PharmaSUG China 2026 - Paper SA-176

Chronic Graft-versus-Host Disease (cGVHD) efficacy criteria exploration and implementation challenges – based on a first line phase 3 case study

Ting Zhang, Sanofi, Chengdu, China

ABSTRACT

Chronic graft versus host disease (cGVHD) is an immune-mediated inflammatory and fibrotic disorder characterized by tissue damage and multiple organ involvement. It is the leading cause of morbidity, late non-relapsed mortality (NRM), and impaired quality of life after an allogeneic hematopoietic cell transplantation (HCT), affecting up to 70% of all allogeneic HCT recipients. Due to its widespread and considerable disease burden, it has accelerated the evolution for international consensus response criteria.

In 2005, the National Institutes of Health (NIH) cGVHD Consensus Response Criteria Working Group initially established the recommendations for measures and response evaluation criteria. And revised by the 2014 Working Group to enhance reliability and practical utility in clinical trials to accelerate development and approval of novel therapeutic agents in cGVHD. Per recent FDA recommendation when initiating first line phase 3 study, based on 2014 NIH consensus response criteria, it's suggested to use progression from nadir instead of baseline after the first complete response or partial response achieved and progression confirmation may need per specific scenario. Based on existing standard and FDA recommendations, the exploration and implementation of the new efficacy criteria embarked on which provided highly valuable experience for future clinical trials development not only for cGVHD indication and also for other indication regarding the challenges and good practice during the initialization, standardization, and implementation of new efficacy criteria.

This paper will present how it works for both current standard and FDA recommendations, and corresponding challenges and good practice. Including data collection, organ response and overall response derivation, progression confirmation, response-based endpoints, time to event endpoints like event free survival and statistical considerations regarding analysis visit time window and programming cut off.

INTRODUCTION

In this paper, we will introduce the guidelines of 2014 National Institutes of Health (NIH) consensus response criteria and FDA recommendation when initiating and implement of the first line phase 3 study, including detailed data collection, organ response and overall response derivation, progression confirmation, response-based endpoints, time to event endpoints, analysis visit time window and programming cut off to provide readers with very valuable reference for how cGVHD efficacy algorithm works and benefiting the future research and development for new cGVHD clinical trials, bringing more renewed hope to patients living with this debilitating disease, who have long awaited new treatment options.

BACKGROUND

Allogeneic hematopoietic cell transplantation (HCT) is an important and often the only curative treatment option for many hematologic malignancies. However, chronic graft versus host disease (cGVHD) remains a major complication of allogeneic HCT, leading cause of morbidity, late non-relapse mortality (NRM), and impaired QoL, affecting up to 70% of all allogeneic HCT recipients. It is estimated that approximately 19 800 allogeneic HCT per year are performed in the European Union (EU) and the number of allogeneic HCT has surpassed 8000 per year in the United States (US) in 2020.

Systemic corticosteroids of prednisone or equivalent is the standard of care (SOC) for first-line treatment. While systemic corticosteroids are an effective treatment, only one-half of patients respond, and the disease progresses in many of these responders. Corticosteroid treatment is also associated with considerable toxicity. In addition, limited progress has been made to influence the SOC for front-line systemic cGVHD treatment over the last 3 decades.

Therefore, there continues to be a very important need for safe and efficacious front-line treatments for patients with cGVHD.

HOW 2014 NATIONAL INSTITUTES OF HEALTH (NIH) CONSENSUS RESPONSE CRITERIA WORKS AND FDA RECOMMENDATIONS?

DATA COLLECTION

Skin, mouth, liver, upper and lower GI, esophagus, lung, eye, and joint/fascia are organs or sites considered in evaluating overall response. Therefore, at each timepoint/assessment, these 9 organs scores or measures, date of assessment, and reason for abnormalities present will be collected and used for analysis. The best practice is to collect these 9 organs scores or measures, date of assessment, and reason for abnormalities present within one eCRF form, which will be easier to link all the data for these 9 organs at each timepoint/assessment. If liver ALT, ALP and total BILI collected under laboratory CRF form once, then need to make sure the information to distinguish which timepoint/assessment this assessment belongs to, besides, date of liver reason for abnormalities present need to be collected. Similar for lung % FEV1, if collected under Pulmonary Function Test CRF form belongs to Respiratory System Findings, also need to make sure the information to distinguish which timepoint/assessment this assessment belongs to, besides, date of lung % FEV1 reason for abnormalities present need to be collected. Below data collection for each organ is based on the best practice.

