Published in final edited form as:
Biol Blood Marrow Transplant. 2013 May ; 19(5): 784–791. doi:10.1016/j.bbmt.2013.02.001.

Analysis of gastrointestinal and hepatic chronic GVHD manifestations on major outcomes: A Chronic GVHD Consortium study

Joseph Pidala, MD, MS[1], **Xiaoyu Chai, MS**[2], **Brenda F. Kurland, PhD**[2], **Yoshihiro Inamoto, MD**[2], **Mary E.D. Flowers, MD**[2], **Jeanne Palmer, MD**[3], **Nandita Khera, MD**[4], **Madan Jagasia, MBBS, MS**[5], **Corey Cutler, MD, MPH**[6], **Mukta Arora, MD**[7], **Georgia Vogelsang, MD**[8], and **Stephanie J. Lee, MD, MPH**[2]

[1]H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL

[2]Fred Hutchinson Cancer Research Center, Seattle, WA

[3]Medical College of Wisconsin, Milwaukee, WI

[4]Mayo Clinic Arizona, Phoenix, AZ

[5]Vanderbilt University Medical Center, Nashville, TN

[6]Dana-Farber Cancer Institute, Boston, MA

[7]University of Minnesota Medical Center, Minneapolis, MN

[8]Johns Hopkins University, Baltimore, MD

Abstract

While data support adverse prognosis of overlap subtype of chronic GVHD, the importance of site of gastrointestinal (GI) and type of hepatic involvement is not known. Using data from the Chronic GVHD Consortium observational cohort study (n=567, total of 2115 visits), we examined whether the site of GI (esophageal, upper GI, lower GI) and type of hepatic (bilirubin, alkaline phosphatase (AP), alanine aminotransferase (ALT)) involvement are associated with overall survival (OS)and non-relapse mortality (NRM), symptoms, quality of life (QOL) and functional status measures. In multivariate analysis utilizing data from enrollment visits only, lower GI involvement (HR 1.67, p=0.05) and elevated bilirubin (HR 2.46, p=0.001) were associated with OS; both were also associated with NRM. In multivariable analysis using all visits (time-dependent covariates), GI score greater than zero (HR 1.69, p=0.02) and elevated bilirubin (HR 3.73, p<0.001) were associated with OS; results were similar for NRM. Any esophageal

Corresponding author: Joseph Pidala, MD, MS Blood and Marrow Transplantation Moffitt Cancer Center Office: 813-745-2556 Fax: 813-449-8248 Joseph.pidala@moffitt.org.

Conflict of Interest/Study Support:

Guarantor of the article: Joseph Pidala, MD, MS

Specific author contributions: JP: Developed the study concept, conducted analyses, wrote the paper
XC, BFK: Contributed to design, statistical analyses, and writing of the paper
YI, MEDF, JP, NK, MJ, CC, MA, GV: Contributed to study design, analyses, and critical review of the paper
SJL: Developed the study concept, conducted analyses, wrote the paper

Potential competing interests: None

involvement and GI score greater than zero were associated with both symptoms and QOL while elevated bilirubin was associated with QOL. We found no consistent evidence that upper GI involvement, AP, ALT, or NIH liver score add prognostic value for survival, overall symptom burden, or quality of life. These data support important differences in patient-reported outcomes according to GI and hepatic involvement among chronic GVHD affected patients, and identify those with elevated bilirubin or higher GI score at any time, or lower GI involvement at cohort enrollment, as patients at greater risk for mortality under current treatment approaches.

Keywords

chronic GVHD; gastrointestinal; hepatic

Introduction

Chronic graft-versus-host disease (GVHD) is a significant source of morbidity, mortality, impaired patient-reported quality of life (QOL), greater symptom burden, and prolonged duration of immune suppressive therapy following allogeneic hematopoietic cell transplantation (HCT).(1-10) Many,(11, 12) but not all retrospective studies,(13, 14) and prospective data from the Chronic GVHD Consortium,(15) have demonstrated that overlap subtype of chronic GVHD, defined as chronic GVHD together with concurrent acute GVHD manifestations,(16) is associated with worse prognosis and inferior patient-reported outcomes.

The proposed NIH Consensus criteria for organ-specific severity grading do not distinguish between the site of GI or hepatic involvement, but rather assign severity according to degree of weight loss or by magnitude of elevation of hepatic laboratory tests over the upper limit of normal, respectively. The impact on major outcomes of each site of gastrointestinal (esophagus, upper and lower GI) or type of hepatic (transaminases, bilirubin, alkaline phosphatase) manifestation of chronic GVHD is unknown.

We analyzed prospectively acquired observational cohort data to examine whether the site of GI involvement and type of hepatic laboratory test abnormality among chronic GVHD affected patients have association with major clinical outcomes (mortality, symptom burden, quality of life, and functional ability).

Methods

Chronic GVHD observational cohort

The Chronic GVHD Consortium is a multi-center observational cohort study of chronic GVHD-affected HCT recipients. The rationale and design of this cohort study has been previously described.(17) In brief summary, included are allogeneic HCT recipients age 2 or greater with chronic GVHD requiring systemic immunosuppressive therapy, both those with classic chronic GVHD and those with overlap subtype.(16) Cases are classified as incident (enrollment less than 3 months after chronic GVHD diagnosis) or prevalent (enrollment three or more months but less than 3 years after chronic GVHD diagnosis). Exclusion criteria include primary disease relapse, and inability to comply with study procedures.

