Paper PP02

The Journey of Deriving the Best Overall Response (BOR) for Solid Tumors – A Programmer's Perspective

Milan Adesara, Merck & Co., Inc., Rahway, NJ, USA[1]
Vishal Gandhi, Merck & Co., Inc., Rahway, NJ, USA[1]

ABSTRACT

The Best Overall Response (BOR) is a critical endpoint in clinical trials for solid tumors, serving as an important measure of treatment efficacy. However, the determination of BOR can be complex while using different response evaluation criteria, such as RECIST 1.1, mRECIST 1.1, and iRECIST 1.1. Each criteria has its own methodology for assessing tumor response, which can lead to variation in how BOR is derived. Given its significance in trial outcomes, it is essential to understand the variation of BOR derivation using different criteria.

Moreover, an often-overlooked factor in BOR derivation is the patient's trial completion status. This status can provide additional information especially when the response confirmation is required in subsequent assessment. For example, "Confirmed BOR – Awaiting for the next scan" may apply to the patients who have not yet completed the trial, whereas "Confirmed BOR – Not confirmed in the next scan" may apply to the patients who have discontinued the trial.

This presentation will thoroughly explore various response criteria used in solid tumor, the differences in BOR derivation across these criteria, and the patient's trial completion status in the BOR determination. By addressing this complexity, we aim to enhance the understanding of how BOR is derived.

INTRODUCTION

Objective evaluation of tumor response has long been central to oncology drug development. In clinical trials, reliable assessment of treatment effect is critical not only for understanding drug activity but also for informing regulatory decisions and guiding subsequent clinical management. Radiographic endpoints, which evaluate changes in tumor burden over time, serve as pivotal indicators of therapeutic activity, particularly in early-phase clinical trials where survival-based endpoints—such as overall survival (OS) or progression-free survival (PFS)—may be immature or influenced by subsequent therapies.

Among response-based endpoints, *Best Overall Response (BOR)* has emerged as one of the most widely reported and influential measures of antitumor efficacy. BOR captures the most favorable tumor response observed for an individual patient from the initiation of therapy until disease progression, death, or treatment discontinuation. As such, it provides a patient-level summary of treatment benefit and forms the basis for calculating *Objective Response Rate (ORR)*, a key endpoint that frequently supports regulatory approvals, including accelerated approval pathways for therapies addressing unmet medical needs.

Despite its apparent simplicity, BOR is a *derived endpoint* that depends on multiple methodological decisions. These include the choice of response evaluation criteria, the timing and frequency of assessments, confirmation requirements, and the handling of missing or incomplete data. The growing diversity of anticancer therapies—including targeted therapies, biologics, and immunotherapies—has prompted the development of specialized response evaluation frameworks. While these frameworks aim to standardize tumor response assessment, methodological differences can introduce variability in BOR determination and interpretation.

Additionally, *patient-level trial dynamics*—such as early treatment discontinuation, loss to follow-up, or ongoing participation at the time of interim or final analysis—add further complexity. This is particularly relevant in trials requiring confirmation of response or progression. A nuanced understanding of how response criteria interact with trial completion status is therefore essential to ensure accurate interpretation of BOR and to maintain consistency across clinical trials and regulatory submissions.

[1] The practices described in this paper are informed by the author's prior professional experience at a former employer.

CONCEPTUAL FRAMEWORK OF BEST OVERALL RESPONSE

BOR as a Derived Endpoint

Unlike endpoints such as overall survival, which are directly observed, BOR is derived through the application of predefined rules to longitudinal tumor assessment data. Each post-baseline assessment contributes to the determination of BOR, and the final classification reflects the most favorable response observed over time rather than the patient's status at a single timepoint.

This retrospective derivation introduces inherent complexities:

- An *unconfirmed response* at an interim analysis may later become confirmed, altering BOR.
- Conversely, an initially observed response can be invalidated by *subsequent progression* before confirmation.

Thus, BOR is inherently dynamic and sensitive to follow-up duration, timing of assessments, and trial-specific confirmation rules.

Relationship Between BOR, ORR, and DoR

BOR serves as the foundation for several critical downstream endpoints:

- Objective Response Rate (ORR): the proportion of patients achieving Complete Response (CR) or Partial Response (PR) based on BOR.
- Duration of Response (DoR): measures the time from first documented response until disease progression or death.

