# Overlap subtype of chronic graft-*versus*-host disease is associated with an adverse prognosis, functional impairment, and inferior patient-reported outcomes: a Chronic Graft-*versus*-Host Disease Consortium study

Joseph Pidala,[1] Georgia Vogelsang,[2] Paul Martin,[3] Xiaoyu Chai,[3] Barry Storer,[3] Steven Pavletic,[4] Daniel J. Weisdorf,[5] Madan Jagasia,[6] Corey Cutler,[7] Jeanne Palmer,[8] David Jacobsohn,[9] Sally Arai,[10] and Stephanie J. Lee[3]

[1]Moffitt Cancer Center, Tampa, FL; [2]Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Baltimore, MD; [3]Fred Hutchinson Cancer Research Center, Seattle, WA; [4]National Cancer Institute, Bethesda, MD; [5]University of Minnesota, Minneapolis, MN; [6]Vanderbilt University, Nashville, TN, USA; [7]Dana Farber Cancer Institute, Boston, MA; [8]Medical College of Wisconsin, Milwaukee, WI; [9]Children's Memorial Hospital, Chicago, IL, and [10]Stanford University, Stanford, CA, USA

*Manuscript received on September 8, 2011. Revised version arrived on October 10, 2011. Manuscript accepted on October 24, 2011.*

*Correspondence: Joseph Pidala, Blood and Marrow Transplantation Moffitt Cancer Center, 12902 Magnolia Drive, FOB 3308 Tampa, Fl 33612 813-745-2556 USA. Fax: international +1.813.4498248. E-mail: joseph.pidala@moffitt.org*

## ABSTRACT

### Background

The National Institutes of Health Consensus Conference proposed the term "overlap" graft-*versus*-host disease to describe the situation when both acute and chronic graft-*versus*-host disease are present.

### Design and Methods

We examined whether the overlap subtype of graft-*versus*-host disease was associated with a different prognosis, functional limitations, or patient-reported outcomes compared to "classic" chronic graft-*versus*-host disease without any acute features.

### Results

Prospective data were collected from 427 patients from nine centers. Patients were classified as having overlap (n=352) or classic chronic (n=75) graft-*versus*-host disease based on reported organ involvement. Overlap cases had a significantly shorter median time from transplantation to cohort enrollment ($P$=0.01), were more likely to be incident cases ($P$<0.001), and had a lower platelet count at onset of the graft-*versus*-host disease ($P$<0.001). Patients with overlap graft-*versus*-host disease had significantly greater functional impairment measured by a 2-minute walk test, higher symptom burden and lower Human Activity Profile scores. Quality of life was similar, except patients with overlap graft-*versus*-host disease had worse social functioning, assessed by the Short Form-36. Multivariable analysis utilizing time-varying covariates demonstrated that the overlap subtype of graft-*versus*-host disease was associated with worse overall survival (HR 2.1, 95% CI 1.1–4.7; $P$=0.03) and higher non-relapse mortality (HR 2.8, 95% CI 1.2–8.3; $P$=0.02) than classic chronic graft-*versus*-host disease.

### Conclusions

These findings suggest that the presence of acute features in patients with chronic graft-*versus*-host disease is a marker of adverse prognosis, greater functional impairment, and higher symptom burden.

Key words: overlap subtype, graft-*versus*-host disease, GVHD, prognosis.

*Citation: Pidala J, Vogelsang G, Martin P, Chai X, Storer B, Pavletic S, Weisdorf DJ, Jagasia M, Cutler C, Palmer J, Jacobsohn D, Arai S, and Lee SJ. Overlap subtype of chronic graft-*versus*-host disease is associated with an adverse prognosis, functional impairment, and inferior patient-reported outcomes: a Chronic Graft-*versus*-Host Disease Consortium study. Haematologica 2012; 97(3)451-458. doi:10.3324/haematol.2011.055186*

## Introduction

Chronic graft-*versus*-host disease (GVHD) is a major late complication of allogeneic hematopoietic cell transplantation associated with transplant-related morbidity and mortality, prolonged immunosuppressive therapy, and impaired quality of life.[1-10] A 2005 National Institutes of Health (NIH) Consensus Conference discarded the previous definition of chronic GVHD based on time after hematopoietic cell transplantation (≥100 days) in favor of diagnostic criteria based on clinical manifestations.[11] According to these proposed NIH criteria, acute GVHD manifestations (erythematous or maculopapular rash, nausea and vomiting or diarrhea, and cholestatic hepatitis) occurring more than 100 days after hematopoietic cell transplantation are classified as "persistent", "recurrent", or "late onset" acute GVHD depending on the antecedent history of acute GVHD and absence of other chronic GVHD manifestations. "Classic chronic" GVHD is defined by diagnostic manifestations of chronic GVHD without characteristic features of acute GVHD, and an "overlap" subtype of chronic GVHD is defined by simultaneous features of both chronic and acute GVHD. The proposed overlap subtype of chronic GVHD was, however, based on consensus opinion without evidence to support its prognostic relevance or practical utility. Since the publication of these proposed criteria, retrospective analyses have not consistently demonstrated that transplantation outcomes differ significantly between patients with classic chronic GVHD and those with overlap GVHD.[7,12-14]

In order to address whether the overlap subtype of chronic GVHD has any clinical, functional or prognostic significance, we analyzed prospectively collected data from a multicenter observational cohort study.[15] Our hypothesis was that identification of differences would support continued recognition of the overlap subtype of chronic GVHD and efforts to distinguish it from late acute and classic chronic GVHD.

