Comparison of Patient-Reported Outcomes in 5-Year Survivors Who Received Bone Marrow vs Peripheral Blood Unrelated Donor Transplantation Long-term Follow-up of a Randomized Clinical Trial

Stephanie J Lee, Fred Hutchinson Cancer Research Center
Brent Logan, Medical College of Wisconsin
Peter Westervelt, Washington University
Corey Cutler, Harvard Medical School
Ann Wooffrey, Fred Hutchinson Cancer Research Center
Shakila P. Khan, Mayo Clinic
Edmund K Waller, Emory University
Richard T. Maziarz, Oregon Health & Science University
Juan Wu, The Emmes Corporation
Bronwen Shaw, Medical College of Wisconsin

Only first 10 authors above; see publication for full author list.

Journal Title: JAMA Oncology
Volume: Volume 2, Number 12
Publisher: American Medical Association | 2016-12-01, Pages 1583-1589
Type of Work: Article | Final Publisher PDF
Publisher DOI: 10.1001/jamaoncol.2016.2520
Permanent URL: https://pid.emory.edu/ark:/25593/s6pbn

Final published version: http://dx.doi.org/10.1001/jamaoncol.2016.2520

Copyright information:
Copyright 2016 American Medical Association. All rights reserved.
Accessed September 19, 2020 4:14 AM EDT
Comparison of Patient-Reported Outcomes in 5-Year Survivors Who Received Bone Marrow vs Peripheral Blood Unrelated Donor Transplantation
Long-term Follow-up of a Randomized Clinical Trial

Stephanie J. Lee, MD, MPH; Brent Logan, PhD; Peter Westervelt, MD; Corey Cutler, MD; Ann Wooffrey, MD; Shakila P. Khan, MD; Edmund K. Waller, MD; Richard T. Maziarz, MD; Juan Wu, MS; Bronwen E. Shaw, MD; Dennis Confer, MD; Mary M. Horowitz, MD; Claudio Anasetti, MD

IMPORTANCE Bone marrow or peripheral blood from unrelated donors may be used for hematopoietic cell transplantation. Information about the relative success of transplantation with these 2 graft sources would help physicians and patients choose between them.

OBJECTIVE To compare patient-reported outcomes between patients randomized to receive 1 of 2 graft types for unrelated donor transplantation.

DESIGN, SETTING, AND PARTICIPANTS This follow-up of a randomized clinical trial included English- or Spanish-speaking patients 16 years or older participating in a multicenter randomized clinical trial of unrelated donor bone marrow (BM) vs peripheral blood (PB) (N = 551) in hematopoietic cell transplantation for hematologic neoplasms. Patient-reported outcomes were collected from patients at enrollment and 0.5, 1, 2, and 5 years after transplantation.

INTERVENTIONS Unrelated donor BM or PB hematopoietic cell transplantation.

MAIN OUTCOMES AND MEASURES Functional Assessment of Cancer Therapy–Bone Marrow Transplant, Mental Health Inventory, occupational functioning, Lee Chronic Graft-vs-Host Disease Symptom Scale.

RESULTS At 5 years after transplantation, 102 BM and 93 PB participants were alive and eligible for assessment (age ≥ 40 years or older: 104 [53.5%] male: 101 [51.8%]). The mean (SE) Mental Health Inventory Psychological Well-Being scores (78.9 [1.7] vs 72.2 [1.9]; P = .01; higher better) and Lee chronic graft-vs-host disease symptom scores (13.1 [1.5] vs 19.3 [1.6]; P = .004; lower better) were significantly better for BM recipients, adjusting for baseline scores and missing data. Recipients of BM were also more likely to be working full or part-time than recipients of PB (odds ratio, 1.5; 95% CI, 1.2-2.0; P = .002), adjusting for work status before transplantation. With a median follow-up of 73 months (range, 30-121 months) for survivors, no differences in survival (40% vs 39%; P = .84), relapse (32% vs 29%; P = .47), or treatment-related mortality (29% vs 32%; P = .44) between BM and PB were observed.

CONCLUSIONS AND RELEVANCE Recipients of unrelated donor BM had better psychological well-being, less burdensome chronic GVHD symptoms, and were more likely to return to work than recipients of PB at 5 years after transplantation. Bone marrow should be the standard of care for these types of transplant procedures.

