

Patient-reported treatment response in chronic graft-versus-host disease

Annie Im,¹ Iskra Pusic,² Lynn Onstad,³ Carrie L. Kitko,⁴ Betty K. Hamilton,⁵ Amin M. Alousi,⁶ Mary E. Flowers,³ Stefanie Sarantopoulos,⁷ Paul Carpenter,³ Jennifer White,⁸ Sally Arai,⁹ Najla El Jurdi,¹⁰ George Chen,¹¹ Corey Cutler,¹² Stephanie Lee³ and Joseph Pidala¹³

¹University of Pittsburgh/UPMC Hillman Cancer Center, Pittsburgh, PA, USA; ²Division of Medicine and Oncology, Washington University, Saint Louis, MO, USA; ³Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA; ⁴Pediatric Blood and Marrow Transplantation Program, Vanderbilt University Medical Center, Nashville, TN, USA; ⁵Blood and Marrow Transplantation, Department of Hematology and Medical Oncology, Taussig Cancer Institute, Cleveland Clinic, Cleveland, OH, USA; ⁶Department of Stem Cell Transplantation and Cellular Therapy, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; ⁷Division of Hematological Malignancies and Cellular Therapy, Duke University Department of Medicine, Duke Cancer Institute, Durham, NC, USA; ⁸Leukemia/Bone Marrow Transplant Program of British Columbia, British Columbia Cancer Agency, Vancouver, British Columbia, Canada; ⁹Department of Medicine, Division of Blood and Marrow Transplantation, Stanford University School of Medicine, Stanford, CA, USA; ¹⁰Division of Hematology, Oncology and Transplantation, University of Minnesota Medical Center, Minneapolis, MN, USA; ¹¹Department of Medicine, Roswell Park Comprehensive Cancer Center, Buffalo, NY, USA; ¹²Division of Stem Cell Transplantation and Cellular Therapy, Dana-Farber Cancer Institute, Boston, MA, USA and ¹³Blood and Marrow Transplantation and Cellular Immunotherapy, H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL, USA

Correspondence: J. Pidala
joseph.pidala@moffitt.org

Received: January 18, 2023.

Accepted: May 17, 2023.

Early view: May 25, 2023.

<https://doi.org/10.3324/haematol.2023.282734>

©2024 Ferrata Storti Foundation

Published under a CC BY-NC license



Abstract

Chronic graft-versus-host disease (GvHD) treatment response is assessed using National Institutes of Health (NIH) Consensus Criteria in clinical trials, and by clinician assessment in routine practice. Patient-reported treatment response is central to the experience of chronic GvHD manifestations as well as treatment benefit and toxicity, but how they correlate with clinician- or NIH-responses has not been well-studied. We aimed to characterize 6-month patient-reported response, determine associated chronic GvHD baseline organ features and changes, and evaluate which patient-reported quality of life and chronic GvHD symptom burden measures correlated with patient-reported response. From two nationally representative Chronic GVHD Consortium prospective observational studies, 382 subjects were included in this analysis. Patient and clinician responses were categorized as improved (completely gone, very much better, moderately better, a little better) versus not improved (about the same, a little worse, moderately worse, very much worse). At six months, 270 (71%) patients perceived chronic GvHD improvement, while 112 (29%) perceived no improvement. Patient-reported response had limited correlation with either clinician-reported (kappa 0.37) or NIH chronic GvHD response criteria (kappa 0.18). Notably, patient-reported response at six months was significantly associated with subsequent failure-free survival. In multivariate analysis, NIH responses in eye, mouth, and lung had significant association with 6-month patient-reported response, as well as a change in Short Form 36 general health and role physical domains and Lee Symptom Score skin and eye changes. Based on these findings, patient-reported responses should be considered as an important complementary endpoint in chronic GvHD clinical trials and drug development.

Introduction

Chronic graft-versus-host disease (GvHD) is the most common cause of late morbidity and mortality after allogeneic hematopoietic cell transplantation (HCT).^{1,2} Prior to the National Institutes of Health (NIH) Consensus Conferences, there was limited standardization of chronic GvHD severity scoring and treatment response assess-

ment. In 2005, the NIH Consensus Conference on Clinical Trials for Treatment of Chronic GVHD provided standardized criteria to assess organs involved in chronic GvHD and to determine response to therapies for use in clinical trials;³⁻⁵ consensus criteria were updated in 2014 based on data that had accumulated since the original guidelines.^{6,7} These criteria have elevated the scientific rigor of current clinical trials in chronic GvHD therapy. However,

substantial improvements in affected organs may be required to achieve response by these criteria, and clinician- or patient-reported response metrics may capture lesser but still important degrees of clinical benefit. For example, while NIH response criteria have been correlated with outcomes,⁸ other studies reported limited correlation between NIH response and clinician-reported response,^{9,10} and certain quality of life measures.¹¹