Clinicians and patients report standardized information on chronic GVHD organ involvement and symptoms at cohort enrollment and at serial follow up visits. Chronic GVHD global severity according to the NIH Chronic GVHD Consensus is scored according to objective criteria for each organ involved, which is summarized for an overall score of mild, moderate or severe.(16) Additional measures examine the impact of chronic GVHD on

patients' functional ability, symptom burden, and QOL. The assessments performed reflect the recommendations of the NIH Consensus Conference, are described briefly in the following sections, and in the published cohort study rationale and design summary.(17)

Functional assessments

Functional measures examined in this analysis include standardized hand grip strength, and 2 minute walk test. In the assessment of grip strength, a series of three measurements are made using a portable electronic dynamometer.(18, 19) In the conduct of the 2 minute walk test, the patient is instructed to walk a 50 foot course with 180 degree turns at each end, and total distance covered is recorded.(19-21)

Patient reported outcomes

The Lee Chronic GVHD Symptom Scale is a 30 item, 7 subscale symptom scale, which evaluates adverse effects of chronic GVHD on skin, vitality, lung, nutritional status, psychological functioning, eye, and mouth symptoms.(22) The Human Activity Profile is a 94-item self reported assessment of energy expenditure and physical fitness. The instrument was first developed in a population with pulmonary disease, and has since been validated in an HCT population.(23),(24) Respondents indicate whether they never did, have stopped or are still performing the listed activities. A maximum activity score (MAS), and adjusted activity score (AAS) are calculated. The FACT-BMT v4.0 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 (PWB), functional (FWB), social/family, and emotional well being, and BMT specific concerns. Individual domain scores can be summarized to give a total FACT-BMT score (including all subscales) or a FACT-TOI (PWB + FWB + BMTS).(25, 26) The SF-36 v2 is a 36 item self-report questionnaire which assesses health and functioning. The instrument examines the following domains: 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). (27-31)

Statistical methods

Patient, transplantation, and chronic GVHD characteristics of the study subjects were summarized with descriptive statistics including median and range, or frequencies according to the nature of the data. The site of GI involvement was characterized as none, esophageal, upper GI, or lower GI either alone or in combination by the treating clinician. Biopsy confirmation was not required for diagnosis. Hepatic involvement was characterized as none, or elevation of bilirubin, alkaline phosphatase (AP), or alanine aminotransferase (ALT) over the upper limit of normal (calculated based on study site-specific laboratory reference ranges). The co-occurrence of GI or hepatic involvement was summarized and graphically represented.

At study enrollment, the association between site of GI involvement and Lee symptom scale items was examined using logistic regression, with $p < 0.05$ considered statistically significant. The type of GI and hepatic abnormalities were dichotomized as involved vs. not involved, with score > 0 as GI involvement, and score > upper limit of normal (ULN) as liver involvement. Multivariate models were constructed to examine the relationship of these variables with Lee symptom overall score, QOL measures (SF-36 PCS, SF-36 MCS, FACT-G, FACT-TOI, FACT-BMT), HAP (MAS, AAS), and functional measures (walk test, grip strength), all with $p < 0.05$ as significance level. Linear mixed models were used to account for within-patient correlation and data missing at random. Covariates adjusted in

these analyses included patient age at HCT (< 50 vs. greater), patient gender, patient education level, months from HCT to cohort enrollment (< 12 months vs. greater), donor-patient gender combination, transplant type, diagnosis, disease status, Karnofsky performance status (KPS) (< 80, ≥ 80, missing), prior history of acute GVHD, case type, platelet count at chronic GVHD onset (< 100 vs. greater), NIH global severity score, and study site.

In the study of overall survival (OS) and non-relapse mortality (NRM), multivariate models were constructed separately utilizing only cohort enrollment data, as well as all available enrollment and follow up data as time-dependent covariates, all with $p < 0.05$ significance level. Additional covariates considered included study site (FHCRC vs. others), months from HCT to cohort enrollment (< 12 months vs. greater), case type (incident vs. prevalent), platelet count (< 100K vs. greater), Karnofsky performance status (KPS) (< 80, ≥ 80, missing), patient age at HCT (< 50 vs. greater), donor match relation (matched related, matched unrelated, mismatched), donor-patient gender combination (female into male vs. others), transplant type (myeloablative vs. not) and prior history of acute GVHD (yes vs. no). NIH severity score was not included, as its scoring contains GI and hepatic severity information. Similarly, overlap subtype vs. classic chronic GVHD was not included, as this analysis aims to address mechanisms of the effect of overlap subtype on outcome.

Results

Study population

From the overall cohort study, we restricted cases for the purpose of this analysis to visit dates ≤ 12/31/2011. This analysis included 567 individual subjects. With 1548 follow up visits, data from a total of 2115 visits was utilized for this analysis. Characteristics of the study population are summarized in Table 1, and chronic GVHD characteristics and assessments are presented in Table 2. At chronic GVHD onset, KPS was < 80% in 17%, bilirubin was > 2mg/dL in 7%, and platelet count was < 100 K/uL in 23%. Overall NIH chronic GVHD severity was mild or less in 9%, moderate in 52%, and severe in 39% at enrollment. Patients with GI involvement (vs. those without) were older, and had greater proportion with Karnofsky performance status (KPS) < 80% at onset (all $p < 0.01$). Patients with hepatic involvement (compared to those without liver involvement) were less likely to have received UCB as the graft source, and had greater proportion with KPS ≥ 80% and total bilirubin >2 mg/dL at onset (all $p < 0.01$). For both comparisons of GI involvement (vs. none) and hepatic involvement (vs. none), there were no significant differences in time from HCT to onset of chronic GVHD. The sites of GI and hepatic involvement at cohort enrollment are graphically represented separately in Figure 1.