## Design and Methods

### *Chronic Graft-versus-Host Disease Consortium cohort*

The Chronic GVHD Consortium is a prospectively assembled, multi-center observational cohort study of hematopoietic cell transplant recipients with chronic GVHD. This protocol was approved by the Institutional Review Board at each study site. All subjects signed informed consent. Patients included were allogeneic hematopoietic cell transplant recipients aged 2 or older with chronic GVHD requiring systemic immunosuppressive therapy, including both those with classic chronic GVHD and those with overlap syndrome.[11] Cases were classified as incident (enrollment less than 3 months after the diagnosis of chronic GVHD) or prevalent (enrollment 3 or more months but less than 3 years after the diagnosis of chronic GVHD). Exclusion criteria were primary disease relapse, and inability to comply with study procedures. The rationale and design of this observational cohort study have been described in more detail elsewhere.[15] Importantly, enrolled patients do not represent the consecutive series of all chronic GVHD-affected patients at the participating Consortium centers.

Clinicians and patients reported standardized information on chronic GVHD organ involvement and symptoms. Chronic GVHD severity according to the NIH Chronic GVHD Consensus was scored using objective criteria for each organ involved, and summarized to give an overall score of mild, moderate or severe.[11] Beyond these measures of chronic GVHD activity, the impact of chronic GVHD on patients' functional ability, symptom burden, and quality of life was comprehensively assessed in this cohort study. The battery of assessments performed reflect the recommendations of the NIH Consensus Conference, and are described in the following sections, as well as in the previously published rationale and design summary of the cohort study.[15] One of the ongoing aims of the Consortium is to define the most parsimonious battery of assessments necessary for chronic GVHD evaluation.

### *Functional assessments*

Functional measures included standardized hand grip strength, portable spirometry measurement of forced expiratory volume in one second ($FEV_1$) or $FEV_1$ measured in formal pulmonary function testing, and the 2-minute walk test. In the assessment of grip strength, a series of three measurements were made using a portable electronic dynamometer.[16,17] $FEV_1$ was assessed using a portable spirometer: three measures of $FEV_1$ were recorded after the patient had been instructed to take a deep breath and then exhale forcefully and rapidly. Formal pulmonary function test results were used preferentially if available. In the 2-minute walk test, the patient was instructed to walk a 50-foot course with 180 degree turns at each end, and total distance covered was recorded.[17-19]

### *Patient-reported outcomes*

The Chronic GVHD Symptom Scale is a 30-item, 7-subscale symptom scale, which evaluates adverse effects of chronic GVHD on skin, vitality, lungs, nutritional status, psychological functioning, eyes, and mouth symptoms.[20] The Human Activity Profile (HAP) 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 hematopoietic cell transplant recipients.[21,22] Respondents indicate whether they have stopped or are still performing activities. A maximum activity score and adjusted activity score are calculated; the adjusted activity score is determined by counting how many activities with lower values than the maximum activity score the respondent has "stopped doing" and subtracting this from the maximum activity score. The Functional Assessment of Cancer Therapy – Bone Marrow Transplantation (FACT-BMT) version 4.0 is a 37-item self-report questionnaire, which includes a ten-item bone marrow transplant subscale. The instrument measures the effect of cancer therapy on multiple quality of life domains including physical well-being, functional well-being, social/family, and emotional well being, and specific bone marrow transplantation concerns. Individual domain scores can be summarized to give a total FACT-BMT score (including all subscales) or a Functional Assessment of Cancer Therapy – Trial Outcome Index (FACT-TOI; physical well-being + functional well-being + bone marrow transplant subscale).[23,24] The Short Form-36 (SF-36) version 2 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 are the physical component score (PCS) and the mental component score (MCS).[25-29]

### *Statistical methods*

Standard algorithms were used to compute summary and subscale scores for the Lee symptom scale,[20] FACT-BMT,[23] SF-36,[26,27] and HAP instruments.[22] Overlap and classic chronic GVHD were classified by the proposed NIH criteria.[11] Acute GVHD manifesta-

tions were defined by three organs: skin involvement by erythematous or maculopapular rash, gastrointestinal manifestations of nausea and vomiting or diarrhea, or liver involvement with abnormal liver function tests. Skin involvement was indicated by an NIH erythema score greater than 0, or Vienna Skin Score grade 3/4 erythema fraction greater than 0.

Patients' socio-demographics, transplantation characteristics, functional test results, clinician-assessed GVHD severity in organs, and patient-reported outcomes were compared in patients with overlap syndrome or classic chronic GVHD at enrollment. Descriptive statistics are presented as medians and ranges for continuous variables, and as frequencies and percentages for categorical variables. Statistical comparisons between groups were made with the two-sample t-test for continuous variables, and the $\chi^2$ test for categorical variables. A type I error was controlled by considering a $P$ value of 0.01 or lower as statistically significant.

Overall survival was calculated as months between enrollment and death, with patients censored at the date they were last known to be alive. Non-relapse mortality was defined as death prior to relapse. Relapsed patients were censored at the relapse date as a competing event, and those without relapse or death were censored at date they were last known to be alive. Survival was estimated by the Kaplan-Meier method. The cumulative incidence of non-relapse mortality was estimated by standard methods.

We first utilized Cox regression to compare overall survival and non-relapse mortality based on enrollment designation of overlap syndrome *versus* classic chronic GVHD. Adjusted hazard ratios were estimated with adjustment for time from transplant to study enrollment, platelet count at enrollment (<100×10$^9$/L *versus* higher), Karnofsky performance status at enrollment (< 80, 80+, missing), patients' age at transplant (<50 years *versus* higher), donor-recipient match and relation (matched related, matched unrelated, mismatched), donor-patient gender combination (female into male *versus* other), prior history of acute GVHD (yes *versus* no), and calculated NIH consensus GVHD global severity score (none/mild, moderate, severe). Hazard ratios were calculated with classic chronic GVHD patients at enrollment as the reference. The contributions of the involvement of each acutely affected organ and organ combinations to overall survival and non-relapse mortality were also evaluated using Cox regression analysis. Acute organ involvement was categorized as skin only, gastrointestinal tract only, liver only, skin and gastrointestinal tract, skin and liver, gastrointestinal tract and liver, and skin, gastrointestinal tract and liver.

