Outcomes of acute leukemia patients transplanted with naive T cell–depleted stem cell grafts

Marie Bleakley,1,3 Shelly Heimfeld,1 Keith R. Loeb,1,3 Lori A. Jones,4 Colette Chaney,1 Stuart Seropian,5 Ted A. Gooley,1 Franziska Sommermeyer,1 Stanley R. Riddell,1,4,7 and Warren D. Shlomchik5,8

1Fred Hutchinson Cancer Research Center (FHCRC), Seattle, Washington, USA. 2Department of Pediatrics and 3Department of Pathology, University of Washington, Seattle, Washington, USA. 4Yale University School of Medicine (YUSM), New Haven, Connecticut, USA. 5Department of Medicine, University of Washington, Seattle, Washington, USA. 6Seattle Cancer Care Alliance, Seattle, Washington, USA. 7Institute for Advanced Study, Technical University of Munich, Munich, Germany. 8University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA.

BACKGROUND. Graft-versus-host disease (GVHD) is a major cause of morbidity and mortality following allogeneic hematopoietic stem cell transplantation (HCT). In mice, naïve T cells (T_N) cause more severe GVHD than memory T cells (T_M). We hypothesized that selective depletion of T_N from human allogeneic peripheral blood stem cell (PBSC) grafts would reduce GVHD and provide sufficient numbers of hematopoietic stem cells and T_M to permit hematopoietic engraftment and the transfer of pathogen-specific T cells from donor to recipient, respectively.

METHODS. In a single-arm clinical trial, we transplanted 35 patients with high-risk leukemia with T_N-depleted PBSC grafts following conditioning with total body irradiation, thiotepa, and fludarabine. GVHD prophylactic management was with tacrolimus immunosuppression alone. Subjects received CD34-selected PBSCs and a defined dose of T_M purged of CD45RA+ T_N. Primary and secondary objectives included engraftment, acute and chronic GVHD, and immune reconstitution.

RESULTS. All recipients of T_N-depleted PBSCs engrafted. The incidence of acute GVHD was not reduced; however, GVHD in these patients was universally corticosteroid responsive. Chronic GVHD was remarkably infrequent (9%; median follow-up 932 days) compared with historical rates of approximately 50% with T cell–replete grafts. T_M in the graft resulted in rapid T cell recovery and transfer of protective virus-specific immunity. Excessive rates of infection or relapse did not occur and overall survival was 78% at 2 years.

CONCLUSION. Depletion of T_N from stem cell allografts reduces the incidence of chronic GVHD, while preserving the transfer of functional T cell memory.

TRIAL REGISTRATION. ClinicalTrials.gov (NCT 00914940).

FUNDING. NIH, Burroughs Wellcome Fund, Leukemia and Lymphoma Society, Damon Runyon Cancer Research Foundation, and Richard Lumsden Foundation.

Introduction

Allogeneic hematopoietic stem cell transplantation (HCT) is often curative for patients with acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), and other hematologic malignancies (1, 2). Donor T cells in the transplanted graft contribute to successful HCT by promoting the establishment of donor hematopoiesis, transferring pathogen-specific immunity, and mediating a graft-versus-leukemia (GVL) effect. Unfortunately in HLA-matched HCT, donor T cells that recognize recipient minor histocompatibility (H) antigens are also central to the development of graft-versus-host disease (GVHD), which remains a major cause of morbidity and mortality after HCT (3, 4).

To prevent or diminish the severity of GVHD in patients receiving T cell–replete allografts, patients receive several months of pharmacologic immunosuppression with calcineurin inhibitor–based regimens. Nonetheless, 30% to 70% and 40% to 63% of patients who receive HLA-matched related donor (MRD) grafts develop acute GVHD (aGVHD) and chronic GVHD (cGVHD), respectively (5–7). GVHD can be substantially reduced by nonselective removal of T cells from the stem cell graft or by early in vivo administration of T cell–depleting antibodies (7–9). Unfortunately, pan-T cell depletion (TCD) is complicated by delayed immune reconstitution and an increased frequency of opportunistic infections (10–12).

 naïve T cells exist in the blood, secondary lymphoid organs, and tissues as distinct naïve (T_N), effector (T_E), and memory (T_M) subsets that can be distinguished by alterations in cell surface phenotype that occur as a consequence of activation with cognate antigen (13). The CD45RA+CD62L+ T_N subset is antigen inexperienced and has a more diverse T cell receptor (TCR) repertoire than T_M (14, 15). After antigen-driven activation, T_N are induced to clonally expand and differentiate into short-lived effector cells and subsets of long-lived T_M that protect the host.

Authorship note: Stanley R. Riddell and Warren D. Shlomchik are co–senior authors.

Conflict of interest: The authors have declared that no conflict of interest exists.

Submitted: February 4, 2015; Accepted: April 30, 2015.

