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

Joseph Pidala,1 Brenda Kurland,2 Xiaoyu Chai,2 Navneet Majhail,3 Daniel J. Weisdorf,3 Steven Pavletic,4 Corey Cutler,5 David Jacobsohn,6 Jeannie Palmer,7 Sally Arai,8 Madan Jagasia,9 and Stephanie J. Lee2

1Moffitt Cancer Center, Tampa, FL; 2Fred Hutchinson Cancer Research Center, Seattle, WA; 3University of Minnesota, Minneapolis, MN; 4National Cancer Institute, Bethesda, MD; 5Dana-Farber Cancer Institute, Boston, MA; 6Children’s Memorial Hospital, Chicago, IL; 7Medical College of Wisconsin, Milwaukee, WI; 8Stanford University, Stanford, CA; and 9Vanderbilt University, Nashville, TN

Quality of life (QOL) after hematopoietic cell transplantation (HCT) is compromised by chronic GVHD. In a prospectively assembled multicenter cohort of adults with chronic GVHD (n = 298), we examined the relationship between chronic GVHD severity defined by National Institutes of Health (NIH) criteria and QOL as measured by the SF-36 and FACT-BMT instruments at time of enrollment. Chronic GVHD severity was independently associated with QOL, adjusting for age. Compared with population normative data, SF-36 scores were more than a SD (10 points) lower on average for the summary physical component score (PCS) and role-physical subscale, and significantly lower (with magnitude 4-10 points) for several other subscales. Patients with moderate and severe cGVHD had PCS scores comparable with scores reported for systemic sclerosis, systemic lupus erythematosus, and multiple sclerosis, and greater impairment compared with common chronic conditions including diabetes, hypertension, and chronic lung disease. Moderate to severe cGVHD as defined by NIH criteria is associated with significant compromise in multiple QOL domains, with PCS scores in the range of other systemic autoimmune diseases. Compromised QOL provides a functional assessment of the effects of chronic GVHD, and may be measured in cGVHD clinical studies using either the SF-36 or the FACT-BMT.

Methods

Chronic GVHD Consortium: description of study cohort and cohort for this analysis

A cohort of HCT recipients with chronic GVHD was prospectively assembled in a multicenter observational study. The protocol was approved by the Institutional Review Board at each of the 5 sites (Fred Hutchinson Cancer Research Center, University of Minnesota, Dana-Farber Cancer Institute, Stanford University, and Vanderbilt University), and all subjects provided informed consent in accordance with the Declaration of Helsinki. Patients enrolled in the cohort were allogeneic HCT recipients age 2 or greater with chronic GVHD requiring systemic immunosuppressive therapy, including both those with classic chronic GVHD and those with overlap syndrome.19 Cases were classified as incident (enrollment < 3 months after...
chronic GVHD diagnosis) or prevalent (enrollment 3 or more months but < 3 years after chronic GVHD diagnosis). Primary disease relapse, and inability to comply with study procedures were exclusion criteria. At enrollment and every 6 months thereafter, physicians and patients report standardized information on chronic GVHD organ involvement and symptoms. Incident cases had an extra assessment time point 3 months after enrollment. Chronic GVHD severity was calculated from individual organ scoring provided by clinicians using the NIH consensus scoring (mild, moderate, severe).19 Standardized chart review after each visit abstracted objective medical data (including ancillary testing and laboratory results), medical complications, and medication profiles. This analysis only examines adult (age 18 + years) patients’ QOL. Only the baseline (time of enrollment) data are analyzed for patients enrolled as of December 2009; enrollment and collection of longitudinal data are ongoing.

QOL instruments
The FACT-BMT and the SF-36 were administered to assess patient-reported QOL. The FACT-BMT Version 4.0 is a 37 item self-report questionnaire, which includes a 10 item Bone Marrow Transplant Subscale (BMTS). The instrument measures the effect of cancer therapy on multiple QOL domains including physical (PW), functional (FW), social/family, and emotional well being, and BMT-specific concerns. Individual domain scores are summarized to give a total FACT-BMT score. As well, a FACT-TOT (total outcome index) score consists of the sum of the physical and functional well being and the BMT subscale (PW + FW + BMTS).21,22