Site of GI involvement and Lee symptom scale items

At study enrollment, higher level of patient-reported difficultly swallowing solids (OR 3.02, 95% CI: 2.34-3.89, $p < 0.001$) and liquids (OR 3.91, 95% CI: 2.66-5.74, $p < 0.001$) were associated with clinician-reported esophageal involvement, and patient-reported vomiting was associated with clinician-reported upper GI involvement (OR 2.98, 95% CI: 1.94-4.58, $p < 0.001$). Higher level of patient-reported weight loss demonstrated significant association with all sites of clinician-reported GI involvement (Esophageal: OR 1.24, 95% CI: 1.02-1.51, p=0.03; Upper GI: OR 1.55, 95% CI: 1.30-1.85, $p < 0.001$; Lower GI: OR 1.40, 95% CI: 1.14-1.72, p=0.001). In addition, higher level of Lee symptom nutrition scale was associated with the involvement of esophageal (OR 1.07, 95% CI: 1.05-1.09, $p < 0.001$), upper GI (OR 1.07, 95% CI: 1.05-1.09, $p < 0.001$), and lower GI (OR 1.03, 95% CI: 1.01-1.05, $p = 0.01$).

GI/hepatic involvement and Lee overall symptom score

In multivariate analysis utilizing data from all visits, clinician-reported esophageal involvement ($p<0.001$) and overall GI (NIH score 0-3) involvement ($p=0.001$) were significantly associated with Lee overall symptom score. Lee overall symptom score was estimated to be 3.04 higher (95% CI: 1.68-4.41) for esophageal involvement, and 1.87 higher (95% CI: 0.73-3.02) for overall GI involvement, after adjusting for other significant covariates. Conversely, upper and lower GI involvement and all of the considered hepatic involvement variables did not have significant association with the Lee overall symptom score.

GI/hepatic involvement and patient-reported QOL and HAP

Table 3 summarizes data on the relationship between sites of GI and hepatic involvement and patient-reported QOL and activity. The overall GI score (NIH severity 0-3 score) had significant association with SF-36 MCS and PCS, as well as FACT-G, FACT-TOI, FACT-BMT, and HAP-MAS and HAP-AAS. Individual sites of GI involvement largely did not show significant association with these studied QOL and activity measures, except esophageal and FACT-G. Elevated bilirubin was associated with significantly worsened SF-36 MCS, FACT-G, FACT-TOI, FACT-BMT, and also HAP-MAS and AAS. Among other measures of hepatic chronic GVHD (AP, ALT, overall liver score 0-3), only AP had significant association with HAP-AAS. No other significant relationships were identified between these measures of hepatic chronic GVHD and the studied QOL outcomes.

GI/hepatic involvement and functional measures

Of the considered GI and hepatic variables, upper GI involvement was associated with an estimated of 31.7 feet less ($p < 0.001$), and overall liver involvement was associated with an estimated of 19.4 feet less ($p=0.001$) achieved in the 2 minute walk test. Only in the case of upper GI involvement did we observe significant association with grip strength ($p=0.04$).

GI/hepatic measure change: Association with clinician perception of change

In this analysis, change in each considered GI and hepatic variable was studied for its association with short-term clinician perception of change in overall chronic GVHD severity. In a multivariate model, change in lower GI ($p=0.03$), overall GI 0-3 score ($p=0.004$), and AP ($p=0.008$) were significantly associated with short-term clinician perception of change in overall chronic GVHD severity.

GI/hepatic involvement and survival

Multivariate analysis results for OS and NRM are presented in detail in Table 4 and Table 5. From these data, the following most consistent findings emerge: First, bilirubin was significantly associated with both OS and NRM based on enrollment data, as well as both OS and NRM in the time-dependent model. As well, lower GI involvement was associated with OS and NRM at cohort enrollment, and overall GI score 0-3 was associated with OS and NRM in the time-dependent model. Additional significant covariates included platelet count < 100, and KPS < 80. Graphical plots for OS and NRM stratified according to bilirubin and lower GI involvement at enrollment are presented in Figure 2.

Separate models were constructed to examine the association of GI and hepatic severity with OS and NRM, rather than according to involvement vs. not. Levels of the severity were defined according to the proposed NIH consensus criteria for organ-specific severity scoring:(16) Increasing lower GI severity at enrollment was significantly associated with overall survival (lower GI score 2/3 vs. 0: HR 2.65, 95% CI 1.24-5.66, $p = 0.01$) and non-relapse mortality (lower GI score 2/3 vs. 0: HR 4.89, 95% CI 2.11-11.34, $p < 0.001$), with

progressively increasing HR for greater severity levels. Increasing bilirubin was also associated with OS (bilirubin score 2/3 vs. 0: HR 3.48, 95% CI 1.60-7.57, p = 0.002) and NRM (bilirubin score 2/3 vs. 0: HR 4.92, 95% CI 2.04-11.85, p < 0.001). A similar trend was observed for bilirubin elevation with OS (bilirubin 2/3 vs. 0: HR 6.58, 95% CI 2.76-15.68, p < 0.001) and NRM in the time-dependent model (bilirubin 2/3 vs. 0: HR 9.13, 95% CI 3.40-24.56, p < 0.001).