We then performed a separate analysis using Cox regression modeling with time-varying covariates. The following variables were considered time-dependent: platelet count (<100×10$^9$/L *versus* higher), Karnofsky performance status (< 80, 80+, missing), NIH calculated severity score (none/mild, moderate, severe), and chronic GVHD subtype (overlap *versus* classic). Missing time-dependent variables were imputed from the 'last observation carry forward' method.

Statistical analyses were conducted using SAS/STAT software, version 9.2 (SAS Institute, Inc., Cary, NC, USA) and R version 2.9.2 (R Foundation for Statistical Computing, Vienna, Austria).

## Results

### Characteristics of overlap and classic chronic graft-versus-host disease

Data from 427 participants were included in this analysis representing those patients enrolled in the cohort study up to September, 2010 from nine centers. The majority (82%, n=352) had overlap syndrome. Among the 352 subjects with overlap syndrome, skin erythema or maculopapular rash was present in 189 (54%), gastrointestinal involvement in 129 (37%), and liver involvement in 216 (62%). In addition, 198 (56%) met criteria for acute features in only one organ, 126 (36%) for acute features in two organs, and 28 (8%) for acute features in all three organs. Only 18% (n=75) had classic chronic GVHD without any acute features.

The subjects were predominantly adults. The median age at enrollment was 51 years (range, 2-79) for patients with overlap syndrome and 49 years (range, 17-69) for the patients with classic chronic GVHD. Baseline socio-demographic, disease and transplantation variables according to overlap *versus* classic status are presented in Table 1.

Patients with overlap syndrome were more likely to be incident cases (59% *versus* 35%, $P$<0.001), to have a shorter median time from transplantation to enrollment (11.9 months *versus* 19.8 months, $P$=0.01), and to have lower platelet counts at the onset of chronic GVHD ($P$<0.001). Neither clinician-reported nor patient-reported global severity of GVDH was associated with overlap designation. Involvement of joints and fascia was more common in classic chronic GVHD. The distribution of mouth, eye, lung and genital involvement did not differ significantly between overlap and classic chronic GVHD (Table 1).

### Association between overlap chronic graft-versus-host disease and morbidity

Patients with overlap syndrome performed significantly less well on the 2-minute walk test. In addition the $FEV_1$ was lower in patients with overlap syndrome (Ttable 2). Differences in patient-reported quality of life between those with overlap syndrome and classic chronic GVHD were minimal, although those with overlap syndrome had inferior SF-36 social functioning scores (5 points lower, in keeping with a clinically meaningful 0.5 standard deviation difference). Patients with overlap syndrome and classic chronic GVHD had similar chronic GVHD symptom scale item and summary scores except that nutrition and skin symptom scale scores were significantly worse in those with overlap syndrome, as expected. HAP maximum and adjusted activity scores and modified maximum and adjusted activity scores were significantly inferior for those with overlap GVHD (Table 2).

### Association between overlap chronic graft-versus-host disease and mortality – analysis according to cohort enrollment data

Of the 427 enrolled patients, 413 had follow-up beyond the baseline visit and could be evaluated for survival outcomes. The median follow-up time after enrollment for patients with overlap GVHD was 17.1 months, while that for patients with classic chronic GVHD was 19.0 months. Figures 1 and 2 illustrate overall survival and non-relapse mortality according to classic GVHD *versus* overlap GVHD status at enrollment. The 2-year overall survival estimate for patients with overlap syndrome was 77% *versus* 94% for those with classic GVHD. The 2-year non-relapse mortality estimates were 20% and 4% for those with overlap GVHD and classic GVHD, respectively. Causes of death of patients with overlap and classic chronic GVHD were, respectively, GVHD-related (n=27, n=0), relapse (n=18, n=0), unknown (n=3, n=1), and other (n=14, n=3). After

