Full-length article
Stem cell therapy

Early immune reconstitution as predictor for outcomes after allogeneic hematopoietic cell transplant; a tri-institutional analysis

Alexandre G. Troullioud Lucas[12], Caroline A. Lindemans[2,3], Senthil Velan Bhoopalan[4], Rana Dandis[2], Susan E. Prockop[5], Swati Naik[6], Dinesh Keerthi[6], Coco de Koning[7], Akshay Sharma[6,*], Stefan Nierkens[2,7,*], Jaap Jan Boelens[1,*,**]

[1] Transplantation and Cellular Therapy, MSK Kids, Department of Pediatrics, Memorial Sloan Kettering Cancer Center, New York, New York, USA
[2] Princess Máxima Center for Pediatric Oncology, Utrecht, the Netherlands
[3] Department of Pediatrics, University Medical Center Utrecht, Utrecht, the Netherlands
[4] Department of Hematology, St. Jude Children's Research Hospital, Memphis, Tennessee, USA
[5] Pediatric Stem Cell Transplantation, Boston Children's Hospital/Dana-Farber Cancer Institute, Boston, Massachusetts, USA
[6] Department of Bone Marrow Transplantation and Cell Therapy, St. Jude Children's Research Hospital, Memphis, Tennessee, USA
[7] Center for Translational Immunology, University Medical Center Utrecht, Utrecht, the Netherlands

ARTICLE INFO

ABSTRACT

Background aims: CD4 immune reconstitution (IR) after allogeneic hematopoietic cell transplant (allo-HCT) correlates with lower non-relapse mortality (NRM), but its impact on leukemia relapse remains less clear, especially in children. We studied the correlation between IR of lymphocyte subsets and HCT outcomes in a large cohort of children/young adults with hematological malignancies.
Methods: We retrospectively analyzed CD4, CD8, B-cell and natural killer (NK) cell reconstitution in patients after first allo-HCT for a hematological malignancy at three large academic institutions (n = 503; period 2008−2019). We used Cox proportional hazard and Fine−Gray competing risk models, martingale residual plots and maximally selected log-rank statistics to assess the impact of IR on outcomes.
Results: Achieving CD4 >50 and/or B cells >25 cells/μL before day 100 after allo-HCT was a predictor of lower NRM (CD4 IR: hazard ratio [HR] 0.26, 95% confidence interval [CI] 0.11−0.62, *P* = 0.002; CD4 and B cell IR: HR 0.06, 95% CI 0.03−0.16, *P* < 0.001), acute graft-versus-host disease (GVHD) (CD4 and B cell IR: HR 0.02, 95% CI 0.01−0.04, *P* < 0.001) and chronic GVHD (CD4 and B cell IR: HR 0.16, 95% CI 0.05−0.49, *P* = 0.001) in the full cohort, and of lower risk of relapse (CD4 and B cell IR: HR 0.24, 95% CI 0.06−0.92, *P* = 0.038) in the acute myeloid leukemia subgroup. No correlation between CD8 and NK-cell IR and relapse or NRM was found.
Conclusions: CD4 and B-cell IR was associated with clinically significant lower NRM, GVHD and, in patients with acute myeloid leukemia, disease relapse. CD8 and NK-cell IR was neither associated with relapse nor NRM. If confirmed in other cohorts, these results can be easily implemented for risk stratification and clinical decision making.

Introduction

Allogeneic hematopoietic cell transplantation (allo-HCT) is a curative treatment option for patients with high-risk hematological malignant diseases. Outcomes after allo-HCT have significantly improved in the last few decades, mainly due to improvement in the use of alternative cell sources, patient and donor selection, supportive care measures as well as conditioning regimens [1]. Despite these advances, complications after allo-HCT, such as opportunistic infections, graft-versus-host disease (GVHD) and disease relapse, remain major causes of treatment failure and mortality. Early CD4 immune reconstitution (IR), defined as a total CD4 count >50 cells/μL before day 100 after allo-HCT (CD4 IR), has been found to be associated with better overall survival (OS), event-free survival (EFS) and decreased non-relapse−related mortality (NRM) in various transplant settings [2,3]. NRM, largely secondary to infections and death due to acute GVHD (aGVHD), is greater in patients who do not achieve CD4 IR before 100 days after allo-HCT [4,5].

** Correspondence: Jaap Jan Boelens, Chief, Transplantation and Cellular Therapies, MSK Kids, Memorial Sloan Kettering Cancer Center, 1275 York Ave., New York, NY 10065.
 E-mail address: boelensj@mskcc.org (J.J. Boelens).
 * These authors contributed equally to this work.

The association between IR and risk of relapse has however been less clear. In one study, cumulative incidence of relapse-related mortality was shown to be significantly lower in patients achieving early CD4 IR after allo-HCT for acute myeloid leukemia (AML; hazard ratio [HR] 0.25, n = 63, P = 0.015) but not for acute lymphoblastic leukemia (ALL) [2]. Another study found no association between CD4 IR and relapse risk in patients after allo-HCT for malignant disease [3]. The discrepancy in results might be due to differences in number of patients included or because there is only a significant impact of CD4 IR on relapse in particular diagnoses.

In this study, we describe our analysis on a large cohort of pediatric and young adult patients who received their first allo-HCT for a hematological malignancy, by combining data from three large academic institutions. Due to the large size of this cohort, we were able to use unique statistical models, such as martingale residual plots and maximally selected log-rank statistics, to explore whether there was an association between reconstitution of specific lymphocyte subsets and post allo-HCT outcomes, mainly NRM and relapse incidence. Thus, using these various statistical models, we wanted to (re)analyze associated immune cell markers with outcomes, including the optimal cell marker cut off to predict these outcomes. We focused these analyses on markers that are available at most transplant centers as Clinical Laboratory Improvement Amendments−approved assays, e.g., CD4, CD8, B and natural killer (NK) cell IR. Hence, they can potentially be used for clinical decision-making.

