Paper TA03

Confirmation and Cancellation - Best Strategies for Data Collection and Handling within Solid Tumor Studies using RECIST 1.1 and iRECIST Guidelines

Abraham Yeh, AbbVie Inc
Sherry Ye, AbbVie Inc

ABSTRACT

As immuno-oncology compounds are getting popular in clinical studies, RECIST 1.1 and iRECIST are the primary guidelines to handle tumor response assessments. From data management, what are the best ways of collecting the assessment without making mistakes? From a statistical perspective, what are the easiest ways to process the confirmations if the protocol is asking for confirmed assessments for the Best Overall Response? What can be done if there are better assessments after the iUPD in iRECIST which need to be cancelled out while deriving iPFS?

This paper presents three main categories from the end-to-end approach on (1) Is it possible to start these processes earlier while collecting data? (2) What is the best approach to handle the confirmation? And finally (3) What is the best approach to handle the cancellation?

We will think through the potential of the collection and metadata-driven data handling along the way.

INTRODUCTION

RECIST (Response Evaluation Criteria in Solid Tumors) provides a simple and pragmatic methodology to evaluate the activity and efficacy of new cancer therapeutics in solid tumors, using validated and consistent criteria to assess changes in tumor burden. The revised RECIST guidelines (version 1.1) are available here for free with permission from the European Journal of Cancer (EJC). The guidelines and accompanying articles were published in a special issue of EJC in January 2009.

A consensus guideline—iRECIST—was developed by the RECIST working group for the use of modified Response Evaluation Criteria in Solid Tumors (RECIST version 1.1) in cancer immunotherapy trials, to ensure consistent design and data collection, facilitate the ongoing collection of trial data, and ultimate validation of the guideline. This guideline describes a standard approach to solid tumor measurements and definitions for objective change in tumor size for use in trials in which immunotherapy is used. Additionally, it defines the minimum datapoints required from future trials and those currently in development to facilitate the compilation of a data warehouse to use to later validate iRECIST.

As the statistical team (usually statisticians and the statistical programmers) worked as the downstream with the data management as their upstream. There are ongoing discussions with the clinical / medical experts within the study on how to best set up and implement these complicated scenarios. We learned the key aspect on these evaluations, especially while there might be some confusions or uncertainties existed, it is best to involve the clinical / medical team in advance so there will be no worries while processing these from the study to the submission. There are certain elements that could be set up well in advance so the downstream activities could be hassle-free. However, in the meantime, certain elements, especially the cancellations, may not be best to utilized too early, as that could hurt the traceability of the clinical data hence that could cause unnecessary difficulties while the agencies review the submission package.

This paper will review and evaluate, as well as discuss what is the best approach for handling two of the most critical operations that will occur when the teams are dealing with the confirmation of the responses, especially the non-PD category except NE (i.e., CR, PR, and SD). We will first discuss how to get the Best Overall Response (BOR) for studies where confirmation is not required, which is a simple case. Then we will use the table to illustrate and explain why the derivation is such, along with some suggestions on the handling and when is best to handle such in the workflow of the lifecycle. The other one area of concern is the cancellation of the responses while the better responses (again, that could be CR, PR, and SD) occurred after the initial immunotherapy drugs kicked in. In the rule that is the time the patient is getting the iUPD (in iRECIST) then the prior iUPD (or iCPD) should be removed according to the guideline while the team is deriving the iPFS. Therefore, the cancellations are trickier as there were more than two responses that could be removed once the conditions as set were met. We will discuss the SAS® code used to operate these cancellations. It could be applied to other software like R and Python once similar logic is utilized.

THE FIRST STEP: TEAM COLLABORATIONS AND EARLY QUERIES

The most important and critical thing to keep in mind, especially for the statistical team is 'No one is a lonely island'. As this could become complex if we just look at those responses in the RS data. There might be cases that did not make sense statistically. The first lifeline we usually should call is to ask our data management colleagues to double check, as soon as such discrepancies are observed or in doubt as there might be entry errors or misinterpretations along the lifecycle in which the data was sent to the statistical team. Even after queries, some conditions might still exist, that is the reason we need to prepare in advance and have the medical and clinical experts to confirm using their expertise. Since those confirmations and cancellations are not straight mathematics, we need to handle them with care.