Any inaccuracy or inconsistency in BOR determination can propagate to these downstream metrics, potentially impacting conclusions about therapeutic efficacy and trial success.

Strengths and Limitations of BOR

Strengths:

- Clinically intuitive and easily interpretable by physicians and regulatory bodies.
- Simpler to calculate compared to time-to-event endpoints like OS or PFS.

Limitations:

- Does not reflect durability of response or timing of response onset.
- Limited ability to capture clinical benefit in patients with stable disease.

Therefore, BOR should always be interpreted alongside complementary endpoints such as DoR, OS, and safety metrics to provide a comprehensive view of treatment effect.

RESPONSE EVALUATION CRITERIA IN SOLID TUMORS

RECIST 1.1

RECIST 1.1 is the most widely used framework for assessing tumor response in solid tumors:

- Focuses on unidimensional measurements of up to 5 target lesions.
- Response classification thresholds:
    - Partial Response (PR): ≥30% decrease in sum of lesion diameters.
    - Progressive Disease (PD): ≥20% increase in sum of lesion diameters or appearance of new lesions.

Limitations: Primarily size-based, potentially missing biologic effects, necrosis, or immune-mediated responses, may fail to detect biological or immunotherapy-related effects.

Modified RECIST (mRECIST) 1.1

Developed for tumor types like *hepatocellular carcinoma,* mRECIST emphasizes *viable tumor tissue* rather than total lesion size:

- Viable tumor is often defined by *arterial enhancement* on imaging.
- More sensitive to *therapies inducing necrosis without immediate shrinkage*, leading to potential differences in BOR classification compared to RECIST 1.1.

This can lead to differences in BOR classification compared with RECIST 1.1, highlighting the importance of selecting criteria aligned with tumor biology and treatment mechanism.

Immune RECIST (iRECIST) 1.1

Designed to capture atypical response patterns in immunotherapy, such as pseudo progression:

- Introduces immune-specific response categories.
- Requires confirmation of progression before assigning definitive PD.
- Allows treatment continuation beyond initial progression in clinically stable patients.

iRECIST increases BOR derivation complexity but provides a more accurate reflection of immunotherapy outcomes.

COMPARISON OF BOR DEFINITIONS ACROSS RESPONSE CRITERIA

Definition of Best Overall Response (BOR) Categories Across Response Criteria:

BOR Category	RECIST 1.1	mRECIST 1.1	iRECIST 1.1
Complete Response (CR)	Disappearance of all lesions	Disappearance of viable tumor	iCR, confirmation required
Partial Response (PR)	≥30% decrease in lesion size	≥30% decrease in viable tumor	iPR, confirmation required
Stable Disease (SD)	Neither PR nor PD	Neither PR nor PD	iSD
Progressive Disease (PD)	≥20% increase or new lesions	Increase in viable tumor	iUPD or iCPD
Not Evaluable (NE)	Insufficient data	Insufficient data	Insufficient data

METHODOLOGICAL DIFFERENCES AFFECTING BOR DERIVATION

Differences in tumor measurement, response thresholds, confirmation rules, and handling of new lesions across response criteria contribute to variability in BOR derivation.

Key Methodological Differences Affecting BOR Derivation:

Aspect	RECIST 1.1	mRECIST 1.1	iRECIST 1.1
Tumor measurement	Size-based	Viable tumor	Size-based
Confirmation requirement	Trial-specific	Usually required	Mandatory
New lesion handling	Immediate PD	PD if viable	iUPD pending confirmation
Sensitivity to pseudo progression	Low	Low	High

VARIABILITY IN BOR ACROSS RESPONSE CRITERIA AND CLINICAL IMPLICATIONS

Variability in BOR across response criteria has important clinical and regulatory implications. A patient classified as a responder under one criterion may be classified as non-responder under another, affecting ORR estimates and perceived treatment efficacy.

Such discrepancies complicate cross-trial comparisons and meta-analyses and underscore the importance of aligning response criteria with therapeutic mechanism of action. Sensitivity analyses using alternative criteria may help contextualize results and enhance interpretability.

Example of BOR Variability Across Response Criteria:

Timepoint	RECIST 1.1	mRECIST 1.1	iRECIST 1.1
Week 8	SD	PR	iUPD
Week 16	PR	PR	iPR
Final BOR	PR	PR	iPR

ROLE OF TRIAL COMPLETION STATUS IN BOR DETERMINATION

Definition and Classification

Trial completion status reflects whether a patient completes the study per protocol, discontinues early, or remains ongoing at the time of analysis. Each scenario influences the completeness of tumor assessment data available for BOR determination.