In chronic GvHD management outside of clinical trials, clinicians consider patient-reported symptom burden, findings from physical examination, and routine laboratory tests to determine clinical response, rather than strictly applying the NIH Response Criteria. In total, this clinician assessment of improvement, stability, or worsening is reduced to a perceived presence or absence of clinical benefit, and informs immunosuppressive (IS) therapy management. Clinician-reported response has been correlated with survival,⁸ and subsequent analyses found that changes in serum bilirubin, NIH 0 to 3-point scores of lower gastrointestinal tract, mouth, joint/fascia, lung, and skin were factors that correlated most with clinician-reported responses.⁹ Thus, while clinician-assessed response is less standardized and rigorous compared to NIH Response Criteria, it appears to be of value.

Patients affected by chronic GvHD experience the physical manifestations, symptoms, functional limitations and impairment in quality of life known to be due to chronic GvHD, and can identify their own benefit and toxicity from chronic GvHD therapies. Thus, a patient's self-report of treatment response may be of critical importance. While patient-reported responses are captured in clinical trials according to the NIH criteria, they are not used to determine overall NIH response. In fact, clinical trials have yet to incorporate patient-reported outcomes into determination of success; however, in chronic GvHD, improvement in quality of life is an important treatment goal.¹² It is unknown whether patient-reported responses correlate with clinician or NIH responses, nor do we know the factors that contribute to patient-reported responses. The aims of this study were to characterize 6-month patient-reported treatment responses, including associated chronic GvHD baseline organ features and changes, and to evaluate which patient-reported quality of life and chronic GvHD symptom measures correlate with patient-reported response. Ultimately, we sought to determine whether patient-reported response measures may capture clinical benefit in a different way to standard NIH or clinician-reported response measures in chronic GvHD.

Methods

Patients

Patients were enrolled in two prospective, multicenter

Chronic GVHD Consortium observational studies. The "Chronic GVHD Consortium Improvement Outcomes Assessment in Chronic GVHD" study enrolled 601 patients between 2007 and 2012.¹³ Patients in this study enrolled at any time after starting systemic treatment for chronic GvHD. At enrollment and every six months thereafter, providers and patients recorded standardized information regarding current chronic GvHD organ involvement and symptoms using forms developed according to the 2005 NIH Chronic GVHD Consensus Criteria. For incident cases, providers and patients also recorded the same information at three months after enrollment. The "Chronic GVHD Consortium Response Measures Validation Study" enrolled 383 patients with chronic GvHD between 2013 and 2017.¹⁴ Patients in this study enrolled within four weeks before or after starting a new systemic treatment for chronic GvHD. At enrollment and 3, 6, and 18 months thereafter, providers and patients recorded standardized information according to the 2014 NIH Consensus Conference Criteria. Exclusion criteria in both studies included primary disease relapse and inability to comply with study procedures. At each assessment, providers and patients rated overall changes in GvHD manifestations from enrollment according to an 8-point scale with categories of "completely gone", "very much better", "moderately better", "a little better", "about the same", "a little worse", "moderately worse", or "a lot worse". This is in the form of a single question that patients and providers complete. In addition, patients completed the Short Form 36 (SF-36) survey, and rated the severity of GvHD symptoms according to the Lee Chronic GVHD Symptom Scale.¹⁵ The protocols were approved by the Institutional Review Board at each site, and all patients provided informed consent in accordance with the Declaration of Helsinki.

For the purposes of this analysis, patients from the two parent cohort studies were included based on the availability of patient-reported response data at cohort enrollment and at six months. Responses were categorized as "improved" (completely gone, very much better, moderately better, a little better) *versus* "not improved" (about the same, a little worse, moderately worse, a lot worse).

Statistical analysis

Comparisons of characteristics by patient perception of chronic GvHD improvement *versus* no improvement were performed using the χ^2 test and Fisher's Exact Test for categorical variables and the Wilcoxon rank sum test for continuous variables. Multivariable logistic regression was used to examine the relationships between patient perception of chronic GvHD improvement, and transplant and chronic GvHD characteristics, as well as 6-month organ responses, using a stepwise procedure with entry and retention criteria of $P \leq 0.1$. The initial set of variables included all those found to be univariately related at the $P \leq 0.1$ level.