The SF-36 Version 2 is a 36-item self-report questionnaire which assesses patient-reported health and functioning. The instrument examines the following domains of QOL: physical functioning (PF), role functioning-physical (RP), bodily pain (BP), general health (GH), vitality (VT), social functioning (SF), role functioning-emotional (RE), and mental health (MH). Two summary scales from the SF-36 include the physical component score (PCS) and the mental component score (MCS).23-27

Statistical methods
Standard algorithms were used to compute total and subscale scores for FACT-BMT21 and SF-36 instruments.24,25 Graphical displays and linear correlation were used to describe the relationships between individual QOL domains within the FACT-BMT and the SF-36. Subjects’ QOL scores were displayed by chronic GVHD severity (mild, moderate, severe) according to NIH consensus criteria for global severity. In univariate analysis, the relationship between chronic GVHD severity and patient-reported QOL was examined using NIH chronic GVHD severity as the independent variable of interest. Linear-mixed models were planned to account for a random effect of study site. However, the variance of this random intercept variable of interest. Linear-mixed models were planned to account for a random effect of study site. However, the variance of this random intercept was estimated as near the boundary value (0) for all models indicating no effect of transplant center, so the results shown are for linear regression models without random effects. Linear contrasts were used to estimate pairwise differences in average QOL scores between mild/moderate, moderate/severe, and mild/severe chronic GVHD severity.

Multivariable models were constructed to determine whether GVHD severity was associated with QOL after controlling for patient and disease characteristics. Covariates considered were age at enrollment, time from HCT to enrollment, disease stage (early/intermediate/advanced), donor type (matched sibling donor vs other), conditioning regimen (myeloablative vs not), chronic GVHD status at enrollment (incident vs prevalent case), subject gender, and education level (5-level scale). Separate models were developed for each QOL composite or subscale score for both the SF-36 and FACT-BMT instruments, although each covariate carried forward to the multivariable model with severity was included in multivariable models for all subscales. Statistical interaction (effect modification) was not investigated, because no interactions were expected and any findings would likely be spurious with small subgroups created by the interaction terms.

To quantify the magnitude of impairment in QOL in chronic GVHD affected subjects, we compared SF-36 total and subscale mean scores (according to chronic GVHD severity per the NIH score) to age- and sex-adjusted US population normative mean SF-36 scores. First, individual scores were subtracted from age- and sex-specific means. These differences were evaluated using a sign test of the null hypothesis that each patient’s score is equally likely to be higher or lower than the age and gender adjusted norm. Means and 95% confidence intervals for SF-36 scores of chronic GVHD cohort subjects were compared graphically to means and 95% confidence intervals for SF-36 scores reported for selected chronic health conditions.

One aim of the observational study is to evaluate measures of QOL in chronic GVHD patients to determine which would be most useful in evaluating status and change in GVHD symptom burden in clinical trials. We therefore compared the association between QOL and severity measures through an extension of receiver operating characteristic (ROC) methods and the concordance index for an ordinal “gold standard” (chronic GVHD severity) classified by levels of a marker (QOL measures).28,29

Statistical analyses were conducted using SAS/STAT software, Version 9.2 (SAS Institute) and R Version 2.9.2 (R Foundation for Statistical Computing). To recognize the multiple tests introduced by comparing several QOL measures and pairwise comparisons among GVHD severity levels, type I error was controlled by considering a P value of 0.01 or lower as statistically significant, and by looking for consistency of results across related constructs.

Results
Chronic GVHD characteristics and baseline QOL scores
A total of 298 subjects meeting analysis criteria were enrolled between August 2007 and December 2009 at 5 centers. Enrollment at each site included Fred Hutchinson Cancer Research Center (n = 158, or 53%), Stanford (n = 48, or 16%), University of Minnesota (n = 35, or 12%), Dana-Farber Cancer Institute (n = 34, or 11%), and Vanderbilt (n = 23, or 8%). NIH severity was mild in 31 (10%), moderate in 175 (59%), and severe in 92 (31%). Median age was 53 (range 20-79). Overall the cohort was 92% white, 58% male with 57% receiving myeloablative conditioning, and 89% peripheral blood. HLA-identical siblings were the donor is 46% of cases and unrelated donors in 51% of cases. Approximately half (54%) of the cases were diagnosed within the previous 3 months. Both acute and chronic GVHD manifestations (overlap syndrome) were present in 44% of cases while 56% had only classic chronic manifestations.