There are explanations on the 'soft' and 'hard' SD and its implications when the condition comes with the explanation in the paper.

DEFINITIONS FOR CONFIRMATIONS

There are two definitions that need to be cleared first before we proceed with the derivation of BOR, especially the confirmed ones.

Minimum SD duration: The minimum SD duration requirement should be counted from randomization date (or first dose date of study drug for non-randomized trials) up to the last evaluable timepoint with non-PD response category (i.e., CR, PR, or SD). For example, minimum SD duration requirement specified in the protocol is 6 weeks, Timepoint 1 has CR at Week 4, and Timepoint 2 has PD at Week 8. BOR for this subject is PD instead of SD since SD duration is 4 weeks (the last evaluable timepoint with non-PD response category is at Week 4) which is < 6 weeks. Please note unconfirmed CR and unconfirmed PR are considered stable disease so can be used to calculate SD duration when defining BOR.

There are two notes worth mentioning about the minimum SD duration:

1. The minimum SD duration is the key to implementing the confirmation logic. It started with the first dose date (or randomization date) to the last evaluable non-PD assessment.

2. This duration is dependent on the response assessment schedule. Based upon our experience, it should be at least 6-8 weeks from study entry depends on the study protocol for response assessments. For phase 1 and shorter response assessments (saying only 6 weeks), we saw 4 weeks in practice. Again, this is critical and must be set in the protocol before all the operations start.

Response confirmation: The minimum time that was needed for the confirmed response (CR and PR). The first timepoint response needs to be confirmed by the second timepoint. A CR/PR at the first timepoint needs to be confirmed by a PR/CR at least 4 weeks later, here we check if the second timepoint is at least 4 weeks after the first timepoint. A confirmed response shorter than 4 weeks cannot confirm the good response at the first timepoint. Like the minimum SD duration, it is a good practice to mention such in the protocol, or at least in the SAP to document such confirmation criterion well in advance.

Please see the table below as illustrated.

Table 1: Explanation for minimum SD duration and response confirmation.

<table>
<tr><td></td><td>Baseline Timepoint 0(Randomization/First Dose Date)</td><td>Overall Response First Timepoint</td><td>Overall Response Subsequent Timepoint</td><td>Best Overall Response</td></tr>
<tr><td></td><td>Week 0</td><td>CR at week 4</td><td>PD at week 8</td><td rowspan="4">PD (3)</td></tr>
<tr><td rowspan="2">The SD duration (1)</td><td colspan="2">4 weeks - To the 1st Timepoint</td><td></td></tr>
<tr><td colspan="3">8 weeks - To the 2nd Timepoint</td></tr>
<tr><td>Duration of the two timepoints post baseline (2)</td><td></td><td colspan="2">4 weeks - Between the 1st & 2nd Timepoints</td></tr>
</table>

(1) The minimum SD duration requirement is 6 weeks as specified in the protocol, counting from the Baseline Timepoint.
(2) To confirm the response (CR or PR) by a subsequent timepoint, the subsequent timepoint must have a response category that is the same as or better than the current timepoint and the two timepoints must be at least 4 weeks apart as specified in the protocol.
(3) This example only confirms the Overall Response at the First Timepoint.

The response confirmation is different than the minimum SD duration assessment in the following ways:

1. The starting point is different. The minimum SD duration is at the first dose (or randomization date) while the response confirmation starts at the first assessment that measured as CR and / or PR and the confirmation is the next one after such response found.

2. The duration is different. While the minimum SD duration is usually 6-8 weeks, the response confirmation is 4 weeks based on RECIST 1.1 suggestions.

3. The measurements required are different. As the name suggested, minimum SD duration considered non-PD responses. And response confirmation only considered CR and PR (note, there is no SD here).

To summarize the above two important concepts. It is a little bit what the medical experts call them "hand waving", because SD is a "lower" level of assessment of activity, a little longer duration seemed reasonable to qualify. Whereas CR and PR are harder to achieve so it is reasonable to have a slightly shorter duration to allow for it to count. It is a pre-set measure as described in the beginning when the clinical trial is starting, and the team agreed on. Once the consensus is made, the following downstream activities and functions do have a reference point to consider and it is a good practice to have these so that all the derivations and considerations would make medical sense.

