

Patient-reported outcomes and health status associated with chronic graft-versus-host disease

Stephanie J. Lee,¹ Lynn Onstad,¹ Eric J. Chow,¹ Bronwen E. Shaw,² Heather S.L. Jim,³ Karen L. Syrjala,¹ K. Scott Baker,¹ Sarah Buckley¹ and Mary E. Flowers¹

¹Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA; ²Center for International Blood and Marrow Transplant Research, Department of Medicine, Medical College of Wisconsin, Milwaukee, WI and ³Department of Health Outcomes and Behavior, Moffitt Cancer Center, Tampa, FL, USA



Haematologica 2018
Volume 103(9):1535-1541

ABSTRACT

Chronic graft-versus-host disease occurs in 20-50% of allogeneic hematopoietic cell transplant survivors. We surveyed patients about their quality of life, symptoms, health status, comorbid conditions and medications. Instruments included the Short-Form-36 (SF-36), the Patient-Reported Outcomes Measurement Information System (PROMIS) Global and PROMIS-29 scales and the Lee Chronic Graft-versus-Host Disease Symptom Scale. Functional status was measured by self-reported Karnofsky performance status and work status. Of 3027 surveys sent to recipients surviving one or more years after transplantation, 1377 (45%) were returned. Among these, patients reported that their chronic graft-versus-host disease was mild (n=257, 18.7%), moderate (n=110, 8.0%) or severe (n=25, 1.8%). Another 377 (27.4%) had never had chronic graft-versus-host disease and 280 (20.3%) had had chronic graft-versus-host disease but it had resolved. We excluded 328 (23.8%) patients who did not answer the questions about chronic graft-versus-host disease. Patients who reported moderate or severe chronic graft-versus-host disease reported worse quality of life, lower performance status, a higher symptom burden and were more likely to be taking prescription medications for pain, anxiety and depression compared to those with resolved chronic graft-versus-host disease. Self-reported measures were similar between patients with resolved chronic graft-versus-host disease and those who had never had it. Our data suggest that the PROMIS measures may be able to replace the SF-36 in the assessment of chronic graft-versus-host disease. Between 26.7-39.4% of people with active chronic graft-versus-host disease were unable to work due to health reasons, compared with 12.1% whose chronic graft-versus-host disease had resolved and 15.4% who had never had chronic graft-versus-host disease. Mouth, eye and nutritional symptoms persisted after resolution of chronic graft-versus-host disease. These results show that better prevention of and treatment for chronic graft-versus-host disease are needed to improve survivorship after allogeneic transplantation.

Introduction

Approximately 20-50% of transplant survivors experience chronic graft-versus-host disease (GvHD), the most common late complication of allogeneic hematopoietic cell transplantation (HCT). Chronic GvHD is an iatrogenic complication of HCT that occurs when the donor's immune system attacks the recipient's tissues. The onset is typically 4-6 months after transplantation and three or more organs are involved in the majority of cases.¹ Risk factors include older age of the recipient, use of peripheral

Correspondence:

sjlee@fredhutch.org.

Received: March 7, 2018.

Accepted: May 25, 2018.

Pre-published: June 1, 2018.

doi:10.3324/haematol.2018.192930

Check the online version for the most updated information on this article, online supplements, and information on authorship & disclosures: www.haematologica.org/content/103/9/1535

©2018 Ferrata Storti Foundation

Material published in *Haematologica* is covered by copyright. All rights are reserved to the Ferrata Storti Foundation. Use of published material is allowed under the following terms and conditions:

<https://creativecommons.org/licenses/by-nc/4.0/legalcode>.

Copies of published material are allowed for personal or internal use. Sharing published material for non-commercial purposes is subject to the following conditions:

<https://creativecommons.org/licenses/by-nc/4.0/legalcode>, sect. 3. Reproducing and sharing published material for commercial purposes is not allowed without permission in writing from the publisher.



blood instead of bone marrow or cord blood, females donating to male patients, lack of *in vivo* or *ex vivo* T-cell depletion or post-transplant cyclophosphamide, and HLA mismatching.² Chronic GvHD may be mild, requiring only topical or local interventions or short-term systemic immunosuppression; it may also be moderate to severe and poorly controlled with available treatments, causing substantial morbidity or even death.³ Overall, chronic GvHD has been associated with worse quality of life (QOL) and functional status, and a higher symptom burden.⁴ Interventions that prevent chronic GvHD have been associated with a lower symptom burden and better functional status in randomized controlled trials.^{5,6} The impact of chronic GvHD on QOL as measured by the National Institutes of Health (NIH) Patient-Reported Outcomes Measurement Information System (PROMIS) instruments has not been reported⁷ and currently the PROMIS measures are not included in the list of recommended scales for chronic GvHD assessment in clinical trials.

Interest in preventing and treating chronic GvHD has increased in the last 5-10 years with a growing recognition of the impact that this complication has on the long-term health of survivors. There are better options to prevent^{6,8-12} and treat chronic GvHD,^{13,14} and an increasing number of reports about biomarkers and insights from animal and human studies on its pathophysiology.¹⁵⁻¹⁷ In August 2017, ibrutinib became the first drug approved by the Food and Drug Administration for the management of chronic GvHD that persists or progresses after treatment with steroids.¹⁸

The goal of this study was to describe the QOL scores and health status of patients with active chronic GvHD of differing severity compared to those with resolved chronic GvHD or those who had never had chronic GvHD. We also aimed to investigate the PROMIS measures in chronic GvHD relative to established measures of QOL in long-term HCT survivors.