Two hundred sixty subjects (87%) completed all or part of the SF-36 and FACT-BMT questionnaires. The other 38 patients (13%) were missing their entire patient questionnaires at baseline. Of those who completed the SF-36, individual items were completed 97%-100% of the time except for sexual functioning, which appeared to be missing out of answering the question if they were not sexually active and a question about concerns about keeping a job, which would not be relevant to someone retired or unemployed (see supplemental Data; available on the Blood Web site; see the Supplemental Materials link at the top of the online article). Reasons for missing surveys most commonly included patient not returning a survey despite 3 attempts to collect it (55%), patient too ill to complete (16%), and patient refusal (8%). The rate of missing questionnaires was similar for patients with mild (5/31, 16%), moderate (22/175, 13%), and severe GVHD (11/92, 12%). Enrollment SF-36 and FACT-BMT total and subscale scores are described according to NIH chronic GVHD severity in Table 1.

To characterize the relationship between QOL domains, Pearson correlation coefficients were calculated for all pairings of SF-36 and FACT-BMT subscales. In the SF-36, theoretically related constructs demonstrated high correlation (physical/role
functioning-physical, $r = 0.70$; PCS/physical, $r = 0.84$; PCS/role functioning-physical, $r = 0.81$; MCS/mental health, $r = 0.87$). Conversely, there was minimal correlation between the MCS and PCS ($r = 0.23$). For the FACT-BMT, small correlation was observed in theoretically dissimilar domains (FACT-TOL/social-family, $r = 0.46$; PWB/social-family, $r = 0.32$; PWB/emotional, $r = 0.58$; FWB/emotional, $r = 0.56$; FWB/social-family, $r = 0.44$).

Impact of chronic GVHD severity on patient-reported QOL

Of the covariates considered for multivariable analysis, only age demonstrated an association with QOL scales, and only with the SF-36 physical functioning scale. Age was included as a covariate in all analyses. However, these comparisons reflect the small sample size for mild GVHD by NIH criteria ($n = 31$), as well as smaller magnitude of effects (all differences in average QOL were $< 5$ points for mild vs moderate).

QOL composite and subscale scores were higher on average for patients with moderate GVHD compared with severe GVHD. The estimated average difference for the SF-36 role-physical was 4.17 points (95% confidence interval 1.13, 7.22), compared with a SD of 10 points for an unselected population. Statistically significant differences were also observed for the FACT total, TOI, and BMT subscale. As expected, differences in QOL between mild and severe GVHD groups were of greater magnitude than between moderate and severe. Figure 1 shows these results graphically, with average QOL scores in moderate and severe GVHD severity categories, with fewer significant differences between mild and moderate GVHD severity. However, these comparisons reflect the small sample size for mild GVHD by NIH criteria ($n = 31$), as well as smaller magnitude of effects (all differences in average QOL were $< 5$ points for mild vs moderate).

QOL composite and subscale scores were higher on average for patients with moderate GVHD compared with severe GVHD. The estimated average difference for the SF-36 role-physical was 4.17 points (95% confidence interval 1.13, 7.22), compared with a SD of 10 points for an unselected population. Statistically significant differences were also observed for the FACT total, TOI, and BMT subscale. As expected, differences in QOL between mild and severe GVHD groups were of greater magnitude than between moderate and severe. Figure 1 shows these results graphically, with average QOL scores in moderate and severe GVHD severity categories, with fewer significant differences between mild and moderate GVHD severity. However, these comparisons reflect the small sample size for mild GVHD by NIH criteria ($n = 31$), as well as smaller magnitude of effects (all differences in average QOL were $< 5$ points for mild vs moderate).

QOL composite and subscale scores were higher on average for patients with moderate GVHD compared with severe GVHD. The estimated average difference for the SF-36 role-physical was 4.17 points (95% confidence interval 1.13, 7.22), compared with a SD of 10 points for an unselected population. Statistically significant differences were also observed for the FACT total, TOI, and BMT subscale. As expected, differences in QOL between mild and severe GVHD groups were of greater magnitude than between moderate and severe. Figure 1 shows these results graphically, with average QOL scores in moderate and severe GVHD severity categories, with fewer significant differences between mild and moderate GVHD severity. However, these comparisons reflect the small sample size for mild GVHD by NIH criteria ($n = 31$), as well as smaller magnitude of effects (all differences in average QOL were $< 5$ points for mild vs moderate).