This same procedure was used to examine the relationships between patient perception of chronic GvHD improvement with transplant and patient-reported outcomes (PRO) of the Lee Symptom Scale and the SF-36.

The log-rank test was used to compare overall survival (OS) and failure-free survival (FFS) (composite outcome including death, malignancy relapse, and start of new line of systemic IS therapy) by patient perception of chronic GvHD improvement. OS was calculated from the 6-month visit until death; FFS was calculated from the 6-month visit until malignancy relapse, death, or addition of a new systemic IS medication for chronic GvHD among those who had not started a new IS medication for chronic GvHD before six months. Analyses were performed in SAS 9.4 (SAS Institute, Cary NC, USA).

Results

From the overall parent cohort (N=605), this study population was limited to 382 patients with baseline and 6-month patient-reported response data. The 223 excluded for missing 6-month patient response were predominantly due to missed visit or missed patient survey (N=166, 74.4%), with a lesser contribution from other reasons (withdrew from study before 6-month visit N=9, 4%; relapsed before six months N=23, 10.3%; or died before six months N=25, 11.2%). A comparison of those with versus without 6-month response data is presented in the *Online Supplementary Appendix*. Table 1 summarizes baseline patient and chronic GvHD characteristics (see also the *Online Supplementary Appendix*). The median time from chronic GvHD to enrollment was 0.5 months (Interquartile Range [IQR] 0-8.6 months). Most patients had moderate (50%) or severe (38%) chronic GvHD, and the most involved organ sites were skin (73%), mouth (60%), and eyes (54%). Thirty-four percent of the cases were prevalent, while 66% were incident cases. Chronic GvHD features are presented in Table 1.

At six months, 71% of patients reported improvement in their chronic GvHD, while 29% reported no improvement. Among those categorized as improved (N=270), patient response included: completely gone (N=31, 11.5%), very much better (N=119, 44.1%), moderately better (N=68, 25.2%), and a little better (N=52, 19.3%). For the not improved group (N=112), patient responses were: about the same (N=46, 41.1%), a little worse (N=37, 33%), moderately worse (N=18, 16.1%), and very much worse (N=11, 9.8%). Notably, there was limited correlation with clinician-reported response (kappa 0.37) and NIH response (kappa 0.18). Among clinicians, 66% reported improvement in chronic GvHD, while 34% reported no improvement. Per NIH response, 46% of patients had improvement in chronic GvHD while 54% had

no improvement. Clinician-reported and NIH responses grouped per patient-reported response are presented in the *Online Supplementary Appendix*. In univariate analysis, enrollment lung involvement was associated with patients' perception of response at six months ($P<0.001$). NIH response in skin ($P=0.003$), eye ($P<0.001$), mouth ($P=0.004$), and lung ($P<0.001$) at six months was also associated with patient-reported response at six months (*Online Supplementary Appendix*).

In multivariable analysis of patient-reported response, moderate-severe lung involvement at enrollment ($P=0.007$) was associated with report of no improvement, and NIH responses in eye ($P=0.009$), mouth ($P=0.01$), and lung ($P=0.03$) were associated with patient report of improvement at six months (Table 2).

In multivariable analysis of patient perception of chronic GvHD improvement with quality of life and symptom burden measures, a higher SF-36 general health score at enrollment ($P<0.001$) (indicating better quality of life) and improvement from enrollment to six months in the SF-36 general health score ($P<0.001$) were associated with patient report of improvement, while worsening in SF-36 role physical domain score was associated with patient report of no improvement ($P=0.04$) (Table 3). A lower LSS skin score at enrollment ($P=0.1$), indicating lower burden of symptoms, and improvement from enrollment to six months in the LSS eye score ($P=0.007$) were associated with patient report of improvement, while worsening of LSS skin score ($P<0.001$) was associated with patient report of no improvement (Table 3).

Failure-free survival after six months was higher among patients who reported improvement compared to those who reported no improvement ($P=0.005$) (Figure 1). The most common cause of failure was starting new treatment for chronic GvHD (Table 4). The cumulative incidence of new systemic IS therapy stratified by 6-month patient-reported response is included in the *Online Supplementary Appendix*. In a multivariable analysis of the association of 6-month response with subsequent time to new systemic IS therapy, we found no significant difference in the Hazard Ratios for patient, clinician, or NIH response (*data not shown*). There was no difference in overall survival at six months from enrollment between the two groups (Figure 2).

Discussion

This study provides new insight into patient-reported treatment response in chronic GvHD, a previously understudied area. Six-month patient-reported response was associated with subsequent FFS, and may represent a clinically meaningful measure. Moreover, given the limited correlation with NIH and clinician-assessed response, it

Table 1. Baseline chronic graft-versus-host disease characteristics by patient perception.