As a secondary analysis approach, multivariate models were constructed to examine change in individual GI and hepatic involvement variables from cohort enrollment to 6 months as predictors of OS and NRM from a 6 month post-enrollment landmark. Weight loss (HR 1.69, 0.94-3.03, p=0.08) demonstrated increased hazard for overall mortality, but this did not reach our pre-specified significance level.

Discussion

While the presence of concurrent acute features such as GI, liver and erythematous skin involvement in the setting of chronic GVHD manifestations confers adverse prognosis, the association of the specific site of GI involvement and type of hepatic laboratory test abnormality with survival, symptom burden, quality of life, and function has not been adequately studied. We report here results of an analysis addressing this question utilizing prospectively acquired observational cohort data. These data provide important information that may guide clinical practice and inform design of clinical trials.

First, we have confirmed the relationship between clinician-reported site of GI involvement and patient-reported symptom burden. Intuitive relationships were discerned, wherein esophageal involvement was associated with difficulty swallowing, upper GI involvement with vomiting, and all sites with weight loss. All sites of GI involvement except lower GI were associated with nutrition. In the analysis utilizing all available data, esophageal and overall GI (0-3 score) were significantly associated with the Lee overall symptom scale. These data support the Lee Symptom Scale as a useful measure among chronic GVHD patients with GI involvement, and suggest that the NIH 0-3 GI severity scale is sensitive to this patient-reported outcome. The studied hepatic involvement variables had no relationship with the patient-reported overall symptom scale, suggesting that this instrument is not a useful measure of hepatic chronic GVHD activity in practice or in clinical trials. This finding mirrors clinical experience where asymptomatic patients may have very abnormal liver function tests.

Second, we report extensive data on the relationship between sites of GI and type of hepatic involvement and patient-reported QOL and functional ability. The principle finding was that overall GI 0-3 score and bilirubin elevation have strong association with patient-reported QOL, while no consistent association was detected between upper GI, lower GI, and hepatic involvement measures (AP, ALT, overall liver 0-3 score) and QOL. Overall GI 0-3 score and bilirubin also had significant association with HAP-MAS and HAP-AAS. Given their sensitivity to patient-reported QOL and functional ability as well as their relative simplicity, we recommend overall GI 0-3 score and bilirubin elevation as useful measures for clinical practice and interventions to improve or maintain patient-reported QOL among chronic GVHD affected patients.

In the study of OS and NRM, the predominant finding of this analysis was the significant association between bilirubin elevation and both OS and NRM. The association of bilirubin elevation and adverse prognosis among patients with chronic GVHD is well established, and supported by prior literature.(10, 14, 32, 33) Alternatively, we could not detect association between AP or ALT with OS and NRM in multivariate analyses. With regard to GI

involvement, lower GI (enrollment data model) and overall GI 0-3 score (time-dependent model) conferred increase hazard for mortality. Thus, these data demonstrate that those with elevated bilirubin or higher GI score at any time, or lower GI involvement at enrollment are at greater risk for mortality under current treatment approaches, and helps explain the higher risks associated with overlap subtype of chronic GVHD compared to classic chronic GVHD.

We acknowledge the following potential limitations of this analysis: First, the observed frequencies of GI and hepatic involvement reflect the characteristics of the study population, and are not a true incidence estimate among all chronic GVHD-affected patients since there may be biases at work in selection of enrolled patients. Second, while the study population is large, relative under-representation of sites of involvement may limit power to detect small but important effects. For example, the relatively infrequent co-occurrence of sites of GI and hepatic involvement limits our ability to examine the potential synergistic effect of multiple concurrent sites of involvement on outcome. Next, we acknowledge that sites and severity of GI and hepatic involvement may vary over time and with changes in intensity of immune suppressive therapy; thus, we have performed multivariate analyses using both enrollment data alone and all data in time-dependent models. Another concern is the lack of standardized chronic GVHD treatment. In this observational study, treatment was not mandated, but rather reflects usual clinical practice. Insufficient data on treatment delivered limits our ability to comment on the impact of immune suppressive therapies delivered on the studied outcomes. Greater immune suppressive therapy delivered may in part explain the observed increased mortality among patients with bilirubin elevation and greater overall GI score. Finally, we acknowledge risk for chronic GVHD misclassification, particularly in the case of hepatic laboratory test abnormalities due to medications. This problem, however, is not particular to this study, but rather true of routine clinical practice, as confirmatory hepatic biopsy is infrequently performed.

In summary, our results do not support the need to capture upper GI involvement, AP, ALT, or NIH liver score separately since they are not associated with survival, overall symptom burden, or quality of life. However, there are important differences in patient-reported outcomes according to GI and hepatic involvement among chronic GVHD affected patients. Those with elevated bilirubin or higher GI score at any time, or lower GI involvement at cohort enrollment, have a greater risk for mortality under current treatment approaches.