Reference information: J Clin Invest. 2015;125(7):2677–2689. doi:10.1172/JCI81229.
from reinfec- tion and include CD45RO−CD62L− central memory (T_{CM}), CD45RO−CD62L− effector memory (T_{EM}), and CD45RO−CD62L−CD69+ tissue-resident memory (T_{TRM}) cells. CD4+FOXP3+ Tregs are a separate subset of T cells that is derived by both thymic and extrathymic pathways and suppresses autoimmunity (16). Based on knowledge of the phenotype, repertoire, and reactivity of T cell subsets, we predicted that a strategy for engineering allogeneic stem cell grafts might be designed to separate the beneficial functions of T cells from detrimental GVHD after HCT. In mouse allogeneic HCT performed without immunosuppression, we and others showed that T_{N} caused severe GVHD, T_{CM} did not cause significant GVHD (17–23). Importantly, T_{EM} transferred antipathogen immunity and had GVL activity in these models (17, 22, 24). Mechanistic studies demonstrated that TCR repertoire–independent and –dependent differences between T_{N} and T_{CM} subsets contributed to differences in GVHD induction (20, 25–27). Consistent with the results in mice, we found, using sensitive in vitro assays, that the frequency of human CD8+ T cells specific for minor H antigens was at least 5- to 20-fold higher in T_{CM} than T_{N} (28).

To test the hypothesis that removing T_{N} from human allogeneic HCT grafts would reduce serious GVHD and allow the transfer of functional pathogen-specific immunity, we developed what we believe to be a novel graft-engineering strategy in which T_{N} were selectively depleted from granulocyte colony-stimulating factor–mobilized peripheral blood stem cells (PBSCs) using immunomagnetic selection with a clinical-grade iron-dextran bead conjugated to a monoclonal antibody targeting CD45RA, which is expressed on the cell surface of all T_{N} but is absent on T_{CM} and most T_{EM} (29). We then designed a single-arm phase II trial in which patients with high-risk acute leukemia or advanced myelodysplastic syndrome (MDS) received a T_{N}-depleted stem cell graft from a HLA-MRD. The use of a MRD provided a margin of safety if unforeseen problems occurred during graft engineering by allowing immediate access to additional granulocyte colony-stimulating factor–mobilized donor PBSCs. We used myeloablative conditioning composed of fludarabine, thiotepa, and total body irradiation (TBI; 1,320 cGy) (30). This regimen was chosen because it results in acceptable engraftment rates in the context of complete TCD HCT and should also permit engraftment following T_{N}-depleted HCT without the need for antibody therapy before HCT with antithymocyte globulin or anti-T cell mAb, which could lead to depletion of T_{CM} and most T_{EM} (30). Tacrolimus alone was used as GVHD prophylaxis and was tapered after 50 days in the absence of GVHD. Here, we report the clinical outcomes of 35 consecutive patients treated in this clinical trial.

Results

Patient characteristics and graft engineering. The patient characteristics are listed in Table 1. Of 108 patients aged 14 to 55 years with an HLA-MRD (>13 years old) who were evaluated for HCT for the treatment of acute leukemia or advanced MDS during the study period (December 2009 to May 2014), 35 were allocated to HCT in the T_{N} depletion clinical trial and treated accordingly (Figure 1).

The major reasons that other patients were not allocated to the trial included allocation to a non-TBI-containing myeloablative HCT at the discretion of the transplant attending physician (in most cases, patients with AML or MDS with better-risk disease and/or comorbidities); allocation to a nonmyeloablative regimen; or allocation to transplantation on a standard myeloablative treatment plan with cyclophosphamide and TBI conditioning. Patients allocated to cyclophosphamide and TBI included patients that lacked insurance coverage for a clinical trial, were unwilling to consent to the trial of T_{N} depletion, or were precluded for participating for logistical reasons or because they did not meet eligibility criteria for the trial.

To ensure consistency in the trial, we established target ranges for the CD34+ cells, CD3+CD45RO CD45RA+ T_{EM} and total CD3+ T cells in the engineered stem cell grafts. A CD34+ cell dose
of $>5.0 \times 10^6$ cells/kg was selected because exceeding this threshold is associated with improved overall survival in recipients of PBSC HCT (31). A target of $7.5 \times 10^4$ T$_N$/kg was chosen based on estimates that a quantity exceeding this number would be sufficient to cause GVHD. A target of $1 \times 10^7$ CD3$^+$ T cells/kg, with an acceptable range of $1 \times 10^6$ to $10 \times 10^6$ CD3$^+$ T cells/kg, was selected because this range of T cells is 10- to 100-fold greater than the number of unselected T cells predicted to cause GVHD after MRD HCT, and we reasoned that this number was likely to be sufficient T$_N$ to facilitate immune reconstitution (32). We were successful in achieving these targets for all patients. As shown in Table 1, the PBSC grafts administered contained a median of $7.4 \times 10^4$ CD3$^+$ cells/kg (range $5.1 \times 10^4$ to $19.9 \times 10^4$ CD3$^+$ cells/kg) and $10^7$ CD3$^+$ T cells/kg that included $3,600$ T$_N$/kg (range 500–74,600 T$_N$/kg; interquartile range 2,200–6,500 T$_N$/kg). Eighty-six percent of patients received $<10,000$ T$_N$/kg.