There are two types of BOR that can be derived directly using RECIST 1.1

DERIVE BOR IN TRIALS WHILE CONFIRMATION OF CR AND PR IS NOT REQUIRED

The first type is the responses without confirmation. Best response in these trials is defined as the best response across all post-baseline timepoints based on the following hierarchical order: CR -> PR -> SD -> PD -> NE. Per RECIST 1.11: "When SD is believed to be best response, it must also meet the protocol specified minimum time from baseline. If the minimum time is not met when SD is otherwise the best timepoint response, the patient's best response depends on the subsequent assessments. For example, a patient who has SD at first assessment, PD at second and does not meet minimum duration for SD, will have a best response of PD. The same patient lost to follow-up after the first SD assessment would be considered unevaluable."

DERIVE BOR IN TRIALS WHILE CONFIRMATION OF CR AND PR IS REQUIRED

The second type is the responses with confirmation.

Complete or partial responses may be claimed only if the criteria for each are met at a subsequent timepoint as specified in the protocol (generally ≥ 4 weeks).

In general, to confirm the response (CR or PR) by a subsequent timepoint, the subsequent timepoint must have a response category that is the same as or better than the current timepoint and the two timepoints must be at least x weeks apart as specified in the protocol. There cannot be any timepoint between those two timepoints with a worse response category except the following scenarios:

1. Response assessment is non-evaluable between those two timepoints:

Per RECIST 11.: "In trials where confirmation of response is required, repeated 'NE' time point assessments may complicate best response determination. In most trials it is reasonable to consider a patient with time point responses of PR-NE-PR as a confirmed response." Cases where there is more than one NE between those two timepoints need to be discussed with the study team to determine BOR.

2. Tumor shrinkage hovers around 30% mark (for PR):

For example, a subject had 32% (PR), 28% (SD), and 33% (PR) decreases in SOD (Sum Of Diameters) from baseline at Timepoints 1, 2, and 3, respectively. Per RECIST 1.11: "It is not infrequent that tumor shrinkage hovers around the 30% mark. In this case, most would consider PR to have been confirmed looking at this overall case. Had there been two or three non- PR observations between the two timepoint PR responses, the most conservative approach would be to consider this case SD."

The following table 2 described scenarios for the first overall response along with the subsequent overall response. We can use this as the basis for the discussions on how we derive the BOR accordingly for each case.

Table 2. Best overall Response when confirmation of CR and PR is required. per RECIST 1.1.

Scenario	*Overall Response First*	*Overall Response Subsequent Timepoint*	*BEST overall response*
1	*CR*	*CR*	*CR*
2	*CR*	*PR*	*SD, PD, or PR*[a]
3	*CR*	*SD*	*SD provided minimum SD duration are met. Otherwise, PD*
4	*CR*	*PD*	*SD provided minimum SD duration are met. Otherwise, PD*
5	*CR*	*NE*	*SD provided minimum SD duration are met. Otherwise, NE*
6	*PR*	*CR*	*PR*
7	*PR*	*PR*	*PR*
8	*PR*	*SD*	*SD*
9	*PR*	*PD*	*SD provided minimum SD duration are met. Otherwise, PD*
10	*PR*	*NE*	*SD provided minimum SD duration are met. Otherwise, NE*
11	*NE*	*NE*	*NE*

a. If a CR is truly met at first time point, then any disease seen at a subsequent time point, even disease meeting PR criteria relative to baseline, makes the disease PD at that point (since disease must have reappeared after CR). The best response would depend on whether minimum duration for SD was met. However, sometimes 'CR' may be claimed when subsequent scans suggest small lesions were likely still present and in fact the patient had PR, not CR at the first time point. Under these circumstances, the original CR should be changed to PR and the best response is PR.

b. For 1, 6, and 7, it is assuming the subsequent timepoint is > = 4 weeks

CLARIFICATIONS FOR TABLE 2:

Scenario 2 (CR followed by a PR):

It is recommended to query the sites to re-evaluate whether CR criteria was truly met at the first timepoint.