Acknowledgments

Grant support: This work was supported by grant CA 118953 (PI: Stephanie J. Lee)

Financial support: Supported by grant CA 118953 (PI: Stephanie J. Lee)

References

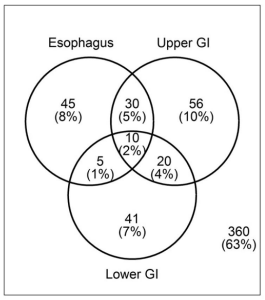
1. Arai S, Jagasia M, Storer B, et al. Global and organ-specific chronic graft-versus-host disease severity according to the 2005 NIH Consensus Criteria. Blood. 2011; 118:4242–4249. [PubMed: 21791424]
2. Fraser CJ, Bhatia S, Ness K, et al. Impact of chronic graft-versus-host disease on the health status of hematopoietic cell transplantation survivors: a report from the Bone Marrow Transplant Survivor Study. Blood. 2006; 108:2867–2873. [PubMed: 16788100]
3. Lee SJ, Flowers ME. Recognizing and managing chronic graft-versus-host disease. Hematology Am Soc Hematol Educ Program. 2008:134–141. [PubMed: 19074071]
4. Lee SJ, Vogelsang G, Flowers ME. Chronic graft-versus-host disease. Biol Blood Marrow Transplant. 2003; 9:215–233. [PubMed: 12720215]
5. Pidala J, Anasetti C, Jim H. Quality of life after allogeneic hematopoietic cell transplantation. Blood. 2009; 114:7–19. [PubMed: 19336756]

6. Pidala J, Anasetti C, Jim H. Health-related quality of life following haematopoietic cell transplantation: patient education, evaluation and intervention. Br J Haematol. 2010; 148:373–385. [PubMed: 19919651]
7. Pidala J, Kim J, Anasetti C, et al. The global severity of chronic graft-versus-host disease, determined by National Institutes of Health consensus criteria, is associated with overall survival and non-relapse mortality. Haematologica. 2011; 96(11):1678–84. [PubMed: 21791465]
8. 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. 117:4651–4657. [PubMed: 21355084]
9. Pidala J, Kurland B, Chai X, et al. Sensitivity of changes in chronic graft vs. host disease activity (measured by National Institute of Health global severity, clinician, and patient assessment) to changes in patient-reported quality of life: results from the chronic graft vs. host disease Consortium. Haematologica. 2011; 96(10):1528–35. [PubMed: 21685473]
10. Stewart BL, Storer B, Storek J, et al. Duration of immunosuppressive treatment for chronic graft-versus-host disease. Blood. 2004; 104:3501–3506. [PubMed: 15292060]
11. Arora M, Nagaraj S, Witte J, et al. New classification of chronic GVHD: added clarity from the consensus diagnoses. Bone Marrow Transplant. 2009; 43:149–153. [PubMed: 18794869]
12. Jagasia M, Giglia J, Chinratanalab W, et al. Incidence and outcome of chronic graft-versus-host disease using National Institutes of Health consensus criteria. Biol Blood Marrow Transplant. 2007; 13:1207–1215. [PubMed: 17889358]
13. Cho BS, Min CK, Eom KS, et al. Feasibility of NIH consensus criteria for chronic graft-versus-host disease. Leukemia. 2009; 23:78–84. [PubMed: 18830253]
14. Vigorito AC, Campregher PV, Storer BE, et al. Evaluation of NIH consensus criteria for classification of late acute and chronic GVHD. Blood. 2009; 114:702–708. [PubMed: 19470693]
15. Pidala J, Vogelsang G, Martin P, et al. Overlap subtype of chronic graft vs. host disease is associated with adverse prognosis, functional impairment, and inferior patient reported outcomes: a chronic graft vs. host disease Consortium study. Haematologica. 2012; 97(3):451–8. [PubMed: 22058206]
16. 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; 11:945–956. [PubMed: 16338616]
17. Rationale and Design of the Chronic GVHD Cohort Study. Improving Outcomes Assessment in Chronic GVHD. Biol Blood Marrow Transplant. 17:1114–1120.
18. Mathiowetz V, Weber K, Volland G, Kashman N. Reliability and validity of grip and pinch strength evaluations. J Hand Surg Am. 1984; 9:222–226. [PubMed: 6715829]
19. 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; 12:252–266. [PubMed: 16503494]
20. Waters RL, Lunsford BR, Perry J, Byrd R. Energy-speed relationship of walking: standard tables. J Orthop Res. 1988; 6:215–222. [PubMed: 3343627]
21. Li Li MD LCM PMH, Lynn H, Gerber MD. Validation of 2-Minute Walk Test as a Measure of Exercise Tolerance and Physical Performance in Patients With Chronic Graft Versus Host Disease. Archives of Physical Medicine and Rehabilitation. 89:e28.
22. 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; 8:444–452. [PubMed: 12234170]
23. Daughton DM, Fix AJ, Kass I, Bell CW, Patil KD. Maximum oxygen consumption and the ADAPT quality-of-life scale. Arch Phys Med Rehabil. 1982; 63:620–622. [PubMed: 7149948]
24. Herzberg PY, Heussner P, Mumm FH, et al. Validation of the Human Activity Profile Questionnaire in Patients after Allogeneic Hematopoietic Stem Cell Transplantation. Biol Blood Marrow Transplant. 2010; 16(12):1707–17. [PubMed: 20541028]
25. McQuellon RP, Russell GB, Cella DF, et al. Quality of life measurement in bone marrow transplantation: development of the Functional Assessment of Cancer Therapy-Bone Marrow