Donor cell engraftment. Successful and sustained engraftment of donor hematopoietic cells was a primary safety endpoint of the study. All patients had neutrophil engraftment, achieved at a median of 13 days (range 9–17 days), and platelet counts exceeded 20,000 per mm$^3$ at a median of 14 days (range 9–111 days). Donor and recipient chimerism was determined by DNA genotyping for short tandem repeat polymorphisms. Myeloid (CD33$^+$) engraftment was 100% donor in all recipients, and CD3$^+$ T cells were 100% donor in most recipients after day 28 (Figure 2). There were no graft rejections, although one patient developed secondary graft failure at day 260 while still having 100% donor cells.

aGVHD. Twenty-three of the thirty-five patients developed clinical symptoms and signs consistent with grades II–IV aGVHD (66%; 95% CI 41%–74%), and the diagnosis was confirmed in all cases by biopsy of an involved site (Figure 3A). A test of the null hypothesis that the true rate of grades II–IV aGVHD is 60% yields $P = 1.0$ (1-sided binomial test), and therefore, we could not reject our prespecified null hypothesis for the second primary endpoint of the study. Three patients (9%; 95% CI 0%–18%) developed grade III aGVHD, and no cases of grade IV aGVHD or liver aGVHD were observed. The clinical pattern and stage of gastrointestinal (GI) and skin aGVHD are shown in Figure 3B–D. Ten patients (29%) had no aGVHD; two (6%) developed grade I aGVHD (skin stage 2 only); thirteen (37%) had grade II GVHD limited to GI stage 1; seven (20%) had GI stage 1 and skin stage 1–2; and three (9%) had grade III aGVHD manifested by diarrhea (GI stage 2 [$n = 1$]; stage 3 [$n = 2$]) (Figure 3B).

All 23 patients with GI aGVHD developed symptoms while receiving prophylactic tacrolimus. Histologic grading of GVHD severity was performed in each case, and the scores were minimal in 9 patients and mild in 14 patients (Figure 3E). Twenty-two patients received systemic corticosteroids (median initial dose of 1 mg/kg; range 0.5–2 mg/kg) in addition to tacrolimus for treatment of grade II–IV aGVHD. Twenty patients had complete resolution of symptoms in <7 days, and the remaining 3 patients achieved a complete response in 7 to 14 days. Importantly, no patient required a second-line agent for treatment of aGVHD.

cGVHD. Three of the thirty-five T$_N$-depleted recipients developed cGVHD, with a median follow-up of 932 days (range 209–1,826 days), for an estimated probability of 9% at 2 years (95% CI 0%–19%; Figure 3F). One patient had a skin rash and dry eyes (mild cGVHD); one patient had late upper GI GVHD, with mild oral changes (moderate cGVHD); and the third patient had upper GI GVHD, mild dry eyes, and developed deep sclerosis of the skin on the lower extremities during the taper of immunosuppression (severe cGVHD). The cGVHD in the latter patient resolved rapidly after initiation of rituximab, and systemic corticosteroids and rituximab were subsequently discontinued without cGVHD recurrence. All of the patients diagnosed with cGVHD had a preceding history of aGVHD. None of the patients diagnosed with grade III aGVHD subsequently developed cGVHD.

Adverse events and nonrelapse mortality. Adverse events were as expected for patients undergoing HCT following a myeloblastic preparative regimen containing TBI (Supplemental Table 1; supplemental material available online with this article; doi:10.1172/JCI81229DS1). Nonrelapse mortality (NRM) occurred in 1 patient.
remained infrequent throughout the first year after HCT (Figure 5). Concurrent with the appearance of T_N, both TCR excision circles (TRECs) and TCR diversity, as measured by Vβ spectratyping, increased after day 180 (Figure 5H and Supplemental Figure 1). The recovery of T cells in the blood occurred well before emergence of thymic emigrants, consistent with a contribution from TME administered with the graft. We evaluated the transfer of virus-specific TME in a subset of 7 HLA-A*0201 CMV-seropositive T_N-depleted HCT recipients with CMV-seropositive donors by measuring T cells in donor and recipient blood that were specific for a CMV pp65 peptide (pp65NLV) using an HLA A*0201/pp65 NLV tetramer. By tetramer staining, T cells specific for pp65NLV were present in donor blood and in the blood of all 7 recipients 28 days after HCT (Figure 6, A and B). The pp65NLV-specific T cells contained CD27+CD28+ TCM and CD28– and/or CD27– TEM subsets and were functional, as demonstrated by an increase in cell frequency and number in temporal association with CMV reactivation and production of IFN-γ and IL-2 after pp65NLV peptide stimulation (Figure 6, A and C).

Analysis of cGVHD in a contemporary T cell-replete cohort. The 9% incidence of cGVHD among T_N-depleted HCT recipients is far lower than the rates of 40% to 63% reported for comparable patients transplanted with MRD T cell-replete HCT and ablative conditioning at our center and others (5–7). In a recent retrospective analysis of the frequency of cGVHD fulfilling NIH cGVHD consensus criteria and requiring systemic immunosuppression, a rate of 44% was found for MRD HCT recipients who received PBSCs between 1992 and 2008 (34). In order to exclude the possibility that the rate of cGVHD after MRD HCT had decreased over time independent of our intervention, we compared the incidence of GVHD in recipients of TN-depleted HCT with that in a cohort of patients that received T cell-replete HCT using a standard cyclophosphamide (120 mg/kg) and TBI (12 Gy) conditioning,
that this disparity could explain the higher rate of cGVHD in T cell–replete HCT recipients, we analyzed GVHD in gender combinations other than male recipients of female grafts and found that cGVHD occurred in 11% (CI 0%–23%) of TN-depleted recipients and 48% (CI 27%–68%) of recipients of T cell–replete grafts. Thus, the apparent benefit of TN depletion for reducing cGVHD occurred independent of donor/recipient gender.