Materials and Methods

Methods

All consecutive patients undergoing their first allo-HCT for any hematological malignancy at Memorial Sloan Kettering Cancer Center (New York, NY, USA), University Medical Center Utrecht/Princess Máxima Center for Pediatric Oncology (Utrecht, the Netherlands) and St. Jude Children's Research Hospital (Memphis, TN, USA) between January 1, 2008, and December 31, 2019, were included in this study. Clinical and IR data were prospectively collected. Immune monitoring was not a standard procedure from the start of this inclusion period but was implemented at different time points in the different centers or was part of certain protocols. For the analyses studying linearity between variables and outcomes, patients without lymphocyte subset measurements before day 100 were excluded. For multivariable (MV) analyses, CD4 IR was defined as HCT recipients achieving a CD4 count of >50 cells/μL on two consecutive measures within 100 days of HCT [2,4,5]. If there was one measurement >50 cells/μL within 2 months of the first measurement, we considered the first time point the time of achieving CD4 IR. If there was only one measurement or no second measurement >50 cells/μL, we considered that CD4 IR was not achieved. Patients without CD4 measurements or no second measurement within 2 months of the first and those who died before the median time of CD4 IR were excluded from multivariable analyses. Any new immune cell marker we found to impact outcomes we would define similarly as we did CD4 IR, based on the found threshold. In this case, we would also adjust exclusion criteria by excluding patients who died before the median time to IR of the marker that would reconstitute the latest. Stem cell donors were defined as matched donors (human leukocyte antigen [HLA] match 10/10, 8/8 [bone marrow {BM} and peripheral blood stem cells {PBSCs}] or 6/6, cord blood [CB]), including matched related donors and matched unrelated donors; or mismatched donors (HLA match ≤9/10, or ≤5/6 for CB). Patients with Fanconi anemia were excluded from these analyses. The study was approved by the institutional review boards or Biobank of all three institutions. For these analyses, the need for a specific informed consent was waived.

Outcomes

Main outcomes of interest were incidence of relapse and NRM. NRM was defined as death due to any cause other than relapse and patients were censored at time of relapse. Other outcomes of interest were OS, EFS and GVHD. OS was defined as time from allo-HCT to death from any cause or last follow-up for survivors, and EFS was defined as time from allo-HCT to graft failure, relapse, death or last follow-up for patients who did not experience an event. We classified aGVHD as defined by Center for International Blood and Marrow Transplant Research criteria [6] and extensive chronic GVHD as defined by the National Institutes of Health [7].

Statistical analysis

Continuous variables are displayed as median and range or interquartile range. Discrete variables are described as counts and proportions. A univariable Cox proportional hazard (PH) model was used to explore the association between continuous variables and outcomes. These models assume linearity between variables and outcomes. To test whether the linearity assumption held up for our lymphocyte IR data, we fitted Cox PH models with maximum total CD4, CD8, B or NK-cell count within 100 days after allo-HCT as single continuous markers and NRM or relapse as outcomes, followed by computing the martingale residuals to visually check for linearity [8]. If linearity was confirmed, we trusted the HRs we got from the Cox PH models. If the martingale plots indicated there was no linearity between subset numbers and outcomes, we used maximally selected log-rank statistics to determine whether there was a specific cut point of subset count, above which the risk of an outcome would change. This outcome-oriented method could provide a cut point value that corresponds to the most significant relationship with the outcome [9].

MV Cox PH models were then used to evaluate the impact of various variables on the outcomes of interest, including cut points of lymphocyte IR determined in our linearity analyses. In addition to IR data, other variables considered based on clinical relevance were age, primary diagnosis, disease status (for AML and ALL), graft source, HLA match, conditioning regimen (total body irradiation−based versus chemotherapy-containing only), use of anti-thymocyte serotherapy, *ex vivo* graft manipulation and transplant before or after the median year of transplant (2015). To avoid overfitting with too many covariates, we also analyzed MV models including different subsets with smaller number of covariates based on the number of events for each outcome studied. The results are presented as HR, 95% confidence intervals (95% CIs) and log-likelihood test P values. For analysis of cumulative incidences of our main outcomes of interest relapse and NRM, as well as other outcomes of interest aGVHD and cGVHD Fine−Gray models for competing risk were used. Kaplan−Meier plots were used to compare the other outcomes of interest OS and EFS.

All statistical analyses were done using R statistical software, version 4.2.1 (R Foundation for Statistical Computing, Vienna, Austria), packages: *tidyverse, survival, survminer, prodlim, cmprsk, riskRegression, forestmodel*.

Results

Patient and transplant characteristics

We identified a total of 890 consecutive patients who had their first allo-HCT for a hematological malignancy between 2008 and 2019. Based on IR data availability per lymphocyte subset, 503 patients with CD4, 491 with CD8, 499 with B cell and 481 with NK cell data before day 100 after allo-HCT were included for the exploration of linearity between lymphocyte subsets and
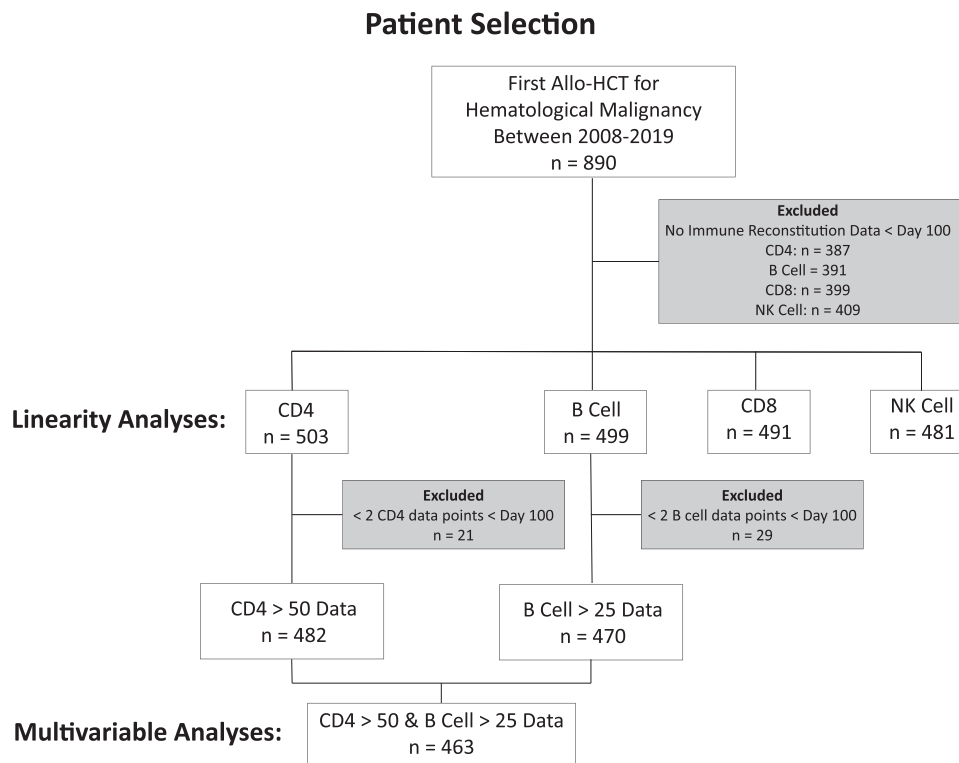
Patient Selection