Variables	Categories	Total N=382	GvHD not improved N=112	GvHD improved N=270	P*
Skin*	0	103 (27.0%)	31 (27.7%)	72 (26.7%)	0.77
	1	66 (17.3%)	18 (16.1%)	48 (17.8%)	
	2	109 (28.5%)	29 (25.9%)	80 (29.6%)	
	3	104 (27.2%)	34 (30.4%)	70 (25.9%)	
Fascia*	0	266 (69.6%)	75 (67.0%)	191 (70.7%)	0.52
	1	92 (24.1%)	30 (26.8%)	62 (23.0%)	
	2	20 (5.2%)	7 (6.3%)	13 (4.8%)	
	3	4 (1.0%)	0 (0.0%)	4 (1.5%)	
GI	0	259 (68.0%)	77 (68.8%)	182 (67.7%)	0.86
	1	84 (22.0%)	24 (21.4%)	60 (22.3%)	
	2	35 (9.2%)	11 (9.8%)	24 (8.9%)	
	3	3 (0.8%)	0 (0.0%)	3 (1.1%)	
Liver	0	294 (78.8%)	88 (81.5%)	206 (77.7%)	0.07
	1	27 (7.2%)	11 (10.2%)	16 (6.0%)	
	2	43 (11.5%)	9 (8.3%)	34 (12.8%)	
	3	9 (2.4%)	0 (0.0%)	9 (3.4%)	
Lung	0	274 (71.7%)	69 (61.6%)	205 (75.9%)	<0.001
	1	82 (21.5%)	26 (23.2%)	56 (20.7%)	
	2	21 (5.5%)	14 (12.5%)	7 (2.6%)	
	3	5 (1.3%)	3 (2.7%)	2 (0.7%)	
Eye	0	176 (46.2%)	45 (40.5%)	131 (48.5%)	0.16
	1	130 (34.1%)	37 (33.3%)	93 (34.4%)	
	2	69 (18.1%)	26 (23.4%)	43 (15.9%)	
	3	6 (1.6%)	3 (2.7%)	3 (1.1%)	
Mouth	0	152 (39.8%)	50 (44.6%)	102 (37.8%)	0.22
	1	173 (45.3%)	51 (45.5%)	122 (45.2%)	
	2	52 (13.6%)	11 (9.8%)	41 (15.2%)	
	3	5 (1.3%)	0 (0.0%)	5 (1.9%)	
Joint	0	236 (61.9%)	68 (60.7%)	168 (62.5%)	0.53
	1	91 (23.9%)	24 (21.4%)	67 (24.9%)	
	2	47 (12.3%)	18 (16.1%)	29 (10.8%)	
	3	7 (1.8%)	2 (1.8%)	5 (1.9%)	
Genital	0	228 (83.8%)	59 (76.6%)	169 (86.7%)	0.19
	1	31 (11.4%)	13 (16.9%)	18 (9.2%)	
	2	8 (2.9%)	3 (3.9%)	5 (2.6%)	
	3	5 (1.8%)	2 (2.6%)	3 (1.5%)	

GvHD: graft-versus-host disease; GI: gastrointestinal. *Skin sclerosis and fascia involvement presented in the *Online Supplementary Appendix*.

appears patient-reported response may capture unique aspects of clinical benefit. We also identified both baseline chronic GvHD organ involvement and organ responses that had association with patient-reported response, and determined which PRO measure changes had greatest association with patient-reported response. Taken together, the findings from this study advance a potential additional approach for assessing clinical benefit in chronic GvHD therapy, and lay a foundation for future research in this area.

We determined that baseline lung and liver involvement, and NIH responses in eyes, mouth, and lungs had the greatest association with patient-reported responses in this cohort. Patient-reported response appeared to be most associated with organs where responses can be profound in their impact on symptom burden or functionality. These results may have been influenced by the baseline frequency and reversibility of specific chronic GvHD organ manifestations, scale definitions in calculated NIH response, as well as patient bias focused on more recent

Table 2. Multivariate analysis of patient characteristics, chronic graft-versus-host disease characteristics, National Institutes of Health organ responses by patient perception.