Discussion

This first-in-human trial demonstrates that engineering PBSC grafts to contain minimal numbers of T_N while retaining CD34+ cells and TM is feasible. The 2-step graft-engineering strategy we developed to deplete TN was highly successful in achieving a profound depletion of TN (<7.5 × 10^4) relative to the usual dose of TN infused in PBSC HCT (1.8 × 10^8 TN CD3+/kg; ref. 29) and in meeting our target doses of CD34+ cells, TN, and total CD3+ T cells. Because 86% of patients received grafts containing <10,000 T_N/kg, we were able to rigorously test the effect of T_N depletion of the PBSC graft on outcome. Importantly, the trial met its primary safety endpoint to achieve durable donor hematopoietic engraftment, which occurred in 34 of 35 patients. The engraftment endpoint was important because pan-TCD has sometimes been associated with higher rates of graft failure and the engraftment potential of TN-depleted grafts had not been previously tested (9). Furthermore, the results in the first 35 patients treated with this approach to HCT provide the first experimental support to our knowledge for the
The severity of aGVHD after TN-depleted HCT was generally consistent with historical aGVHD data from FHCRC. The frequency and hence, the importance of comparing our results to concurrent patients who received myeloablative allogeneic MRD HCT at our institution between 2001 and 2008 and 2008 and 2014, respectively patients who received GVHD prophylaxis with a calcineurin antagonist. Second, the intensity of conditioning, the TBI dose in particular, is a risk factor for aGVHD (36, 37), and we used a conditioning regimen that is more intense than the cyclophosphamide/TBI regimen used in the comparison groups.

A notable finding in our study was the remarkably low rate (9%) of cGVHD, which is similar to the rate of 19% reported after CD34+-selected (TCD) HCT and lower than the rates of 40% to 63% reported after T cell-replete MRD HCT (5–7, 33) and of 56% in our concurrent comparison group. Although our study was single armed, it is unlikely that the cGVHD incidence was significantly underestimated. Patients undergo structured comprehensive long-term follow-up evaluations to screen for cGVHD by experts not associated with the study using standardized NIH criteria. The median time from HCT to the onset of cGVHD, as per the NIH criteria, after T cell-replete HCT is 162 days, and greater than 90% of cGVHD diagnoses are made in the first year after HCT (34). Furthermore, most patients who develop cGVHD do so while on pharmacological immunosuppression or within 3 months of its discontinuation. The patients in our trial have been followed for a median of 2.5 years, and 25 patients are beyond 1 year after HCT and have been off immunosuppression for many months in most cases. Patients diagnosed with cGVHD have an approximately 20% risk of NRM, inferior overall survival, and a greatly compromised quality of life (34, 38). The use of allogeneic PBSCs as a stem cell source has increased over the past decade and has remained high for logistical reasons, despite PBSCs being associated with a higher risk of cGVHD than bone marrow transplantation (5, 39–41). Because we observed a very low rate of cGVHD, even in the context of PBSC transplantation, our results suggest that TN depletion is likely to be a powerful strategy for reducing the substantial cGVHD-related morbidity, mortality, and disability of allogeneic HCT.

It may seem surprising that we observed such a low rate of cGVHD in T<sub>N</sub>-depleted recipients, without a major effect on the occurrence of aGVHD. However, this is not inconsistent with our preclinical work. In mouse models, T<sub>EM</sub> consistently cause little to no GVHD, but we found that T<sub>CM</sub> could cause GVHD, albeit milder than T<sub>N</sub>, and the grafts infused in our clinical trial did contain both T<sub>EM</sub> and T<sub>CM</sub> (17, 22, 23). The pathogenesis of aGVHD and its relationship to cGVHD are not fully understood, and our clinical trial results suggest that the aGVHD syndrome that occurs after T<sub>N</sub>-depleted HCT may differ biologically from that occurring after T cell-replete HCT, in that it less frequently presages the development of cGVHD. There are several possibilities that alone or in combination could account for biologic differences in GVHD after T<sub>N</sub> depletion that are the subject of ongoing work in our laboratories. The more limited TCR repertoire of T<sub>N</sub> could result in fewer minor H antigens being recognized by fewer T cell clones. The affinity of cross-reactive T<sub>N</sub> for minor H antigens may also be lower than that of alloreactive T<sub>N</sub> recruited into the response. T<sub>CM</sub> progeny may also traffic differently, have a reduced ability to undergo sustained division, or have fewer pathogenic effector functions, including different cytokine profiles, as have been suggested in
mouse studies (21, 25, 26). Thymic damage induced by aGVHD has been suggested to contribute to cGVHD pathogenesis through defective negative selection and impaired Treg development, and it is conceivable that T₈ᵦ depletion results in less thymus GVHD (42–47). An interesting hypothesis to explain the rarity of sustained alloreactivity after TN-depleted HCT is that the T₈ᵦ involved in the GVH syndrome do not primarily target minor H antigens but are instead directed against nonpolymorphic self antigens or microbial antigens. CD45RA⁺ Tregs are also removed by our graft manipulation, and transplantation of autologous Treg-depleted T cells in rodents induces GI inflammation, thought to be in part due to dysregulated immunity to microbial antigens (48–51). Finally, it is possible that the aGVHD syndrome observed in TN-depleted HCT recipients is primarily driven by cytokines such as IL-6 (52).

