Biol Blood Marrow Transplant 23 (2017) 840–844

Biology of Blood and Marrow Transplantation

journal homepage: www.bbmt.org

ASBMT American Society for Blood and Marrow Transplantation

Clinical Research: Analysis

Prognostic Limitations of Donor T Cell Chimerism after Myeloablative Allogeneic Stem Cell Transplantation for Acute Myeloid Leukemia and Myelodysplastic Syndromes

Eric Wong [1,2,*], Kate Mason [1], Jenny Collins [1], Barbara Hockridge [1], Janis Boyd [1], Alexandra Gorelik [2,3], Jeffrey Szer [1,2], David S. Ritchie [1,2]

[1] *Bone Marrow Transplant Service, Royal Melbourne Hospital, Victoria, Australia*
[2] *Bone Marrow Transplant Service, University of Melbourne, Victoria, Australia*
[3] *Melbourne EpiCentre, Royal Melbourne Hospital, Victoria, Australia*

Article history:
Received 9 December 2016
Accepted 31 January 2017

Key Words:
Chimerism
Biomarker
Acute myeloid leukemia
Myelodysplastic syndrome

A B S T R A C T

Donor T cell chimerism is associated with relapse outcomes after allogeneic stem cell transplantation (alloSCT) for acute myeloid leukemia (AML) and myelodysplastic syndromes (MDS). However, measures of statistical association do not adequately assess the performance of a prognostic biomarker, which is best characterized by its sensitivity and specificity for the chosen outcome. We analyzed donor T cell chimerism results at day 100 (D100chim) after myeloablative alloSCT for AML or MDS in 103 patients and determined its sensitivity and specificity for relapse-free survival at 6 months (RFS6) and 12 months (RFS12) post-alloSCT. The area under the receiver operating characteristic curve for RFS6 was .68, demonstrating only modest utility as a predictive biomarker, although this was greater than RFS12 at .62. Using a D100chim threshold of 65%, the specificity for RFS6 was 96.6%; however, sensitivity was poor at 26.7%. This equated to a negative predictive value of 88.5% and positive predictive value of 57.1%. Changing the threshold for D100chim to 75% or 85% modestly improved the sensitivity of D100chim for RFS6; however, this was at the expense of specificity. D100chim is specific but lacks sensitivity as a prognostic biomarker of early RFS after myeloablative alloSCT for AML or MDS. Caution is required when using D100chim to guide treatment decisions including immunologic manipulation, which may expose patients to unwarranted graft-versus-host disease.

INTRODUCTION

Disease relapse is the most common cause of treatment failure and mortality after allogeneic stem cell transplantation (alloSCT) for acute myeloid leukemia (AML) and myelodysplastic syndromes (MDS). Therapeutic options for post-alloSCT relapse of AML or MDS include chemotherapy; hypomethylating agents, either alone or in combination with donor lymphocyte infusions; or a second alloSCT. However, overall, the outcomes of these patients are poor; overall survival at 1 year is approximately 20%, and fewer than 10% survive beyond 1 year if they are unable to proceed with intensive salvage measures, including a second alloSCT [1-5].

Given the lack of effective salvage therapies for relapsed AML and MDS after alloSCT, strategies are required to identify patients at high risk of relapse, which may allow the initiation of pre-emptive therapy to prevent relapse. These biomarkers should be reproducible, highly sensitive, and specific for disease relapse and validated in independent cohorts. The 2010 International Workshop on the Biology, Prevention and Treatment of Relapse after AlloSCT identified several candidate biomarkers that warranted further investigation as prognostic tools for post-transplant relapse, including donor chimerism [6]. Several reports have described incomplete donor T cell chimerism as a risk factor for relapse of AML and MDS after myeloablative conditioning alloSCT [7,8]. However, descriptions of statistical associations are usually inadequate to describe the ability of a biomarker to accurately distinguish between patients who will relapse and those who will not [9]. For a biomarker that has been identified from regression analyses to be effective in predicting future outcome, the associated odds ratio must be of a magnitude rarely observed in such analyses [10]. Instead, the performance of a biomarker is best described by its sensitivity and

Financial disclosure: See Acknowledgments on page 844.
* Correspondence and reprint requests: Eric Wong, MBBS, FRACP, FRCPA, Royal Melbourne Hospital, Grattan Street, Parkville, Vic. 3050, Australia.
E-mail address: eric.wong@mh.org.au (E. Wong).

http://dx.doi.org/10.1016/j.bbmt.2017.01.086
1083-8791/

specificity. To date, the performance characteristics of donor T cell chimerism as a predictor for relapse has not been well described. In this report we validate the observation that donor T cell chimerism is associated with relapse-free survival (RFS) after myeloablative alloSCT for AML or MDS in an independent cohort and describe the performance of donor T cell chimerism as a predictive biomarker of early relapse outcomes post-alloSCT, in particular with regards to its sensitivity and specificity.