TRIAL REGISTRATION clinicaltrials.gov Identifier: NCT00075816
More transplants are performed from unrelated donors than related donors, according to statistics from the Center for International Blood and Marrow Transplant Research (CIBMTR). More than 80% of transplants from unrelated donors currently use peripheral blood (PB) instead of bone marrow (BM). In 2012, the primary results of a large, multicenter, randomized study from the Blood and Marrow Transplant Clinical Trials Network (BMT CTN) showed similar survival, nonrelapse mortality, and relapse using the 2 graft sources. There was a higher rate of graft failure with BM (9% vs 3%; \( P = .002 \)) and a higher rate of chronic graft-vs-host disease (GVHD) with PB (53% vs 41%; \( P = .01 \)).

Since 2012, there has been no change in the proportion of PB grafts being used for unrelated donor transplantation, with greater than 80% PB, including for patients with early-stage disease such as acute leukemia in first complete remission (M. Horowitz, MD, MS, CIBMTR data, written communication, November 2015).

This study also included a patient-reported outcomes (PRO) component that has not been reported before. The PRO data collection was incorporated in recognition that knowledge about the quality of life (QOL), symptoms, and functional well-being associated with each graft source would be valuable information to help inform future patients and their physicians in choosing a graft source.

## Methods

### Participants

Patients enrolled in the parent BMT CTN 0201 study were eligible for participation in the patient-reported end points substudy if they were 16 years or older, could communicate in English or Spanish, and had access to a telephone. Exclusion criteria included inability to participate in interviews due to cognitive, linguistic, or emotional difficulties and current uncontrolled psychiatric illness. Five hundred seventeen trial participants met these criteria and were included in the QOL study.

### Study Design

The parent study was a randomized, open-label, phase 3, multicenter trial with a primary end point of 2-year survival by intent-to-treat analysis. Enrollment began on March 31, 2004, and ended on September 9, 2009. Eligibility included age up to 66 years, acceptable organ function, and a diagnosis of acute or chronic leukemia, myelodysplasia, or myelofibrosis. Unrelated donors were 5/6 or 6/6 matched at HLA-A, HLA-B, and HLA-DRB1. Exclusion criteria included donor-specific antibodies, prior receipt of allogeneic or autologous transplants, human immunodeficiency virus infection, pregnancy or breastfeeding, active infection, or concomitant enrollment in phase 1 studies. Patients received 1 of 4 myeloablative or reduced-intensity conditioning regimens and 1 of 2 GVHD prophylaxis regimens (methotrexate plus tacrolimus or cyclosporine), specified prior to randomization. Randomization was stratified based on transplant center and disease risk.

After written consent was obtained, contact information for participants was faxed to a central site for contact by the QOL data collection team. Participants were mailed a packet of materials from a central interview center, then contacted by telephone to collect responses. Interviews were conducted in English or Spanish. There was no allowance for paper or online completion of the instruments. Data were entered electronically by interviewers. Several attempts were made to contact participants prior to transplantation. Post-hematopoietic cell transplantation (HCT) assessment times were 6 months, 1 year, 2 years, and 5 years.

The research protocol (Supplement 1) was approved by a protocol review committee of the National Heart, Lung, and Blood Institute (NHLBI), and by the relevant institutional review boards or ethics committees. A Data and Safety Monitoring committee of the NHLBI provided oversight. All prospective donors and recipients gave written informed consent. This study is registered with ClinicalTrials.gov (NCT00075816).

### Data Collection Instruments

Participants completed several validated patient-reported measures at all assessment points except as noted herein. The Functional Assessment of Cancer Therapy–Bone Marrow Transplant (FACT-BMT) is a 37-item instrument composed of the FACT-G and transplant-specific subscale. The FACT-G comprises 4 domains, physical (7 items), social (7 items, including sexual satisfaction), emotional (6 items), and functional (7 items, including work, sleep, and leisure activities). The transplant-specific module has 10 scored items, addressing appetite, appearance, mobility, and fatigue.\(^2,3\) Higher scores indicate better functioning. The published clinically important difference is 7 to 9 points based on the distribution method.\(^3,4\) The Trial Outcome Index is composed of the physical, functional, and transplant-specific modules.\(^5\)