Figure 1. Retrospective selection of patients for the analyses. This diagram depicts that from the 890 patients who underwent first allo-HCT for a hematologic malignancy, 387 (CD4), 391 (B cell), 399 (CD8) and 409 (NK cells) were excluded from linearity analyses, due to lack of immune reconstitution data before day 100 after allo-HCT. Moreover, from the 503 patients with CD4 data and 499 with B-cell data, 21 and 29 patients were excluded from the multivariable analyses, respectively, due to having fewer than two datapoints before day 100. Combining only patients who had both CD4 and B-cell data for the multivariable analyses, left us with a total of 463 patients.

outcomes (Figure 1). In total, 463 patients had sufficient CD4 and B-cell lymphocyte subset measurements to use in MV models analyzing cut points found in the maximally selected log-rank methods as described in the results below. Subsets of patients have been reported previously; from the Princess Máxima Center, 109 patients by Admiraal *et al.* [4], and 139 (overlapping the 109 described in Admiraal *et al.*) by de Koning *et al.* [5], while from Memorial Sloan Kettering, 220 patients by van Roessel *et al.* [3] and de Koning *et al.* [5]. After we excluded patients who died before the median time to CD4 and B-cell IR, patients who were excluded due to insufficient IR data had no significantly different cumulative incidence of NRM in MV Fine−Gray analysis (P = 0.23, supplementary Figure 1 and supplementary Table 1). Patient and HCT characteristics for the entire cohort included for MV analyses are summarized in Table 1. Median age at transplant was 12.1 (interquartile range 5.6−16) years. Most of the patients included were transplanted for ALL (n = 202 [43.6%]) or AML (n = 171 [36.9%]).

Testing for linearity between absolute lymphocyte numbers and outcomes

The martingale residual plots depicting excess risk of NRM related to the greatest obtained absolute CD4 count before day 100 after HCT, suggested that there was no linear association between CD4 count and risk of NRM (Figure 2a). Maximally selected log-rank statistics indicated that there might be a transition point around the CD4 count of 57 cells/μL (range 20−80 cells/μL), demonstrating that achieving a CD4 count above this within 100 days was associated with decreased risk of NRM (Figure 2b,c). Competing risk analyses separating patients who achieved CD4 count of 57 or more, compared with the patients who did not, confirmed increased cumulative

incidence of NRM in the latter (HR 0.27, 95% CI 0.17−0.44, P < 0.001; Figure 2d).

Martignale residual plots for the excess risk or NRM related to maximum absolute B-cell count before day 100 after HCT also indicated non-linearity between early B-cell reconstitution and NRM (Figure 3a). Here maximally selected log-rank statistics indicated that there might be a transition point around the B-cell count of 25 cells/μL, demonstrating that achieving a B-cell count above this within 100 days was associated with decreased risk of NRM (Figure 3b,c). Competing risk analyses separating patients who achieved B cell count of 25 or more, compared with the patients who did not, confirmed increased NRM in the latter (HR 0.26, 95% CI 0.16−0.43, P < 0.001; Figure 3d).

The martingale plots depicting excess risk of NRM related to maximum absolute CD8 and NK cell counts in the first 100 days after transplant showed a linear association between variables and outcome (Figure 4a,b). Univariable Cox PH models with these lymphocyte values as continuous variable and NRM as outcome, showed no statistically significant association (CD8: HR 0.999, 95% CI 0.998−1, P = 0.046; NK: HR 0.9997, 95% CI 0.999−1, P = 0.381), indicating that IR of these cell subsets within 100 days after allo-HCT was not associated with NRM.

Excess risk of relapse related to maximum absolute CD4, CD8, B and NK cells in the first 100 days after transplant showed a linear association between variables and outcome in the martingale plots (Figure 5a-d). Cox PH models with these lymphocyte counts as continuous variable and relapse as outcome, showed no statistical correlation (CD4: HR 1, 95% CI 0.9996−1.001, P = 0.407; CD8: HR 1, 95% CI 0.9997−1, P = 0.943; B cell: HR 0.998, 95% CI 0.997-0.999, P = 0.003, NK cell: HR 1, 95% CI 0.9996−1, P = 0.926), indicating that IR of these cell subsets within 100 days after allo-HCT was not associated with relapse.

Table 1
Patient and HCT characteristics for the entire cohort included for MV analyses.

	Included patients (N = 463)
Age, y	
Median [IQR]	12.1 [5.6, 16]
Treatment center	
MSKCC	227 (49.0%)
PMC	149 (32.2%)
SJCRH	87 (18.8%)
Sex	
Male	274 (59.2%)
Female	189 (40.8%)
Diagnosis	
ALL	202 (43.6%)
AML	171 (36.9%)
MDS	41 (8.9%)
NHL	17 (3.7%)
Other[a]	32 (6.9%)
Cell source	
BM	155 (33.5%)
Cord	120 (25.9%)
PBSC	188 (40.6%)
HLA matching	
Matched	245 (52.9%)
Mismatched	214 (46.2%)
Missing	4 (0.9%)
Conditioning regimen	
Only chemotherapy	282 (60.9%)
Chemotherapy and TBI	181 (39.1%)
***Ex vivo* TCD**	
No	271 (58.5%)
Yes	192 (41.5%)
Serotherapy	
No	212 (45.8%)
Yes	251 (54.2%)
Year of HCT	
Median [min, max]	2015 [2008, 2019]

ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; BM, bone marrow; HCT, hematopoietic cell transplant; HLA, human leukocyte antigen; IQR, interquartile range; MDS, myelodysplastic syndrome; MSKCC, Memorial Sloan Kettering Cancer Center; MV, multivariable; NHL, non-Hodgkin lymphoma; PBSC, peripheral blood stem cells; PMC, Princess Maxima Center for Pediatric Oncology; SJCRH, St. Jude Children's Research Hospital; TBI, total body irradiation; TCD, T-cell depletion.

[a] Other diagnoses included were acute undifferentiated leukemia, chronic myelogenous leukemia, chronic neutrophilic leukemia, Hodgkin lymphoma, juvenile myelomonocytic leukemia, mixed phenotype acute leukemia and myeloid sarcoma.