METHODS

Adult patients (≥18 years) who underwent myeloablative conditioning followed by peripheral blood or bone marrow alloSCT from a matched sibling or unrelated donor for AML or MDS at the Royal Melbourne Hospital between 2000 and 2016 were included in this retrospective analysis. Patients who relapsed or died before day 100 post-alloSCT were excluded. This study was approved by the Royal Melbourne Hospital institutional human research ethics committee.

The myeloablative conditioning regimens used over this period included cyclophosphamide 120 mg/kg in combination with either intravenous busulfan 12.8 mg/kg or total body irradiation 12 Gy. The choice of conditioning regimen was not protocolized and was according to physician discretion. Graft-versus-host disease (GVHD) prophylaxis consisted of cyclosporine 3 mg/kg i.v commencing day −1, with transition to an equivalent oral dose upon recommencement of oral intake, and short-course methotrexate 15 mg/m^2 day +1 followed by 10 mg/m^2 days 3, 6, and 11.

AlloSCT recipients from sibling donors also received prednisolone .5 mg/kg/d days 14 to 34 decreasing to .25 mg/kg/d days 35 to 48, whereas transplants from unrelated donors received T cell depletion with antithymocyte globulin 4 mg/kg pre-alloSCT in divided doses (Thymoglobulin; Genzyme Corporation, Cambridge, MA). Cyclosporine taper commenced from day +100 or earlier depending on physician discretion due to perceived risk of relapse or toxicity.

Chimerism Analysis

Donor–recipient chimerism analysis was performed routinely on day 100 post-alloSCT on peripheral blood samples collected in EDTA. Recipient samples were separated into CD3-positive (T cell) and CD3-negative (non–T cell) fractions using immunomagnetic cell separation with density gradient centrifugation (RosetteSep; StemCell Technologies, VIC, Australia). PCR was performed using oligonucleotide primers specific for short tandem repeats that were identified pretransplant to distinguish donor from recipient DNA (Invitrogen, Thermo Fisher Scientific Inc., Waltham, MA). Products were analyzed using a 3730 DNA Sequencer (Applied Biosystems, Foster City, CA), and chimerism results were expressed as the percentage of donor-specific DNA present.

Endpoints and Statistical Analysis

The primary objective of the study was to evaluate the sensitivity and specificity of day 100 donor T cell chimerism (D100chim) as a predictor of RFS 6 months (RFS6) and 12 months (RFS12) post-alloSCT, treated as binary outcomes. Relapse of AML or MDS was defined as >5% blasts or unequivocal morphologic dysplasia not attributable to other causes on bone marrow aspirate and trephine biopsy morphology. Logistic regression was used to investigate the relationships between RFS6 and RFS12, and D100chim; age; sex; Disease Risk Index (DRI); cytogenetics; graft type; donor relation; and use T cell depletion. D100chim was investigated as both a continuous variable and using predefined thresholds of 55%, 65%, 75%, 85%, and 95%. DRI was separated into 2 categories of low/intermediate or high/very high, as previously defined [11]. Cytogenetics risk categories were defined according to the refined Medical Research Council classification [12]. Variables found to be statistically significant in univariate analyses were included in a multivariate model. Significance level was set at $\alpha < .05$. Statistical analysis was performed using R analysis software (Comprehensive R Archive Network Project, Vienna, Austria).

RESULTS

In total, 103 patients were included in the analysis (Table 1). Of the patients with AML, 72 (80%) were in complete remission at the time of alloSCT, whereas 11 patients (12.2%) underwent transplantation in early morphologic relapse. Twelve patients with MDS (92.3%) had not received any disease-modifying therapy before alloSCT, and 1 patient had received the hypomethylating agent azacitidine without response. Of the 49 patients (54.4%) who received transplants from unrelated donors, 5 patients were mismatched at a single HLA class II loci (DR or DQ), whereas the remainder were 10/10 matches. Thirty-eight patients (42.2%) received glucocorticosteroids for GVHD prophylaxis in addition to cyclosporine and short-course methotrexate.