The Mental Health Inventory (MHI) is a 38-item scale divided into 2 summary scores (Psychological Well-Being [14 items, higher scores are better] and Psychological Distress [24 items, lower scores are better]) and 5 subscales that measure anxiety, depression, positive affect, emotional ties, and loss of behavioral and emotional control; it includes the entire mental subscale of the SF-36. As transplantation is associated with both positive and negative psychoemotional sequelae, it is important that the instrument detects both.\(^6\)

The Lee Chronic GVHD symptom scale is a 30-item scale that measures degree of bothersome symptoms in skin, energy, lungs, nutrition, psychological status, eyes, and mouth.
and provides a summary score. Scores for each domain are converted to a 0-to-100 scale in which higher scores indicate more bother. The published clinically important difference is 6 to 7 points based on the distribution method.\textsuperscript{7,8} The chronic GVHD symptom scale was not collected at enrollment because no patient had chronic GVHD before HCT.

Additional questions collected self-assessed Karnofsky performance status, overall health and overall QOL (excellent, very good, good, fair, poor), and a rating scale for overall QOL (in which 0 equals death and 100 equals perfect QOL). Occupational health was measured using 6 items that assess current job status, type of work (captured using Hollingshead categories), number of hours of paid and unpaid work, school attendance, importance of work, and change in work goals.\textsuperscript{9} Additional information on late effects and outcomes after 3 years for participants was retrieved from the CIBMTR database.

Statistical Analyses
Four primary PRO measures were prespecified for analysis: the FACT Trial Outcome Index, the MHI Psychological Well-Being, MHI Psychological Discomfort, and the Chronic GVHD Symptom Score. An analysis of the mean QOL among survivors accounting for missingness was performed using the method of Kurland and Heagerty.\textsuperscript{10} The missing data were modeled among survivors at each time point using stepwise logistic regression considering baseline demographic characteristics, QOL at baseline or at the prior time point, and concurrent clinical issues including development of chronic GVHD or relapse prior to the assessment time, as well as occurrence of death or relapse within 6 months after assessment time. The treatment group was forced into the model, and interactions with each variable were checked and included if significant. These logistic regression models were used to predict probabilities of missing data for each patient; these were used as inverse weights in a weighted least-squares analysis of QOL at each time point. For time points after baseline, the baseline QOL measurement was included as a predictor in the outcome model, except for the chronic GVHD symptom tool, which was not available at baseline. The resulting estimated means, standard errors, and P values are reported. Robustness of the results to missing data was examined by also considering several multiple-imputation approaches. First we considered a missing at random (MAR) assumption for imputation. Next we identified the mean shift in imputation distributions that would change the significance of the results under a not missing at random assumption (a tipping point analysis).

For each scale, the “clinically meaningful difference” was calculated using a distribution-based approach, which is half the standard deviation of the population.\textsuperscript{11-13} The clinically meaningful difference is the difference in QOL that would prompt adoption of an intervention or a change in practice because it has been established to be “noticeable” or “meaningful” to patients.

A weighted logistic regression model was used to compare the likelihood of working full-time or part-time between the graft types, adjusted for baseline work status and using inverse weights from the missing data model.

Results
Participant Characteristics and Enrollment Assessments
Table 1 shows the characteristics of responders and nonresponders at enrollment and 5 years. The response rate at enrollment was 71% to 73% (eTable 1 in Supplement 2). At enrollment, functional well-being (P = .01), FACT-G total (P = .02), and the Health Rating Scale (P = .03) were all significantly higher in the BM group.

The QOL database is complete for 5-year follow-up and was sealed as of May 26, 2015. The patient clinical database was updated using long-term follow-up data as of June 30, 2015.

Six-Month Through Two-Year Assessments
Response rates at 0.5, 1, and 2 years were 28% to 43% of survivors. There were no differences in any of the primary scores in the first 2 years after HCT using univariate comparisons, although missing data limit the conclusions that can be drawn during this period. The chronic GVHD skin score was higher in the BM group (P = .009) at 6 months, and the chronic GVHD mouth score was higher in the PB group at 2 years (P = .03). There were no additional differences (data not shown).