**Table 1. Baseline characteristics of patients with overlap syndrome *versus* classic chronic GVHD.**

| Characteristics | Overlap (N = 352) N. | Count | Percent | Classic (N = 75) N. | Count | Percent | *P* value* |
|---|---|---|---|---|---|---|---|
| Site | 352 | | | 75 | | | 0.84 |
| Fred Hutchinson Cancer Research Center | | 164 | 47 | | 39 | 52 | |
| University of Minnesota | | 41 | 12 | | 8 | 11 | |
| Dana-Faber Cancer institute | | 42 | 12 | | 9 | 12 | |
| Stanford University Medical Center | | 50 | 14 | | 8 | 11 | |
| Northwestern Childrens Hospital | | 12 | 3 | | 1 | 1 | |
| Vanderbilt University Medical Center | | 25 | 7 | | 8 | 11 | |
| Medical College of Wisconsin | | 12 | 3 | | 1 | 1 | |
| Washington University Medical Center | | 2 | 1 | | 0 | 0 | |
| Moffitt Cancer Center | | 4 | 1 | | 1 | 1 | |
| Case type | 352 | | | 75 | | | <0.001 |
| Incident | | 206 | 59 | | 26 | 35 | |
| Prevalent | | 146 | 41 | | 49 | 65 | |
| Adults | 352 | 339 | 96 | 75 | 74 | 99 | 0.30 |
| Male | 352 | 200 | 57 | 75 | 48 | 64 | 0.25 |
| Race | 352 | | | 75 | | | 0.82 |
| White | | 316 | 90 | | 68 | 91 | |
| Non-white | | 36 | 10 | | 7 | 9 | |
| Hispanic | 352 | 19 | 5 | 75 | 1 | 1 | 0.16 |
| Diagnosis at transplant | 352 | | | 75 | | | 0.89 |
| Acute myeloid leukemia | | 114 | 32 | | 28 | 38 | |
| Acute lymphoblastic leukemia | | 45 | 13 | | 9 | 12 | |
| Chronic myeloid leukemia | | 16 | 5 | | 4 | 5 | |
| Chronic lymphocytic leukemia | | 26 | 7 | | 4 | 5 | |
| Myelodysplastic syndrome | | 56 | 16 | | 7 | 10 | |
| Non-Hodgkin's lymphoma | | 52 | 15 | | 11 | 15 | |
| Hodgkin lymphoma | | 12 | 3 | | 4 | 5 | |
| Multiple myeloma | | 17 | 5 | | 4 | 5 | |
| Aplastic anemia or other | | 14 | 4 | | 4 | 5 | |
| Disease status | 350 | | | 74 | | | 0.94 |
| Early | | 114 | 33 | | 25 | 34 | |
| Intermediate | | 154 | 44 | | 33 | 44 | |
| Advanced | | 82 | 23 | | 16 | 22 | |
| Transplant source | 352 | | | 75 | | | 0.73 |
| Peripheral blood | | 307 | 87 | | 67 | 90 | |
| Bone marrow | | 28 | 8 | | 4 | 5 | |
| Cord blood | | 17 | 5 | | 4 | 5 | |
| Transplant type | 351 | | | 75 | | | 0.62 |
| Myeloablative | | 195 | 56 | | 44 | 59 | |
| Non-myeloablative | | 156 | 44 | | 31 | 41 | |
| Patient CMV positive | 351 | 200 | 57 | 73 | 43 | 59 | 0.76 |
| Donor CMV positive | 348 | 137 | 39 | 73 | 29 | 40 | 0.95 |
| Donor match | 352 | | | 74 | | | 0.13 |
| Matched related | | 150 | 43 | | 39 | 53 | |
| Mismatched | | 63 | 18 | | 7 | 9 | |
| Matched unrelated | | 139 | 39 | | 28 | 38 | |
| Donor gender | 348 | | | 74 | | | 0.82 |
| Female into male | | 99 | 28 | | 22 | 30 | |
| Other | | 249 | 72 | | 52 | 70 | |
| Had prior acute GVHD | 352 | 235 | 67 | 75 | 46 | 61 | 0.37 |
| NIH consensus severity score | 352 | | | 75 | | | 0.36 |
| None | | 0 | 0 | | 2 | 3 | |
| Mild | | 36 | 10 | | 7 | 9 | |
| Moderate | | 197 | 56 | | 47 | 63 | |
| Severe | | 119 | 34 | | 19 | 25 | |
| Clinician-assessed sev. score | 352 | | | 75 | | | 0.19 |
| Mild | | 166 | 47 | | 43 | 58 | |
| Moderate | | 150 | 43 | | 28 | 37 | |
| Severe | | 36 | 10 | | 4 | 5 | |
| Patient-reported severity score | 304 | | | 60 | | | 0.30 |
| None | | 9 | 3 | | 1 | 2 | |
| Mild | | 159 | 52 | | 33 | 55 | |
| Moderate | | 108 | 36 | | 24 | 40 | |
| Severe | | 28 | 9 | | 2 | 3 | |
| Skin score | 352 | | | 75 | | | NA** |
| None | | 118 | 34 | | 42 | 56 | |
| Mild | | 96 | 27 | | 7 | 9 | |
| Moderate | | 85 | 24 | | 19 | 26 | |
| Severe | | 53 | 15 | | 7 | 9 | |
| Mouth score | 352 | | | 75 | | | 0.02 |
| None | | 134 | 38 | | 31 | 41 | |
| Mild | | 161 | 45 | | 41 | 55 | |
| Moderate | | 48 | 14 | | 3 | 4 | |
| Severe | | 9 | 3 | | 0 | 0 | |
| GI tract score | 352 | | | 75 | | | NA** |
| None | | 223 | 63 | | 75 | 100 | |
| Mild | | 102 | 29 | | 0 | 0 | |
| Moderate | | 25 | 7 | | 0 | 0 | |
| Severe | | 2 | 1 | | 0 | 0 | |
| Eye score | 352 | | | 75 | | | 0.46 |
| None | | 182 | 51 | | 38 | 50 | |
| Mild | | 101 | 29 | | 26 | 35 | |
| Moderate | | 60 | 17 | | 9 | 12 | |
| Severe | | 9 | 3 | | 2 | 3 | |
| Joints and fascia score | 352 | | | 75 | | | <0.001 |
| None | | 262 | 75 | | 37 | 49 | |
| Mild | | 64 | 18 | | 23 | 31 | |
| Moderate | | 22 | 6 | | 14 | 19 | |
| Severe | | 4 | 1 | | 1 | 1 | |
| Lung score | 352 | | | 75 | | | 0.51 |
| None | | 177 | 50 | | 36 | 48 | |
| Mild | | 119 | 34 | | 30 | 40 | |
| Moderate | | 48 | 14 | | 8 | 11 | |
| Severe | | 8 | 2 | | 1 | 1 | |
| Liver score | 350 | | | 75 | | | NA** |
| None | | 134 | 38 | | 75 | 100 | |
| Mild | | 161 | 46 | | 0 | 0 | |
| Moderate | | 53 | 15 | | 0 | 0 | |
| Severe | | 2 | 1 | | 0 | 0 | |
| Genital score | 314 | | | 71 | | | 0.12 |
| None | | 282 | 90 | | 58 | 82 | |
| Mild | | 18 | 6 | | 6 | 8 | |
| Moderate | | 10 | 3 | | 5 | 7 | |
| Severe | | 4 | 1 | | 2 | 3 | |
| Platelet count ($\times 10^9$/L) at enrollment | 344 | | | 73 | | | <0.001 |
| $< 100 \times 10^9$/L | | 92 | 27 | | 5 | 7 | |
| $\geq 100 \times 10^9$/L | | 252 | 73 | | 68 | 93 | |
| Karnofsky performance status at enrollment | 306 | | | 62 | | | 0.06 |
| < 80 | | 118 | 39 | | 16 | 26 | |
| ≥ 80 | | 188 | 61 | | 46 | 74 | |
| Months from transplant to enrollment | | | | | | | |
| Median (range) | 352 | 11.9 (2.9-61.7) | | 75 | 19.8 (3.1-299.1) | | 0.01 |