Variable*	Class	OR (95% CI)	P	Global P-value
Liver (at enrollment)	Not involved	Reference	-	0.06
	Mild	0.3 (0.1-0.9)	0.04	-
	Moderate/severe	1.5 (0.6-3.7)	0.41	-
Lung (at enrollment)	Not involved	Reference	-	0.02
	Mild	0.7 (0.3-1.6)	0.46	-
	Moderate/severe	0.1 (0.0-0.6)	0.007	-
NIH eye response	Not involved/stable	Reference	-	0.009
	CR	1.7 (0.6-4.7)	0.32	-
	PR	0.7 (0.2-2.2)	0.5	-
	Prog	0.3 (0.1-0.7)	0.003	-
NIH mouth response	Not involved/stable	Reference	-	0.01
	CR	2.3 (1.1-4.8)	0.03	-
	PR	3.8 (1.6-9.1)	0.003	-
	Prog	1.4 (0.5-4.0)	0.51	-
NIH lung response	Not involved/stable	Reference	-	0.03
	CR	2.7 (0.8-9.1)	0.1	-
	PR	2.2 (0.3-15.9)	0.43	-
	Prog	0.4 (0.1-1.0)	0.05	-

OR: Odd's Ratio; CI: Confidence Interval; NIH: National Institutes of Health; CR: complete response; PR: partial response; Prog: progression. *Model adjusted for disease diagnosis and parent cohort study.

Table 3. Multivariate analysis of patient-reported outcome by patient perception.

Variable*	OR (95% CI)**	P
SF-36 general health (enrollment)	2.7 (1.9-3.8)	<0.001
LSS skin (enrollment)	0.9 (0.7-1.0)	0.1
SF-36 general health (change score)	3.3 (2.0-5.3)	<0.001
SF-36 role-physical (change score)	1.4 (1.0-1.9)	0.04
LSS skin (change score)	0.7 (0.6-0.8)	<0.001
LSS eye (change score)	0.8 (0.7-1.0)	0.007

OR: Odd's Ratio; CI: Confidence Interval; SF-36: Short-Form 36; LSS: Lee Symptom Scale. *Model adjusted for parent cohort study. **OR presented for each variable represents OR per 10-point change.

changes versus change from baseline, as previously reviewed.⁹ The organ site-specific LSS change measures most associated with patient-reported response demonstrated some, but not complete, agreement with these same organ-specific NIH response findings. Finally, we acknowledge that the association of 6-month patient-reported response with subsequent FFS was primarily driven by changes in systemic IS treatment, and that this did not impact OS. The dominance of treatment change in the composite FFS outcome was well-established in prior studies.^{16,17} In chronic GvHD therapy, it represents a significant failure where additional lines of therapy lead to additional risk of infectious complications, treatment toxicity, medication costs, and associated healthcare costs.

Additional investigation is needed in larger patient populations, both for overall validation of this work, and to

further refine conclusions regarding specific organ manifestations. Following validation, patient-reported response could be analyzed in clinical trials, potentially as a secondary outcome measure to complement NIH responses. This would allow for a formal determination of benefit that could lead to approvals for treatments in the situations where NIH responses do not capture the full clinical benefit in the setting of a clinical trial. Patient-reported response could also be formally captured in routine clinical care through use of a simple patient-reported ordinal response scale to provide structured assessment above and beyond information currently shared by patients regarding their perceived treatment benefit. While inviting patients' impressions of benefit of treatments is routine in clinical practice, this structured assessment may provide insight into clinical benefit not captured through usual response measures.

This analysis has some limitations. First, this is an unplanned retrospective analysis of existing data from two parent national cohort studies. While these parent studies were rigorously designed, this analysis was facilitated using only those subjects with available baseline and 6-

month patient-reported response data. Second, we note heterogeneity in the included subjects, both regarding differing parent cohort enrollment criteria, and other factors, most notably variation in the number of lines of systemic therapy and the actual agents used to treat the chronic