Five-Year Primary End Points
The response rate was 74% to 78% of survivors for each instrument (eTable 1 in Supplement 2). Missing PRO data at 5 years was associated with younger age and high-risk disease but not graft source (OR, 1.05; 95% CI, 0.52-2.14; P = .89). At 5 years, the FACT-Trial Outcome Index, the MHI Psychological Well-Being, and the chronic GVHD symptom scale scores were all significantly better for BM compared with PB patients in univariate comparisons, although only the latter two were still significant after adjustment for multiple testing (P < .0125 because of 4 primary QOL variables) (Table 2). Results were similar when tested in multivariate models adjusting for baseline patient-reported scores and also imputing missing values based on patient characteristics (Table 3). Table 3 also provides the clinically significant difference for each measure. Although none of the observed differences in the scales clearly exceed the clinically significant difference when missing data were handled by inverse probability weighting, the limited overlap in the confidence intervals suggests that the observed differences are clinically meaningful. Multiple imputation of missing data under MAR generally led to slightly larger effects than inverse probability weighting, and the tipping points are fairly large for MHI Psychological Well-Being (~5.1 points) and for the chronic GVHD symptom scale score (12.6 points), indicating that these findings are robust (Table 3).

We explored whether inclusion of chronic GVHD variables in the multivariate models abrogated the statistical differences, which, if true, would suggest that the observed differences may be due to chronic GVHD. Patient-reported outcome differences remained significant in these models when chronic GVHD incidence and moderate-to-severe severity were included in the models (data not shown). However, when additional details about transplant center–reported organ involvement (specifically chronic GVHD of the skin, eye,
and musculoskeletal system, and avascular necrosis) were included in the model, none of the PROs remain significantly different between BM and PB (eTable 2 in Supplement 2). These results suggest that chronic GVHD-associated complications accounted for the observed PRO differences between BM and PB but that overall incidence and global severity information does not fully capture these deficits. A diagnosis of chronic GVHD was associated with patient-reported chronic GVHD symptoms but not with QOL or psychological status (eTable 3 in Supplement 2).

We explored predictors of PRO scores at 5 years. Results suggested that some baseline variables (age, sex, coexisting diseases, primary diagnosis, and baseline scores) were predictive of 5-year PROs, but no potentially modifiable factors other

### Table 1. Characteristics of Responders and Nonresponders at Enrollment Before Hematopoietic Cell Transplantation and at Five Years