(1) When CR was truly met at the first timepoint: When the first timepoint was confirmed to have met the CR criteria, it means the subject did not have any disease at the first timepoint. When the second timepoint met the PR criteria (i.e., failed to maintain CR criteria), it means the disease had reappeared at the second timepoint. Therefore, the overall response for the second timepoint should be PD and BOR would depend on whether the minimum duration for SD was met or not (i.e., BOR would be SD if the minimum duration for SD was met and PD if the minimum duration for SD was not met). In the situation where the site was queried and confirmed the first timepoint was CR but did not change the overall response from PR to PD for the second timepoint, BOR for this subject would be SD or PD depending on whether the minimum duration for SD was met or not (see above).

(2) When CR was not truly met at the first timepoint: When the scan at the subsequent timepoint(s) suggested the subject had PR instead of CR at the first timepoint, the site should change overall response at the first timepoint from CR to PR. The best overall response for this subject should be PR since PR was confirmed by a second scan.

Table 3. Example for scenario 2

SUBJID[a]	ADT[a]	AVALC[a]	TRTSDT[a]
001	2023-03-01	CR	2023-01-01
001	2023-05-01	PR	2023-01-01
002	2023-03-10	CR	2023-03-01
002	2023-03-15	PR	2023-03-01
002	2023-03-30	PD	2023-03-01

[a] SUBJID=subject ID, ADT=Date of tumor assessment, AVALC=Overall response for each tumor assessment, TRTSDT=First dose date

Table 3.1 BOR for scenario 2 if the CR at the first timepoint not changed after re-evaluation.

SUBJID[a]	ADT[a]	AVALC[a]	TRTSDT[a]	BOR[a]
001	2023-03-01	CR	2023-01-01	SD
001	2023-05-01	~~PR~~PD[b]	2023-01-01	
002	2023-03-10	CR	2023-03-01	PD
002	2023-03-15	~~PR~~PD[b]	2023-03-01	
002	2023-03-30	PD	2023-03-01	

[a] SUBJID=subject ID, ADT=Date of tumor assessment, AVALC=Overall response for each tumor assessment, TRTSDT=First dose date, BOR=Confirmed Best Overall response

[b] The overall response for the second timepoint should be PD if the CR at the first timepoint not changed after re-evaluation

Table 3.2 BOR for scenario 2 if the CR at the first timepoint changed to PR after re-evaluation.

SUBJID[a]	ADT[a]	AVALC[a]	TRTSDT[a]	BOR[a]
001	2023-03-01	~~CR~~PR	2023-01-01	PR
001	2023-05-01	PR	2023-01-01	
002	2023-03-10	~~CR~~PR	2023-03-01	PD
002	2023-03-15	PR	2023-03-01	
002	2023-03-30	PD	2023-03-01	

[a] SUBJID=subject ID, ADT=Date of tumor assessment, AVALC=Overall response for each tumor assessment, TRTSDT=First dose date, BOR=Confirmed Best Overall response

A note from the medical expert: In solid tumor it happens. If a lymph node goes under 10mm, it is CR (especially if that is the only target lesion). Next scan, lymph node is up and meets PR criteria. The medical team had seen it happened before on their experiences. It's a little 'soft CR' as an easy way to interpret this.

The scenario is wildly different if there are multiple metastases >20+mm and they all disappear (CR), that is a "hard" CR (as compared with the soft one described above). Lymph nodes are impossible to interpret.

Scenario 3 (CR followed with SD):
Like Scenario 2, sites will be queried to re-evaluate whether CR was truly met at the first timepoint.

(1) If CR was truly met at the first timepoint and the second timepoint met SD criteria, it means the subject failed to maintain CR criteria and the disease had reappeared at the second timepoint. Therefore, the overall response for the second timepoint should be PD and BOR would depend on whether the minimum duration for SD was met or not (i.e., BOR would be SD if the minimum SD duration was met and PD if that duration was not met – please see table 4.1).

(2) If CR was not truly met at the first timepoint, overall response for the first timepoint should be updated accordingly. BOR should be determined based on the updated overall response at the first timepoint.

Table 4. Example for scenario 3

SUBJID[a]	ADT[a]	AVALC[a]	TRTSDT[a]
001	2023-03-01	CR	2023-01-01
001	2023-05-01	SD	2023-01-01
002	2023-03-10	CR	2023-03-01
002	2023-03-15	SD	2023-03-01
002	2023-03-30	PD	2023-03-01

[a] SUBJID=subject ID, ADT=Date of tumor assessment, AVALC=Overall response for each tumor assessment, TRTSDT=First dose date

Table 4.1 BOR for scenario 3 if the CR at the first timepoint not changed after re-evaluation.