CD4 and B-cell IR and outcomes

Based on the aforementioned results, we continued using the previously described definition for CD4 IR as a total CD4 count >50 cells/μL before day 100 after HCT. In addition, we also defined B-cell IR as a total B-cell count >25 cells/μL before day 100 after HCT. For the aGVHD work, we analyzed patients who achieved CD4 and/or B-cell IR after developing aGVHD as if they had not achieved IR. For the overall cohort, the cumulative incidence of CD4 IR at day 100 after transplant was 77.4% (95% CI 73.6−81.2%). On MV analyses with the age 0−3 years group as reference, ages between 10 and 18 years (HR 0.47, 95% CI 0.34−0.65, P < 0.001) and >18 years (HR 0.47, 95% CI 0.31−0.73, P < 0.001) were associated with a lower cumulative incidence of CD4 IR (Figure 6a). The 100-day cumulative incidence of B cell IR was 74.4% (95% CI 70.4−78.4%). Compared with age group 0−3 years, age groups 3−10 (HR 0.7, 95% CI 0.50−0.98, P = 0.04), 10−18 (HR 0.54, 95% CI 0.39−0.75, P < 0.001) and >18 years (HR 0.62, 95% CI 0.41−0.94, P = 0.02) were all associated with a lower cumulative incidence of B-cell IR (Figure 6b).The 2-year cumulative incidence of NRM was 3.7% (95% CI 1.5−6.0%) for patients who

achieved both CD4 and B-cell IR, 14.2% (95% CI 6.7−21.7%) for only CD4 IR, 22.3% (95% CI 12.6−32%) for only B-cell IR and 34.4% (95% CI 17.6−51.2%) for no IR at all (Figure 7a). In MV analyses, compared with no IR at all, CD4 IR (HR 0.26, 95% CI 0.11−0.62, P = 0.002) and CD4 and B-cell IR together (HR 0.06, 95% CI 0.03−0.16, P < 0.001) were significant predictors for decreased incidence of NRM (supplementary Table 2). For relapse, no statistically significant difference was seen based on CD4 and/or B-cell IR in the full cohort (Figure 7b and supplementary Table 3). Age >18 years (HR 0.32, 95% CI 0.13−0.81, P = 0.02) and transplant after 2015 (HR 0.63, 95% CI 0.41−0.97, P = 0.03) were found to be the only predictors for a decreased relapse risk (supplementary Table 3).

In the subgroups of patients with AML achieving CD4 and B-cell IR together was found to be a significant predictor of decreased incidence of relapse (HR 0.24, 95% CI 0.06−0.92, P = 0.038; supplementary Table 4). In the group of patients with ALL, CD4 and/or B-cell IR was not found to be a predictor of relapse (supplementary Table 5). For the AML subgroup, active disease (HR 3.58, 95% CI 1.46-8.78, P = 0.005) and use of serotherapy alone (HR 2.48, 95% CI 1.02−6, P = 0.045) were predictive of increased risk of relapse, whereas transplant after the year 2015 (HR 0.35, 95% CI 0.16−0.78, P = 0.011) was predictive for decreased incidence of relapse (supplementary Table 4). For ALL only, disease status of complete remission 2 or greater (HR 3.21, 95% CI 1.51−6.83, P = 0.002) was predictive of greater risk of relapse (supplementary Table 5). Patients who achieved CD4 IR (HR 0.17, 95% CI 0.11-0.27, P < 0.001), B-cell IR (HR 0.10, 95% CI 0.05−0.19, P < 0.001) or both (HR 0.02, 95% CI 0.01−0.04, P < 0.001) had lower risk of aGVHD grade II-IV (Figure 7c and supplementary Table 6). Graft source PBSC and serotherapy with or without *ex vivo* TCD were also predictive of lower incidence of aGVHD grade II-IV (supplementary Table 6). For patients who achieved CD4 IR or B-cell IR, there appeared to be a lower cumulative incidence of extensive cGVHD, compared with patients who did not, but this did not reach statistical significance at a 0.05 threshold, whereas patients who achieved CD4 and B-cell IR together did have a statistically significant decreased risk of extensive cGVHD (HR 0.16, 95% CI 0.05−0.49, P = 0.001; Figure 7d and supplementary Table 7). None of the other variables analyzed were found to be predictors of extensive cGvHD (supplementary Table 7).

Discussion

In our large tri-institutional cohort, we confirmed, using various refined statistical models, the strong association between CD4 IR and risk of NRM in pediatric/young adult patients with hematological malignancies. To that knowledge, we added that early B-cell reconstitution, defined as a total B-cell count of >25 cells/μL before day 100 after allo-HCT is associated with decreased NRM, and that achieving both CD4 and B-cell IR correlates with even better outcomes with very low cumulative incidence of NRM (at 2 years 3.7%, 95% CI 1.5−6%), which is around a 15- to 20-fold lower risk compared with patients with no CD4 and B cell IR at all and 4- to 5-fold lower risk compared with patients with only CD4 or B-cell IR seperately. In the AML subgroup, we found that having patients reconstitute both CD4 and B cells correlated to a 4-fold decrease in risk of relapse. We also found that patients who achieved CD4 or B-cell IR had a more than 5-fold decreased risk of aGVHD grade II-IV, and that this risk becomes even lower in patient who reconstitute both. IR of both CD4 and B cells together was also correlated to more than 5-fold lower risk of extensive cGVHD. No correlation between CD8 or NK cell IR and NRM or relapse was found.

The results we report are of significance for the field. However, our analyses are limited by their retrospective nature, with a high number of patients excluded from analyses due to the lack of IR data points. We did show that the cumulative incidence of NRM was not different between the group of included versus excluded patients. A

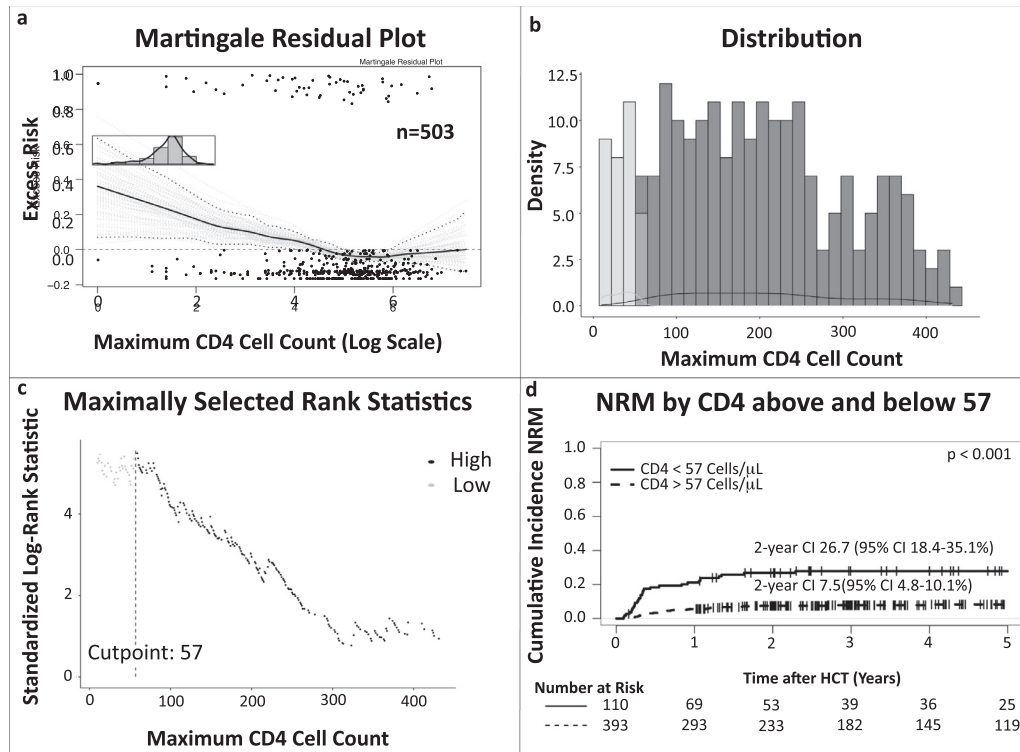## CD4 Immune Reconstitution and Non-Relapse Mortality