QOL composite and subscale scores were higher on average for patients with moderate GVHD compared with severe GVHD. The estimated average difference for the SF-36 role-physical was 4.17 points (95% confidence interval 1.13, 7.22), compared with a SD of 10 points for an unselected population. Statistically significant differences were also observed for the FACT total, TOI, and BMT subscale. As expected, differences in QOL between mild and severe GVHD groups were of greater magnitude than between moderate and severe. Figure 1 shows these results graphically, with average QOL scores in moderate and severe GVHD severity categories, with fewer significant differences between mild and moderate GVHD severity. However, these comparisons reflect the small sample size for mild GVHD by NIH criteria ($n = 31$), as well as smaller magnitude of effects (all differences in average QOL were $< 5$ points for mild vs moderate).

QOL composite and subscale scores were higher on average for patients with moderate GVHD compared with severe GVHD. The estimated average difference for the SF-36 role-physical was 4.17 points (95% confidence interval 1.13, 7.22), compared with a SD of 10 points for an unselected population. Statistically significant differences were also observed for the FACT total, TOI, and BMT subscale. As expected, differences in QOL between mild and severe GVHD groups were of greater magnitude than between moderate and severe. Figure 1 shows these results graphically, with average QOL scores in moderate and severe GVHD severity categories, with fewer significant differences between mild and moderate GVHD severity. However, these comparisons reflect the small sample size for mild GVHD by NIH criteria ($n = 31$), as well as smaller magnitude of effects (all differences in average QOL were $< 5$ points for mild vs moderate).

QOL composite and subscale scores were higher on average for patients with moderate GVHD compared with severe GVHD. The estimated average difference for the SF-36 role-physical was 4.17 points (95% confidence interval 1.13, 7.22), compared with a SD of 10 points for an unselected population. Statistically significant differences were also observed for the FACT total, TOI, and BMT subscale. As expected, differences in QOL between mild and severe GVHD groups were of greater magnitude than between moderate and severe. Figure 1 shows these results graphically, with average QOL scores in moderate and severe GVHD severity categories, with fewer significant differences between mild and moderate GVHD severity. However, these comparisons reflect the small sample size for mild GVHD by NIH criteria ($n = 31$), as well as smaller magnitude of effects (all differences in average QOL were $< 5$ points for mild vs moderate).

QOL composite and subscale scores were higher on average for patients with moderate GVHD compared with severe GVHD. The estimated average difference for the SF-36 role-physical was 4.17 points (95% confidence interval 1.13, 7.22), compared with a SD of 10 points for an unselected population. Statistically significant differences were also observed for the FACT total, TOI, and BMT subscale. As expected, differences in QOL between mild and severe GVHD groups were of greater magnitude than between moderate and severe. Figure 1 shows these results graphically, with average QOL scores in moderate and severe GVHD severity categories, with fewer significant differences between mild and moderate GVHD severity. However, these comparisons reflect the small sample size for mild GVHD by NIH criteria ($n = 31$), as well as smaller magnitude of effects (all differences in average QOL were $< 5$ points for mild vs moderate).

QOL composite and subscale scores were higher on average for patients with moderate GVHD compared with severe GVHD. The estimated average difference for the SF-36 role-physical was 4.17 points (95% confidence interval 1.13, 7.22), compared with a SD of 10 points for an unselected population. Statistically significant differences were also observed for the FACT total, TOI, and BMT subscale. As expected, differences in QOL between mild and severe GVHD groups were of greater magnitude than between moderate and severe. Figure 1 shows these results graphically, with average QOL scores in moderate and severe GVHD severity categories, with fewer significant differences between mild and moderate GVHD severity. However, these comparisons reflect the small sample size for mild GVHD by NIH criteria ($n = 31$), as well as smaller magnitude of effects (all differences in average QOL were $< 5$ points for mild vs moderate).
fitted average QOL and 95% confidence interval for the mean for chronic GVHD severity levels according to the NIH criteria severity score. Age is held constant at the average value of about 51 years.