SUBJID[a]	ADT[a]	AVALC[a]	TRTSDT[a]	BOR[a]
001	2023-03-01	CR	2023-01-01	SD
001	2023-05-01	~~SD~~PD[b]	2023-01-01	
002	2023-03-10	CR	2023-03-01	PD
002	2023-03-15	~~SD~~PD[b]	2023-03-01	
002	2023-03-30	PD	2023-03-01	

[a] SUBJID=subject ID, ADT=Date of tumor assessment, AVALC=Overall response for each tumor assessment, TRTSDT=First dose date, BOR=Confirmed Best Overall response

[b] The overall response for the second timepoint should be PD if the CR at the first timepoint not changed after re-evaluation

Table 4.2 BOR for scenario 3 if the CR at the first timepoint changed to SD after re-evaluation.

SUBJID[a]	ADT[a]	AVALC[a]	TRTSDT[a]	BOR[a]
001	2023-03-01	~~CR~~SD	2023-01-01	SD
001	2023-05-01	SD	2023-01-01	
002	2023-03-10	~~CR~~SD	2023-03-01	PD
002	2023-03-15	SD	2023-03-01	
002	2023-03-30	PD	2023-03-01	

[a] SUBJID=subject ID, ADT=Date of tumor assessment, AVALC=Overall response for each tumor assessment, TRTSDT=First dose date, BOR=Confirmed Best Overall response

BEST STRATEGY FOR THE BOR DERIVATIONS FOR THE CONFIRMATIONS

Let us think about what the best strategy is to collect these confirmations. At the first place, we would imagine if we can get these confirmations as early as possible – that would save a lot of time and resources on the statistical team, is not the case?

As we think further, as there are four elements involved for the confirmation, which are:

(1) Is the minimum duration of the SD met?

(2) Once non–PD response was observed, we need to look at the next response- one to check if that goes equal or better to confirm, second is to check the elapsed time between the non-PD exceed the response confirmation duration as defined above.

(3) If unfortunately, the subsequent response is worse, then the queries as well as the minimum duration of SD, perhaps the previous assessment needs to be re-considered.

(4) The implementations of the 'NE' need to be considered.

To summarize, there are time related derivations which might be best to handle in the ADaM level according to CDISC guides. For the comparisons on those overall response assessments, it would be ideal to have a flag to

denote if that is better or worse, depending on the preset order per the RECIST 1.1 guideline. This will make the team easier to review, especially if we have a non-PD response and then getting worsened (not to mention that will enhance the traceability accordingly). As these were complicated for the SDTM to handle and it was a risk these records might be lost during the transformation and process, we recommend performing these using the ADaM, which provides better and transparent ways to get those properly and correctly, at the first time. During the conference, we can discuss the pros and cons of what is the best workflow during the lifecycle to make this optimal.

CANCELLATIONS OF BETTER ASSESSMENTS AFTER IUPD IN IRECIST

In addition to the BOR confirmation derivations as described using RECIST 1.1, the iRECIST guides kicked in once the study contained immunotherapy oncologic drugs in addition to traditional ones. In most protocols, when such design was utilized and both RECIST 1.1 and iRECIST are utilized, the team will have PD kicked in after the first dose intake and then immunotherapy will be added. Such PD will be equivalent to iUPD in iRECIST.

The way of handling BOR in iRECIST will not be discussed here as it was like what we discussed above plus some necessary considerations.

However, the cancellations of these better assessments (iCR, iPR, iSD) after the iUPD would need to be considered here as our focus. The iRECIST usually kicked in by the protocols definition and the first PD would be ideally equivalent to iUPD which need a subsequent confirmed iCPD to confirm. This is used to calculate the censoring and duration of the iPFS. The iRECIST guideline suggested to cancel out the iUPD if the subsequent is a better assessment (iCR, iPR, iSD).

(1) The progression event date would be the same as RECIST 1.1 date (i.e., first iUPD date) unless iSD, iPR or iCR intervenes which is ultimately date at which progression criteria are met provided that iCPD is confirmed at the next assessment.

