontinuation • Date of death or end of follow-up 	<p style="text-align: center;">Oral Drugs</p> <p>Strengths</p> <ul style="list-style-type: none"> • Integrates total amount of time a patient is treated on therapy, regardless of reason for discontinuation, whether due to effectiveness or tolerability. • Not subject to bias associated with variable tumor burden assessments (e.g., time intervals, assessment methodologies). • May be associated proxy for PFS. <p>Limitations</p> <ul style="list-style-type: none"> • Treatment cycles may vary, both on a treatment and a per patient basis, impacting the ability to define discontinuation. • Date of treatment discontinuation is not always (and often not) available, and patient adherence may be unknown. Therefore, might need to rely on assumptions based on start dates or other algorithms. • Discontinuation may be due to tolerability or causes other than treatment ineffectiveness. 	<ul style="list-style-type: none"> • In combination therapies, discontinuation should be considered when both therapies are discontinued (one therapy in the combination can be discontinued and still be considered on index treatment), however, timing of discontinuation of the one therapy in the combination should be noted. • Consider sensitivity analyses on the 120-day period, as this criterion may differ based on clinical opinion in the disease of interest or data source.
	<ul style="list-style-type: none"> • Date of index treatment initiation • Date of last administration • Date of death or end of follow-up 	<p style="text-align: center;">Infused Drugs</p> <p>Strengths</p> <ul style="list-style-type: none"> • Integrates total amount of time a patient is treated on therapy, regardless of reason for discontinuation, whether due to effectiveness or tolerability. • Often based on structured data and easier to implement in an EMR system. • Not subject to bias associated with variable tumor burden assessments (e.g., time intervals, assessment methodologies). <p>Limitations</p> <ul style="list-style-type: none"> • Discontinuation may be due to tolerability or causes other than treatment ineffectiveness. 	<ul style="list-style-type: none"> • In combination therapies, discontinuation should be considered when both therapies are discontinued (one therapy in the combination can be discontinued and still be considered on index treatment), however, timing of discontinuation of the one therapy in the combination should be noted. • Consider sensitivity analyses on the 120-day period, as this criterion may differ based on clinical opinion in the disease of interest or data source.

rw-Endpoint and Definition	Minimum Data Elements Needed	Strengths & Limitations	Statistical Analysis Considerations
<p>rw-Time to Next Treatment (rwTTNT) Length of time from the index treatment date to the date the patient received their next systemic therapy (next-line therapy) or date of death. If patient has not received subsequent treatment or has not died, patients will be censored at their last known activity or end of follow-up.</p>	<ul style="list-style-type: none"> • Date of index treatment initiation • Date of index treatment discontinuation • Date of next line treatment • Date of death or end of follow-up 	<p>Strengths</p> <ul style="list-style-type: none"> • A proxy for disease control or potential benefit (e.g., duration of effect). • Integrates total amount of time a patient is treated on therapy, regardless of reason for discontinuation, whether due to effectiveness or tolerability. • Captures possible response durability of initial treatment. • Advantage over rwTTD given the data availability of timing of next therapy initiation compared to discontinuation. • Not subject to bias associated with variable tumor burden assessments (e.g., time intervals, assessment methodologies). <p>Limitations</p> <ul style="list-style-type: none"> • Endpoint as defined is specific to next line systemic therapy and is not inclusive of other interventions such as surgery or radiation which could result in bias. • Censoring is likely not independent of prognosis (violation of censoring assumption). • Missingness of data if patients receive treatment outside of the system. 	<ul style="list-style-type: none"> • Consider analyses that account for intercurrent events, if data are available.

EMR, electronic medical record; rwCR, real-world complete response; rwPR, real-world partial response; rwPD, real-world progressive disease; rwMR, real-world mixed response; rwSD, real-world stable disease.