*continued in next column*

**To get valid Chi-square tests, none and mild groups were collapsed together for NIH consensus and patient-reported severity scores; moderate and severe groups were collapsed together for individual organ scores. **Comparisons for these organs are not meaningful because they are used to define overlap syndrome. CMV, cytomegalovirus; HLA, human leukocyte antigen; GI: gastrointestinal.*

adjusting for known chronic GVHD risk factors, the overlap subtype of chronic GVHD at enrollment was associated with inferior overall survival (HR 2.8, 95% CI 1.1–9.3; *P*=0.03) and higher non-relapse mortality (HR 4.4, 95% CI 1.3–27.2; *P*=0.01). Among the covariates included in these models, platelet count at enrollment and NIH global chronic GVHD severity at enrollment were also significantly associated with overall survival and non-relapse mortality.

We explored whether the higher non-relapse mortality and lower overall survival observed in the patients with the overlap subtype of GVHD were associated with particular combinations of acute features. A total of 424 patients were included in this analysis, after excluding three patients with missing organ scores. Compared with classic chronic GVHD, statistically significant differences were found among all organ involvement combinations of overlap except when skin and gastrointestinal tract were involved together, and when liver and gastrointestinal tract were involved together. The involvement of all three organs was associated with the greatest hazard for overall survival (HR 6.8, 95% CI 2.1–22.8; *P*=0.002) and non-relapse mortality (HR 11.9, 95% CI 2.5–57.6; *P*=0.002). Table 3 shows the overall survival analysis. Results of the non-relapse mortality analysis were similar and are not shown.

**Table 2.** Functional characteristics of overlap *versus* classic chronic GVHD.

| Characteristics | Overlap (N = 352) | | | Classic (N = 75) | | | *P*-value |
|---|---|---|---|---|---|---|---|
| | n | Median | Range | n | Median | Range | |
| Walk test (feet) | 298 | 495 | 170-1150 | 69 | 540 | 300-1140 | <0.001 |
| Grip strength (lb) | 337 | 59.7 | 2.0-167.0 | 73 | 63.3 | 15.3-139.0 | 0.43 |
| Portable spirometer $FEV_1$ (L/sec) | 316 | 2.4 | 0.2-5.6 | 68 | 2.7 | 0.2-5.2 | 0.02 |
| SF36-physical functioning | 292 | 42.3 | 14.9-57.0 | 60 | 44.4 | 17.0-57.0 | 0.07 |
| SF36-role physical | 294 | 37.3 | 17.7-56.9 | 61 | 37.3 | 17.7-56.9 | 0.02 |
| SF36-bodily pain | 294 | 46.1 | 19.9-62.1 | 62 | 43.9 | 24.9-62.1 | 0.20 |
| SF36-general health | 291 | 40.1 | 18.6-63.9 | 62 | 41.7 | 16.2-63.9 | 0.52 |
| SF36-vitality | 295 | 45.8 | 20.9-70.8 | 62 | 47.4 | 24.0-70.8 | 0.13 |
| SF36-social functioning | 295 | 40.5 | 13.2-56.8 | 62 | 45.9 | 13.2-56.8 | 0.01 |
| SF36-role emotional | 293 | 48.1 | 9.2-55.9 | 61 | 48.1 | 9.2-55.9 | 0.18 |
| SF36-mental health | 295 | 50.0 | 13.4-64.1 | 62 | 51.4 | 16.2-64.1 | 0.75 |
| SF36-physical component scale | 286 | 39.3 | 15.5-59.4 | 60 | 40.1 | 16.6-60.7 | 0.25 |
| SF36-mental component scale | 286 | 49.8 | 15.3-65.8 | 60 | 53.0 | 25-68.4 | 0.25 |
| FACT-physical | 294 | 22 | 1-28 | 62 | 23 | 6-28 | 0.39 |
| FACT-social/family | 293 | 23 | 0-28 | 62 | 24 | 6-28 | 0.37 |
| FACT-emotional | 293 | 20 | 6-24 | 62 | 19 | 4-24 | 0.20 |
| FACT-functional | 293 | 16 | 2-28 | 62 | 18 | 4-27 | 0.12 |
| FACT-BMT subscale | 293 | 27 | 11-39 | 62 | 28 | 11.1-40 | 0.23 |
| FACT-BMT trial outcome index | 293 | 64.7 | 22-94 | 62 | 65.7 | 25-93 | 0.18 |
| FACT-G | 291 | 79 | 29-108 | 62 | 81.5 | 23-101 | 0.37 |
| FACT-BMT total score | 291 | 106 | 49-146 | 62 | 108 | 36-138 | 0.30 |
| Symptom scale-skin | 308 | 15 | 0-100 | 63 | 10 | 0-80 | 0.005 |
| Symptom scale-energy | 307 | 32.1 | 0-100 | 63 | 32.1 | 3.6-82.1 | 0.66 |
| Symptom scale-lung | 308 | 5 | 0-70 | 63 | 5 | 0-40 | 0.47 |
| Symptom scale-eye | 307 | 25 | 0-100 | 63 | 25 | 0-100 | 0.93 |
| Symptom scale-nutrition | 307 | 5 | 0-70 | 63 | 0 | 0-20 | <0.001 |
| Symptom scale-psychological | 305 | 16.7 | 0-83.3 | 63 | 25 | 0-100 | 0.16 |
| Symptom scale-mouth | 308 | 12.5 | 0-100 | 63 | 0 | 0-100 | 0.09 |
| Symptom scale summary | 308 | 20.7 | 0-65.3 | 63 | 18.1 | 4.1-56.6 | 0.25 |
| HAP-maximum activity score | 293 | 70 | 36-94 | 62 | 78 | 52-94 | <0.001 |
| HAP-adjusted activity score | 293 | 59 | 14-94 | 62 | 70 | 30-94 | <0.001 |
| Modified HAP-maximum activity score | 293 | 52 | 19-74 | 62 | 58 | 38-74 | <0.001 |
| Modified HAP-adjusted activity score | 293 | 46 | 11-74 | 62 | 54 | 22-74 | <0.001 |