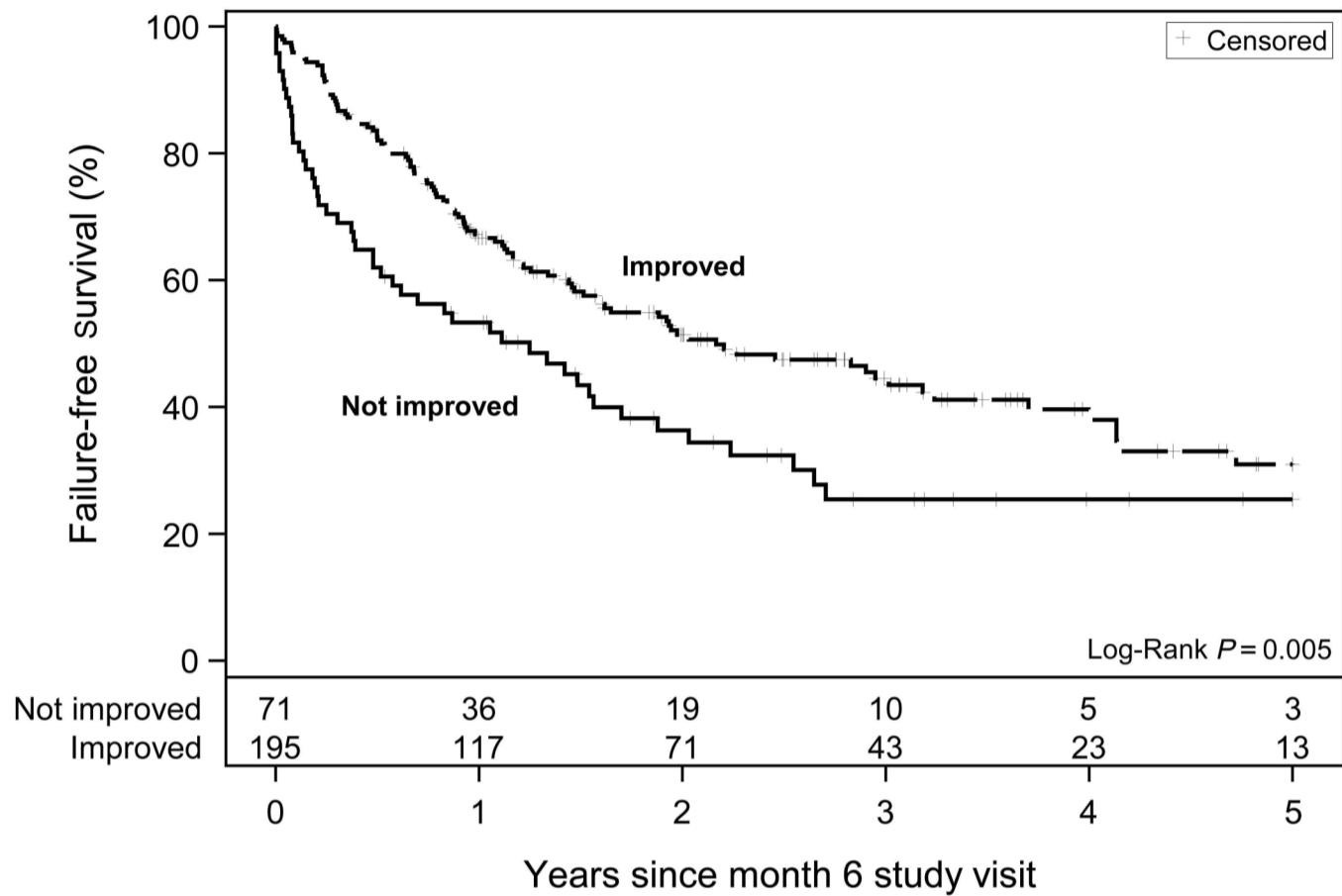


Figure 1. Failure-free survival and patient-reported response. 5-year failure-free survival by patient perception of chronic graft-versus-host disease at six months among those failure free at six months.