Methods

Participants

Questionnaires are sent annually to transplant survivors as part of the Fred Hutchinson Cancer Research Center (FHCRC) long-term follow-up research program. Non-responders to the initial survey are sent a reminder letter 1 month later. All patients who underwent allogeneic transplantation from any graft source at FHCRC and who consented to allow their information to be used for research purposes were included. Patients were aged 18 or greater at the time of the survey. The analyzed data were from surveys administered by mail or online, according to the participants' choice, from July 1, 2015 to June 30, 2016. The dataset was sealed as of September 30, 2016. The Institutional Review Board of FHCRC approved the study, and all patients provided written informed consent at the time of their transplants.

Data collection instruments

The core survey includes questions about QOL, symptoms, medical complications, medications, work status and chronic GvHD. In the chronic GvHD section, patients report having never had chronic GvHD, having had chronic GvHD but it resolved, or having current mild, moderate or severe involvement. Definitions were not provided for these categories. Patients also report their work/school status, whether they have had certain medical complications, and whether they are taking medications for specified conditions. An abbreviated comorbidity score consists of self-

reported pulmonary disease, avascular necrosis, adrenal insufficiency and diabetes since these co-morbid conditions were previously associated with QOL.¹⁹ The core survey also includes the Medical Outcomes Study Short Form-36 (SF-36) health survey. The SF-36 is a 36-item multidimensional QOL instrument with population norms.^{20,21} Two summary scales, the physical component score (PCS) and mental component score (MCS), are normalized to a T score of 50 with a standard deviation of 10. Higher scores indicate better QOL. A clinically meaningful difference is 0.5 times the standard deviation, so scores <45 or >55 are considered to have clinical significance. There are also eight subscales: physical functioning, role physical functioning, emotional functioning, role emotional functioning, fatigue, social functioning, pain and general health.

Each year, a different module is added to the long-term follow-up core survey and, for this analysis, included the PROMIS Global Health 10, PROMIS-29 v2.0, and the Lee chronic GvHD symptom scale (LSS). The PROMIS Global Health 10 measure comprises ten items with two summary scores for physical and mental functioning. The PROMIS-29 contains 29 scored items and seven subscales: for the physical and social functioning scales, higher scores indicate better functioning; for fatigue, pain, anxiety, depression, and sleep scales, higher scores indicate a greater symptom burden.²² Similar to the SF-36, scores are normalized to 50 with a standard deviation of 10 with higher scores indicating better functioning, and scores greater than 0.5 times the standard deviation (i.e., <45 or >55, compared to the general population) are considered clinically meaningful.

The LSS is a 30-item measure with one summary score and seven domains: skin, mouth, eye, lung, psycho-emotional, vitality and nutrition.^{23,24} Higher scores indicate greater symptom burden. Half of the cases received the LSS with a 1-month recall period and half received a 1-week version; results were combined for analysis because there was no statistical difference between the aggregate scores for the two versions (*data not shown*). Scores may be calculated if more than half of the items in a subscale are answered. Scores range from 0-100 with higher scores indicating greater symptom burden. A difference of 6-7 points on the summary score is considered clinically meaningful but since this scale measures chronic GvHD symptoms, general population norms are not available.

Clinical and transplant variables

Clinical information retrieved from the institutional database included patient's age, sex, diagnosis, disease stage, conditioning regimen, donor type, graft source, GvHD prophylaxis, diagnosis of chronic GvHD, post-transplant relapse and vital status.

Statistical analysis

Patient, transplant and chronic GvHD characteristics and known characteristics of responders and non-responders were compared using chi-square or Fisher exact tests and *t*-tests for categorical and continuous variables, respectively. Only patients who answered the chronic GvHD question on the survey were included in this analysis because this question was used to classify whether chronic GvHD was currently active and its severity. This information is not captured in the transplant database; only the onset date and maximum severity are routinely recorded.

Patient-reported surveys were scored according to the directions of the developers, including methods for handling missing data. Multivariate linear regression models were used to examine the associations between the SF-36, PROMIS and LSS scores and potential explanatory variables including those from the medical record: patient's age and sex, conditioning intensity, donor type and graft source, diagnosis of chronic GvHD, post-transplant relapse,

time since transplant, and variables from the patient-self-reported survey: severity of chronic GVHD, and presence of any of four comorbidities (pulmonary disease, avascular necrosis, adrenal insufficiency or diabetes) previously associated with QOL.¹⁹

Work or school status was categorized as: (i) work, school or homemaking full or part time, or retired by choice; (ii) not working, not in school or retired due to health problems, and unemployed.

Because of multiple comparisons, *P*-values less than 0.01 were considered statistically significant.