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Enrollment No. (%)</th>
<th>5-Year Assessment No. (%)</th>
<th>P Value</th>
<th>Enrollment No. (%)</th>
<th>5-Year Assessment No. (%)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graft source</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bone marrow</td>
<td>199 (50)</td>
<td>60 (49)</td>
<td>.82</td>
<td>80 (53)</td>
<td>22 (51)</td>
<td>.87</td>
</tr>
<tr>
<td>Peripheral blood</td>
<td>196 (50)</td>
<td>62 (51)</td>
<td>.52</td>
<td>72 (47)</td>
<td>21 (49)</td>
<td>.73</td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute myeloid leukemia</td>
<td>193 (49)</td>
<td>54 (44)</td>
<td>.56</td>
<td>70 (46)</td>
<td>19 (44)</td>
<td>.99</td>
</tr>
<tr>
<td>Acute lymphoblastic leukemia</td>
<td>71 (18)</td>
<td>29 (24)</td>
<td>.73</td>
<td>29 (19)</td>
<td>7 (16)</td>
<td>.76</td>
</tr>
<tr>
<td>Chronic myeloid leukemia</td>
<td>51 (13)</td>
<td>14 (11)</td>
<td>.73</td>
<td>24 (16)</td>
<td>8 (19)</td>
<td>.76</td>
</tr>
<tr>
<td>Myelodysplastic syndrome</td>
<td>68 (17)</td>
<td>23 (19)</td>
<td>.73</td>
<td>27 (18)</td>
<td>8 (19)</td>
<td>.76</td>
</tr>
<tr>
<td>Chronic myelomonocytic leukemia</td>
<td>12 (3)</td>
<td>2 (2)</td>
<td>.73</td>
<td>2 (1)</td>
<td>1 (2)</td>
<td>.73</td>
</tr>
<tr>
<td>High-risk disease</td>
<td>102 (26)</td>
<td>47 (39)</td>
<td>.007</td>
<td>24 (16)</td>
<td>14 (33)</td>
<td>.01</td>
</tr>
<tr>
<td>Age ≥40 y</td>
<td>245 (62)</td>
<td>72 (59)</td>
<td>.55</td>
<td>89 (59)</td>
<td>15 (35)</td>
<td>.006</td>
</tr>
<tr>
<td>Male sex</td>
<td>221 (56)</td>
<td>70 (57)</td>
<td>.78</td>
<td>79 (52)</td>
<td>22 (51)</td>
<td>.93</td>
</tr>
<tr>
<td>White race</td>
<td>339 (86)</td>
<td>101 (83)</td>
<td>.41</td>
<td>136 (89)</td>
<td>36 (84)</td>
<td>.30</td>
</tr>
<tr>
<td>Karnofsky performance score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥90%</td>
<td>245 (62)</td>
<td>55 (45)</td>
<td>&lt;.001</td>
<td>101 (66)</td>
<td>31 (72)</td>
<td>.82</td>
</tr>
<tr>
<td>&lt;90%</td>
<td>108 (27)</td>
<td>32 (26)</td>
<td>.95</td>
<td>39 (26)</td>
<td>9 (21)</td>
<td>.82</td>
</tr>
<tr>
<td>Missing</td>
<td>42 (11)</td>
<td>35 (29)</td>
<td>.41</td>
<td>12 (8)</td>
<td>3 (7)</td>
<td>.41</td>
</tr>
<tr>
<td>Conditioning regimen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyclophosphamide/total body irradiation</td>
<td>173 (44)</td>
<td>64 (52)</td>
<td>.73</td>
<td>66 (43)</td>
<td>15 (35)</td>
<td>.73</td>
</tr>
<tr>
<td>Cyclophosphamide/busulfan</td>
<td>128 (32)</td>
<td>35 (29)</td>
<td>.73</td>
<td>57 (38)</td>
<td>17 (40)</td>
<td>.73</td>
</tr>
<tr>
<td>Fludarabine/melphalan</td>
<td>27 (7)</td>
<td>11 (9)</td>
<td>.73</td>
<td>8 (5)</td>
<td>3 (7)</td>
<td>.73</td>
</tr>
<tr>
<td>Fludarabine/busulfan/antithymocyte globulin</td>
<td>67 (17)</td>
<td>12 (10)</td>
<td>.73</td>
<td>21 (14)</td>
<td>8 (19)</td>
<td>.73</td>
</tr>
<tr>
<td>Graft-vs-host disease prophylaxis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyclosporine/methotrexate</td>
<td>83 (21)</td>
<td>23 (19)</td>
<td>.65</td>
<td>36 (24)</td>
<td>8 (19)</td>
<td>.65</td>
</tr>
<tr>
<td>Tacrolimus/methotrexate</td>
<td>275 (70)</td>
<td>90 (74)</td>
<td>.65</td>
<td>104 (68)</td>
<td>31 (72)</td>
<td>.65</td>
</tr>
<tr>
<td>Other</td>
<td>37 (9)</td>
<td>9 (7)</td>
<td>.65</td>
<td>12 (8)</td>
<td>4 (9)</td>
<td>.65</td>
</tr>
<tr>
<td>Donor mismatches at HLA-A, HLA-B, HLA-C, HLA-DRB1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None; 8/8 matched</td>
<td>313 (79)</td>
<td>100 (82)</td>
<td>.51</td>
<td>130 (86)</td>
<td>41 (95)</td>
<td>.51</td>
</tr>
<tr>
<td>≥1; 6/8 or 7/8 matched</td>
<td>82 (21)</td>
<td>22 (18)</td>
<td>.82</td>
<td>22 (14)</td>
<td>2 (5)</td>
<td>.82</td>
</tr>
</tbody>
</table>