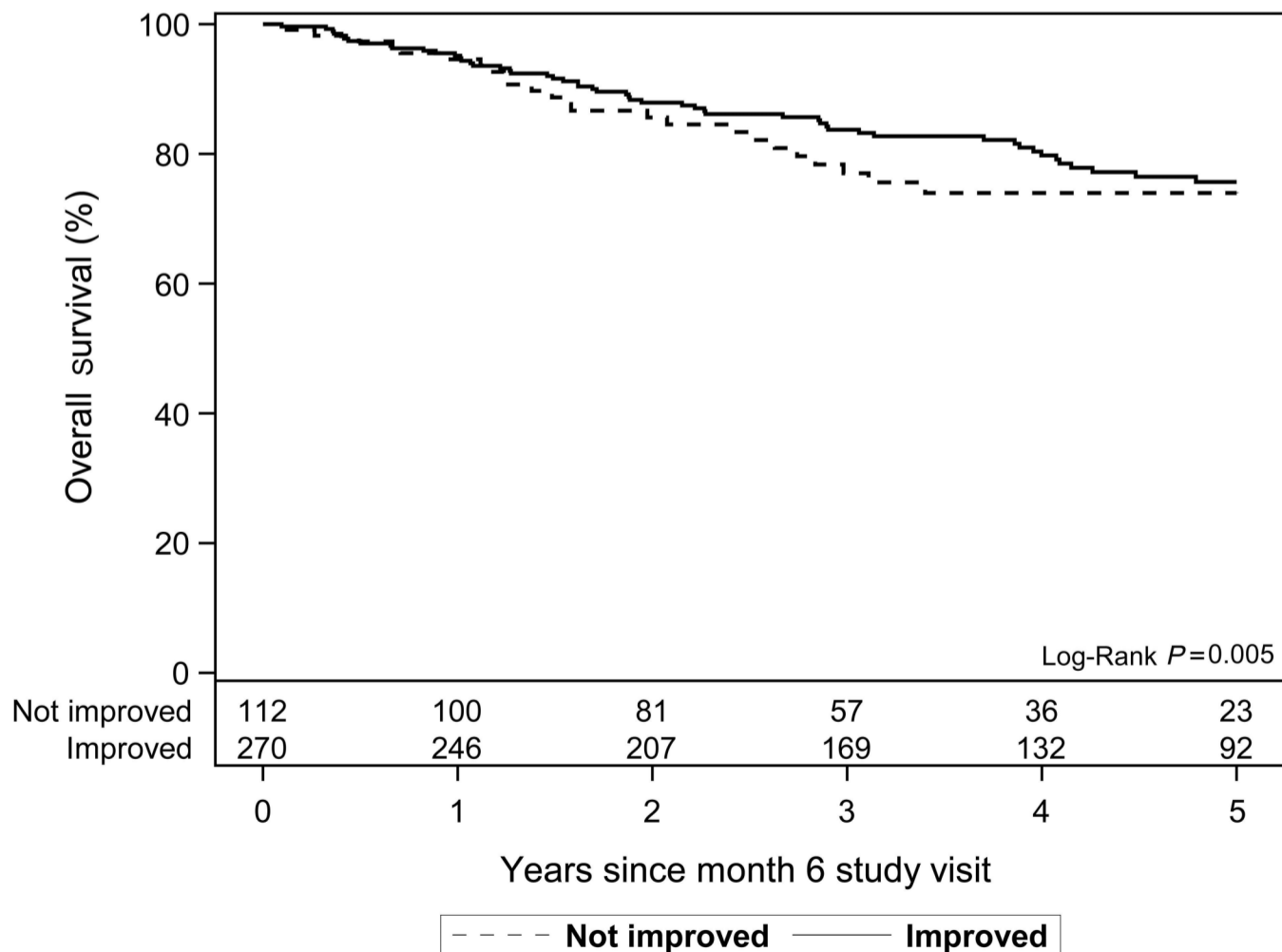


Figure 2. Overall survival and patient-reported response.

Table 4. Causes of failure in failure-free survival.

GvHD	At 1 year		At 2 years		At 3 years		At 5 years	
	Not improved	Improved	Not improved	Improved	Not improved	Improved	Not improved	Improved
New medication	42.4%	26.5%	54.2%	39.4%	65.1%	43.6%	65.1%	51.2%
Relapse	4.3%	5.6%	6.0%	6.7%	6.0%	7.8%	6.0%	9.8%
Non-relapse mortality	0.0%	1.6%	3.5%	2.8%	3.5%	4.6%	3.5%	7.8%

GvHD: graft-versus-host disease.

GvHD after cohort. Third, in the study of association between NIH responses and patient-reported response, we acknowledge certain organ sites were better represented, and that these may differ in treatment responsiveness overall. Given this, we were unable to conduct more detailed analyses according to each type of organ involvement per affected organ site.

Patient-reported responses are an important response measure that is associated with FFS in chronic GvHD and should be considered as a complementary response outcome for clinical trials and drug development in chronic GvHD.

Disclosures

IM reports research funding from Incyte, and consultancy for Abbvie and CTI Biopharma. IP sits on the advisory board for Incyte. CLK sits on the advisory board for Horizon Therapeutics. BKH sits on the advisory board for NKARTA, Equillium, Incyte, Kadmon, and Syndax, and has received speaker fees from Mallinckrodt. AMA is a consultant for Prolacta and Genentech, has received research support from Incyte, and has been a speaker for Johnson and Johnson. MEF has received research support from Pharmacyclics Inc. and Novartis/Incyte Corp., has been a Speaker Honorarium for Janssen, Pharmaceutical Companies of Johnson & Johnson, Astellas, Mallinckrodt Pharmaceuticals, and is a consultant for Pharmacyclics Inc., CSL Behring, and Fresenius Kabi. SS has received drug support, sits on the advisory board, and offers consultancy services to Rigol. PC has been involved