Results

The survey was sent to 3027 recipients surviving one or more years after allogeneic HCT and 1377 responded (overall response rate 45%). The median age of the respondents was 54 years [interquartile range (IQR) 42-63 years] and 55.2% were male. Non-respondents were younger (mean age 48.5 versus 56.4, *P*<0.001), more likely to have received bone marrow (55.9% versus 45.7%, *P*<0.001), myeloablative conditioning (86.6% versus 78.5%, *P*<0.001), or high dose total body irradiation (44.5% versus 34.7%, *P*<0.001) and to have survived longer since their transplant (mean 15.8 versus 13.8 years, *P*<0.001) than respondents. The median time since HCT was 11 years (IQR 4-20) for all respondents whose chronic GvHD status was known and 6 years (IQR 2-13) for those with current chronic GvHD.

Table 1 shows the characteristics of the groups of patients divided according to their self-reported current chronic GvHD status, excluding the 328 who responded to the survey but did not answer the chronic GvHD questions. There were 377 who had never had chronic GvHD and 280 whose chronic GvHD had resolved by the time of the survey. Of those with current chronic GvHD, mild involvement was reported by 257, moderate involvement by 110, and severe involvement by 25. Overall, the groups were significantly different for all characteristics examined except for age. Patients who had never had chronic GvHD or whose chronic GvHD had resolved were more likely to have had matched related donors, myeloablative conditioning, high dose total body irradiation and to have survived longer since HCT. They were less likely to have received peripheral blood as the source of their graft.

Because of these cohort differences, all of these factors were included in the multivariate models. Starting in 1992, the database captures date last seen at FHCRC. Among the 1043 patients who had a date of last visit recorded, the median time since last being seen at FHCRC was 9.9 years (IQR 2.8-22.2) for patients with no chronic GvHD, 13.3 (IQR 7.0-20.9) for those whose GvHD had resolved, 1.2 (IQR 0.3-8.4) for mild, 0.8 (IQR 0.2-4.8) for moderate and 0.7 (IQR 0.2-9.5) for severe chronic GvHD.

Table 2 shows the raw scores of the different instruments for the chronic GvHD subgroups, as well as the comparisons between never versus resolved chronic GvHD and, for those with active chronic GvHD, between self-identified mild, moderate, and severe chronic GvHD. Patients whose chronic GvHD had resolved reported scores similar to those who reported never having had chronic GvHD on nearly all subscales. Exceptions were seen with the LSS in which patients with resolved chronic GvHD still reported more mouth, eye and nutritional symptoms and had higher summary scores than those never affected by chronic GvHD, although none of these differences was considered clinically meaningful (i.e., all <0.5 standard deviations). Patients with mild chronic GvHD had statistically significant worse scores than patients with resolved chronic GvHD in many domains (*data not shown*) but these differences were not considered clinically significant by virtue of the 0.5 standard deviation criterion. In contrast, among patients with current chronic GvHD, scores varied as expected between those with self-reported mild and those with moderate and severe chronic GvHD for all scales. Patients with moderate or severe chronic GvHD reported scores in multiple physical, mental and symptom domains that were both statistically and clinically worse than those of patients whose chronic GvHD had resolved, as shown by the bold values in Table 2. However, we did not detect any statistically significant differences between patients with moderate and severe chronic GvHD (*data not shown*).

A similar pattern was seen for the physical functioning scales of the SF-36 and PROMIS Global Health (Figure 1A), the mental functioning scales of the SF-36 and PROMIS Global Health (Figure 1B) and the physical and social subscales of the PROMIS-29 (Figure 1C).

Presence and severity of Chronic GvHD remained pre-

Table 1. Population characteristics.

	Never (n=377)	Resolved (n=280)	Mild (n=257)	Moderate (n=110)	Severe (n=25)	<i>P</i> -value
Female, n (%)	209 (55.4)	135 (48.2)	113 (44.0)	49 (44.5)	9 (36.0)	0.02
Age, mean years (SD)	53.9 (13.5)	58.8 (12.3)	57.2 (12.6)	58.3 (10.8)	57.6 (15.0)	<0.001
Matched related	212 (56.2)	136 (48.6)	88 (34.2)	40 (36.4)	7 (28.0)	<0.001
Mismatched related	16 (4.2)	19 (6.8)	4 (1.6)	1 (0.9)	1 (4.0)	
Haplo-identical related	11 (2.9)	7 (2.5)	10 (3.9)	2 (8.1)	1 (4.0)	
Matched unrelated	95 (25.2)	81 (28.9)	114 (44.4)	54 (49.1)	11 (44.0)	
Mismatched unrelated	17 (4.5)	33 (11.8)	31 (12.1)	13 (11.8)	4 (16.0)	
Cord blood	19 (5.0)	4 (1.4)	9 (3.5)	0	0	
Syngeneic	6 (1.6)	0	1 (0.4)	0	0	
Peripheral blood, n (%)	150 (39.8)	110 (39.3)	187 (72.8)	97 (88.2)	20 (80.0)	<0.001
Myeloablative, n (%)	309 (82.0)	242 (86.4)	172 (66.9)	64 (58.2)	14 (56.0)	<0.001
High dose TBI, n (%)	133 (35.3)	131 (46.8)	59 (23.0)	18 (16.4)	3 (12.0)	<0.001
Years since HCT, mean, (SD)	14.9 (11.3)	17.5 (8.4)	8.4 (7.5)	7.2 (5.8)	9.0 (8.1)	<0.001

TB: total body irradiation; HCT: hematopoietic cell transplant.

dictors of QOL scores even after adjustment for sex, age, conditioning intensity, total body irradiation, prior relapse, donor and graft type and an abbreviated comorbidity score. Patients with any degree of chronic GvHD reported worse health and a greater symptom burden than patients with resolved chronic GvHD (Table 3).