Table 1
Patient and Treatment Characteristics

Characteristic	No. of Patients
Median age, yr (range)	43 (18-60)
Sex	
Male	52
Female	51
Disease	
AML	90
MDS	13
Cytogenetic group	
Favorable	10
Intermediate	72
Adverse	21
DRI	
Low/intermediate	80
High/very high	23
Stem cell source	
Peripheral blood	80
Bone marrow	23
Donor type	
Sibling	54
Unrelated	49
T cell depletion	
Yes	45
No	58
Conditioning regimen	
CyTBI	18
BuCy	85

Cy indicates cyclophosphamide; TBI, total body irradiation; Bu, busulfan.

The median duration of follow-up was 48.0 months (interquartile range, 29.1 to 76.0). Rates of overall survival at 6 and 12 months post-alloSCT were 97.1% (95% confidence interval [CI], 91.2% to 98.5%) and 91.0% (95% CI, 83.4% to 95.2%), respectively. RFS6 and RFS12 were 85.4% (95% CI, 77.0% to 91.0%) and 73.2% (95% CI, 63.4% to 80.8%), respectively. The cumulative incidence of relapse, with nonrelapse mortality as a competing risk at 6 and 12 months post-alloSCT, was 13.6% (95% CI, 7.8% to 21.0%) and 23.8% (95% CI, 16.0% to 32.5%) respectively. The cumulative incidence of extensive chronic GVHD at 6 and 12 months was 36.9% (95% CI, 27.6% to 46.2%) and 45.7% (95% CI, 35.8% to 55.0%), respectively.

Associations between D100chim and RFS6 and RFS12

The median D100chim was 96% (interquartile range, 89% to 100%), and complete D100chim (defined as ≥95%) was achieved by 62.2% of patients. Patients who received T cell depletion paradoxically had a higher median D100chim compared with patients who did not (100% versus 91.5%), although this was confounded by the observation that a greater number of patients who did not receive T cell depletion had bone marrow grafts (34.5% versus 6.8%). On univariate analysis, covariates significantly associated with RFS6 included D100chim as a continuous variable as well as D100chim as a binary variable dichotomized by prespecified thresholds of 65%, 75%, and 85% (Table 2). DRI was also significantly associated with RFS6 (odds ratio [OR], 3.94; 95% CI, 1.22 to 12.60; $P = .020$). Cytogenetic classification, donor relation, stem cell source, or administration of T cell–depleting antibodies were not significantly associated with RFS6. In a multivariate model including DRI and D100chim, DRI (OR, 3.616; 95% CI, 1.03 to

Table 2
Univariate Logistic Regression Analysis for Associations with RFS6 and RFS12

	RFS6		RFS12	
	OR (95% CI)	*P*	OR (95% CI)	*P*
Age	.99 (.95-1.04)	.76	1.00 (.97-1.05)	.85
DRI				
Low/intermediate	Ref*		Ref	
High/Very high	3.94 (1.22-12.60)	.02	2.55 (.91-7.05)	.07
Cytogenetics				
Favorable	Ref		Ref	
Intermediate	1.62 (.26-31.45)	.66	3.20 (.54-61.34)	.29
Adverse	1.50 (.16-32.74)	.74	2.29 (.28-48.93)	.49
Stem cell source				
Peripheral blood	Ref		Ref	
Bone marrow	1.94 (.55-6.23)	.27	1.02 (.33-2.87)	.97
Donor type				
Sibling	Ref		Ref	
Unrelated	.50 (.15-1.53)	.24	.79 (.31-1.96)	.61
T cell depletion	.84 (.26-2.53)	.76	1.35 (.54-3.41)	.52
D100chim				
Continuous†	1.06 (1.02-1.11)	<.01	1.05 (1.01-1.09)	<.01
≤65%	10.30 (2.03-58.49)	<.01	8.75 (1.74-64.41)	.01
≤75%	6.83 (1.71-27.10)	<.01	6.64 (1.80-27.66)	<.01
≤85%	6.13 (1.83-20.65)	<.01	4.50 (1.50-13.86)	<.01
≤95%	2.76 (.90-9.48)	.09	1.52 (.61-3.83)	.37

* Reference value.
† D100chim as a continuous variable decreasing from 100%.