Skin

If Skin score by BSA and Skin feature score both reported with abnormalities, the highest score should be considered for organ score response.

- Skin score by BSA
  - 0-No BSA involved
  - 1-1-18% BSA
  - 2-19-50% BSA
  - 3->50% BSA
- Skin features score
  - 0-No sclerotic features
  - 2-Superficial sclerotic features 'not hidebound' (able to pinch)
  - 3-Deep sclerotic features, 'Hidebound' (unable to pinch), Impaired mobility, Ulceration
- Date of assessment (Skin)
- Explain reason for abnormalities present (Skin)
  - Abnormalities present but explained entirely by non-GVHD documented cause(s)
  - Abnormalities thought to represent GVHD plus other causes
  - Abnormalities thought to be due to GVHD only
  - No abnormalities present

Mouth

- Erythema
  - 0-None
  - 1-Mild erythema or moderate erythema (<25%)
  - 2-Moderate (>=25%) or Severe erythema (<25%)

    - 3-Severe erythema (>=25%)
- Lichenoid
    - 0-None
    - 1-Lichen-like changes (<25%)
    - 2-Lichen-like changes (25-50%)
    - 3-Lichen-like changes (>50%)
- Ulcers
    - 0-None
    - 3-Ulcers involving (<=20%)
    - 6-Severe ulcerations (>20%)
- Total score for all mucosal changes
- Date of assessment (Mouth)
- Explain reason for abnormalities present (Mouth)
    - Abnormalities present but explained entirely by non-GVHD documented cause(s)
    - Abnormalities thought to represent GVHD plus other causes
    - Abnormalities thought to be due to GVHD only
    - No abnormalities present

Liver

- ALT and ULN
- Alkaline Phosphatase and ULN
- Total Bilirubin and ULN
- Date of assessment (Liver)
- Explain reason for abnormalities present (Liver)
    - Abnormalities present but explained entirely by non-GVHD documented cause(s)
    - Abnormalities thought to represent GVHD plus other causes
    - Abnormalities thought to be due to GVHD only
    - No abnormalities present

Upper GI

- Gastrointestinal-Upper GI
    - 0-No symptoms
    - 1-Mild, occasional symptoms, with little reduction in oral intake during the past week
    - 2-Moderate, intermittent symptoms, with some reduction in oral intake during the past week
    - 3-More severe or persistent symptoms throughout the day, with marked reduction in oral intake, on almost every day of the past week
- Date of assessment (Upper GI)

- Explain reason for abnormalities present (Upper GI)
  - Abnormalities present but explained entirely by non-GVHD documented cause(s)
  - Abnormalities thought to represent GVHD plus other causes
  - Abnormalities thought to be due to GVHD only
  - No abnormalities present

Lower GI

- Gastrointestinal-Lower GI
  - 0-No loose or liquid stools during the past week
  - 1-Occasional loose or liquid stools, on some days during the past week
  - 2- Intermittent loose or liquid stools, throughout the day, on almost every day of the past week, without requiring intervention to prevent or correct volume depletion
  - 3-Voluminous diarrhea on almost every day of the past week, requiring intervention to prevent or correct volume depletion
- Date of assessment (Lower GI)
- Explain reason for abnormalities present (Lower GI)
  - Abnormalities present but explained entirely by non-GVHD documented cause(s)
  - Abnormalities thought to represent GVHD plus other causes
  - Abnormalities thought to be due to GVHD only
  - No abnormalities present