Table 4 shows health status according to chronic GvHD category. In general, the groups of patients differed significantly in self-reported Karnofsky performance status, organs still affected by chronic GvHD, diagnosis of osteopenia or osteoporosis and adrenal insufficiency, use of prednisone, other immunosuppressive medications and antibiotics to prevent infections, use of medications to prevent or treat osteoporosis, and prescription medication use for pain, anxiety and depression. Patients with active mild, moderate or severe chronic GvHD were less likely to work, be in school or be homemakers due to their health problems. No differences were detected in the proportions of patients with self-reported pulmonary disease, avascular necrosis or diabetes.

Discussion

The goal of this study was to describe the QOL scores and health status among patients with active chronic

GvHD of differing severity and compare them to those of patients with resolved chronic GvHD or those who had never had chronic GvHD, with an emphasis on the PROMIS measures relative to the SF-36. Patients with resolved chronic GvHD reported QOL similar to patients who had never had chronic GvHD for most instruments, although they continued to have more mouth, eye, nutritional and overall symptoms. Among patients with current chronic GvHD, moderate or severe chronic GvHD was associated with worse patient-reported outcomes than mild disease for all QOL and symptom measures, confirming that PROMIS measures can detect these differences and are comparable to established measures of the impact of chronic GvHD on QOL.

A previous 2006 study found similar QOL in patients without chronic GvHD and those whose chronic GvHD had resolved.²⁵ Our findings confirm this observation and extend it by including a number of validated multidimensional QOL instruments and other measures of health status. These observations suggest that prevention and treatment of moderate-severe chronic GvHD rather than mild disease should be targeted in clinical trials. In support of this recommendation is our finding that patients with moderate or severe chronic GvHD were much more likely to still require immunosuppressive treatment and to be taking prescription medication for pain, anxiety and depression. Our

Table 2. Mean scores (standard deviation) for summary and subscale measures.

Instrument	Subscale	Type of chronic graft-versus-host disease			P-value: never vs. resolved	Mild	Moderate	Active chronic GVHD Severe	P-value: Mild vs. moderate vs. severe	Clinically meaningful difference*
		No active chronic GVHD Never	Resolved							
N		375	282		257	110	25			
PROMIS Global Health	Physical	52.6 (9.3)	52.9 (8.9)	0.750	48.4 (8.0)	42.3 (6.8)	42.0 (8.9)	<0.001	4.7	
	Mental	52.9 (9.2)	53.0 (9.4)	0.913	49.4 (8.6)	44.7 (7.6)	42.2 (8.7)	<0.001	4.7	
PROMIS 29	Physical	50.4 (8.4)	51.1 (8.0)	0.378	47.1 (9.2)	41.7 (7.7)	40.9 (11.1)	<0.001	4.5	
	Social	55.3 (9.1)	55.7 (8.8)	0.593	51.1 (9.0)	46.3 (7.7)	46.2 (11.5)	<0.001	4.7	
	Anxiety	47.1 (8.2)	46.5 (8.0)	0.459	48.7 (8.1)	52.9 (9.0)	54.6 (9.4)	<0.001	4.2	
	Depression	46.4 (7.6)	45.3 (6.5)	0.074	47.0 (7.2)	52.2 (8.8)	51.4 (9.2)	<0.001	3.8	
	Fatigue	46.5 (10.5)	46.6 (9.7)	0.858	49.8 (10.1)	55.1 (8.3)	55.5 (10.3)	<0.001	5.2	
	Sleep	47.1 (9.1)	46.9 (8.6)	0.761	48.5 (8.3)	51.0 (8.3)	52.8 (7.2)	0.010	4.4	
	Pain	48.0 (8.4)	47.9 (8.3)	0.876	49.7 (9.0)	53.9 (8.3)	51.7 (8.9)	<0.001	4.3	
Short Form - 36	PCS	49.6 (10.1)	49.3 (10.2)	0.655	44.5 (10.8)	38.1 (10.1)	36.2 (12.1)	<0.001	5.5	
	MCS	52.9 (9.2)	53.4 (8.8)	0.516	50.7 (9.8)	47.8 (10.6)	42.9 (13.6)	<0.001	4.9	
	Physical	49.7 (9.6)	49.4 (9.5)	0.760	46.0 (10.5)	40.5 (10.6)	36.1 (13.2)	<0.001	5.3	
	Role physical	49.0 (10.5)	49.2 (10.3)	0.830	43.6 (11.9)	38.7 (11.6)	34.0 (12.6)	<0.001	5.8	
	Mental health	54.1 (8.1)	53.6 (8.6)	0.442	51.5 (9.1)	47.9 (9.3)	44.8 (11.1)	<0.001	4.5	
	Role emotional	50.2 (9.9)	50.6 (9.2)	0.616	48.0 (11.1)	44.9 (12.0)	40.2 (15.8)	0.002	5.3	
	Social	50.5 (9.5)	50.4 (9.7)	0.893	46.3 (11.3)	41.9 (11.3)	37.9 (13.0)	<0.001	5.4	
	Vitality	53.4 (10.6)	53.8 (10.0)	0.605	49.9 (10.9)	44.8 (9.4)	39.7 (12.1)	<0.001	5.5	
	Pain	52.5 (9.8)	51.8 (9.7)	0.342	49.7 (10.0)	43.8 (9.7)	46.5 (11.9)	<0.001	5.1	
	General health	49.8 (11.1)	49.8 (10.7)	0.958	43.8 (10.9)	37.3 (10.1)	35.1 (11.3)	<0.001	5.9	
Lee symptom scale	Summary	8.5 (7.4)	11.1 (8.7)	<.001	16.2 (9.6)	26.2 (11.1)	26.9 (11.6)	<0.001	5.3	
	Skin	5.4 (9.3)	6.1 (9.8)	0.387	11.1 (14.4)	20.4 (17.6)	22.3 (18.7)	<0.001	6.6	
	Mouth	2.0 (7.7)	4.5 (12.5)	0.004	13.7 (20.2)	26.0 (28.7)	12.5 (19.4)	<0.001	8.9	
	Eye	15.5 (20.4)	28.3 (27.4)	<.001	36.5 (30.0)	56.5 (30.9)	57.5 (31.5)	<0.001	14.7	
	Vitality	18.0 (17.4)	19.0 (18.0)	0.468	24.9 (17.5)	37.4 (18.8)	43.9 (22.6)	<0.001	9.5	
	Lung	3.2 (7.3)	3.6 (7.2)	0.454	4.9 (8.6)	8.3 (10.3)	13.0 (16.3)	<0.001	4.2	
	Nutrition	1.3 (3.3)	2.4 (6.2)	<.001	3.7 (6.7)	8.0 (12.3)	6.5 (12.8)	<0.001	3.5	
	Psychological	14.5 (17.8)	14.2 (17.1)	0.296	18.0 (17.5)	26.0 (21.7)	32.5 (23.6)	<0.001	9.3	