Impact on Response Confirmation

Response confirmation requirements mean BOR determination may be incomplete for *ongoing or discontinued patients:*

- Ongoing patients may eventually achieve confirmation.
- Discontinued patients may never receive confirmation, impacting the final BOR.

Interaction With Immune-Specific Criteria

In iRECIST, trial completion status plays a particularly critical role, as confirmation of progression is required. Patients who discontinue after an initial iUPD may never receive definitive classification.

Statistical, Analytical, and Regulatory Implications

- Interim ORR estimates can change substantially as follow-up matures.
- Regulatory agencies emphasize *transparent reporting* of trial completion status and its influence on BOR.

Influence of Trial Completion Status on BOR Determination:

Trial Status	Initial Response	Confirmation	BOR
Completed	PR	Yes	Confirmed PR
Ongoing	PR	Pending	Unconfirmed PR
Early discontinuation	PR	No	PR not confirmed
Lost to follow-up	PR	No	NE

OPERATIONAL AND ANALYTICAL CHALLENGES IN BOR DERIVATION

Common challenges include:

- Missing data due to early discontinuation.
- Inter-reader variability in imaging interpretation.
- Deviations from assessment windows.
- Interim analyses that may yield provisional BORs.

Common BOR Scenarios Requiring Special Handling:

Scenario	Impact on BOR
Early discontinuation	Missing confirmation
Interim analysis	Provisional BOR
Pseudo progression	Delayed PD classification
Missing baseline imaging	NE classification

IMPLICATIONS FOR CLINICAL TRIAL DESIGN AND REPORTING

To ensure *accurate and consistent BOR derivation*:

- Pre-specify *response criteria and confirmation rules* in protocols and analysis plans.
- Align response criteria with *therapeutic mechanism of action.*
- Clearly define handling of *trial completion status.*
- Report *sensitivity analyses* using alternative criteria.

Recommended BOR Reporting Elements:

Element	Description
BOR definition	Criteria and derivation rules
Confirmation status	Confirmed vs unconfirmed
Trial completion status	Completed, ongoing, discontinued
Sensitivity analyses	Alternative criteria

FUTURE DIRECTIONS

- Incorporation of functional imaging, circulating tumor biomarkers, and real-world evidence can refine response assessment.
- Harmonization of BOR derivation across trials will improve comparability, regulatory confidence, and clinical decision-making.
- Future frameworks may integrate tumor biology and immune dynamics, enabling more precise and patient-centered efficacy evaluations.

CONCLUSION

Best Overall Response is a foundational endpoint in oncology trials, reflecting the most favorable tumor response observed over time. Its derivation is influenced by response criteria selection and patient trial completion status, which can introduce variability in interpretation.

- RECIST 1.1, mRECIST 1.1, and iRECIST 1.1 differ methodologically, resulting in potential variability in BOR classification.
- Accurate interpretation requires careful consideration of methodology, confirmation rules, and operational factors.
- Transparent reporting ensures robust clinical trial conclusions and supports regulatory decision-making.

REFERENCES

1. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). European Journal of Cancer. 2009;45(2):228–247.
2. Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. Seminars in Liver Disease. 2010;30(1):52–60.
3. The evaluation of immune therapy activity in solid tumors: Immune-related response criteria. Clinical Cancer Research. 2009;15(23):7412–7420.
4. OpenAI. (2026). ChatGPT-5 (version 5.2) [Large language model]. OpenAI. https://chatgpt.com/

ACKNOWLEDGEMENT

The authors would like to thank Amy Gillespie, Changhong Shi, Erica Davis, Sandeep Meesala, Jeff Xia and Abhilash Chimbirithy for reviewing the paper and providing valuable suggestions.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Author Name: Milan Adesara
Company: Merck & Co., Inc., Rahway, NJ, USA[1]
Email: milan.adesara@merck.com

Co-Author Name: Vishal Gandhi
Company: Merck & Co., Inc., Rahway, NJ, USA[1]
Email: vishal.gandhi@merck.com

The views and opinions presented in the paper are our own and do not necessarily reflect the official position of the company.

[1] The practices described in this paper are informed by the author's prior professional experience at a former employer.