Esophagus

- Gastrointestinal- Esophagus
  - 0-No esophageal symptoms
  - 1-Occasional dysphagia or odynophagia with solid food or pills during the past week
  - 2-Intermittent dysphagia or odynophagia with solid foods or pills, but not for liquids or soft foods, during the past week
  - 3-Dysphagia or odynophagia for almost all oral intake, on almost every day of the past week
- Date of assessment (Esophagus)
- Explain reason for abnormalities present (Esophagus)
  - Abnormalities present but explained entirely by non-GVHD documented cause(s)
  - Abnormalities thought to represent GVHD plus other causes
  - Abnormalities thought to be due to GVHD only
  - No abnormalities present

Lung

- Lung symptom score
  - 0-No symptoms
  - 1-Mild symptoms (shortness of breath after climbing one flight of steps)

    - 2-Moderate symptoms (shortness of breath after walking on flat ground)
    - 3-Severe symptoms (shortness of breath at rest requiring O2)
- Date of assessment (Lung symptom score)
- Explain reason for abnormalities present (Lung symptom score)
    - Abnormalities present but explained entirely by non-GVHD documented cause(s)
    - Abnormalities thought to represent GVHD plus other causes
    - Abnormalities thought to be due to GVHD only
    - No abnormalities present
- % FEV1 (Percent Predicted Forced Expiratory Volume in 1 Second)
- Date of assessment (% FEV1)
- Explain reason for abnormalities present (% FEV1)
    - Abnormalities present but explained entirely by non-GVHD documented cause(s)
    - Abnormalities thought to represent GVHD plus other causes
    - Abnormalities thought to be due to GVHD only
    - No abnormalities present

Eye

- Eye
    - 0-No symptoms
    - 1-Mild dry eye symptoms not affecting ADL (requirement of lubricant eye drops <=3x per day)
    - 2-Moderate dry eye symptoms partially affecting ADL (requiring lubricant eye drops >3x per day or punctual plugs), WITHOUT new vision impairment due to KCS
    - 3-Severe dry eye symptoms significantly affecting ADL (special eyewear to relieve pain) or unable to work because of ocular symptoms or loss of vision due to KCS
- Date of assessment (Eye)
- Explain reason for abnormalities present (Eye)
    - Abnormalities present but explained entirely by non-GVHD documented cause(s)
    - Abnormalities thought to represent GVHD plus other causes
    - Abnormalities thought to be due to GVHD only
    - No abnormalities present

Joints and Fascia

- Joints and Fascia
    - 0-No symptoms
    - 1-Mild tightness of arms or legs, normal or mild decreased range of motion (ROM) and not affecting ADL
    - 2-Tightness of arms or legs or joint contractures, erythema thought due to fasciitis, moderate decreased ROM and mild to moderate limitation of ADL
    - 3-Contractures with significant decrease of ROM and significant limitation of ADL (unable to tie

shoes, button shirts, dress self etc.)

- Date of assessment (Joints and Fascia)
- Explain reason for abnormalities present (Joints and Fascia)
  - Abnormalities present but explained entirely by non-GVHD documented cause(s)
  - Abnormalities thought to represent GVHD plus other causes
  - Abnormalities thought to be due to GVHD only
  - No abnormalities present
- P-ROM
  - P-ROM Shoulder score, P-ROM Elbow score and P-ROM Wrist/finger score
    - 1-worst
    - 2
    - 3
    - 4
    - 5
    - 6
    - 7-normal
  - P-ROM Ankle score
    - 1-worst
    - 2
    - 3
    - 4-normal
  - P-ROM score
- Date of assessment (P-ROM)
- Explain reason for abnormalities present (P-ROM)
  - Abnormalities present but explained entirely by non-GVHD documented cause(s)
  - Abnormalities thought to represent GVHD plus other causes
  - Abnormalities thought to be due to GVHD only
  - No abnormalities present

BASELINE ORGAN INVOLVEMENT

The best practice for baseline organ involvement information could be collected within CRF, and for analysis could use collected information per CRF or derived information by Sponsor per table 1 below. In the study ongoing period, could monitor the correction of the collected data.