PCS: physical component score; MCS: mental component score. * 0.5 x standard deviation of the entire study population. Bold values are clinically significant, as defined by a greater than 0.5 x standard deviation difference compared to patients whose chronic GvHD had resolved.

data further suggest that current interventions are insufficient in controlling the impact of chronic GvHD on QOL.

The fact that chronic GvHD remained a significant predictor of QOL despite inclusion of comorbidities suggests that chronic GvHD is associated with additional detrimental comorbidities or causal pathways beyond those included in the analysis. In addition to organ dysfunction and symptoms directly caused by chronic GvHD, patients with chronic GvHD require medications that often have side effects, need frequent appointments for medical monitoring, have functional deficits and environmental limitations, and suffer from frequent infections. These factors can all have a negative impact on QOL.

We found that the NIH PROMIS measures seemed to perform well in distinguishing patients who had never had chronic GvHD or whose chronic GvHD had resolved from those with mild disease and those with moderate-severe disease, suggesting that they could be used in place of the SF-36. The NIH PROMIS measures have a number of advantages over other commonly used instruments in HCT. First, they are available at no cost and feature valid translations in multiple languages. Second, the scales are available in various lengths and can be combined as needed allowing significant flexibility to investigators. Third, the scales can be administered with computer-assisted technology that uses item response theory to minimize

Table 3. Multivariate model of clinical characteristics predicting patient-reported outcomes.¹

Factor	SF-36 PCS	SF-36 MCS	PROMIS-Global Physical: T score	PROMIS-Global Mental: T score	Lee symptom scale summary
Age, per decade	-1.4 (-1.9, -0.9) <0.0001	0.8 (0.3, 1.3) 0.003	-0.5 (-1.0, -0.02) 0.04	0.1 (-0.4, 0.6) 0.76	0.5 (0.05, 1.0) 0.03
Comorbidity burden, any vs. none ²	-4.5 (-5.5, -3.4) <0.0001	-1.5 (-2.6, -0.4) 0.005	-3.6 (-4.6, -2.6) <0.0001	-2.0 (-3.1, -1.0) 0.0002	2.1 (1.0, 3.1) < 0.0001
Chronic GvHD					
Never vs. resolved	-0.5 (-2.1, 1.1)	-0.4 (-1.9, 1.2)	-0.8 (-2.2, 0.7)	-0.4 (-1.9, 1.1)	-2.3 (-3.8, -0.8)
Mild vs. resolved	-4.2 (-6.0, -2.4)	-3.0 (-4.7, -1.2)	-4.3 (-6.0, -2.7)	-3.6 (-5.3, -1.9)	5.3 (3.6, 7.0)
Moderate vs. resolved	-9.6 (-11.9, -7.2)	-5.8 (-8.0, -3.6)	-9.6 (-11.7, -7.5)	-7.7 (-10.0, -5.5)	14.8 (12.6, 17.0)
Severe vs. resolved	-13.0 (-17.4, -8.7) <0.0001	-9.0 (-13.1, -4.8) <0.0001	-10.8 (-14.9, -6.7) <0.0001	-9.7 (-14.1, -5.4) <0.0001	14.9 (10.6, 19.2) < 0.0001