in research finding/consulting for Abbvie, for Incyte, and Seres. SA is a Member on the Board of Directors or advisory committees for Kadmon. CC has received consultancy/advisory honoraria from Sanofi, Equillium, CTI Biopharma, Mallinckrodt, Omeros, CSL Behring, Incyte, InhibRx, Cellarity Scientific Advisory Board with Equity: Cimeio, Oxford Algorhythmics. SL has offered consultancy services, and has received honoraria, and research funding from Kadmon, research funding from Pfizer, Syndax, Amgen, Incyte, and Novartis, and consultancy services and honoraria from Equillium, Consultancy from Mallinckrodt, and honoraria from AstraZeneca, sits on the Steering Committee of Novartis, and is a member of the Board of Directors or advisory committees of the National Marrow Donor Program. JP reports consulting and advisory board membership for Syndax, CTI Biopharma, Amgen, Regeneron, and Incyte, and clinical trial support from Novartis, Amgen, Takeda, Janssen, Johnson & Johnson, Pharmacyclics, Abbvie, CTI Biopharma, and BMS. LO, JW, NEJ and GC have no conflicts of interest to disclose.

Contributions

AI, IP, LO, SJL and JP designed and performed the research, analyzed data, and wrote the paper. All other authors contributed to the data analysis and writing of the paper.

Data-sharing statement

For original data or further information, contact the corresponding author.

References

- Lee SJ, Vogelsang G, Flowers ME. Chronic graft-versus-host disease. *Biol Blood Marrow Transplant.* 2003;9(4):215-233.
- Arai S, Arora M, Wang T, et al. Increasing incidence of chronic graft-versus-host disease in allogeneic transplantation: a report from the Center for International Blood and Marrow Transplant Research. *Biol Blood Marrow Transplant.* 2015;21(2):266-274.
- Martin PJ, Weisdorf D, Przepiorka D, et al. National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: VI. Design of Clinical Trials Working Group report. *Biol Blood Marrow Transplant.* 2006;12(5):491-505.
- Filipovich AH, Weisdorf D, Pavletic S, et al. National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: I. Diagnosis and staging working group report. *Biol Blood Marrow Transplant.* 2005;11(12):945-956.
- Pavletic SZ, Martin P, Lee SJ, 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. Response Criteria Working Group report. *Biol Blood Marrow Transplant.* 2006;12(3):252-266.

6. Lee SJ, Wolff D, Kitko C, 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. *Biol Blood Marrow Transplant.* 2015;21(6):984-999.
7. Jagasia MH, Greinix HT, Arora M, 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.
8. Palmer J, Chai X, Pidala J, et al. Predictors of survival, nonrelapse mortality, and failure-free survival in patients treated for chronic graft-versus-host disease. *Blood.* 2016;127(1):160-166.
9. Martin PJ, Storer BE, Palmer J, et al. Organ changes associated with provider-assessed responses in patients with chronic graft-versus-host disease. *Biol Blood Marrow Transplant.* 2019;25(9):1869-1874.
10. Palmer JM, Lee SJ, Chai X, et al. Poor agreement between clinician response ratings and calculated response measures in patients with chronic graft-versus-host disease. *Biol Blood Marrow Transplant.* 2012;18(11):1649-1655.
11. Inamoto Y, Martin PJ, Chai X, et al. Clinical benefit of response in chronic graft-versus-host disease. *Biol Blood Marrow Transplant.* 2012;18(10):1517-1524.
12. Pidala J, Kurland B, Chai X, et al. Patient-reported quality of life is associated with severity of chronic graft-versus-host disease as measured by NIH criteria: report on baseline data from the Chronic GVHD Consortium. *Blood.* 2011;117(17):4651-4657.
13. Chronic GC. Rationale and design of the chronic GVHD cohort study: improving outcomes assessment in chronic GVHD. *Biol Blood Marrow Transplant.* 2011;17(8):1114-1120.
14. Lee SJ, Hamilton BK, Pidala J, et al; The Chronic GVHD Consortium. Design and patient characteristics of the Chronic Graft-versus-Host Disease Response Measures Validation Study. *Biol Blood Marrow Transplant.* 2018;24(8):1727-1732.
15. Lee S, Cook EF, Soiffer R, Antin JH. Development and validation of a scale to measure symptoms of chronic graft-versus-host disease. *Biol Blood Marrow Transplant.* 2002;8(8):444-452.
16. Inamoto Y, Flowers ME, Sandmaier BM, et al. Failure-free survival after initial systemic treatment of chronic graft-versus-host disease. *Blood.* 2014;124(8):1363-1371.
17. Inamoto Y, Storer BE, Lee SJ, et al. Failure-free survival after second-line systemic treatment of chronic graft-versus-host disease. *Blood.* 2013;121(12):2340-2346.