12.58; $P = .04$) and D100chim (OR, 1.06; 95% CI, 1.02 to 1.11; $P = .006$) remained significantly associated with RFS6.

Ninety-seven patients had a complete 12-month follow-up and were included in the analysis of associations with RFS12. At 12 months, D100chim as a continuous variable, as well as D100chim at thresholds of 65%, 75%, and 85%, was significantly associated with RFS12. However, only the continuous variable remained significantly associated with RFS12 (OR, .95; 95% CI, .92 to .99; $P = .01$) in a multivariable model also including DRI. There was only moderate correlation between D100chim (ie, CD3-positive T cell) and CD3-negative chimerism (Spearman correlation coefficient R = .46). There was also no significant association between CD3-negative chimerism and RFS6 or RFS12, suggesting that D100chim was not merely a surrogate for myeloid chimerism.

Sensitivity and Specificity of D100chim as a Predictor of RFS6 and RFS12

We analyzed the performance of D100chim as a predictor of RFS6 and RFS12 by generating receiver-operating characteristic curves and calculating the sensitivity and specificity of D100chim at different predetermined threshold values. The area under the receiver-operating characteristic curve (AUC) for D100chim as a predictor of RFS6 was .68 (Figure 1). Using a D100chim threshold of 65%, the specificity for RFS6 was 96.6%; however, sensitivity was poor at 26.7%. In our study cohort this equated to a negative predictive value of 88.5% and positive predictive value (PPV) of 57.1%. Changing the threshold for D100chim to 75% or 85% modestly improved the sensitivity of D100chim for RFS6; however, this was at the expense of specificity (Table 3). In our cohort, PPV decreased when the D100chim was increased from 65% to 75% or 85%. The AUC for D100chim as a predictor of RFS12 was .62, which was lower than that observed for RFS6. In keeping with this, D100chim at thresholds of 65%, 75%, and 85% demonstrated poorer sensitivity for RFS12 compared with RFS6.

Given that DRI and D100chim were independently associated with RFS6 in our multivariable model, we investigated the utility of combining both variables for the prediction of RFS6. The AUC of the multivariable regression model was .77, demonstrating an improved ability to predict RFS6 compared with D100chim alone. Accordingly, we devised a scoring algorithm where high/very high DRI scored 1 point and D100chim less than 65% scored 1 point. The specificity of a total score of 1 or greater for RFS6 was 79.5% and sensitivity 66.7%. Using an alternative D100chim threshold of 75% improved sensitivity at the expense of specificity. There was no further benefit in sensitivity or specificity of using a D100chim threshold of 85%.

Kinetics of T Cell Chimerism up to Day 100

T cell engraftment after alloSCT is a dynamic process that may be reflected in the change in T cell chimerism over time, particularly in the first few months after donor cell infusion. Of the 103 patients in the D100chim cohort, there were 80 patients in whom T cell chimerism was examined between day 30 and day 60 post-alloSCT in addition to D100chim. The

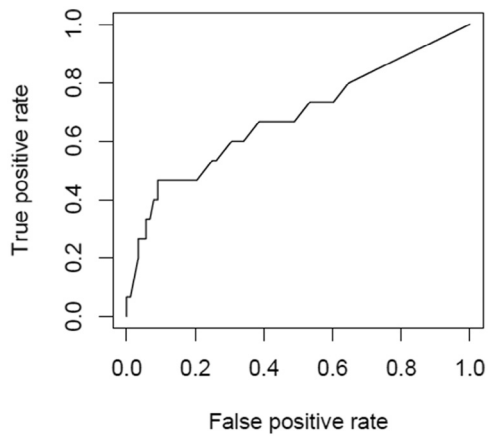

Figure 1. Receiver operating characteristic curve of D100chim as a predictor of RFS6.