### Table 2. Baseline Scores Before Hematopoietic Cell Transplantation and Five-Year Unadjusted Results

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Enrollment Mean (SE) [No. of Respondents]</th>
<th>5-Year Assessment Mean (SE) [No. of Respondents]</th>
<th>P Value</th>
<th>Enrollment Mean (SE) [No. of Respondents]</th>
<th>5-Year Assessment Mean (SE) [No. of Respondents]</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>FACT-BMT Trial Outcome Index*</td>
<td>63.4 (1.2) [n = 190]</td>
<td>60.7 (1.2) [n = 184]</td>
<td>.12</td>
<td>76.0 (1.7) [n = 79]</td>
<td>69.1 (2.2) [n = 69]</td>
<td>.01</td>
</tr>
<tr>
<td>MHI-Psychological Well-Being*</td>
<td>72.5 (1.0) [n = 186]</td>
<td>70.1 (1.1) [n = 182]</td>
<td>.09</td>
<td>77.8 (1.7) [n = 80]</td>
<td>70.2 (2.1) [n = 72]</td>
<td>.005</td>
</tr>
<tr>
<td>MHI-Psychological Distress*</td>
<td>19.0 (0.7) [n = 186]</td>
<td>20.0 (0.8) [n = 182]</td>
<td>.35</td>
<td>17.3 (1.5) [n = 80]</td>
<td>21.4 (1.5) [n = 71]</td>
<td>.05</td>
</tr>
<tr>
<td>Chronic graft-vs-host disease symptoms*</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td>13.0 (1.5) [n = 80]</td>
<td>21.2 (1.7) [n = 72]</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Abbreviations: FACT-BMT, Functional Assessment of Cancer Therapy–Bone Marrow Transplant; MHI, Mental Health Inventory; NA, not applicable.

* Lower scores better. ** Higher scores better.
than graft type (eTable 4 in Supplement 2). In particular, interval from diagnosis to transplantation, GVHD prophylaxis, conditioning regimen, HLA mismatching, and use of antithymocyte globulin were not significant predictors.

With a median follow-up of 73 months (range, 30-121 months) for survivors, no differences between BM and PB were observed for all trial participants in survival (40% vs 39%; \( P = .84 \)) (Figure), relapse (32% vs 29%; \( P = .47 \)), or treatment-related mortality (29% vs 32%; \( P = .44 \)).

### Five-Year Secondary End Points

At 5 years after transplantation, the proportion of return to full- or part-time work for surviving and responding participants was 42 of 80 (52%) for BM and 29 of 72 (40%) for PB. The likelihood of working full or part time outside the home at 5 years was higher for BM recipients (OR, 1.5; 95% CI, 1.2-2.0; \( P = .002 \)) adjusted for work status before HCT and missing data based on graft source, disease risk, and age, using inverse probability weighting. Of the 7 chronic GVHD subscales, symptoms in the eyes (\( P < .001 \)) and lungs (\( P = .004 \)) and energy (\( P = .003 \)) were significantly better with BM (eTable 5 in Supplement 2).

There were nonsignificant findings of better functional well-being (\( P = .02 \)), FACT summary scores (\( P = .02 \) for both), higher health rating scale scores (\( P = .02 \)), and perception of overall QOL (\( P = .04 \)) in the BM group.

Of the 5-year survivors, 72 (71%) receiving BM had no active chronic GVHD compared with 46 (49%) of PB recipients, with lower rates of mild (17 [17%] vs 21 [23%]), moderate (9 [9%] vs 16 [17%]), and severe (4 [4%] vs 8 [9%]) favoring BM (\( P = .03 \)). A higher proportion of PB recipients experienced skin sclerosis (17 [18%] vs 8 [8%]; \( P = .03 \)), eye GVHD (31 [33%] vs 15 [15%]; \( P = .002 \)), and musculoskeletal involvement (14 [15%] vs 3 [3%]; \( P = .003 \)). There were no statistically significant differences in the reported incidence of mouth (\( P = .18 \)), lung (\( P = .26 \)), or gastrointestinal (\( P = .06 \)) chronic GVHD. There was a higher incidence of avascular necrosis with PB (14 [15%] vs 7 [8%]; \( P = .003 \)).