(2) However, iUPD would be disregarded if iUPD only occurs before later iSD, iPR, or iCR, that iUPD date should not be used as the event date.

SAMPLE CODES FOR SUCH CANCELLATIONS

```
    /* Use combined data to find PD, iUPD and iCPD */
    data iucpd(keep=subjid rscatn adt rsstresc rsstresn);
        set iprog;
        if rsstresn in (5 15 16); /* PD (5) for RECIST, iUPD (15) and iCPD (16) for
iRECIST */
    RUN;

    proc sort data=iucpd(keep=subjid adt rename=(adt=ipdcutdt)) out=ipd nodupkey;
    by subjid;
    run;

    data ipdcut;
        merge iprog(in=a) ipd(in=i);
        by subjid;
        if i and adt>=ipdcutdt;
        keep subjid rsstresn rsstresc visitnum avisitn rsdt rscat ipdcutdt ;
    RUN;

    /* For subjects have better assessment than PD, delete these records for iPFS */
    data btr_pd;
        set ipdcut;
        if rsstresn in (1 2 3 11 12 13); /* CR:1, PR:2, SD:3, iCR:11, iPR: 12, iSD: 13
*/
    RUN;

    proc sort data=btr_pd(keep=subjid rsdt rename=(rsdt=btrlstdt)) out=btr_pd1;
    by subjid;
    run;
```

```
data btr_pd2;
    set btr_pd1;
    by subjid;
    if last.subjid;
RUN;

/* Filter the data into better and other */
data btr_pd3 other;
    merge iprog(in=i0) btr_pd2(in=b2);
    by subjid;
    if i0 and b2 then output btr_pd3;
    if i0 and not b2 then output other;
RUN;

/* Delete PD, iUPD and iCPD if better before the last better assessment */
data btr_pd4;
    set btr_pd3;
if rsdt<=btrlstdt and rsstresn in (5 15 16) then delete;
run;

/* Adding back the processed above with other */
data rir_ecist;
    set btr_pd4 other;
drop btrlstdt;
RUN;
```

STRATEGY FOR DATA HANDLING ON SUCH CANCELLATIONS

While the derivation of the BOR remained the same when the immunotherapy drugs added when the normal drug got the patient to progressive disease. The algorithm for taking the iPFS will be different as all the better assessments post these iUPD and iCPD will need to cancel as per the algorithm rule.

While we need to recall the compliance rules within CDISC, it is not recommended to delete these records through the data collection. Since the deletion of records is not a good practice in the earlier life cycles on data management or even while implementing the SDTM process. It seemed like the ADaM is the best place to operate such special operations. As suggested earlier while we discussed the confirmation part, if we have a record to compare with the previous one, then it will better help the team to operate on these records (and make necessary flags for these records so later when the team is calculating the iPFS, those records will be excluded per the rule given). Then the statistical team will have the flags for the records after the first PD, any better response assessment after that and then it will be easy to mark those record to post process. As the code shown, these records will be cancelled while the team is deriving the proper censoring and calculating the iPFS duration as per the statistical analysis plan.

CONCLUSION

While the studies of oncological compounds are very hot in the past several decades, the overall survival has not improved a lot. From recent research, it was only improving in less than three months in terms of overall survival. As the cancer cells are so different by biological nature, there are various guidelines like LUGANO and RECIST, just to name two, that will help to evaluate the efficacy of the compound.

This paper focuses on the two most confusing and difficult operations that are often required in oncology clinical studies. From the RECIST 1.1, along with the medical and statistical experiences, the confirmation is quite complicated. If you only looked at the guideline itself, as a member of the statistical team, it might not make too much sense to you as these were not black or white (or you can call it solid science) in absolute terms, since various responses might occur, and we do need a good understanding here. Complicate scenarios led to challenges for the statisticians, with proper and thorough understanding, we could then formulate statistical plans along with the corresponding codes to handle them with care. The other one is cancellation when we have both traditional drugs along with the immune-oncology drugs (e.g., PD-1, PD-L1, …etc.), once better response assessments were observed after the first occurrence of PD, then the PD itself plus these better ones need to be cancelled out when the team is trying to derive the analysis for iPFS. This surely caused some complications while we understood the data as well as we need to exclude these records when it came to the iPFS analysis.