Figure 2. Linearity and maximally selected log-rank statistics analyses on the correlation between CD4 immune reconstitution and NRM. (a) Martingale residual plot demonstrating excess risk of NRM by highest achieved absolute CD4 count in the first 100 days after allo-HCT. The non-linear trajectory of the curve indicates there is no linear correlation between CD4 cell count and non-relapse mortality, (b) distribution of the CD4 values, (c) maximally selected log-rank statistics determining an absolute CD4 count of 57 as the cut off for NRM, (d) cumulative incidence curves of NRM comparing patients who achieved a maximum CD4 count >57 (dotted line) and those who did not (solid line), confirming a significant difference between the two groups. *P* value depicts significance of CD4 IR >57 as predictor in univariable competing risk analyses by the Fine−Gray method.

limitation of MV analyses is that including many covariates might lead to overfitting, when analyzing endpoints with limited number of events. However, by analyzing the MV analyses with smaller numbers of covariates for the outcomes with lower number of events, we confirmed that our overall results did not change because of having too many covariates (supplementary data). For this reason, and because of the clinical relevance of the covariates we included, we decided to report on the variables mentioned in the Materials and Methods. The strength of our work is that it is representative of a heterogenous large cohort of pediatric and young adult patients from three independent pediatric transplant centers with very diverse patient bases as well as clinical practices. Due to the size of the cohort, we have been able to use more refined statistical analyses as described above, to support our conclusions related to IR and outcomes.

Using several orthogonal statistical models, we consistently demonstrated that CD4 IR is associated with NRM. The previously described CD4 >50 cells/μL before day 100 is in line with what we found (CD4 >57 cells/μL) in the present analyses using different statistical methods. In contrast to CD4 IR, B-cell IR after allo-HCT has not been found to be predictive of outcomes in the context of viral reactivations [4], reason for which the focus previously has been more on CD4 IR. Ando *et al.* [10] described the impact of IR on survival outcomes in a cohort of 358 adult patients who underwent allo-HCT for hematological malignancies and also found that CD20+ B-cell counts correlated with lower NRM. However, they did not describe a threshold. In that study subset analyses were done on samples from day 28, day 100 and beyond. We confirmed that achieving a CD4 count of 50 cells/μL or greater and B-cell count of 25 cells/μL or greater within 100 days after transplant is a good and easy to use predictor for NRM. The lower incidence of NRM is in part due to less deaths secondary to

viral reactivations but is also associated with a decreased risk of aGVHD grade II-IV and extensive cGVHD. Interestingly, we showed combined CD4 and B-cell IR to be a very strong predictor of decreased incidence of aGVHD grade II-IV in our cohort, suggesting that IR of these subsets (inter)plays an important role in preventing aGVHD. Also, combined CD4 and B cell IR was a statistically significant predictor of decreased incidence of extensive cGVHD, and CD4 and B-cell IR individually had *P* values close to significance at a threshold of 0.05, with an HR with broad CI, not centered around 1, suggesting a lack of power, rather than effect.

It has been previously shown that pediatric patients who developed grade II-IV, even grade III-IV aGVHD, and had >50 CD4 cells/μL within 100 days after allo-HCT or before aGVHD onset had significantly decreased cumulative incidence of NRM (similar to patients who did not develop GVHD) [5]. The found association between CD4 IR and reduced probability of aGVHD was also recently described in a large T-cell−depleted cohort [11]. This may be explained by (i) fewer infections, which can be a trigger for aGVHD and/or (ii) better immune regulation secondary to the early IR, possibly with the involvement of T-regulatory cells in the CD4 population. This highlights the importance to further study the role of specific immune cell subsets on risk of developing and regulating GVHD.

The anti-leukemic effect of allo-HCT has been of interest to investigators for decades. Although this has been clearly demonstrated in patients with chronic myeloid leukemia, there is no strong convincing data for AML and ALL [12,13]. Moreover, no early cellular immune marker has been described yet. In these analyses as well, we were not able to find an association between any of the early immune cell markers (CD4, CD8, B and NK-cell reconstitution) and relapse, except for the subgroup of patients with AML who achieved both CD4 and B-cell IR. These results highlight the importance to

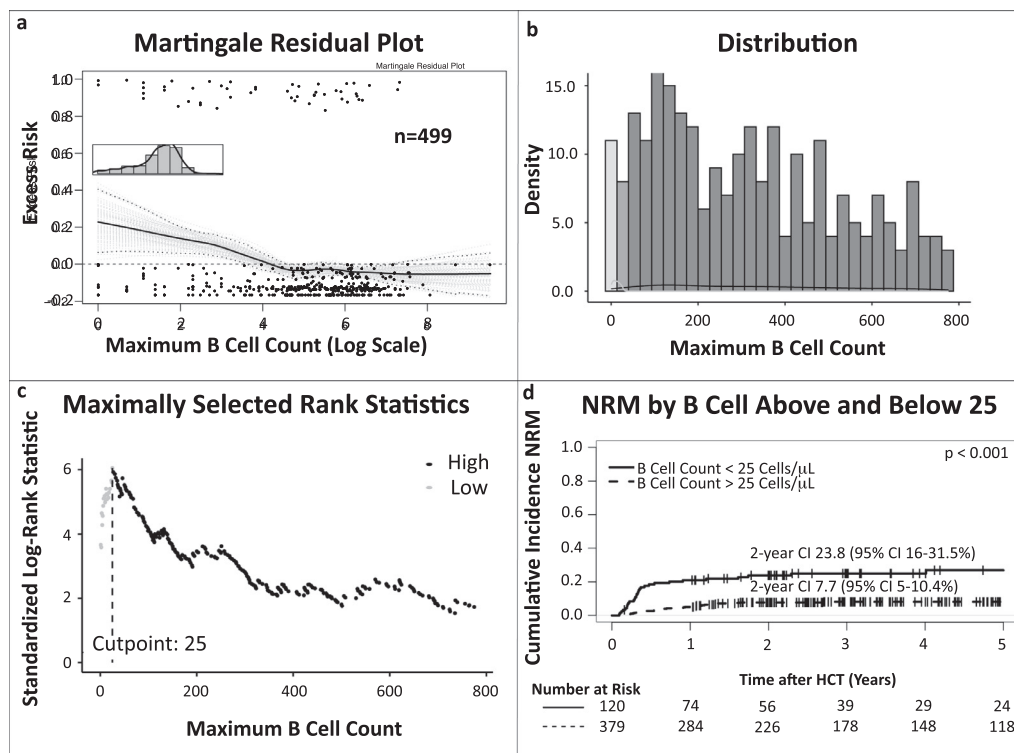## B Cell Immune Reconstitution and Non-Relapse Mortality