Organ	Organ involvement
Skin	NIH Skin Score ≥1, not entirely explained by non-cGVHD cause
Mouth	NIH Modified Oral Mucosal Rating Score ≥1, not entirely explained by non-cGVHD cause

Organ	Organ involvement
Liver	Elevation >2 x ULN of one or more (ALT, ALP, Total Bilirubin), not entirely explained by non-cGVHD cause
Upper GI	NIH Upper GI Score ≥1, not entirely explained by non-cGVHD cause
Lower GI	NIH Lower GI Score ≥1, not entirely explained by non-cGVHD cause
Esophagus	NIH Esophagus Score ≥1, not entirely explained by non-cGVHD cause
Lungs	FEV1 <75% predicted, not entirely explained by non-cGVHD cause If no PFTs available, use NIH lung score ⩾1, not entirely explained by non-cGVHD cause
Eyes	NIH Eye Score ≥1, not entirely explained by non-cGVHD cause
Joints and Fascia	P-ROM Score <25 or NIH Joints/Fascia Score ≥1, not entirely explained by non-cGVHD cause

Table 1. Baseline organ involvement

ORGAN RESPONSE EVALUATION

Organ	Complete Response	Partial Response	Unchanged	Progression	Non evaluable
Skin	NIH Skin Score 0 after previous involvement	Decrease in NIH Skin Score from baseline by 1 or more points	No change or increase from 0 to 1 in NIH Skin Score	Increase in NIH Skin Score from baseline by 1 or more points, except 0 to 1	NIH skin score is Non evaluable
Mouth	NIH Modified OMRS 0 after previous involvement	Decrease in NIH Modified Oral Mucosa Rating Score of 2 or more points	Increase/decrease in NIH Modified OMRS by less than 2 points or equal score at both time points	Increase in NIH Modified Oral Mucosa Rating Score of 2 or more points	NIH mouth score is Non evaluable
Liver	Normal ALT, ALP and total bilirubin after elevation of one or more at baseline	50%improvement in ALT, ALP, or total bilirubin without increase by 2 x ULN for any other parameters of ALT, ALP or total bilirubin	Not in the other categories	Increase by 2 x ULN for any one of ALT, ALP or total bilirubin	If ALT, ALP and total bilirubin are missing
Upper GI	NIH Upper GI Score 0 after Previous involvement	Decrease in NIH Upper GI Score by 1 or more points	No change or increase from 0 to 1 in NIH Upper GI Score	Increase in Upper GI Score by 1 or more points, except 0 to 1	NIH upper GI score is Non evaluable
Lower GI	NIH Lower GI Score 0 after previous involvement	Decrease in NIH Lower GI Score by 1 or more points	No change or increase from 0 to 1 in NIH Lower GI Score	Increase in Lower GI Score by 1 or more points, except 0 to 1	If NIH Lower GI score is Non evaluable

Organ	Complete Response	Partial Response	Unchanged	Progression	Non evaluable
Esophagus	NIH Esophagus Score 0 after previous involvement	Decrease in NIH Esophagus Score by 1 or more points	No change or increase from 0 to 1 in NIH Esophagus Score	Increase in NIH Esophagus Score by 1 or more points, except 0 to 1	NIH esophagus score is Non evaluable
Lungs	%FEV1 ⩾80% after previous involvement; or NIH Lung Symptom Score 0 after previous involvement if PFTs not available	Increase by 10% absolute value of %FEV1; or decrease in NIH Lung Symptom Score by 1 or more points if PTSs not available	Not in the other categories	Decrease by 10% absolute value of %FEV1 and %FEV1 < 65%; or increase in NIH Lung Symptom Score by 1 or more points, except 0 to 1 if PFTs not available	NIH Lung score is Non evaluable and %FEV1 is missing
Eyes	NIH Eye Score 0 after previous involvement	Decrease in NIH Eye Score by 1 or more points	No change or increase from 0 to 1 in NIH Eye Score	Increase in NIH Eye Score by 1 or more points, except 0 to 1	NIH eye score is Non evaluable
Joints and Fascia	NIH Joint and Fascia Score 0 and P-ROM score 25 after previous involvement in at least one measure	Decrease in NIH Joint and Fascia Score by 1 or more points or increase in P-ROM score by 1 or more points for any site	Not in the other categories	Increase in NIH Joint and Fascia Score by 1 or more points or decrease in P-ROM score by 1 or more points for any site	Both in NIH Joint and Fascia and P- ROM scores are Non evaluable

Table 2. Response Determination for Chronic GVHD Based on Clinician Assessments

For organ response derivation, baseline involvement is considered. If baseline and post baseline are not involved, then the organ response should be set to “Not Involved”.