Let us summarize in the following discussions,

1. Understand what you need to understand, also reach out if you do not understand. As a representative from different function areas, we all have expertise within our own areas. As members of the statistical team, it is tough to realize there are some strange realities when it comes to these confirmations and cancellations. Several keys to success are suggested here: Things might get complicated, so early planning and agreement are critical and important. No questions are stupid questions as things might not go the way that you wish. Under such scenarios, besides stretching on our own limitation from the functional expertise, the exchange with different backgrounds is very helpful and we can learn and exchange from each other.

2. Where is the best workflow to handle such complicated operations in the lifecycle of our data? From our own experiences, we felt the downstream activities to handle such complicated confirmation and cancellations make more sense as it will both flag as needed and when we need more derivations based on a certain subset (e.g., we do not want this better assessment after the first PD (or equivalently iUPD) along with that PD, we can use records outside of that specific selections). It is open to the teams that ask if it is possible to apply these in the upper stream (e.g., SDTM) without hurting the traceability of the data itself.

3. Are we doing these in a way that is already too complicated? Are there any enhancements that can be suggested to add in the data flow to make the whole process more streamlined? As the confirmations and cancellations seemed not that straight forward, is there any enhancement we can add / modify in the current workflow on the TU, TR, and RS – from raw data, SDTM to the ADaM that we can enhance the traceability? In the meantime, from collecting to interpreting, along with the analyzing effort can be simpler.

Please note that we are using AbbVie standard on the interpretation to the RECIST guidelines from the solid tumor teams itself, any suggestions and discussions for different interpretations are open to discussion and welcomed.

During the conference, we hope to discuss on these topics more in-depth to see if there are better ways to streamline the whole process with the proper team members involved. Ultimately, we would like to help these patients in the disease, and hopefully with more novel therapies available with a better data science flow, we can help out and eventually find many treatments available which can cure the patients.

GLOSSARIES

Abbreviation Terms	Explanations
BOR	Best Overall Response
CR	Complete Response
iCPD	Immune Confirmed Progressive Disease
iCR	Immune Complete Response
iPFS	Immune Progression Free Survival
iPR	Immune Partial Response
iSD	Immune Stable Disease
iUPD	Immune Un-confirmed Progressive Disease
NE	Not Evaluable
PD	Progressive Disease
PFS	Progression Free Survival
PR	Partial Response
SD	Stable Disease

REFERENCES

1. Xiangchen (Bob) Cui, and Sri Pavan Vemuri, "*Simplifying the Derivation of Best Overall Response per RECIST 1.1 and iRECIST in Solid Tumor Clinical Studies*", Pharma SUG 2020, paper DV-66

2. *RECIST 1.1 – Update and Clarification: From the RECIST Committee; 2016* https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5737828/pdf/nihms926904.pdf.

3. *New response evaluation criteria in solid tumors: Revised RECIST guideline* (version 1.1), 2009 https://ctep.cancer.gov/protocolDevelopment/docs/recist_guideline.pdf.

4. *Revised RECIST Guideline Version 1.1: What Oncologists Want to Know and What Radiologists Need to Know* https://www.ajronline.org/doi/pdf/10.2214/AJR.09.4110.

5. Abraham Yeh, *"From Sea to Shining Sea- End to End discussion on PR and LB data ",* Pharma SUG 2023, paper DS-62

6. Zhongfei Lei, Xiaoxuan Zhu, "How to better understand iRECIST – a new wave in immunotherapeutic ", Pharma SUG China, 2022, Paper ST-124

ACKNOWLEDGMENTS

We want to thank Runxin Huang as the primary supporter of this paper. In addition, Duyen Ngo as the medical expert and friend of ours to give her medical inputs especially for why and how these confirmations went. The AbbVie implementation guide provides us solid background for various lessons on related topics as discussed in the paper.

CONTACT INFORMATION

(In case a reader wants to get in touch with you, please put your contact information at the end of the paper.)
Your comments and questions are valued and encouraged. Contact the author at:
Author Name: Abraham Yeh
Company: AbbVie Pharmaceutical Inc.
Email: abraham.yeh@abbvie.com
Website: www.abbvie.com