Figure 3. Linearity and maximally selected log-rank statistics analyses on the correlation between B-cell immune reconstitution and NRM. (a) Martingale residual plot demonstrating excess risk of NRM by highest achieved absolute B cell count in the first 100 days after allo-HCT. The non-linear trajectory of the curve indicates there is no linear correlation between B-cell count and NRM, (b) distribution of the B-cell values, (c) maximally selected log-rank statistics determining an absolute B-cell count of 25 as the cut off for NRM, and (d) cumulative incidence curves of NRM comparing patients who achieved a maximum B cell count >25 cells/μl (dotted line) and those who did not (solid line), confirming a significant difference between the two groups. P value depicts significance of B cell IR > 25 cells/μl as predictor in univariable competing risk analyses by the Fine−Gray method.

confirm our findings in other cohorts of patients with AML and to further analyze more rare lymphocyte subsets within the CD4 and B-cell populations, to have a better understanding of a possible mechanism behind prevention of relapse. Others have found (in a relatively small cohort) that higher IR of gamma/delta T cells, which, in contrast to CD4, CD8 T cells and NK cells, don't make use of MHC class molecules for cell recognition and destruction, has been associated with decreased cumulative incidence of relapse of malignant disease [14]. To decipher the role of more specific subsets, future studies should include big

discovery panels (30+ color flow cytometry, including functional markers), in large cohorts of patients.

In conclusion, in this large tri-institutional cohort we confirmed the strong association between CD4 IR and NRM, but also showed that B-cell IR, in combination with CD4 IR, leads to even lower NRM. In patients with AML, we showed a 4-fold decreased incidence of relapse with combined CD4 and B cell IR. There was no association between CD8 or NK cell IR and NRM. We also showed a significant decrease in probability of aGVHD grade II-IV, with combined CD4 and B-cell IR having the greatest impact. For extensive cGVHD we showed

CD8 and NK Cell Immune Reconstitution and Non-Relapse Mortality
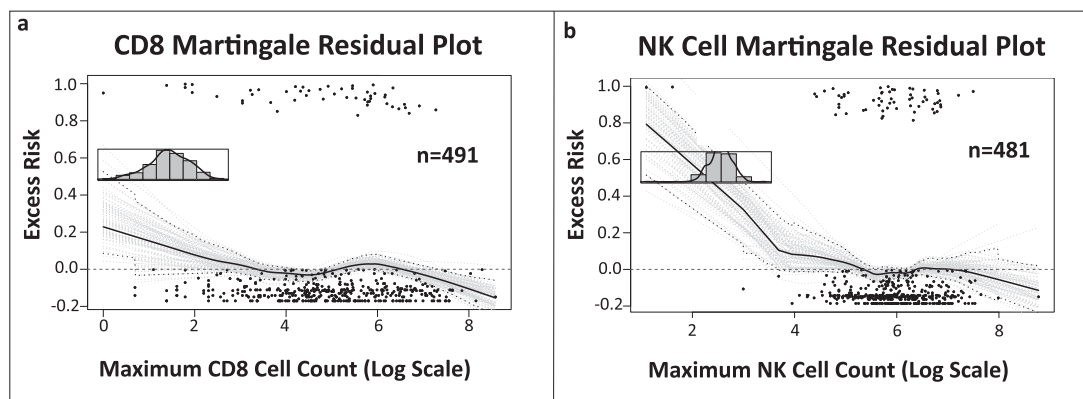

Figure 4. Martingale residual plots depicting linearity between CD8 and NK cell immune reconstitution and NRM. Martingale residual plots demonstrating excess risk of NRM by (a) highest achieved absolute CD8 and (b) NK-cell count in the first 100 days after allo-HCT. Although both curves seem to indicate non-linearity, the initial slope downwards of the curves has a wide confidence interval due to small number of measurements. When looking at the confidence interval in the area with most measurements, the 0.0 excess risk line continues to be within the confidence interval, indicating a linear correlation.