¹Regression coefficients (95% CI) and associated P-values adjusted for conditioning intensity, donor type, prior relapse, sex, graft type, and total body irradiation dose. ²includes pulmonary disease, avascular necrosis, adrenal insufficiency and diabetes, from patient self-reported data

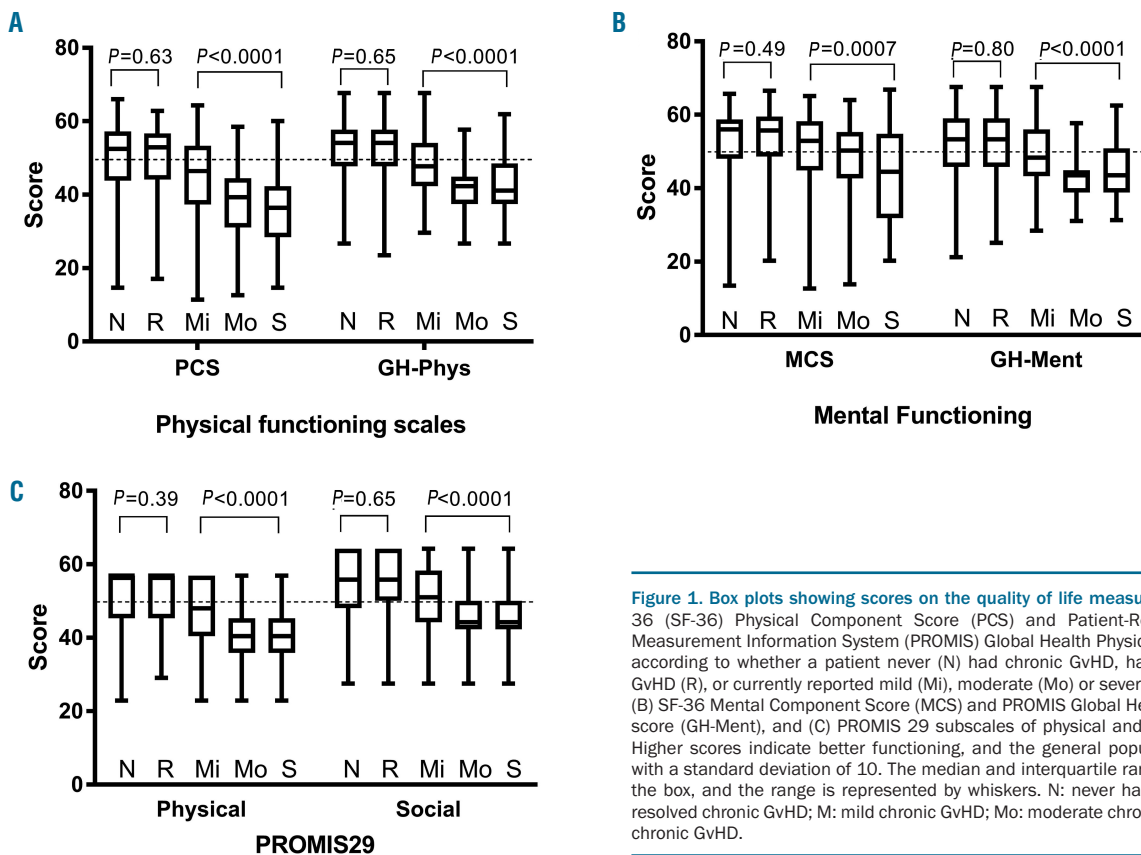


Figure 1. Box plots showing scores on the quality of life measures. (A) Short-Form 36 (SF-36) Physical Component Score (PCS) and Patient-Reported Outcomes Measurement Information System (PROMIS) Global Health Physical score (GH-Phys), according to whether a patient never (N) had chronic GvHD, had resolved chronic GvHD (R), or currently reported mild (Mi), moderate (Mo) or severe (S) chronic GvHD. (B) SF-36 Mental Component Score (MCS) and PROMIS Global Health Mental score (GH-Ment), and (C) PROMIS 29 subscales of physical and social functioning. Higher scores indicate better functioning, and the general population mean is 50 with a standard deviation of 10. The median and interquartile range are depicted by the box, and the range is represented by whiskers. N: never had chronic GvHD; R: resolved chronic GvHD; M: mild chronic GvHD; Mo: moderate chronic GvHD; S: severe chronic GvHD.

respondent burden while delivering robust, reliable QOL scores. Finally, all scores are standardized for comparisons, and there are general population normative values for all the scales. Algorithms exist to convert between other commonly utilized and validated instruments using conversion factors.²⁶

We recognize a number of caveats regarding our study. This is a single institution, cross-sectional study and

reflects the patient population and practices of one institution. The response rate was 45%, which is lower than ideal but adequate to conduct analyses given the sample size. This is a late survivorship population with a mean of 13.8 years since HCT, and only 25 participants reported severe chronic GvHD. We hypothesize that chronic GvHD may have already exacted a toll so that those with the highest morbidity died prior to the survey or did not