### *Association between overlap chronic graft-versus-host disease and mortality – time-dependent analysis*

Acknowledging that overlap and classic chronic GVHD are not fixed categories and that disease manifestations vary with time and intensity of immune suppressive therapy, we performed an analysis in which chronic GVHD subtype (overlap *versus* classic) was considered a time-varying covariate. The overlap subtype of GVHD was associated with worse overall survival (HR 2.1, 95% CI 1.1–4.7; *P*=0.03) and higher non-relapse mortality (HR 2.8, 95% CI 1.2–8.3; *P*=0.02) than classic chronic GVHD. Among other covariates considered in the overall survival analysis, platelet count less than $100\times10^9$/L (HR 3.4, 95% CI 1.9–6; *P*<0.001), Karnofsky performance status less than 80 (HR 2.4, 95% CI 1.4–4.1; *P*=0.002), and severe NIH global severity score (HR 5.1, 95% CI 1.5–32.1, *P*=0.03) were associated with overall survival. In the analysis of non-relapse mortality, platelet count less than $100\times10^9$/L (HR 3.1, 95% CI 1.6–5.9, *P*<0.001), and Karnofsky performance status less than 80 (HR 2.2, 95% CI 1.2–4.2, *P*=0.01) were associated with non-relapse mortality.

## Discussion

From the 2005 NIH Consensus Development Project on GVHD, a proposed overlap subtype of chronic GVHD emerged. This recommendation was controversial, with some doubting whether recognition of the overlap subtype of GVHD had practical clinical utility or any association with meaningful transplant outcomes. Others pointed out that prior studies in chronic GVHD had identified extensive skin involvement,[30-32] elevated bilirubin,[10,33] gastrointestinal tract involvement,[32] and progressive onset from acute GVHD[30,33,34] as poor prognostic findings. Studies using retrospective review of medical records failed to settle the controversy. Our findings based on prospectively collected data support the recognition of the overlap subtype of chronic GVHD as an important condition with an adverse prognosis, functional impairment, and significantly higher symptom burden.

Recognition of the overlap subtype of chronic GVHD appears to have several implications: as patients with

acute features have significantly higher non-relapse mortality and lower overall survival, they represent a vulnerable group who may benefit from different or more aggressive therapeutic interventions. The optimal management of these patients is not known, but they are expected to have high non-relapse mortality and poor overall survival under current practices.

We found that patients with overlap syndrome suffer significant and diverse functional impairments compared to those with classic chronic GVHD, including significantly worse performance on the 2-minute walk test, and lower self-reported HAP scores. These data suggest a systemic functional impairment beyond the more direct ramifications of concurrent acute GVHD manifestations such as skin erythema or maculopapular rash, gastrointestinal tract and liver involvement. The similar incidence of prior acute GVHD in the overlap and classic chronic GVHD groups argues that this functional impairment is not due to extended or more intensive immunosuppressive treatment for acute GVHD in the overlap group.