A rationale for our approach was that the transfer of T₈ᵦ would improve immune reconstitution relative to pan-TCD, which results in prolonged lymphopenia and an increase in opportunistic infections (10–12). T cell numbers recovered much earlier after TN-depleted HCT relative to patients that received TCD HCT (33), and EBV reactivation and post-HCT lymphoproliferative disease (PTLD), which occur in 18% and 2% of TCD HCT recipients, respectively (7, 33), were not observed at all after TN-depleted HCT, consistent with transfer of protective EBV-specific immunity. CMV reactivation is common after allogeneic HCT, independent of whether T cells are depleted or not, and the frequency of viral reactivation among patients at risk in the TN-depleted cohort (73%) and T cell–replete control group (87%) is similar to that reported in the literature for T cell–replete grafts (53, 54). Current practice dictates preemptive treatment of CMV reactivation at low levels of viremia after HCT, and progression to CMV disease is consequently now rare in HLA-matched HCT, precluding assessment of whether a particular HCT approach would change the natural history of CMV reactivation (55). However, we show that prior to the emergence of TRECs, CD8⁺ CMV-specific T cells were detected in the blood of TN-depleted HCT recipients and expanded in response to CMV reactivation, consistent with the transfer of functional anti-CMV immunity with the graft.

Within the limitations of a single-arm trial of this size, the reduction in cGVHD did not come with apparent decrements in other important clinical outcomes. The 2-year DFS rate of 70% in TN-depleted HCT recipients compares favorably to DFS rates of 50% to 65% after TCD or T cell–replete MRD HCT for acute leukemia (7, 33, 56, 57). Potentially fatal EBV reactivation and PTLD did not occur in TN-depleted recipients, and other serious infections were rare. Additionally, relapse was uncommon. In particular, the relapse rate of 28% observed among patients that received HCT with residual leukemia or a history of previous relapses compares favorably with the expected relapse rates of 37% to 60% with T cell–replete or TCD MRD HCT for this subset of patients, implying that the GVL effect may not be abrogated by TN depletion and/or that relapse is suppressed due to an earlier withdrawal of immunosuppressive drugs (7, 33, 57, 58).
Although we observed a large reduction in the rate of cGVHD relative to historical and concurrent patients that received T cell-replete HCT, this comparison of GVHD rates cannot substitute for a randomized trial. The T cell-depleted and concurrent T cell-replete groups appear well balanced for known factors affecting HCT outcomes, but we cannot be sure that unrecognized risk factors were evenly distributed, so we did not conduct formal comparisons of survival. However, we were encouraged by the fact that there were no apparent increases in relapse or NRM among T cell-depleted recipients. A prospective randomized controlled clinical trial will be required to prove that T cell depletion reduces cGVHD without compromising immune reconstitution or increasing relapse rates compared with T cell-replete HCT and to determine whether a reduction in cGVHD results in meaningful improvements in survival, quality of life, or transplantation-related health care costs. Ideally, one would like to compare T cell-replete HCT, T cell-depleted HCT, and T cell–depleted HCT in a single clinical trial. The trial could be designed such that there is a statistical comparison of the rates of aGVHD and cGVHD between T cell-replete and TN cell-depleted HCT and a descriptive prospective evaluation of immune reconstitution, CMV reactivation, and other opportunistic infections between T cell-depleted and T cell–depleted HCT. It is likely that it will be feasible to use a uniform myeloablative conditioning, such as cyclophosphamide and TBI, for the T cell–replete
and T<sub>n</sub>-depleted treatment arms, although the more intense TBI, fludarabine, and thiopeta conditioning would be required for recipients of T cell–depleted HCT.

If randomized trials confirm the substantial reduction in cGVHD with T<sub>n</sub> depletion observed in our trial, without impairment of other important HCT outcomes, this approach could be used at most HCT centers. During the period of patient accrual in our trial, 3 small pilot studies of T<sub>n</sub> depletion in recipients of HLA-mismatched or haploidentical HCT were initiated at other centers and published (59–61). Two of these studies administered antithymocyte globulin or alemtuzumab to provide additional TCD (59, 60), and the third study incorporated total lymphoid irradiation in the conditioning regimen and NK cell infusions after HCT (61). The small sizes of these studies; differences in the patient populations, donor sources, and conditioning; and the inclusion of T cell–depleting antibodies, total lymphoid irradiation, or additional cell infusions preclude comparisons of the reported patient outcomes with the results of our trial. However, these investigators achieved reproducible rigorous depletion of T<sub>n</sub> from the graft, confirming the reliability and general applicability of the new technology.

An important question is how T<sub>n</sub> depletion will compare to other approaches being developed for GVHD prevention. Single-arm studies of the administration of cyclophosphamide after a graft composed of CD34-selected PBSCs (≥5 × 10<sup>6</sup>/kg) and CD45RA-depleted PBSCs containing a target dose of 10<sup>7</sup> CD3<sup>+</sup> T cells/kg and ≤7.5 × 10<sup>4</sup> TN/kg. The cells were infused into the patient the same day as cell selection over 1 to 4 hours, with infusion of the CD45RA-depleted cell populations were each formulated in 100 ml Normosol-R (Hospira) with 1% human serum albumin prior to infusion.

Patients and treatment. Thirty-five patients, aged 19 to 55 years, with AML, ALL, or refractory anemia with excess blasts, who were candidates for myeloablative HCT and had a MRD, were enrolled on the phase II clinical trial at FHCRCT (n = 33) or YUSM (n = 2) between December 2009 and July 2014. Eligible patients were considered by their referring physicians to require allogeneic HCT because they were judged to be at high risk of leukemia relapse following chemotherapy alone. Inclusion and exclusion criteria are detailed in the clinical trial protocol (Supplemental Methods). Patients at a very high risk of relapse after HCT due to a history of previous relapses or detectable disease immediately prior to HCT were designated “poor risk,” and those who had leukemia with high-risk cytogenetic or molecular characteristics but no prior history of relapse or detectable disease at the time of HCT were designated “better risk.” The conditioning regimen was composed of fludarabine (125 mg/m<sup>2</sup>/d), thiopeta (10 mg/kg), and TBI (1,320 cGy) (30).