Transplant (FACT-BMT) scale. Bone marrow transplantation. 1997; 19:357–368. [PubMed: 9051246]

26. Fehse N, Fehse B, Kroger N, et al. Influence of anti-thymocyte globulin as part of the conditioning regimen on immune reconstitution following matched related bone marrow transplantation. J Hematother Stem Cell Res. 2003; 12:237–242. [PubMed: 12804182]
27. McHorney CA, Ware JE Jr. Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Medical care. 1993; 31:247–263. [PubMed: 8450681]
28. Ware, JE.; Kosinski, M.; Keller, SD. SF-36 Physical and Mental Health Summary Scales: A User's Manual. The Health Institute; Boston: 1994.
29. Ware, JE.; Snow, KK.; Kosinski, M.; Gandek, B. SF-36 Health Survey: Manual and Interpretation Guide. The Health Institute; Boston: 1993.
30. Ware, JEKM.; Dewey, J. How to Score Version 2 of the SF-36Ó Health Survey. QualityMetric Incorporated; Lincoln, RI: 2000.
31. Feinstein LC, Sandmaier BM, Hegenbart U, et al. Non-myeloablative allografting from human leucocyte antigen-identical sibling donors for treatment of acute myeloid leukaemia in first complete remission. Br J Haematol. 2003; 120:281–288. [PubMed: 12542488]
32. Arora M, Klein JP, Weisdorf DJ, et al. Chronic GVHD risk score: a Center for International Blood and Marrow Transplant Research analysis. Blood. 117:6714–6720. [PubMed: 21493797]
33. Wingard JR, Piantadosi S, Vogelsang GB, et al. Predictors of death from chronic graft-versus-host disease after bone marrow transplantation. Blood. 1989; 74:1428–1435. [PubMed: 2670000]

(a)



(b)

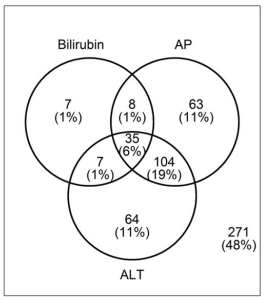


Figure 1.

Site of (a) GI and (b) hepatic involvement in study population at cohort enrollment

*GI involvement – defined by score > 0 on 0-3 NIH scale from clinician survey

*Hepatic involvement – defined by > upper limit of normal reference range according to cohort site-specific reference ranges (8 missing data)

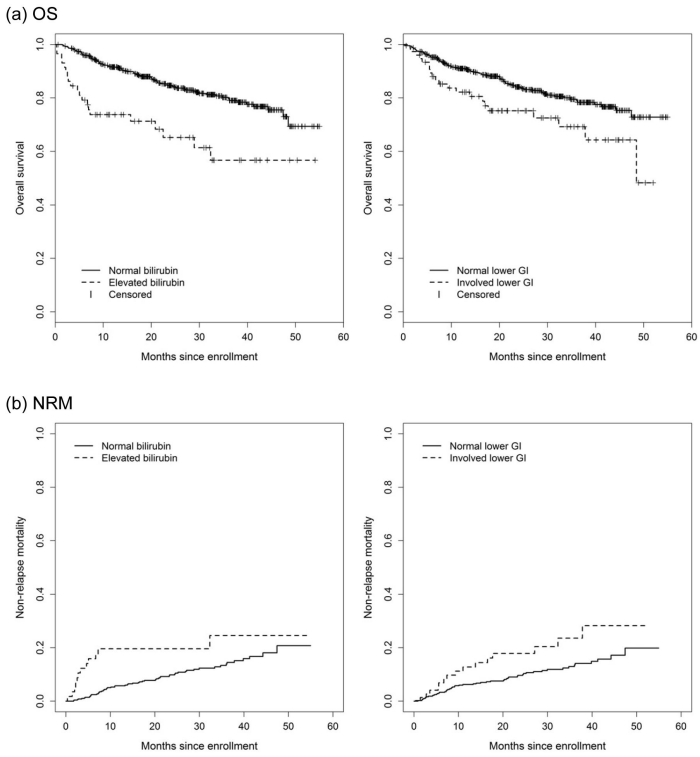


Figure 2.
Overall survival and non-relapse mortality stratified by bilirubin, and lower GI involvement at enrollment

Table 1

Summary of patient and transplantation characteristics

Characteristics	Category	n	Count (%)	Median	Min	Max
Site	Fred Hutchinson Cancer Research Center	567	247 (44%)			
	University of Minnesota		59 (10%)			
	Dana-Faber Cancer institute		65 (11%)			
	Stanford University Medical Center		72 (13%)			
	Northwest Children's Hospital		13 (2%)			
	Vanderbilt University Medical Center		47 (8%)			
	Medical College of Wisconsin		23 (4%)			
	Washington University Medical Center		4 (1%)			
	Moffitt Cancer Center		35 (6%)			
	Memorial Sloan-Kettering Cancer Center		2 (1%)			
Case type	Incident	567	336 (59%)			
	Prevalent		231 (41%)			
Adult or children	Adult (18+)	567	553 (98%)			
	Ped (2-17)		14 (2%)			
Patient age at registration (years)		567		51.0	2.0	79.0
Patient age at transplant (years)		567		50.2	1.3	78.9
Patient gender	Female	567	241 (43%)			
	Male		326 (57%)			
Patient race	Black	567	16 (3%)			
	American Indian/Alaskan Native		2 (<1%)			
	Asian		25 (4%)			
	Native Hawaiian/Pacific Islander		2 (<1%)			
	White		510 (90%)			
	Multi-race		7 (1%)			
	Unknown		5 (1%)			
Ethnicity	Hispanic	565	29 (5%)			