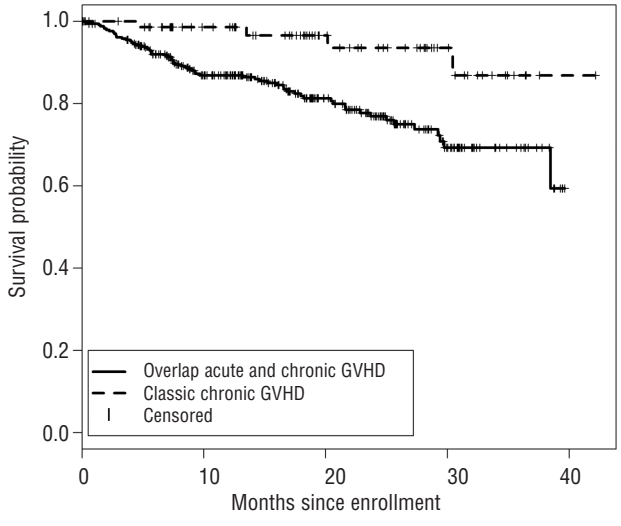


**Figure 1.** Overall survival according to overlap *versus* classic chronic GVHD.

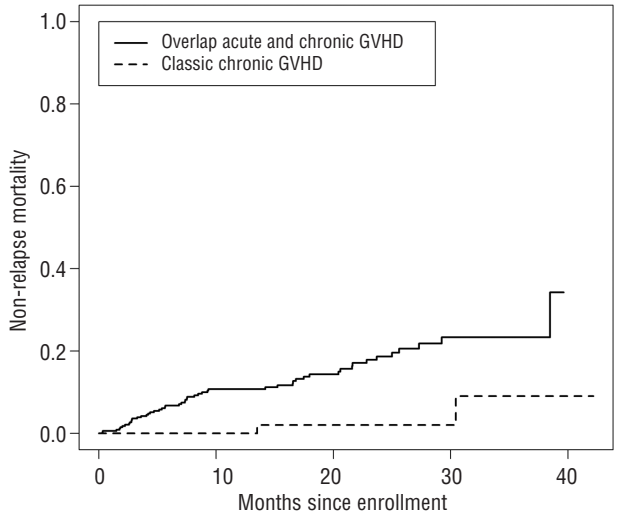


**Figure 2.** Non-relapse mortality according to overlap *versus* classic chronic GVHD.

While overall symptom burden was comparable between patients with overlap GVHD and classic chronic GVHD, significantly worse symptoms were reported by patients with overlap syndrome in intuitive domains including skin and nutrition. Individual domain and summary scores from both the SF-36 and FACT-BMT instruments largely demonstrated similar reported quality of life. Those with overlap syndrome had worse SF-36 social functioning. As items within this domain largely encompass emotional support, relationships, and communication, the data suggest another particular area of vulnerability for such patients, and a potential key area for targeted support and intervention.

Finally, we classified patients as having overlap or classic GVHD in this analysis according to the presence of objective manifestations of concurrent acute GVHD, as we found that clinicians were not accurately applying the NIH definition of overlap. At enrollment, clinicians identified 152 as having overlap chronic GVHD (39%), and 242 as having classic chronic GVHD (61%), after excluding 32 patients reported as having late acute GVHD and one patient reported as not having GVHD. Of the 152 reported as having overlap syndrome, 131 (86%) had acute GVHD manifestations recorded, while 21 (14%) did not. Of the 242 classified as having classic chronic GVHD by clinicians, 189 (78%) objectively had acute manifestations, while 53 (22%) did not. Agreement between clinician designations and classification based on whether or not acute GVHD manifestations were reported was poor, with a simple kappa statistic of 0.07. These data speak for the need for further education and consistency in adherence to the proposed NIH Consensus definition of overlap subtype of chronic GVHD. We acknowledge that inaccurate classification by clinicians may affect treatment decisions and outcomes. However, there are no current guidelines for tailored therapy for classic or overlap subtype of chronic GVHD, and hence clinicians provided what constitutes best clinical care based on current evidence. These data, however, indicate an excess risk associated with the overlap subtype, and argue for rigorous studies to define the best treatment approaches for these patients.

The major strength of this analysis results from the large size of the cohort and the study design in which standardized comprehensive data were gathered prospectively.

**Table 3.** Effect of acute organ involvement on overall survival (N=424).

| Organ(s) involved | Estimate | *P* value | Hazard Ratio* | 95% Hazard Ratio CI |
|---|---|---|---|---|
| Skin only | 1.2 | 0.03 | 3.4 | 1.1-10.6 |
| GI only | 1.4 | 0.03 | 4.0 | 1.2-13.7 |
| Liver only | 1.3 | 0.02 | 3.8 | 1.2-11.6 |
| Skin and GI | 0.8 | 0.35 | 2.2 | 0.4-12.3 |
| Skin and liver | 1.5 | 0.008 | 4.5 | 1.5-13.4 |
| GI and liver | 1.2 | 0.06 | 3.4 | 0.9-12.0 |
| Skin, GI and liver | 1.9 | 0.002 | 6.8 | 2.1-22.8 |

**reference (HR of 1) is classic chronic GVHD, with no features of acute GVHD. GI: gastrointestinal tract.*

These data demonstrate the relevance of the overlap designation, and address major shortcomings of previously reported retrospective analyses. The analysis highlights the importance of the overlap subtype, and calls for advances in the management of these patients. However, we acknowledge the following potential limitations of this analysis. First, patients participating in this cohort study under-represent the total burden of chronic GVHD cases at each of the respective participating centers. While bias in selection of cases for participation in this study could influence results, this concern must be considered in the context of the availability of comprehensive data which is only possible in a prospective study. Additionally, it is important to recognize that the distribution of classic and overlap cases observed in this cohort study reflects only those enrolled, and thus cannot be extrapolated to the entire population of patients with chronic GVHD. Second, we acknowledge that overlap and classic chronic GVHD are not fixed, but rather may change over time according to disease activity and intensity of immune suppressive therapy. This concern may threaten the integrity of the overall survival and non-relapse mortality analyses based on cohort enrollment data. We addressed this concern through the conduct of separate analyses in which chronic GVHD subtype (overlap *versus* classic) was treated as a time-varying covariate. Results of these time-dependent analyses were similar to those using enrollment data, hence providing uniformity in our conclusions. Next, we acknowledge that the majority of subjects were enrolled at four centers. We performed the major study analyses separately on data from only these centers, and did not detect major changes in our reported conclusions. Finally, we recognize a potential risk for GVHD misclassification that may affect these results. Specifically, objective manifestations scored as chronic GVHD by clinicians were not all biopsy-confirmed, and we recognize that some manifestations (e.g. abnormal liver function tests) may result from alternative sources. This is, however, in keeping with challenges inherent in clinical practice in transplantation and should not be considered a particular weakness of this study.

We conclude that the overlap subtype of chronic GVHD, defined by the presence of features of both acute and chronic GVHD, is significantly associated with higher morbidity and mortality. These findings support the importance of the proposed designation, and argue for increased awareness and accurate classification in clinical practice and research.