Comparison to population normative data

Figure 1 also shows the population norm for SF-36 subscales (50 points), marked as a vertical line, for comparison to fitted average QOL scores. To quantify the magnitude of impairment in QOL, we compared chronic GVHD cohort members’ SF-36 mean scores to age- and gender-matched US population normative data.

Comparison to chronic health conditions

To further ascertain the clinical magnitude of QOL impairment observed in chronic GVHD cohort members, mean SF36 scores (PCS and MCS) of chronic GVHD cohort members according to NIH severity criteria were compared with those of other chronic health conditions.24,25,30-34 As demonstrated in Figure 2, those with moderate to severe chronic GVHD (rows 4 and 5) had decrement from expected population normative PCS scores comparable with that previously reported for systemic sclerosis, systemic lupus erythematosus, and multiple sclerosis (rows 6-9), but greater impairment compared with several common chronic health conditions including chronic lung disease, hypertension, diabetes and arthritis. Patients with mild or moderate chronic GVHD had MCS scores in keeping with population normative data and similar to the

Table 3. Comparison of mean SF-36 scores between chronic GVHD (cGVHD) cohort members and US population normative data

<table>
<thead>
<tr>
<th>Measure</th>
<th>Scale</th>
<th>QOL variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIH</td>
<td>PCS</td>
<td>MCS</td>
</tr>
<tr>
<td>Mild (n=31)</td>
<td>▲</td>
<td>▲</td>
</tr>
<tr>
<td>Moderate (n=175)</td>
<td>▲</td>
<td>▲</td>
</tr>
<tr>
<td>Severe (n=92)</td>
<td>▲</td>
<td>▲</td>
</tr>
<tr>
<td>NIH</td>
<td>Physical Functioning</td>
<td>Role-Physical</td>
</tr>
<tr>
<td>Mild (n=31)</td>
<td>▲</td>
<td>▲</td>
</tr>
<tr>
<td>Moderate (n=175)</td>
<td>▲</td>
<td>▲</td>
</tr>
<tr>
<td>Severe (n=92)</td>
<td>▲</td>
<td>▲</td>
</tr>
<tr>
<td>NIH</td>
<td>General Health</td>
<td>Vitality</td>
</tr>
<tr>
<td>Mild (n=31)</td>
<td>▲</td>
<td>▲</td>
</tr>
<tr>
<td>Moderate (n=175)</td>
<td>▲</td>
<td>▲</td>
</tr>
<tr>
<td>Severe (n=92)</td>
<td>▲</td>
<td>▲</td>
</tr>
<tr>
<td>NIH</td>
<td>Role-Emotional</td>
<td>Mental Health</td>
</tr>
<tr>
<td>Mild (n=31)</td>
<td>▲</td>
<td>▲</td>
</tr>
<tr>
<td>Moderate (n=175)</td>
<td>▲</td>
<td>▲</td>
</tr>
<tr>
<td>Severe (n=92)</td>
<td>▲</td>
<td>▲</td>
</tr>
</tbody>
</table>

Figure 1. Fitted average QOL values and 95% confidence intervals for a 51-year-old with mild, moderate, or severe GVHD according to NIH criteria severity. Normal population mean is 50 (vertical dotted line) for SF-36 subscales.

Table 3. Comparison of mean SF-36 scores between chronic GVHD (cGVHD) cohort members and US population normative data

<table>
<thead>
<tr>
<th>Measure</th>
<th>Norm-based scores (mean = 50, SD = 10 based on the 1998 general US population).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical functioning</td>
<td>Mean: 40.94, SD: 10.83</td>
</tr>
<tr>
<td>Role-physical</td>
<td>Mean: 36.80, SD: 11.46</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>Mean: 45.44, SD: 10.83</td>
</tr>
<tr>
<td>General health</td>
<td>Mean: 41.10, SD: 9.59</td>
</tr>
<tr>
<td>Vitality</td>
<td>Mean: 45.65, SD: 10.86</td>
</tr>
<tr>
<td>Social functioning</td>
<td>Mean: 40.28, SD: 12.30</td>
</tr>
<tr>
<td>Role-emotional</td>
<td>Mean: 44.20, SD: 12.80</td>
</tr>
<tr>
<td>Mental health</td>
<td>Mean: 49.47, SD: 9.90</td>
</tr>
<tr>
<td>PCS</td>
<td>Mean: 38.99, SD: 9.69</td>
</tr>
<tr>
<td>MCS</td>
<td>Mean: 48.05, SD: 10.59</td>
</tr>
</tbody>
</table>