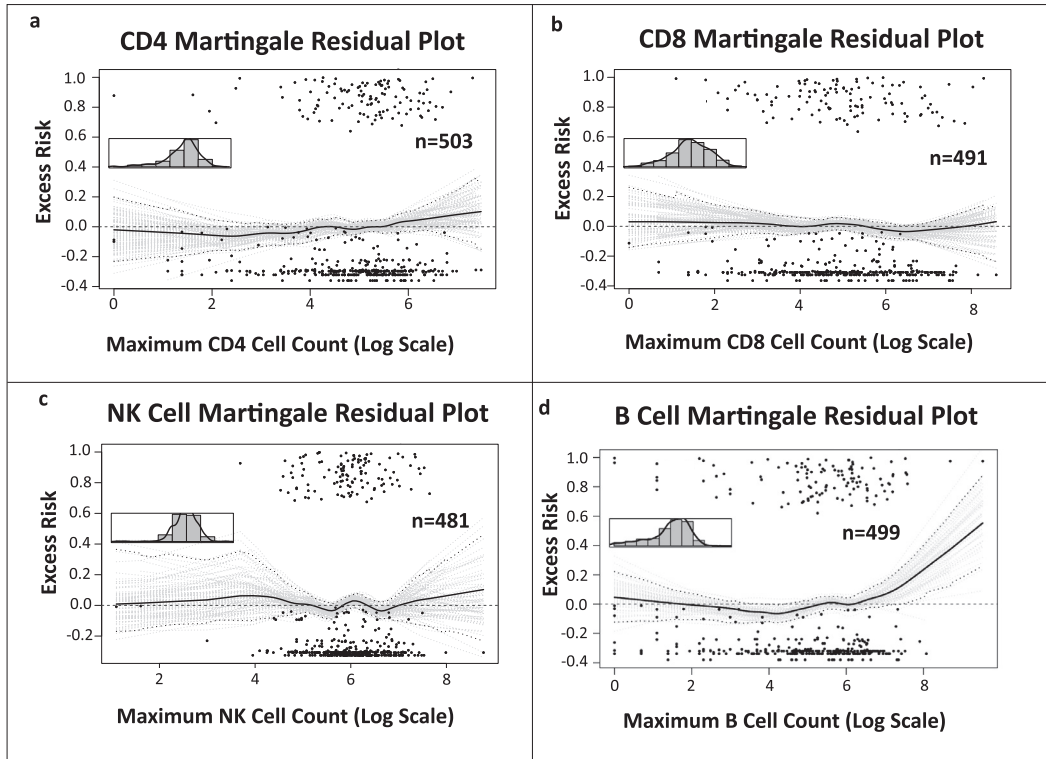## Immune Reconstitution and Relapse

Figure 5. Linearity analyses on the correlation between CD4, CD8, NK and B-cell immune reconstitution and relapse. Martingale residual plots demonstrating excess risk of relapse by highest achieved absolute (a) CD4, (b) CD8, (c) NK and (d) B-cell count in the first 100 days after allo-HCT. The curves in (a-c) show the 0.0 excess risk line falls within the confidence interval throughout, indicating linearity between variables and relapse. Although in panel d the curve seems to have an upward slope, that is due to two outliers. In the area of most datapoints, the 0.0 excess risk line falls within the confidence interval, indicating linearity.

a more than 5-fold decreased incidence when patients had both CD4 and B cell IR. No associations between CD4, CD8, B or NK cell IR and relapse were found in our full cohort or patients with ALL. Given the huge impact on NRM, strategies that promote early IR must be prioritized to make HCT safer by reducing NRM (better infectious control and less GVHD). The newly reported impact of B-cell IR in this work should be confirmed in other cohorts, to potentially serve as an easily applicable prediction tool for clinical practice. Prospective validation studies and detailed evaluation of specific, less common/rare lymphocyte subsets during the early and later post-transplant period, may help decipher the mechanisms of leukemia control after allo-HCT.

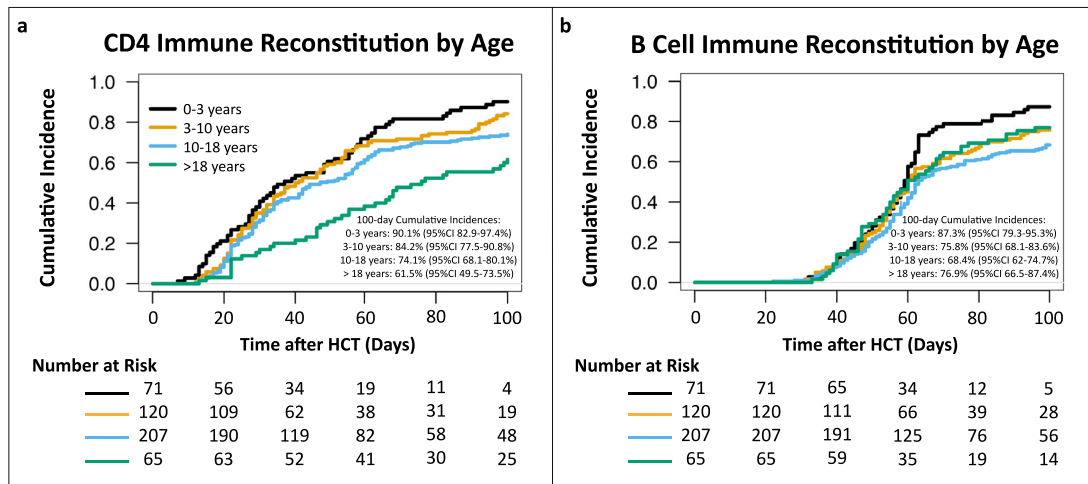## CD4 and B Cell Immune Reconstitution by Age Groups

Figure 6. Cumulative incidences of CD4 and B-cell IR by age groups. CD4 IR was defined as absolute CD4 count >50 cells/μL before day 100 after allo-HCT and B-cell IR was defined as absolute B-cell count >25 cells/μL before day 100 after allo-HCT. Ages between 10−18 and >18 years were found to be associated with decreased cumulative incidence of CD4 IR compared with age 0−3 years in multivariable Cox PH models. For B-cell IR age groups 3−10, 10−18 and >18 years were all found to be associated with lower cumulative incidence of IR, compared with age groups 0−3 years, in multivariable Cox PH models.
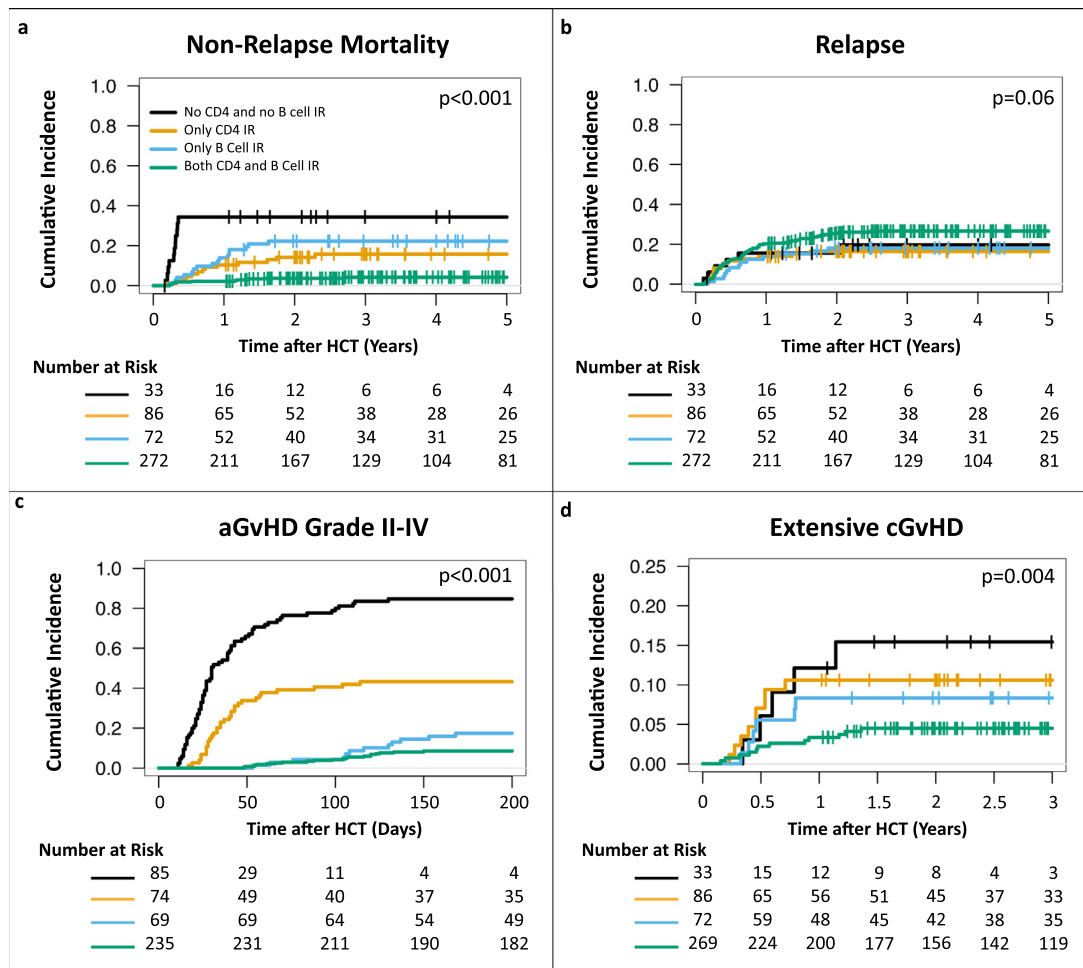
CD4 and B Cell Immune Reconstitution and Outcomes