For organ response derivation, if the reason of abnormality is non cGVHD related, for score used organs, then consider as score equal to 0 for Mouth, Esophagus, Upper GI, Lower GI, Skin, Eyes, Lung symptom score, Joints and Fascia score and consider as score equal to 25 for P-ROM score.

Per 2014 NIH consensus response criteria, the organ response derivation taking the reference of baseline, per FDA recommendations, if the participants achieved first overall response CR or PR, then after first overall response CR or PR assessments, taking the reference of nadir (best prior status on or after first overall response CR or PR assessments) for progression calculation (if compare with nadir not progression, then still compare with baseline for organ response calculation).

As reason of abnormality included for organ response derivation, how to identify the nadir becomes the most challenging part. The nadir will include both nadir score or measure and nadir reason of abnormality.

Best status order for reason of abnormality is as below.

1) No abnormalities present

2) Abnormalities present but explained entirely by non-GVHD documented cause(s)

3) Abnormalities thought to represent GVHD plus other causes

4) Abnormalities thought to be due to GVHD only

For Mouth, Esophagus, Upper GI, Lower GI, Skin, Eyes, will follow rule 1 as below.

If the current reason of abnormality is the same compared with prior nadir reason of abnormality, then if current score or measure is smaller than prior nadir score or measure then current score will be the nadir score or measure, and current score or measure and reason of abnormality will be used for subsequent assessment as nadir score or measure and nadir reason of abnormality.

If the reason of abnormality is not the same will follow below rules.

1) If current reason is 'No abnormalities present' and prior nadir reason is 'Abnormalities thought to be due to GVHD only'/'Abnormalities thought to represent GVHD plus other causes'/'Abnormalities present but explained entirely by non-GVHD documented cause(s)', then current score or measure and reason of abnormality will be the nadir score or measure and nadir reason for subsequent response assessment.
2) Else if current reason is 'Abnormalities present but explained entirely by non-GVHD documented cause(s)' and nadir reason is 'Abnormalities thought to be due to GVHD only'/ 'Abnormalities thought to represent GVHD plus other causes', then current score or measure and reason of abnormality will be the nadir score and nadir reason for subsequent response assessment.
3) Else if current reason is 'Abnormalities thought to represent GVHD plus other causes' and nadir reason is 'Abnormalities thought to be due to GVHD only',
    a. If current score or measure is the same as prior nadir score or measure, then nadir will be current score or measure and current reason.
    b. If current score or measure is smaller than nadir score or measure, then nadir will be current score or measure and nadir reason will be 'Abnormalities thought to represent GVHD plus other causes'.
    c. If current score or measure is bigger than nadir score or measure, then there is no change for nadir and nadir reason.
4) Else if current reason is 'Abnormalities thought to be due to GVHD only'' and nadir reason is 'Abnormalities thought to represent GVHD plus other causes',
    a. If current score or measure is the same as nadir score or measure, then there is no change for nadir and nadir reason.
    b. If current score or measure is smaller than nadir score or measure, then nadir will be current score or measure, and nadir reason will be current reason.
    c. If current score or measure is bigger than nadir score, then there is no change for nadir score or measure and nadir reason.

For Lung, Lung symptom score and corresponding reason of abnormality data, will follow rule 1. Lung % FEV1 data and corresponding reason of abnormality, follow rule 1, except the % FEV1 value bigger corresponding to better status.

For Joints and Fascia score and corresponding reason of abnormality data, follow rule 1. And for P-ROM score and corresponding reason of abnormality data, follow rule 1, except the P-ROM score bigger corresponding to better status, and P-ROM score equal to 25 corresponding to best status.