Table 3
Sensitivity, Specificity, PPV, and Negative Predictive Value of D100chim for RFS6 and RFS12

	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
RFS6				
D100chim				
≤65%	26.7	96.5	57.1	88.5
≤75%	33.3	93.2	45.5	89.1
≤85%	46.7	87.5	38.9	90.6
DRI + D100chim				
≤65%	66.7	79.5	35.7	93.3
≤75%	73.3	78.4	36.7	94.5
≤85%	73.3	73.9	32.4	94.2
RFS12				
D100chim				
≤65%	20.0	97.5	71.4	77.8
≤75%	28.0	94.4	63.6	79.1
≤85%	36.0	88.9	52.9	80.0
DRI + D100chim				
≤65%	52.0	80.6	48.1	82.9
≤75%	60.0	51.7	51.7	85.3
≤85%	60.0	76.4	46.9	84.6

characteristics of this subset of patients did not differ significantly from the original cohort with respect to age, sex, disease type, DRI, donor relation, or use of T cell depletion. The median T cell chimerism between days 30 and 60 was 93.5% (interquartile range, 85% to 98.3%). In 4 patients (5%) T cell chimerism increased by 10% or greater from days 30 to 60 to day 100 post-alloSCT, and in another 4 patients T cell chimerism decreased by 10% or greater over the same time period. There was an association between patients who had a falling chimerism (by 10% or greater) and RFS6 (OR, 18.8; 95% CI, 2.2 to 400.1; $P = .01$). The sensitivity of a decrease in T cell chimerism by 10% or greater for RFS6 was 21.4% and specificity was 98.5%. The PPV of this dynamic measure of T cell chimerism was 75%; however, this was hampered by a high rate of false negative predictions (78.6%).

DISCUSSION

The ability to accurately predict early relapse after alloSCT is an area of significant clinical need, given the poor prognosis of established relapse and the potential for pre-emptive immunologic manipulation to mitigate relapse risk. Donor T cell chimerism is particularly attractive as a candidate biomarker given its ability to be performed on peripheral blood samples and relatively widespread availability to the extent that it is routinely performed post-alloSCT. In addition, from a mechanistic point of view, donor T cell chimerism reflects the degree of donor T cell engraftment and by extension may be used to infer the availability of donor T cells to exert a graft-versus-leukemia effect. Several groups have described an association between donor T cell chimerism and relapse risk post-alloSCT; however, the interpretation of these reports have been hampered by patient and disease heterogeneity [13,14]. In the largest homogeneous cohort of AML/MDS patients reported to date with respect to post-transplant T cell chimerism, Lee et al. [8] reported an association between day 100 T cell chimerism and relapse after myeloablative alloSCT. In patients in first or second complete remission at the time of alloSCT, donor T cell chimerism ≤ 85% was significantly and independently associated with the cumulative incidence of relapse at 3 years post-alloSCT, with a hazard ratio of 2.4 ($P = .02$). In our study we likewise demonstrate that D100chim is significantly associated with RFS6 and RFS12, both as a continuous variable and using thresholds of between 65% and 85%.

However, the utility of D100chim as a predictive biomarker of relapse post-alloSCT is inadequately characterized by measures of association such as ORs in logistic regression models. For a biomarker to accurately predict prognosis, the distributions of the biomarker in patients that have disease/relapse or not must be sufficiently separated to allow dichotomous classification [9]. This frequently, although indirectly, corresponds to a strength of association rarely observed in logistic regression models [10]. Therein lies the considerable difference between statistical methodology for investigation of classification rather than association. Thus, rather than measures of association, the performance of a diagnostic or predictive tool is best described by the AUC of a receiver-operating-characteristic curve or, when using a binary classification tool, the sensitivity and specificity. We demonstrate that the AUC of D100chim as a predictor of RFS6 was .68, where only values of greater than .7 would be considered a reasonably predictive tool. When using D100chim as a binary classification tool using a threshold of 65%, this demonstrated high specificity (96.6%) for RFS6 but poor sensitivity (26.7%). Increasing the D100chim threshold to 75% or 85% only modestly improved sensitivity, at the expense of specificity. In our patient cohort the high specificity but low sensitivity of D100chim corresponded to a high NPV but poor PPV. In practical terms, 43% of patients with D100chim of less than 65% will be incorrectly classified as at risk of relapse (false-positive predictions). This suggests that patients with a low D100chim may not universally require potent immunomodulatory therapy, such as donor lymphocyte infusions, which are associated with a high rate of GVHD toxicity.

In clinical practice, physicians often use a combination of biomarkers that increase the power of prediction compared with 1 marker alone. DRI has been validated in several cohorts as independently associated with overall survival and RFS after alloSCT [11,15]. When we combined DRI and D100chim, sensitivity was improved; however, in our cohort this did not translate to an improved PPV because of the presence of an increased number of false-positive predictions.