Following the completion of conditioning, patients received a graft composed of CD34-selected PBSCs (≥5 × 10<sup>7</sup>/kg) and CD45RA-depleted PBSCs containing a target dose of 10<sup>7</sup> CD3<sup>+</sup> T cells/kg and ≤7.5 × 10<sup>4</sup> TN/kg. The cells were infused into the patient the same day as cell selection over 1 to 4 hours, with infusion of the CD34<sup>+</sup>-enriched cells followed immediately by infusion of the CD45RA-depleted cells. All patients received GVHD prophylaxis, with tacrolimus titrated to a serum level of 5 to 15 ng/ml and tapered after day 50 in the absence of GVHD or subsequently after GVHD.

### Methods

**Graft engineering.** The rationale and methodology for depleting T<sub>n</sub> using anti-CD45RA mAb-conjugated beads have been published previously (29). To allow precise T cell dosing and because a minor subset of CD34<sup>+</sup> stem/progenitor cells express CD45RA, we used a 2-step immunomagnetic selection procedure involving positive selection of CD34<sup>+</sup> progenitor cells, followed by depletion of CD45RA<sup>+</sup> cells from the CD34-negative fraction (29). In brief, CD34<sup>+</sup> selections were performed using the CliniMACS CD34 reagent system (Miltenyi Biotec) (33, 64), followed by depletion of CD45RA<sup>+</sup> cells from the CD34<sup>+</sup> fraction using anti-CD45RA immunomagnetic beads (Miltenyi Biotec) (see Supplemental Methods for a list of reagents). The CD34<sup>+</sup>-enriched and the CD45RA<sup>+</sup>-depleted cell populations were each formulated in 100 ml Normosol-R (Hospira) with 1% human serum albumin prior to infusion.

Patients and treatment. Thirty-five patients, aged 19 to 55 years, with AML, ALL, or refractory anemia with excess blasts, who were candidates for myeloablative HCT and had a MRD, were enrolled on the phase II clinical trial at FHCRCT (n = 33) or YUSM (n = 2) between December 2009 and July 2014. Eligible patients were considered by their referring physicians to require allogeneic HCT because they were judged to be at high risk of leukemia relapse following chemotherapy alone. Inclusion and exclusion criteria are detailed in the clinical trial protocol (Supplemental Methods). Patients at a very high risk of relapse after HCT due to a history of previous relapses or detectable disease immediately prior to HCT were designated “poor risk,” and those who had leukemia with high-risk cytogenetic or molecular characteristics but no prior history of relapse or detectable disease at the time of HCT were designated “better risk.” The conditioning regimen was composed of fludarabine (125 mg/m<sup>2</sup>/d), thiopeta (10 mg/kg), and TBI (1,320 cGy) (30).

Following the completion of conditioning, patients received a graft composed of CD34-selected PBSCs (≥5 × 10<sup>7</sup>/kg) and CD45RA-depleted PBSCs containing a target dose of 10<sup>7</sup> CD3<sup>+</sup> T cells/kg and ≤7.5 × 10<sup>4</sup> TN/kg. The cells were infused into the patient the same day as cell selection over 1 to 4 hours, with infusion of the CD34<sup>+</sup>-enriched cells followed immediately by infusion of the CD45RA-depleted cells. All patients received GVHD prophylaxis, with tacrolimus titrated to a serum level of 5 to 15 ng/ml and tapered after day 50 in the absence of GVHD or subsequently after GVHD.
resolution. We chose to use tacrolimus monotherapy, rather than a more intensive GVHD prophylaxis regimen, such as a combination of a calcineurin inhibitor with methotrexate or mycophenolate mofetil, because our intent was to evaluate T<sub>n</sub> depletion as an alternative rather than additional form of GVHD prophylaxis.

GVHD was treated according to institutional standard practice with systemic and/or topical corticosteroid administration and continuation of tacrolimus. Additional second-line GVHD therapies were permitted for the management of corticosteroid-resistant GVHD if necessary. The duration of full-dose systemic corticosteroids (0.5-2 mg/kg/d prednisone) and subsequent taper schedule were determined by the treating physician within the scope of institutional practice. Antimicrobial prophylaxis, infection definitions, monitoring and preemptive management, chimerism testing, and minimal residual disease evaluation are described in the Supplemental Methods. The clinical trial is registered at ClinicalTrials.gov (NCT 00914940). The full clinical trial protocol is available in the Supplemental Methods.

Endpoints. Primary endpoints were engraftment and the cumulative incidence of aGVHD grades II-IV. Predefined secondary endpoints were cGVHD, relapse, and day 100 NRM. An additional secondary objective was to evaluate immune reconstitution. Independent physician experts graded the peak stage of aGVHD and organ involvement and diagnosed and classified cGVHD according to the NIH consensus criteria (Supplemental Methods). Endoscopic GI and skin biopsies were performed to confirm GVHD and were graded by histological criteria (Supplemental Methods). Time to engraftment was defined as the first of 3 consecutive days with an absolute neutrophil count of ≥500 cells per mm<sup>3</sup>. Donor chimerism was monitored by molecular techniques, and infections were defined and monitored as outlined (Supplemental Methods).