### Figure. Overall Survival

![Overall Survival](https://example.com/figure.png)

Bone marrow recipients Peripheral blood stem cell recipients

<table>
<thead>
<tr>
<th>Patient-Reported Measure</th>
<th>Adjusted for Enrollment Values and Missing Data Using Inverse Probability Weighting Using Significant Clinical Characteristics</th>
<th>Adjusted for Missing Data Using Multiple Imputation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bone Marrow (( n = 102 ))</td>
<td>Peripheral Blood (( n = 93 ))</td>
</tr>
<tr>
<td></td>
<td>Mean (SE) [No. of Respondents]</td>
<td>P Value</td>
</tr>
<tr>
<td>FACT-BMT Trial Outcome Indexa</td>
<td>76.7 (1.6) [n = 79]</td>
<td>70.5 (1.9) [n = 69]</td>
</tr>
<tr>
<td>MHI–Psychological Well-beingb</td>
<td>78.9 (1.7) [n = 80]</td>
<td>72.2 (1.9) [n = 72]</td>
</tr>
<tr>
<td>MHI–Psychological Distressb</td>
<td>16.0 (1.3) [n = 80]</td>
<td>19.0 (1.5) [n = 71]</td>
</tr>
<tr>
<td>Chronic graft-vs-host disease symptomsa</td>
<td>13.1 (1.5) [n = 80]</td>
<td>19.3 (1.6) [n = 72]</td>
</tr>
</tbody>
</table>

Abbreviations: FACT-BMT, Functional Assessment of Cancer Therapy–Bone Marrow Transplant; MAR, missing at random; MHI, Mental Health Inventory.

a Clinically significant difference = 0.5 \times standard deviation.

b Tipping point = shifted difference in mean outcome between BM and PB for missing values that would lead to a different conclusion at \( a = .0125 \).

c Higher scores better.

d Lower scores better.

---

Bone Marrow vs Peripheral Blood for Unrelated Donor Transplantation

Original Investigation Research

jamaoncology.com

Copyright 2016 American Medical Association. All rights reserved.
5 [5%]; \( P = .02 \)). No differences were detected in myocardial infarction, congestive heart failure, diabetes, hypothyroidism, dialysis, seizures, or cataracts. There were no secondary end points for which PB was better.

### Discussion

This large, multicenter, randomized trial showed similar 2- and 5-year survival between patients randomized to receive unrelated BM compared with those randomized to receive PB. While there was an increase in graft failure with BM (9% vs 3%; \( P = .002 \)), there was a higher rate of chronic GVHD with PB at 2 years (53% vs 41%; \( P = .01 \)). Analyzing PROs collected from participants throughout the 5 years, we observed better psychological well-being, less burdensome chronic GVHD symptoms, and a 50% higher likelihood of returning to full- or part-time work in the BM recipients compared with PB recipients. There were no patient-reported end points for which PB was superior.

The reason for the improved self-reported health and well-being of BM recipients is unclear. We hypothesized that chronic GVHD might explain the difference. However, within the limits of the trial data collection and statistical models, we were not able to attribute the differences either to the occurrence or overall severity of chronic GVHD as measured in the study. However, patient-reported symptoms, center-reported chronic GVHD organ involvement, and avascular necrosis were greater for PB than BM, and a model that adjusted for chronic GVHD organ involvement and avascular necrosis abrogated the difference between BM and PB, suggesting that clinical chronic GVHD morbidity was worse with PB and may account for the observed PRO differences.

Patients randomized to BM were more likely to return to full- or part-time work than patients receiving PB. The importance of work to family finances, insurance coverage, societal reintegration, and perceptions of health has been reported in transplantation survivors, and this metric adds an important dimension to how successful transplantation is defined.\(^1\)\(^-\)\(^21\)