Characteristics	Category	n	Count (%)	Median	Min	Max
	Not Hispanic		536 (95%)			
Months from transplant to enrollment		567		11.9	2.9	294.2
Months from transplant to chronic GVHD onset		567		7.3	1.2	291
Months from chronic GVHD onset to enrollment		567		1.8	0	32.5
Diagnosis	AML	567	190 (34%)			
	ALL		66 (12%)			
	CML		29 (5%)			
	CLL		46 (8%)			
	MDS		84 (15%)			
	NHL		80 (14%)			
	HD		17 (3%)			
	MM		29 (5%)			
	AA		7 (1%)			
	Other		19 (3%)			
Disease status at transplant	Early	563	184 (33%)			
	Intermediate		241 (43%)			
	Advanced		138 (24%)			
Graft source	Bone marrow	567	38 (7%)			
	Cord blood		26 (4%)			
	Peripheral blood		503 (89%)			
Conditioning type	Myeloablative	564	326 (58%)			
	Non-myeloablative		238 (42%)			
Donor-patient CMV status	Patient and donor CMV both negative	562	188 (33%)			
	Patient or donor CMV positive		374 (67%)			
Donor-patient gender combination	Female into Male	562	164 (29%)			
	Others		398 (71%)			
Donor match	Matched related	565	240 (42%)			
	Matched unrelated		236 (42%)			

Characteristics	Category	n	Count (%)	Median	Min	Max
	Mismatched		89 (16%)			
Prior acute GVHD?	Yes	567	376 (66%)			
	No		191 (34%)			
Karnofsky performance score at onset	80+	567	348 (61%)			
	<80		95 (17%)			
	Missing		124 (22%)			

Table 2

Chronic GVHD characteristics and assessments

Clinician 0-3 GI tract score	**None**	**567**	**390 (69%)**			
	Mild		137 (24%)			
	Moderate		38 (7%)			
	Severe		2 (<1%)			
Clinician GI esophagus score	None	567	477 (84%)			
	Mild		67 (12%)			
	Moderate		13 (2%)			
	Severe		10 (2%)			
Clinician upper GI score	None	567	451 (80%)			
	Mild		76 (13%)			
	Moderate		28 (5%)			
	Severe		12 (2%)			
Clinician lower GI score	None	567	491 (87%)			
	Mild		52 (9%)			
	Moderate		19 (3%)			
	Severe		5 (1%)			
Clinician 0-3 liver score	None	563	273 (48%)			
	Mild		155 (28%)			
	Moderate		89 (16%)			
	Severe		46 (8%)			
NIH 0-3 chronic GVHD global severity score	Less than Mild	567	53 (9%)			
	Moderate		293 (52%)			
	Severe		221 (39%)			
Total serum bilirubin (mg/dL)		562		0.6	0.1	17.9
Alkaline Phosphatase (units/L)		564		96.0	0	936
ALT (units/L)		564		43.0	2.0	972
Walk test (feet)		480		500	170	1150
Grip strength (lb)		534		59.9	2.0	167
Lee symptom skin score		483		15.0	0	100
Lee symptom energy score		481		32.1	0	100
Lee symptom lung score		483		5.0	0	70.0
Lee symptom eye score		481		25.0	0	100

Clinician 0-3 GI tract score	**None**	**567**	**390 (69%)**			
Lee symptom nutrition score		481		5.0	0	70.0
Lee symptom psychological score		478		25.0	0	100
Lee symptom mouth score		483		12.5	0	100
Lee symptom overall score		483		20.3	0	65.3
FACT physical well-being score		466		22.0	1.0	28.0
FACT social/family well-being score		466		23.2	0	28.0
FACT emotional well-being score		467		19.0	4.0	24.0
FACT functional well-being score		466		16.0	2.0	28.0
FACT-BMT total score		466		27.0	10.0	40.0
FACT-BMT trial outcome index (TOI)		464		64.0	22.0	95.0
FACT-G score		461		80.0	23.0	108
FACT-BMT total score		461		106	36.0	146
SF36 physical component scale (PCS)		454		39.2	15.3	60.7
SF36 mental component scale (MCS)		454		49.8	15.3	68.4
HAP maximum activity score		466		73.0	36.0	94.0
HAP adjusted activity score		466		62.0	14.0	94.0

Table 3

Site of GI and hepatic involvement and QOL and HAP scores ("---" indicates variables that were lack of significance and dropped from multivariate models)

	SF36 PCS		SF36 MCS		FACT G		FACT TOI		FACT BMT		HAP MAS		HAP AAS	
	Estimate	**p-value**	**Estimate**	**p-value**	**Estimate**	**p-value**	**Estimate**	**p-value**	**Estimate**	**p-value**	**Estimate**	**p-value**	**Estimate**	**p-value**
Esophagus	---	---	---	---	−1.68	0.04	---	---	---	---	---	---	---	---
Upper GI	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Lower GI	---	---	---	---	---	---	---	---	---	---	---	---	---	---
GI 0-3	−1.54	0.002	−2.86	<0.001	−3.81	<0.001	−4.29	<0.001	−5.70	<0.001	−2.19	0.001	−2.10	0.004
Bili	---	---	−3.44	0.001	−2.66	0.04	−3.24	0.009	−3.76	0.03	−3.78	0.002	−4.36	0.003
ALP	---	---	---	---	---	---	---	---	---	---	---	---	−2.29	0.002
ALT	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Liver 0-3	---	---	---	---	---	---	---	---	---	---	---	---	---	---