For liver, nadir measure corresponding to ALT/1*ULN and ALP/1*ULN and Total BILI/1*ULN. For ALT/1*ULN or ALP/1*ULN or Total BILI/1*ULN will find smallest 1*ULN value separately for each parameter. If the reason of abnormality is the same, then will find smallest 1*ULN value separately for ALT and ALP and Total BILI. If the reason is not the same, follow best status order for reason of

abnormality to get nadir reason, and for ALT/1*ULN or ALP/1*ULN or Total BILI/1*ULN find smallest 1*ULN value separately for each parameter among all ALT/1*ULN or ALP/1*ULN or Total BILI/1*ULN on or after first overall response achieved CR or PR and until last assessments before current assessment.

OVERALL RESPONSE EVALUATION

Table 3 below is the overall response derivation rule per 2014 NIH consensus response criteria, per FDA recommendation, for the participants if achieved first overall response CR or PR, on or before first overall response CR or PR achieved follow table 3 below for overall response calculation, after the first overall response CR or PR assessments, overall response calculation for progression will follow below cGVHD progression rule. The date of overall response will follow the rule, if overall response is progression, then use the initial progression date from the organ (s) leading to progression, if overall response is not progression, then use the maximum date of all organ response date.

Response	Definitions
Complete Response (CR)	Resolution of all manifestations in each organ or site
Partial Response (PR)	Improvement in at least 1 organ or site without progression in any other organ or sites
Mixed Response	PR or CR in at least one organ accompanied by progression in another organ
Progression	Progression in one or more organ(s) without a response in any other organ or site
Unchanged response	Outcome that does not meet the criteria for CR, PR, progression or mixed response

Table 3. cGVHD overall response definitions based on 2014 NIH consensus response criteria

CGVHD PROGRESSION

- If during cGVHD response assessment, 2 or more organs meet the criteria for disease progression that will be considered that patient has progressed.
- However, if only 1 organ shows progression, there should be reassessment after at least 28 days (4 weeks) from last assessment visit through the date of the next assessment visit.
    - The start of new systemic therapy between visit dates would be considered a positive reassessment of PD and the initial date of observation of PD is assigned as the PD date.
    - If a new systemic therapy is not started, and the organ status is still PD at the reassessment visit, then the initial date of observation of PD is assigned as the PD date without regard to status in other organs at the reassessment visit.
    - If a new systemic therapy is not started, and the organ status improved at the reassessment visit, and there is no progression in other organs, then it is not counted as a PD at the initial observation date.
    - If a new systemic therapy is not started, and the organ status improved at the reassessment visit, but there is progression in other organ, this may trigger a new cycle of assessment for PD for the other organ starting at the reassessment date.
    - If a new systemic therapy is not started, and the patient is lost to follow-up before reassessment or there is a data cut before reassessment, then it is considered PD, and the initial date of observation of PD is assigned as the PD date.