## Authorship and Disclosures

*The information provided by the authors about contributions from persons listed as authors and in acknowledgments is available with the full text of this paper at www.haematologica.org.*

*Financial and other disclosures provided by the authors using the ICMJE (www.icmje.org) Uniform Format for Disclosure of Competing Interests are also available at www.haematologica.org.*

## References


1. Arai S, Jagasia M, Storer B, Chai X, Pidala J, Cutler C, et al. Global and organ-specific chronic graft-versus-host disease severity according to the 2005 NIH Consensus Criteria. Blood. 2011;118(15):4242-9.
2. Fraser CJ, Bhatia S, Ness K, Carter A, Francisco L, Arora M, 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):2867-73.
3. Lee SJ, Flowers ME. Recognizing and managing chronic graft-versus-host disease. Hematology Am Soc Hematol Educ Program. 2008:134-41.
4. Lee SJ, Vogelsang G, Flowers ME. Chronic graft-versus-host disease. Biol Blood Marrow Transplant. 2003:9(4):215-33.
5. Pidala J, Anasetti C, Jim H. Quality of life after allogeneic hematopoietic cell transplantation. Blood. 2009;114(1):7-19.
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(3):373-85.
7. Pidala J, Kim J, Anasetti C, Nishihori T, Betts B, Field T, 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.
8. Pidala J, Kurland B, Chai X, Majhail N, Weisdorf DJ, Pavletic S, 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(17):4651-7.
9. Pidala J, Kurland B, Chai X, Vogelsang G, Weisdorf D, Pavletic S, et al. Sensitivity of changes in chronic graft-versus-host disease activity to changes in patient-reported quality of life: results from the Chronic Graft-versus-Host Disease Consortium. Haematologica. 2011;96(10):1528-35.
10. Stewart BL, Storer B, Storek J, Deeg HJ, Storb R, Hansen JA, et al. Duration of immunosuppressive treatment for chronic graft-versus-host disease. Blood. 2004;104 (12):3501-6.
11. Filipovich AH, Weisdorf D, Pavletic S, Socie G, Wingard JR, Lee SJ, et al. National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: I. Diagnosis and staging working group report. Biol Blood Marrow Transplant. 2005;11(12): 945-56.
12. Arora M, Nagaraj S, Witte J, DeFor TE, MacMillan M, Burns LJ, et al. New classification of chronic GVHD: added clarity from the consensus diagnoses. Bone Marrow Transplant. 2009;43(2):149-53.
13. Cho BS, Min CK, Eom KS, Kim YJ, Kim HJ, Lee S, et al. Feasibility of NIH consensus criteria for chronic graft-versus-host disease. Leukemia. 2009;23(1):78-84.
14. Jagasia M, Giglia J, Chinratanalab W, Dixon S, Chen H, Frangoul H, et al. Incidence and outcome of chronic graft-versus-host disease using National Institutes of Health consensus criteria. Biol Blood Marrow Transplant. 2007;13(10):1207-15.
15. [No authors listed]. Rationale and design of the chronic GVHD cohort study: improving outcomes assessment in chronic GVHD. Biol Blood Marrow Transplant. 2011;17(8): 1114-20.
16. Mathiowetz V, Weber K, Volland G, Kashman N. Reliability and validity of grip and pinch strength evaluations. J Hand Surg Am. 1984;9(2):222-6.
17. Pavletic SZ, Martin P, Lee SJ, Mitchell S, Jacobsohn D, Cowen EW, et al. Measuring therapeutic response in chronic graft-versus-host disease: National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: IV. Response Criteria Working Group report. Biol Blood Marrow Transplant. 2006;12(3):252-66.
18. Waters RL, Lunsford BR, Perry J, Byrd R. Energy-speed relationship of walking: standard tables. J Orthop Res. 1988;6(2):215-22.
19. Li L, Chan L, Gerber LH. Validation of 2-minute walk test as a measure of exercise tolerance and physical performance in patients with chronic graft versus host disease. Arch Phys Med Rehab. 2008; 89(11):e28.
20. 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):444-52.
21. 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(12):620-2.
22. Herzberg PY, Heussner P, Mumm FH, Horak M, Hilgendorf I, von Harsdorf S, 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.
23. McQuellon RP, Russell GB, Cella DF, Craven BL, Brady M, Bonomi A, 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 Transplant. 1997;19(4): 357-68.
24. Functional Assessment of Chronic Illness Therapy www.facit.org [cited; Available from: www.facit.org <http://www.facit.org>
25. 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. Med Care. 1993;31(3): 247-63.
26. Ware JE, Kosinski M, Keller SD. SF-36 Physical and Mental Health Summary Scales: A User's Manual. Boston: The Health Institute, 1994.
27. Ware JE, Snow KK, Kosinski M, Gandek B. SF-36 Health Survey: Manual and Interpretation Guide. Boston: The Health Institute, 1993.
28. Ware JE KM, Dewey J. How to Score Version 2 of the SF-36® Health Survey. Lincoln, RI: QualityMetric Incorporated, 2000.
29. Quality Metric: SF Health Surveys www.qualitymetric.com.
30. Akpek G, Lee SJ, Flowers ME, Pavletic SZ, Arora M, Lee S, et al. Performance of a new clinical grading system for chronic graft-versus-host disease: a multicenter study. Blood. 2003;102(3):802-9.
31. Shulman HM, Sullivan KM, Weiden PL, McDonald GB, Striker GE, Sale GE, et al. Chronic graft-versus-host syndrome in man. A long-term clinicopathologic study of 20 Seattle patients. Am J Med. 1980; 69(2):204-17.
32. Lee SJ, Klein JP, Barrett AJ, Ringden O, Antin JH, Cahn JY, et al. Severity of chronic graft-versus-host disease: association with treatment-related mortality and relapse. Blood. 2002;100(2):406-14.
33. Wingard JR, Piantadosi S, Vogelsang GB, Farmer ER, Jabs DA, Levin LS, et al. Predictors of death from chronic graft-versus-host disease after bone marrow transplantation. Blood. 1989;74(4):1428-35.
34. Arora M, Burns LJ, Davies SM, Macmillan ML, Defor TE, Miller WJ, et al. Chronic graft-versus-host disease: a prospective cohort study. Biol Blood Marrow Transplant. 2003;9(1):38-45.