Table 4. Self-reported health status.¹

Organs still affected by chronic GvHD	Never		Resolved		Type of chronic GvHD				P-value ²		
	n	%	n	%	Mild	Moderate	Severe				
	n	%	n	%	n	%	n	%	n	%	
Skin	5	1.3	6	2.1	123	47.9	77	70.0	14	56.0	<0.0001
Mouth	3	0.8	4	1.4	128	49.8	63	57.3	10	40.0	<0.0001
Eyes	3	0.8	6	2.1	115	44.7	71	64.5	15	60.0	<0.0001
Intestine	4	1.1	2	0.7	34	13.2	24	21.8	3	12.0	<0.0001
Esophagus	0	--	0	--	21	8.2	7	6.4	5	20.0	<0.0001
Liver	0	--	1	0.4	24	9.3	5	4.5	3	12.0	<0.0001
Lungs	0	--	0	--	28	10.9	22	20.0	8	32.0	<0.0001
Joints	2	0.5	2	0.7	30	11.7	32	29.1	6	24.0	<0.0001
Genitals	0	--	1	0.4	16	6.2	11	10.0	2	8.0	<0.0001
Comorbidities											
Pulmonary disease	19	5.1	25	9.2	18	7.1	10	9.2	1	4.0	0.27
Avascular necrosis	19	5.2	33	12.4	24	9.7	10	9.9	3	13.6	0.03
Osteopenia/ osteoporosis	107	29.2	104	39.0	107	44.8	51	49.0	12	50.0	<0.0001
Diabetes	41	10.9	39	14.2	27	10.8	24	22.4	4	16.7	0.02
Adrenal insufficiency	5	1.4	9	3.4	18	7.5	12	11.6	0	--	<0.0001
Current medications											
Prednisone	14	3.7	9	3.3	90	35.7	64	58.7	17	68.0	<0.0001
Other immuno- suppressive drugs	23	6.1	16	5.8	91	37.0	55	53.9	14	56.0	<0.0001
Prophylactic antimicrobials	47	12.6	29	10.6	127	50.6	71	66.4	18	72.0	<0.0001
Anti-hypertensives	115	30.7	112	40.3	85	33.6	44	40.4	9	36.0	0.09
Treatment of arrhythmias	20	5.3	15	5.5	22	8.7	9	8.4	4	16.7	0.11
Anxiolytics	45	12.0	42	15.3	47	19.0	27	25.2	6	25.0	0.007
Antidepressants	46	12.2	43	15.6	48	19.0	27	25.2	8	32.0	0.002
Prevention or treatment of osteoporosis	47	12.5	40	14.7	64	25.7	31	28.7	7	29.2	<0.0001
Prescription pain medication	35	9.4	29	10.5	50	19.8	43	40.2	9	37.5	<0.0001
Work/school status											<0.0001
Work, school or homemaking full or part time, retired by choice	308	84.6	239	87.9	184	73.3	63	60.6	17	70.8	
Not working, not in school, unemployed, or retired due to health	56	15.4	33	12.1	67	26.7	41	39.4	7	29.2	
KPS (median, IQR)	100	(90-100)	100	(90-100)	90	(80-100)	80	(70-90)	70	(70-90)	<0.0001

¹Based on endorsement of each item. Missing responses and "don't know" were excluded. ²Karnofsky performance score (KPS) comparison based on the Wilcoxon rank sum test. All other factors were compared based on chi-squared or Fisher exact tests for patients across all five categories.

participate, thus potentially biasing our results and leading to an inability to differentiate moderate from severe chronic GvHD. This hypothesis is supported by a lack of difference in the frequency of well-known steroid-associated complications such as avascular necrosis and diabetes. Additional studies to show that the range of the PROMIS measures is sufficiently sensitive for patients with extreme levels of dysfunction are necessary. Nevertheless, we were able to show significantly compromised QOL and functional status for patients with moderate or severe chronic GvHD, and the PROMIS measures were able to detect differences between chronic GvHD categories. Another limitation is that the analysis is based on self-reported chronic GvHD severity since most patients did not have a recent clinical severity assessment by a transplant expert. Unfortunately we do not have concurrent clinician-assessed or NIH-classified severity information for this population. Finally, because this is a cross-sectional study, patients were analyzed according to their self-reported category at the time of the survey. A longitudinal study is needed to confirm that patients who move from moderate-severe to resolved chronic GvHD have improved QOL, ideally, similar to those who had no or mild chronic GvHD. This information would be

encouraging to patients currently being treated for chronic GvHD.

In conclusion, among patients with current chronic GvHD, moderate to severe chronic GvHD is associated with worse QOL scores whether measured by the SF-36, PROMIS Global Health, PROMIS-29 or LSS, and the SF-36 and PROMIS measures provide comparable score profiles. Prevention or better treatment of moderate to severe chronic GvHD is important in order to minimize adverse effects on QOL and health status for allogeneic transplant survivors. Patients whose chronic GvHD resolves appear similar to those who have never had chronic GvHD, except that they continue to have statistically but not clinically worse mouth, eye, nutritional and overall chronic GvHD symptoms. The freely available, concise and flexible PROMIS Global Health and PROMIS-29 measures seem to perform as well as the SF-36 in capturing the multi-dimensional QOL of these patients.²⁶ Thus any of these multi-dimensional measures could be used along with the LSS for assessing QOL in HCT clinical trials.