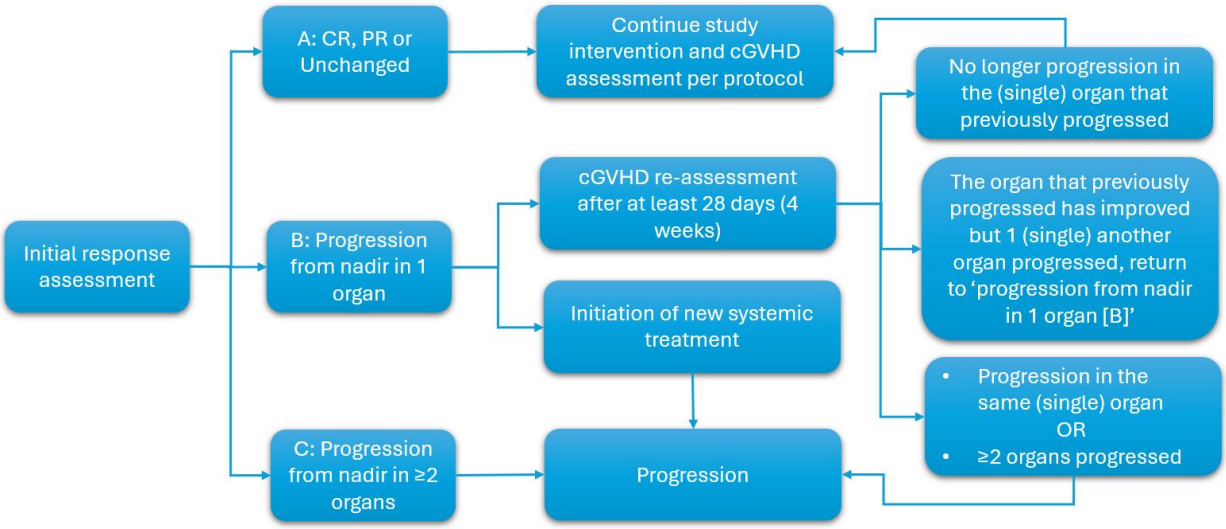

Figure 1. Assessment/Re-assessment Flow of Overall Response/Progression

RESPONSE-BASED ENDPOINTS

BEST OVERALL RESPONSE

The BOR is the best overall response observed from the date of randomization until new systemic treatment.

OVERALL RESPONSE RATE (ORR)

Proportion of participants who achieve an overall response (CR or PR) as per 2014 NIH consensus response criteria at any time on or before the start of new systemic treatment for cGVHD.

RESPONSE BY ORGAN

Proportion of participants who achieve CR or PR as per NIH consensus response criteria (2014) at any time point in each involved organ and on or before the start of a new systemic therapy for cGVHD.

TIME-TO-EVENT ENDPOINTS

Time to response (TTR)

Time to Response is defined as the time from randomization to the date the patient has first response (CR or PR).

Duration of response (DOR)

Time from the date of the first response to the date of cGVHD progression, start of new systemic treatment for cGVHD, or death, whichever occurs first. DOR is determined only for participants who achieved overall response (PR or CR) as per 2014 NIH consensus response criteria.

Event-Free Survival (EFS)

EFS defined as time from the date of randomization to 1) day one after randomization in case of failure to achieve an overall response (partial response [PR] or complete response [CR]) as per 2014 National Institutes of Health (NIH) consensus response criteria by 24 weeks, or 2) the date of cGVHD progression from nadir in any organ in participants who achieve CR or PR by 24 weeks, or 3) the date of initiation of new systemic treatment for cGVHD in participants who achieve CR or PR by 24 weeks, or 4) the date of death in participants who achieve CR or PR by 24 weeks, whichever occurs first.

STATISTICAL AND PROTOCOL CONSIDERATION

ANALYSIS VISIT TIME WINDOW

As it's a fixed visit with allowable time window for the efficacy assessment, analysis visit time window applied for the analysis. Below table 4 example, shows the scheduled visits for efficacy assessment on baseline, and then every 8 weeks, with +/- 7 days' allowable time window. All the efficacy assessments will be included for analysis for both scheduled visits and unscheduled visits.

For organ scores or measures or reason of abnormality or organ response, if multiple assessments within same visit window, for scheduled and unscheduled visits, add .01, .02, .03, .04 etc. after the analysis visit value (e.g. WEEK 8.01, UNSCHEDULED WEEK 4.01).

For overall response, if multiple assessments within same visit window, for unscheduled visits, will be the same as organ level parameters, while for scheduled visits, will keep unique overall response, follow the rule, if overall response with progression, will keep the first progression overall response and date, if overall response without progression, will keep maximum date overall response and date.