Expected values for each cGVHD patient based on age and sex norm averages for the 1998 general US population. Expected physical functioning is slightly lower than 50 on average, and more variable for the patients, presumably due to associations with age.

Sign test (null hypothesis that it is equally likely that each cGVHD patient’s score will be higher or lower than the age and sex norm).

\[ P \equiv \text{Sign test (null hypothesis that it is equally likely that each cGVHD patient’s score will be higher or lower than the age and sex norm).} \]
reported chronic health conditions. Interestingly, those with severe chronic GVHD had MCS scores comparable with depression.

Comparison of FACT-BMT and SF36 instruments

The graphical displays in Figure 2 demonstrate that, while average QOL differs by GVHD severity, there is overlap of QOL values between levels of GVHD severity for all QOL measures. This graphical result is supported by the findings of the diagnostic accuracy analysis (Table 4). Taking all possible pairings of patients and excluding ties (same GVHD severity), the area under the ROC curve (AUC) is the proportion of pairs for which the QOL measure for the patient with less severe GVHD is higher than the QOL measure for the patient with more severe GVHD. The concordance index was modest (0.60) for all QOL scales examined. Using estimated variance and covariance of the AUCs to test the null hypothesis of no difference in accuracy, we conclude that there were no significant differences between QOL instruments' ability to discriminate between levels of chronic GVHD severity. Weighting schemes to allow less penalty for a 1-level difference (QOL measure for patient with moderate GVHD higher than for patient with mild GVHD, or severe higher than moderate) than for a 2-level difference (severe higher than mild) did not affect the conclusion that the QOL measures compared did not differ in performance for classifying GVHD severity.

Discussion

QOL is routinely cited by cancer patients as a concern of central importance. Chronic GVHD threatens QOL after HCT, with previous studies demonstrating moderate to large impairments in multiple domains of QOL compared with those not affected by chronic GVHD. However, the impact of chronic GVHD severity according to the proposed NIH consensus criteria on patient-reported QOL among a cohort of exclusively chronic GVHD affected HCT recipients has not been examined to date. We report the baseline QOL data of chronic GVHD affected HCT recipients at the time of enrollment in the Chronic GVHD Consortium, a multicenter, prospective observational cohort study.

Several important findings emerge from this analysis. First, we have demonstrated that chronic GVHD severity according to the NIH criteria is significantly associated with patient-reported QOL, independent of other disease, transplantation, and socio-demographic variables. This effect was observed across multiple domains of QOL, indicating a wide-reaching impact on chronic GVHD patients' reported QOL. Interestingly, none of the examined covariates, excepting the impact of age on physical functioning, was significantly associated with patient-reported QOL. Of particular relevance, we did not detect differences according to chronic GVHD status (incident vs prevalent) or time from HCT to enrollment: controlling for age at enrollment, the average PCS was estimated as only 0.3 points higher for incident than for prevalent cases, and not statistically different from no difference (P = .82). This is of importance, as the anticipated normal trajectory after

![Figure 2: Comparison of SF-36 PCS and MCS mean scores (and 95% confidence intervals for the mean) from chronic GVHD cohort members according to NIH severity score and chronic health conditions. Normal population mean is 50 (vertical dotted line).](image-url)
HCT is one of recovery and return to normal functioning. Second, we have further demonstrated the magnitude of impairment in QOL by comparison of chronic GVHD cohort subjects’ mean scores to those of age- and sex-matched US population normative data; these findings complement previously reported data. Those affected by chronic GVHD had significant impairment in QOL across multiple domains. As well, we have for the first time examined the magnitude of impairment in QOL in chronic GVHD affected individuals in reference to that previously reported in the setting of other chronic health conditions. This frames the clinical relevance of the impairment observed in the context of the NIH severity staging, and demonstrates the marked impairment in physical functioning (PCS) but relatively preserved mental health domain (MCS) in chronic GVHD–affected individuals. A potentially important exception to this was the marked impairment in MCS in those with severe chronic GVHD, which rivaled that of depression.