Contemporary comparison cohort. A cohort of patients that underwent HCT at FHCRC on a standard treatment plan during the same time period served as a comparison group for analysis of GVHD. This cohort represented all other patients aged 14 to 55 years who received TBI-containing myeloablative T-cell-replete PBSC HCT from a HLA-MRD for the treatment of ALL, AML, or refractory anemia with excess blasts at FHCRC on a standard HCT treatment plan between April 2008 and March 2014. These patients received a standard conditioning regimen consisting of TBI (12 Gy) followed by cyclophosphamide (120 mg/kg) and GVHD prophylaxis, consisting of short-course methotrexate (15 mg/m<sup>2</sup> day 1, 10 mg/m<sup>2</sup> days 3, 6, and 11) and either tacrolimus (n = 28) or cyclosporine (n = 5). For inclusion in the comparison cohort, the patients had to meet the same age, disease status, and organ function eligibility criteria as patients treated in the T<sub>n</sub>-depleted PBSC clinical trial. The comparison group consisted primarily of patients who were unwilling or lacked insurance company approval to participate in an experimental clinical trial. Clinical trial patients and the comparison cohort received the same supportive care according to standard practice guidelines at FHCRC. Evaluations of clinical aGVHD and cGVHD were performed for T<sub>n</sub>-depleted clinical trial patients and the comparison cohort by the same expert evaluators who were not associated with the study. Histopathology of endoscopic GI and skin biopsies was assessed and graded by an expert GVHD pathologist in-house who was blinded to the transplant protocol.

Lymphocyte enumeration. Lymphocyte enumerations were performed by the Hematopathology Laboratory at the University of Washington using multicolor flow cytometry. Briefly, 100 μl of whole blood was labeled with mAbs, red blood cells were lysed (TQ-Prep; Beckman Coulter), and 10,000 mature lymphocyte events were acquired on an FC500 flow cytometer. TruCount beads (Becton Dickinson) were included and used to generate absolute counts for each population. The lymphocyte subsets were defined as follows using CXP software (Beckman Dickinson): CD8<sup>+</sup> T cells (CD8<sup>+</sup>CD3<sup>+</sup>), CD4<sup>+</sup> T cells (CD4<sup>+</sup>CD3<sup>+</sup>), B cells (CD19<sup>+</sup>), and NK cells (CD3<sup>+</sup>CD56<sup>+</sup> and/or CD16<sup>+</sup>). A separate 100-μl aliquot of sample was labeled with appropriately titrated antibodies, red blood cells were lysed using NH<sub>4</sub>Cl containing 0.25% ultra-pure formaldehyde (Polysciences) and washed once with PBS-BSA, and up to 200,000 total events were acquired on a LSRII flow cytometer (Becton Dickinson). The additional lymphocyte subsets investigated using WoodList software were as follows: CD8<sup>+</sup> T<sub>n</sub> (CD8<sup>+</sup>CD3<sup>+</sup>CD45RA<sup>+</sup>CD45RO<sup>+</sup>CD62L<sup>+</sup>), CD4<sup>+</sup> T<sub>n</sub> (CD4<sup>+</sup>CD3<sup>+</sup>CD45RA<sup>+</sup>CD45RO<sup>+</sup>CD62L<sup>+</sup>), and Tregs (CD4<sup>+</sup>CD3<sup>+</sup>CD25<sup>+</sup>CD127<sup>+</sup>). CD127 expression has been demonstrated to correlate inversely with FOXP3 expression on CD4<sup>+</sup>CD25<sup>+</sup> cells, and the CD4<sup>+</sup>CD25<sup>+</sup>CD127<sup>+</sup> phenotype is considered to be an acceptable surrogate marker for human Tregs and a practical alternative to intracellular staining for FOXP3 (65, 66). Antibodies were obtained from Beckman Coulter or Becton Dickinson. (See Supplemental Methods for a list of reagents).

TREC analysis. TREC analysis was performed in the FHCRC Immune Monitoring Shared Resource Facility. Signal-joint TREC<sup>+</sup> T cells were evaluated in peripheral blood samples obtained on days 28, 56, 80, 180, and 360 after HCT. Primers and probes for TCR β locus signal-joint TREC<sup>+</sup> T cells were synthesized according to the methods described by Douek et al. (67). DNA was extracted from PBMCs and used as a template for real-time qPCR. To generate standard curves, plasmid DNA containing TREC and β-actin DNA segments was made into serial dilutions containing 10<sup>2</sup>-10<sup>6</sup> copies per PCR reaction. Triplicate PCR reactions were run for each sample on the StepOnePlus real-time PCR system (Applied Biosystems).

Spectratyping. TCR spectratyping was performed in the FHCRC Immune Monitoring Shared Resource Facility. To assess TCR β<sub>β</sub> repertoire diversity in peripheral blood samples obtained at 6 and 12 months after HCT from T<sub>n</sub>-depleted HCT recipients, we used a multiplex PCR spectratyping method that amplifies 46 functional genes, comparing 23 TCR<sub>β</sub>β families in 5 reactions in which each reaction contains 4 to 7 specific primers, together with a single fluorescence-taggerd TCR β constant region primer (68).