Acknowledgment

This work was supported by the U.S. National Institutes of Health (grants CA018029, CA215134, and CA118953).

References

- Lee SJ, Flowers ME. Recognizing and managing chronic graft-versus-host disease. *Hematology Am Soc Hematol Educ Program*. 2008;134-141.
- Flowers ME, Inamoto Y, Carpenter PA, et al. Comparative analysis of risk factors for acute graft-versus-host disease and for chronic graft-versus-host disease according to National Institutes of Health consensus criteria. *Blood*. 2011;117(11):3214-3219.
- Lee SJ. Classification systems for chronic graft-versus-host disease. *Blood*. 2017;129(1):30-37.
- 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.
- Lee SJ, Logan B, Westervelt P, et al. Comparison of patient-reported outcomes in 5-year survivors who received bone marrow vs peripheral blood unrelated donor transplantation: long-term follow-up of a randomized clinical trial. *JAMA Oncol*. 2016;2(12):1583-1589.
- Walker I, Panzarella T, Couban S, et al. Pretreatment with anti-thymocyte globulin versus no anti-thymocyte globulin in patients with haematological malignancies undergoing haemopoietic cell transplantation from unrelated donors: a randomised, controlled, open-label, phase 3, multicentre trial. *Lancet Oncol*. 2016;17(2):164-173.
- Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. *J Clin Epidemiol*. 2010;63(11):1179-1194.
- Bashey A, Zhang MJ, McCurdy SR, et al. Mobilized peripheral blood stem cells versus unstimulated bone marrow as a graft source for T-cell-replete haploidentical donor transplantation using post-transplant cyclophosphamide. *J Clin Oncol*. 2017;35(26):3002-3009.
- Bleakley M, Heimfeld S, Loeb KR, et al. Outcomes of acute leukemia patients transplanted with naive T cell-depleted stem cell grafts. *J Clin Invest*. 2015;125(7):2677-2689.
- Finke J, Bethge WA, Schmoor C, et al. Standard graft-versus-host disease prophylaxis with or without anti-T-cell globulin in haematopoietic cell transplantation from matched unrelated donors: a randomised, open-label, multicentre phase 3 trial. *Lancet Oncol*. 2009;10(9):855-864.
- Kanakry CG, Bolanos-Meade J, Kasamon YL, et al. Low immunosuppressive burden after HLA-matched related or unrelated BMT using posttransplantation cyclophosphamide. *Blood*. 2017;129(10):1389-1393.
- Kroger N, Solano C, Wolschke C, et al. Antilymphocyte globulin for prevention of chronic graft-versus-host disease. *N Engl J Med*. 2016;374(1):43-53.
- Im A, Hakim FT, Pavletic SZ. Novel targets in the treatment of chronic graft-versus-host disease. *Leukemia*. 2017;31(3):543-554.
- MacDonald KPA, Betts BC, Couriel D. Emerging therapeutics for the control of chronic graft-versus-host disease. *Biol Blood Marrow Transplant*. 2018;24(1):19-26.
- Cooke KR, Luznik L, Sarantopoulos S, et al. The biology of chronic graft-versus-host disease: a task force report from the National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease. *Biol Blood Marrow Transplant*. 2017;23(2):211-234.
- MacDonald KP, Hill GR, Blazar BR. Chronic graft-versus-host disease: biological insights from preclinical and clinical studies. *Blood*. 2017;129(1):13-21.
- Zeiser R, Blazar BR. Pathophysiology of chronic graft-versus-host disease and therapeutic targets. *N Engl J Med*. 2017;377(26):2565-2579.
- Miklos D, Cutler CS, Arora M, et al. Ibrutinib for chronic graft-versus-host disease after failure of prior therapy. *Blood*. 2017;130(21):2243-2250.
- Khera N, Storer B, Flowers ME, et al. Nonmalignant late effects and compromised functional status in survivors of hematopoietic cell transplantation. *J Clin Oncol*. 2012;30(1):71-77.
- Ware JE, Kosinski M, Keller SD. SF-36 physical and mental health summary scales: a user's manual. Boston: The Health Institute, New England Medical Center; 1994.
- Ware JE, Snow KK, Kosinski M, Gandek B. SF-36 Health Survey: a Manual and Interpretation Guide. Boston: The Health Institute, New England Medical Center; 1993.
- Jensen RE, Potosky AL, Moinpour CM, et al. United States Population-Based Estimates of Patient-Reported Outcomes Measurement Information System symptom and functional status reference values for individuals with cancer. *J Clin Oncol*. 2017;35(17):1913-1920.
- 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.
- Merkel EC, Mitchell SA, Lee SJ. Content validity of the Lee Chronic Graft-versus-Host Disease Symptom Scale as assessed by cognitive interviews. *Biol Blood Marrow Transplant*. 2016;22(4):752-758.
- Fraser CJ, Bhatia S, Ness K, et al. Impact of chronic graft-versus-host disease on the health status of hematopoietic cell transplantation survivors: a report from the Bone Marrow Transplant Survivor Study. *Blood*. 2006;108(8):2867-2873.
- Shaw BE, Syrjala KL, Onstad LE, et al. PROMIS measures can be used to assess symptoms and function in long-term hematopoietic cell transplantation survivors. *Cancer*. 2018;124(4):841-849.