We have also examined the discriminative accuracy of competing QOL instruments in this analysis, with the intention of determining which is most useful in evaluating the status of chronic GVHD severity. Using an extension of ROC methods, we found no significant differences between the SF-36 and FACT-BMT in discrimination of chronic GVHD severity. This cross-sectional analysis therefore does not support the superiority of one instrument over the other, and does not assist in the selection of QOL instrument for the purpose of investigation or clinical practice in chronic GVHD. We conclude that, while physical components of self-reported QOL are lower on average for patients with more severe cGVHD, the extent of impairment and symptom burden represented by cGVHD severity are not solely captured by differences in QOL. Future analyses will evaluate sensitivity to change and may help identify the better instrument to use in this population.

Several points are worthy to note as potential limitations to these findings. First, the power to characterize mild chronic GVHD category is limited by the small sample size. Small, but meaningful, differences (eg, differences in mean scores across the mild and moderate chronic GVHD categories) may not have been detected. Closer evaluation of the mild group awaits more patient data. Next, the composition of certain sociodemographic and transplant variables are notable. In regard to patient sociodemographic variables, cohort members were disproportionately white, non-Hispanic (mild 94%, moderate 87%, severe 90% of totals), thus limiting the ability to generalize to other ethnic groups. Stem cell source used in HCT was near-uniformly peripheral blood mobilized stem cells (mild 81%, moderate 90%, and severe 90% of totals), but this reflects current US practice.

Finally, there are several relevant future directions moving beyond this analysis. Longitudinal assessment of QOL is ongoing and will provide more complete information on the duration of impairment, and trajectory of recovery or worsening. These data will permit analysis concerning the impact of resolved chronic GVHD on QOL. 2 prior studies have come to divergent conclusions regarding the durable impact of resolved chronic GVHD on QOL, but have suffered from relatively small numbers of chronic GVHD–affected subjects and methodologic limitations including the assessment of chronic GVHD activity by retrospective medical record abstraction. Longitudinal data can be used to assess whether the SF-36 or FACT-BMT are sensitive to changes in chronic GVHD severity to know whether it is realistic to expect that QOL can be used to measure treatment response.

In summary, the NIH global severity scoring system is correlated with patient-reported QOL, particularly in the physical domains, as detected by both the SF-36 and the FACT-BMT. These deficits are quite profound relative to the general population, and comparable with other chronic immune-mediated disorders.

Acknowledgment

This work was supported by National Institutes of Health/National Cancer Institute grant CA 118953-03 (PI: S.J.L.).

Authorship

Contribution: J.P. proposed the study concept, analyzed data, and wrote the manuscript; B.K. and X.C. performed statistical analyses and contributed to writing the manuscript; N.M., D.J.W., S.P., C.C., D.J., J.P., S.A., and M.J. contributed to data analysis and critical review of the manuscript; and S.J.L. contributed to the development of study concept, data analysis, and writing of the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Joseph Pidala, MD, MS, Blood and Marrow Transplantation, Moffitt Cancer Center, 12902 Magnolia Dr, FOB 3308, Tampa, FL 33612; e-mail: joseph.pidala@moffitt.org.

References


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

Joseph Pidala, Brenda Kurland, Xiaoyu Chai, Navneet Majhail, Daniel J. Weisdorf, Steven Pavletic, Corey Cutler, David Jacobsohn, Jeanne Palmer, Sally Arai, Madan Jagasia and Stephanie J. Lee

Updated information and services can be found at: http://www.bloodjournal.org/content/117/17/4651.full.html

Articles on similar topics can be found in the following Blood collections

- Clinical Trials and Observations (4525 articles)
- Free Research Articles (4463 articles)
- Transplantation (2219 articles)

Information about reproducing this article in parts or in its entirety may be found online at: http://www.bloodjournal.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at: http://www.bloodjournal.org/site/misc/rights.xhtml#reprints

Information about subscriptions and ASH membership may be found online at: http://www.bloodjournal.org/site/subscriptions/index.xhtml