There are number of limitations to the study. Although the response rates were high (>75%) for the enrollment and 5-year assessments, a quarter of patients did not provide patient-reported information at these time points. We addressed this limitation through modeling of missing data and a tipping point analysis, which showed that the conclusions for MHI well-being and chronic GVHD symptoms were robust, in that the missing data would have to be completely opposite to the available data to change the conclusions. In contrast, for MHI Distress, the tipping point analysis was less robust, only partially assuaging concern about potential nonresponder bias because a smaller shift in nonresponders’ reports could result in lack of significant findings. Second, response rates at 6 months, 1 year, and 2 years were much lower, at 28% to 43%, compromising our ability to confidently comment on these time points. In retrospect, the central data collection mechanism during this time was suboptimal and exacerbated by the demands on patients of the early recovery period, particularly at 6 months. Thus, we are only confident about our results at enrollment and 5 years, leaving a large period without accurate patient-reported experiences. Third, we administered validated instruments to participants. While these have the advantage of demonstrated reliability and validity, some information is inevitably lost by not collecting qualitative information about participants’ QOL and functioning. The interviews already took 30 to 45 minutes and we did not believe that we could extend the time longer to collect more detailed information. Fourth, allowed GVHD prophylaxis only included calcineurin inhibitors, methotrexate, and antithymocyte globulin. Novel approaches such as posttransplant cyclophosphamide and novel agents were not studied. Clinical trials of novel conditioning and GVHD prophylaxis regimens may require PB. Finally, this study focuses on PROs reported by transplant recipients. Studies of donors suggest that BM harvests result in a longer recovery period than PB, although with time, donors of both graft types achieve similar recovery.\(^22\)

To date, the primary 2-year results of the clinical trial have not resulted in any noticeable change in clinical practice, with more than 80% of unrelated donor transplantations using PB. Explorations of potential reasons for continued frequent use of PB are the subject of ongoing studies. We hypothesize that several factors may contribute to the continued preference for PB despite a higher rate of chronic GVHD. First, ongoing clinical trials may require PB based on historical data for time to engraftment and GVHD because that has been the standard for the past few years. Second, clinicians may be worried about the higher rate of graft failure or delayed engraftment, both observed in our trial, or a higher rate of relapse, which was not observed in our trial but was seen with some prior studies. Third, donating PB is associated with a shorter duration of symptoms than donating BM so donors may prefer to give PB, although large donor studies suggest that almost all donors recover completely with time.\(^22\)

### Conclusions

Patients randomized to receive BM instead of PB had better psychological well-being, fewer chronic GVHD symptoms, and were more likely to return to work, although survival, relapse, and treatment-related mortality are similar. The failure to see an increase in the proportion of HCTs using BM means that the clinical results published were not compelling enough to change treatment of these patients. It remains to be seen whether the additional information provided by this study will be judged sufficient to recommend BM instead of PB for unrelated donor transplantation when performed for the indications and using the approaches included in this study.
Bone Marrow vs Peripheral Blood for Unrelated Donor Transplantation

Drafting of the manuscript: Maziarz, Wu, Shaw, Confer, Horowitz.

Obtained funding:

Statistical analysis:

intellectual content:

Critical revision of the manuscript for important accuracy of the data analysis.

Study concept and design:

Moffitt Cancer Center, Tampa, Florida (Anasetti).

Department of Blood and Marrow Transplantation, Transplant Research, Medical College of Wisconsin, National Marrow Donor Program, Minneapolis, Minnesota (Confer); Department of Blood and Marrow Transplantation, Moffitt Cancer Center, Tampa, Florida (Anasetti).

Author Contributions: Drs. Lee and Logan had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Lee, Logan, Woolfrey, Confer, Horowitz, Anasetti.

Acquisition, analysis, or interpretation of data: Lee, Logan, Westervelt, Cutter, Woolfrey, Khan, Waller, Maziarz, Wu, Shaw, Confer, Horowitz.

Drafting of the manuscript: Lee.

Critical revision of the manuscript for important intellectual content: All authors.

Statistical analysis: Logan, Woolfrey, Wu.

Obtained funding: Confer, Horowitz, Anasetti.

Administrative, technical, or material support: Shaw, Confer, Horowitz, Anasetti.

Study supervision: Woolfrey, Anasetti.

Conflict of Interest Disclosures: None reported.

Funding/Support: For this study was provided by grant UT01HL069294 from the National Heart, Lung, and Blood Institute and the National Cancer Institute to support the Blood and Marrow Transplant Clinical Trials Network and the Department of the Army, Office of Naval Research, and the National Marrow Donor Program. Enrollment support was provided by DKMS Germany.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: Any views, opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views or the official policy or position of the aforementioned parties.

Previous Presentation: An abstract based on this work was presented at the 57th American Society of Hematology annual meeting; December 6, 2015; Orlando, Florida.

REFERENCES