Antigen-specific T cell evaluation. MHC-tetramer analysis for CMV pp65 NLPXMVATV-specific (pp65<sub>64-72</sub>-specific) T cells was performed by flow cytometry using iTag MHC tetramers (Beckman Coulter) and mAbs specific for CD3, CD8, CD28, CD27, IFN-γ<sub>γ<sub>γ</sub></sub>, and IL-2 (Becton Dickinson) (see Supplemental Methods for a list of reagents). Dead cell exclusion was performed using DAPI (Sigma-Aldrich) or Live/Dead Fixable Violet (Molecular Probes). PBMCs were surface labeled with antibodies and tetramers for 30 minutes at 4°C and evaluated on a LSRII flow cytometer. Analysis was performed using FlowJo software (Treestar). To assess function, aliquots of PBMCs were stimulated with pp65<sub>64-72</sub> peptide in the presence of anti-CD28 and anti-CD49a costimulatory mAbs (5 μl/ml; BD Biosciences). Brefeldin A (1 μl/ml Goligplug; BD Biosciences) was added 1.5 hours into the stimulation. After 6 hours, cells were stained with Live/Dead Fixable Violet, fixed, and permeabilized (Cytofix/Cytoperm, BD Biosciences) and then stained with fluorescent protein-conjugated mAbs against IFN-γ<sub>γ<sub>γ</sub></sub>, IL-2,
CD4, and CD8 (BD Biosciences) (see Supplemental Methods for a list of reagents) in Perm/Wash buffer (BD Biosciences), before washing and analysis on the flow cytometer.

**Statistics.** Data were analyzed as of December 2014. The protocol was designed with enrollment and grades II–IV aGVHD as the primary endpoints. We reviewed the FHCRC clinical research databases and derived estimates of the incidence of grade II–IV (60%) and III–IV (19%) aGVHD in patients undergoing HLA-MRD myeloablative HCT. Thirty-five patients provided 92% power to observe a statistically significant (1-sided significance level of 0.05) reduced probability of GVHD relative to the fixed rate of 60%, under the assumption that the true probability of grades II–IV GVHD is 35%. A 1-sided binomial test was performed in order to test the null hypothesis that the true rate of grade II–IV aGVHD is equal to the fixed rate of 60%. A P value of less than or equal to 0.05 was considered significant. Stopping rules were created such that the trial would stop prior to the accrual of 35 patients if the true probability of graft failure exceeded 5%. aGVHD was a predetermined secondary endpoint of the study. Probabilities of overall survival and DFS were estimated with the Kaplan-Meier method. Probabilities of death not preceded by relapse, recurrent malignancy, and GVHD were summarized with the use of cumulative incidence estimates, with recurrent malignancy viewed as a competing risk for death not preceded by relapse, with death not preceded by relapse viewed as a competing risk for recurrent malignancy, and with death without GVHD viewed as a competing risk for GVHD. Probabilities of discontinuation of systemic corticosteroids and of discontinuation of all systemic immune suppression were also summarized with cumulative incidence estimates, with death while still on corticosteroids or on any systemic immunosuppression viewed as a competing risk for discontinuation of corticosteroids or any systemic immune suppression, respectively. Statistical analyses of clinical outcomes were conducted using SAS 9.3 for Windows (SAS Institute).

**Study approval.** FHCRC and YUSM IRBs and the US FDA (Investigational Device Exemption 14160) approved the trial. Patients and donors provided written informed consent in accordance with the Declaration of Helsinki. A data safety monitoring board and an independent clinical trial monitor provided additional oversight. A concurrent cohort of patients undergoing HCT at FHCRC on a standard HCT treatment plan served as a comparison group for analysis of GVHD. The comparison group patients consented to review of their medical records and pathology.

**Acknowledgments**

The authors would like to thank Mary Flowers, Paul Martin, Brent Wood, Diane Krause, Taiga Nishihori, Lazarus Lekakis, Jan Storek, Steven Devine and the Blood and Marrow Transplant Clinical Trials Network 0303 study team, Andrew Mackie, Su Yi, Wendy Haskell, Julia Richard, Sarah Click, Tanya Cunningham, Nick Calores, Jianhong Cao, Mason Lai, Tim Waters, Susan Lemmon, Amanda Peffer, Caroline McCallor, the Yale Clinical Trials Office staff, and members of the Data and Safety Monitoring Boards for their assistance. This work was supported by NIH grants CA18029, CA 136551, CA15704, HL121568-01A1, and DK56465 and by the NIH Rapid Access to Intervention Development (project 298). W. Shlomchik was the recipient of support from the Burroughs Wellcome Fund and was a Leukemia and Lymphoma Society Clinical Scholar. M. Bleakley is the Damon Runyon-Richard Lumsden Foundation Clinical Investigator, supported in part by the Damon Runyon Cancer Research and Richard Lumsden Foundations (CI-57-11), and received support from the National Cancer Institute (K23CA154532) and previously from a Special Fellowship in Clinical Research from the Leukemia and Lymphoma Society. Support for Blood and Marrow Transplant Clinical Trials Network 0303 (TCD) data shown in Supplemental Table 4 was provided by NIH (National Heart, Lung, and Blood Institute, National Cancer Institute) grant U10HL069294.

**Address correspondence to:** Marie Bleakley, Program in Immunology, Clinical Research Division, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. N., Seattle, Washington 98109, USA. Phone: 206.667.6572; E-mail: mbleakle@fredhutch.org.

16. Sakaguchi S, Miyara M, Costantino CM,