Analysis Visit	Analysis Visit (N)	Target day	Window
BASELINE	0	1	[NEGATIVE, 1]
UNSCHEDULED WEEK 4	4	29	[2,49]
WEEK 8	8	57	[50,64]
UNSCHEDULED WEEK 12	12	85	[65,105]
WEEK 16	16	113	[106,120]
UNSCHEDULED WEEK 20	20	141	[121,161]
WEEK 24	24	169	[162,176]
UNSCHEDULED WEEK 28	28	197	[177,217]
WEEK 32	32	225	[218,232]
UNSCHEDULED WEEK 36	36	253	[233,273]
WEEK 40	40	281	[274,288]
UNSCHEDULED WEEK 44	44	309	[289,329]
WEEK 48	48	337	[330,344]

Table 4. Analysis visit time window

PROGRAMMING CUT OFF

The programming cut off will be implemented for efficacy analysis when there is an interim analysis. The major point is that for each timepoint/assessment, there are 9 organs assessment with corresponding date of assessment may be different, if we cut the date of assessment less than or equal to cut off date, for the timepoint/assessment around cutoff date, maybe some organs' data entered while some other organs' data not entered which may lead inaccuracy for the efficacy results and may leading wrong decisions. So, if we have this kind of case, we will keep the complete assessment for all 9 organs' data if at least one organ date of assessment is less than or equal to cut off date.

CONCLUSION

Based on the exploration and implementation for 2014 National Institutes of Health (NIH) consensus response criteria and FDA recommendations, this paper provides with readers best practice for the data collection, challenges and reference for organ response, overall response and progression confirmation derivation per FDA recommendations for first line clinical trial research and development, statistical considerations for analysis visit time window and programming cut off. These practices, challenges, and attempts have provided highly valuable experience and laid a solid foundation for future clinical trials research and development for new cGVHD drug. Therefore, bringing patients new hope and treatment options faster.

REFERENCES

Penack O, Marchetti M, Ruutu T, Aljurf M, Bacigalupo A, Bonifazi F, et al. Prophylaxis and management of graft versus host disease after stem-cell transplantation for haematological malignancies: updated consensus recommendations of the European Society for Blood and Marrow Transplantation. Lancet Haematol. 2020;7(2):e157-67.

Gonzalez R, Pidala J. Evolving therapeutic options for chronic graft‐versus‐host disease. Pharmacotherapy. 2020;40(8):756-72.

Passweg J, Baldomero H, Chabannon C, Basak G, de la Cámara R, Corbacioglu S, et al. Hematopoietic cell transplantation and cellular therapy survey of the EBMT: monitoring of activities and trends over 30 years. Bone Marrow Transplant. 2021;56(7):1651-64.

Auletta JJ, Kou J, Chen M, Shaw BE. Current use and outcome of hematopoietic stem cell transplantation: CIBMTR US summary slides. 2021.

Lee SJ, Onstad L, Chow EJ, Shaw BE, Jim HSL, Syrjala KL, et al. Patient-reported outcomes and health status associated with chronic graft-versus-host disease. Haematologica. 2018;103(9):1535-41.

Flowers ME, Martin PJ. How we treat chronic graft-versus-host disease. Blood. 2015;125(4):606-15.

Lee S, Wolff D, Kitko C, Koreth J, Inamoto Y, Jagasia M, et al. Measuring therapeutic response in chronic graft-versus-host disease. National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: IV. The 2014 Response Criteria Working Group Report. Bio Blood Marrow Transplant. 2015;21(6):984-99.

Jagasia MH, Greinix HT, Arora M, Williams KM, Wolff D, Cowen EW, et al. National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: I. The 2014 Diagnosis and Staging Working Group report. Biol Blood Marrow Transplant. 2015;21(3):389-401.

ACKNOWLEDGMENTS

I would like to acknowledge Cindy Wen (Sanofi, Beijing, China), Sufang Huang (Sanofi, Morristown, NJ, US), Zhini Wang (Sanofi, Shanghai, China), Flora Yao (Sanofi, Morristown, NJ, US) and Agnes Brunette (Sanofi, Gentilly, France) for guidance and review for the paper. Moreover, I would especially like to thank my husband Hekan Wang, and my little daughter Yiqi Wang have been incredibly supportive.

RECOMMENDED READING

- *National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: I. The 2014 Diagnosis and Staging Working Group Report, available at National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: I. The 2014 Diagnosis and Staging Working Group Report - PMC*

CONTACT INFORMATION

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

Ting Zhang
Sanofi
Ting12.Zhang@sanofi.com

Conflict of Interest statement: Ting Zhang is a Sanofi employee and may hold shares and/or stock options in the company.