Figure 7. Cumulative incidences of (a) NRM, (b) relapse (c) acute GVHD grade II-IV and (d) extensive chronic GVHD by CD4 and B-cell IR. CD4 and B-cell IR were associated with significantly lower cumulative incidence of NRM, aGVHD and cGVHD but not relapse. CD4 IR was defined as CD4 $>$50 cells/μL before day 100 after allo-HCT and B-cell IR was defined as B cells $>$25 cells/μL before 100 days after allo-HCT. P values were obtained by univariable Fine−Gray method.

Author Contributions

Conception and design of the study: ATL, CL, RD, SP, CK, AS, SNi and JJB. Acquisition of data: ATL, CL, SB, DK and CK. Analysis and interpretation of data: ATL, CL, RD, SP, CK, AS, SNi and JJB. Drafting or revising the manuscript: ATL, CL, SB, RD, SP, SNa, DR, CK, AS, SNi and JJB. All authors have approved the final article.

Data Sharing Statement

For original data, please contact troullia@mskcc.org.

Conflicts of Interest

CL reports honoraria for a lecture from Genzyme and is on a Data Safety Monitoring Board of ExCellThera (related to this topic). SP reports support for the conduct of clinical trials: Jasper, Atara, AlloVir, Consulting/Advisory Board Participation CellEvolve, Smart Immune, Regeneron and IP related to VSTs licensed to Atara with all rights assigned to Memorial Sloan Kettering Cancer Center. AS has received consultant fee from Spotlight Therapeutics, Medexus Inc. and Vertex Pharmaceuticals. He has also received research funding from CRISPR Therapeutics and honoraria from Vindico Medical Education. AS is the St. Jude Children's Research Hospital site principal investigator of clinical trials for genome editing of sickle cell disease sponsored by

Vertex Pharmaceuticals/CRISPR Therapeutics, Novartis Pharmaceuticals and Beam Therapeutics. The industry sponsors provide funding for the clinical trial, which includes salary support paid to his institution. AS has no direct financial interest in these therapies. None of these conflicts are related to the work presented here. JJB reports honoraria from Avrobio, BlueRock, Bluebird Bio, Sanofi, Sobi, SmartImmune, Advanced Clinical consulting (not related to this topic). All other authors report no conflicts of interest.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.jcyt.2023.05.012.

References

[1] Jenq RR, van den Brink MRM. Allogeneic haematopoietic stem cell transplantation: individualized stem cell and immune therapy of cancer. Nat Rev Cancer 2010;10:213–21.

[2] Admiraal R, van Kesteren C, Jol-van der Zijde CM, Lankester AC, Bierings MB, Egberts TCG, et al. Association between anti-thymocyte globulin exposure and CD4+ immune reconstitution in paediatric haemopoietic cell transplantation: a multicentre, retrospective pharmacodynamic cohort analysis. Lancet Haematol 2015;2:e194–203.

[3] van Roessel I, Prockop S, Klein E, Boulad F, Scaradavou A, Spitzer B, et al. Early CD4 + T cell reconstitution as predictor of outcomes after allogeneic hematopoietic cell transplantation. Cytotherapy 2020;22:503–10.

[4] Admiraal R, de Koning CCH, Lindemans CA, Bierings MB, Wensing AMJ, Versluys AB, et al. Viral reactivations and associated outcomes in the context of immune reconstitution after pediatric hematopoietic cell transplantation. J Allergy Clin Immunol 2017;140:1643–1650.e9.

[5] de Koning C, Prockop S, van Roessel I, Kernan N, Klein E, Langenhorst J, et al. CD4+ T-cell reconstitution predicts survival outcomes after acute graft-versus-host-disease: a dual-center validation. Blood 2021;137:848–55.

[6] Przepiorka D, Weisdorf D, Martin P, Klingemann HG, Beatty P, Hows J, et al. 1994 consensus conference on acute GVHD grading. Bone Marrow Transplant 1995;15:825–8.

[7] Jagasia MH, Greinix HT, Arora M, Williams KM, Wolff D, Cowen EW, et al. National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: I. The 2014 Diagnosis and Staging Working Group report. Biol Blood Marrow Transplant 2015;21:389–401.e1.

[8] Therneau TM, Grambsch PM, Fleming TR. Martingale-based residuals for survival models. Biometrika 1990;77:147–60.

[9] Hothorn T, Lausen B. On the exact distribution of maximally selected rank statistics. Comput Stat Data Anal 2003;43:121–37.

[10] Ando T, Tachibana T, Tanaka M, Suzuki T, Ishiyama Y, Koyama S, et al. Impact of graft sources on immune reconstitution and survival outcomes following allogeneic stem cell transplantation. Blood Adv 2020;4:408–19.

[11] Lakkaraja M, Scordo M, Mauguen A, Cho C, Devlin S, Ruiz JD, et al. Antithymocyte globulin exposure in CD34+ T-cell-depleted allogeneic hematopoietic cell transplantation. Blood Adv 2022;6:1054–63.

[12] Horowitz MM, Gale RP, Sondel PM, Goldman JM, Kersey J, Kolb HJ, et al. Graft-versus-leukemia reactions after bone marrow transplantation. Blood 1990;75:555–62.

[13] Dickinson AM, Norden J, Li S, Hromadnikova I, Schmid C, Schmetzer H, et al. Graft-versus-leukemia effect following hematopoietic stem cell transplantation for leukemia. Front Immunol 2017;8:496.

[14] Minculescu L, Marquart HV, Ryder LP, Andersen NS, Schjoedt I, Friis LS, et al. Improved overall survival, relapse-free-survival, and less graft-vs.-host-disease in patients with high immune reconstitution of TCR gamma delta cells 2 months after allogeneic stem cell transplantation. Front Immunol 2019 1997;10.