*Data represent results from multivariate analyses using all available visit data. Covariates adjusted included: patient age at transplant (<50 vs. higher), patient gender, patient education level, month from HCT to cohort enrollment (<12 months vs. higher), donor-patient gender combination (female into male vs. other), transplant type (myeloablative vs. not), source, Karnofsky performance status at onset (<80, 80+, missing), case (incident, prevalent), platelet count at enrollment (<100K vs. higher), site (FHCRC vs. other)

Table 4

Multivariate analysis results for OS/NRM since enrollment

		OS				NRM			
Parameter	**Category**	**p-value**	**HR**	**95% HR CI**		**p-value**	**HR**	**95% HR CI**	
Lower GI	Involved	0.05	1.67	1.01	2.77	0.05	1.84	1.01	3.37
	Not involved		1.00				1.00		
Bilirubin	Involved	0.001	2.46	1.48	4.09	0.02	2.15	1.13	4.11
	Not involved		1.00				1.00		
Site	FHCRC	0.21	0.77	0.51	1.16	0.34	0.78	0.47	1.30
	Other sites		1.00				1.00		
Case type	Incident	0.37	0.80	0.49	1.31	0.62	0.86	0.48	1.56
	Prevalent		1.00				1.00		
Time from HCT to enrollment	< 12 months	0.34	1.28	0.77	2.11	0.49	0.81	0.44	1.48
	≥ 12 months		1.00				1.00		
Platelet	< 100K	0.03	1.70	1.05	2.78	0.001	2.56	1.45	4.52
	≥ 100K				1.00		1.00		
KPS	< 80	0.005	1.87	1.21	2.89	<0.001	2.77	1.59	4.83
	Missing	0.18	1.46	0.84	2.56	0.01	2.48	1.25	4.93
	≥ 80		1.00				1.00		
Age at transplant, years	≥ 50	0.72	1.08	0.71	1.66	0.78	0.93	0.55	1.57
	< 50		1.00				1.00		
Donor match	Matched unrelated	0.75	1.08	0.68	1.70	0.34	1.32	0.75	2.33
	Mismatched	0.57	1.17	0.68	2.01	0.56	1.22	0.63	2.36
	Matched related		1.00				1.00		
Donor/patient gender combination	Female donor male patients	0.73	0.93	0.60	1.43	0.51	0.83	0.48	1.44
	Others		1.00				1.00		
Conditioning type	Myeloablative	0.42	0.84	0.55	1.28	0.14	0.68	0.41	1.14
	Non-myeloablative		1.00				1.00		
Prior acute GVHD	Yes	0.66	0.91	0.60	1.39	0.57	0.86	0.51	1.45
	No		1.00				1.00		

Table 5

Multivariate analysis results for OS/NRM using all data as time-dependent covariates

		OS				NRM			
Parameter	**Category**	**p-value**	**HR**	**95% HR CI**		**p-value**	**HR**	**95% HR CI**	
NIH GI 0-3	Involved	0.02	1.69	1.10	2.60	0.02	1.89	1.13	3.15
	Not involved		1.00				1.00		
Bilirubin	Involved	<0.001	3.73	2.16	6.46	<0.001	4.44	2.32	8.52
	Not involved		1.00				1.00		
Site	FHCRC	0.13	0.72	0.47	1.10	0.27	0.74	0.44	1.26
	Other sites		1.00				1.00		
Case type	Incident	0.77	0.93	0.56	1.53	0.83	0.94	0.51	1.73
	Prevalent		1.00				1.00		
Time from HCT to enrollment	< 12 months	0.17	1.43	0.86	2.36	0.90	0.96	0.52	1.80
	≥ 12 months		1.00				1.00		
Platelet	< 100K	<0.001	2.70	1.64	4.44	<0.001	3.57	1.98	6.45
	≥ 100K		1.00				1.00		
KPS	< 80	<0.001	3.10	1.93	4.98	<0.001	3.43	1.87	6.29
	Missing	0.07	1.67	0.97	2.89	0.03	2.11	1.06	4.18
	≥ 80		1.00				1.00		
Age at transplant, years	≥ 50	0.90	1.03	0.67	1.57	0.68	0.90	0.53	1.51
	< 50		1.00				1.00		
Donor match	Matched unrelated	0.59	1.14	0.72	1.81	0.29	1.37	0.77	2.45
	Mismatched	0.27	1.36	0.79	2.35	0.33	1.39	0.71	2.72
	Matched related		1.00				1.00		
Donor gender combination	Female donor male patients	0.72	0.92	0.59	1.43	0.43	0.80	0.46	1.39
	Others		1.00				1.00		
Conditioning type	Myeloablative	0.34	0.82	0.54	1.24	0.09	0.64	0.39	1.07
	Non-myeloablative		1.00				1.00		
Prior acute GVHD	Yes	0.47	0.85	0.56	1.31	0.65	0.89	0.52	1.51
	No		1.00				1.00		