The strength of our analysis is the use of a homogenous patient cohort of AML/MDS and myeloablative conditioning. In addition, our analysis of the sensitivity and specificity of D100chim is unique. This was facilitated by the selection of primary outcomes as dichotomous outcome variables. Furthermore, our choice of RFS6 and RFS12 as the primary endpoints of the analysis recognized that the prognostic impact of D100chim is likely to be time-dependent; that is, the predictive ability of D100chim for relapse is likely to decrease over time. This was confirmed by our observation that sensitivity of D100chim for predicting RFS6 was greater than RFS12. We hypothesized that extending the duration of minimum follow-up (eg, RFS24 or RFS36) would reveal a similar pattern. Our study had a relatively modest sample size, which likely contributed to the lack of association observed between cytogenetic risk group and RFS. However, we doubt that our analysis of D100chim would be significantly altered in a larger cohort because we were able to replicate the association between D100chim and RFS as previously reported in the literature. Our cohort also demonstrated some heterogeneity with regards to disease risk and donor source; however, it is unlikely that these factors would have significantly impacted the results.

Biomarkers predictive of relapse post-alloSCT are desperately required to identify patients who may benefit from post-transplant immunologic strategies to prevent relapse. Other than donor chimerism, minimal residual disease is another candidate biomarker increasingly described as being associated with relapse in patients with AML [16-18]. However, as we and others have described, the performance of these biomarkers as prognostic tools requires further characterization, particularly with regards to sensitivity and specificity. For the moment we encourage caution when using D100chim as a predictor of early relapse and mortality after myeloablative alloSCT in AML and MDS, because it demonstrates a high specificity but lacks sensitivity, which in our cohort corresponded to a poor PPV and a high rate of false-positive predictions. Clinicians should thus tread carefully when using D100chim to guide post-transplant immunologic therapy such as donor lymphocyte infusions, which are associated with substantial rates of GVHD.

ACKNOWLEDGMENTS

Financial disclosure: The authors have nothing to disclose.

Conflict of interest statement: There are no conflicts of interest to report.

REFERENCES

1. Levine JE, Braun T, Penza SL, et al. Prospective trial of chemotherapy and donor leukocyte infusions for relapse of advanced myeloid malignancies after allogeneic stem-cell transplantation. *J Clin Oncol*. 2002;20:405-412.
2. Eefting M, von dem Borne PA, de Wreede LC, et al. Intentional donor lymphocyte-induced limited acute graft-versus-host disease is essential for long-term survival of relapsed acute myeloid leukemia after allogeneic stem cell transplantation. *Haematologica*. 2014;99:751-758.
3. Schmid C, Labopin M, Nagler A, et al. Donor lymphocyte infusion in the treatment of first hematological relapse after allogeneic stem-cell transplantation in adults with acute myeloid leukemia: a retrospective risk factors analysis and comparison with other strategies by the EBMT Acute Leukemia Working Party. *J Clin Oncol*. 2007;25:4938-4945.
4. Schroeder T, Rachlis E, Bug G, et al. Treatment of acute myeloid leukemia or myelodysplastic syndrome relapse after allogeneic stem cell transplantation with azacitidine and donor lymphocyte infusions—a retrospective multicenter analysis from the German Cooperative Transplant Study Group. *Biol Blood Marrow Transplant*. 2015;21:653-660.
5. Craddock C, Labopin M, Robin M, et al. Clinical activity of azacitidine in patients who relapse after allogeneic stem cell transplantation for acute myeloid leukemia. *Haematologica*. 2016;101:879-883.
6. Kroger N, Bacher U, Bader P, et al. NCI First International Workshop on the Biology, Prevention, and Treatment of Relapse after Allogeneic Hematopoietic Stem Cell Transplantation: report from the Committee on Disease-Specific Methods and Strategies for Monitoring Relapse following Allogeneic Stem Cell Transplantation. Part I. Methods, acute leukemias, and myelodysplastic syndromes. *Biol Blood Marrow Transplant*. 2010;16:1187-1211.
7. Huisman C, de Weger RA, de Vries L, Tilanus MG, Verdonck LF. Chimerism analysis within 6 months of allogeneic stem cell transplantation predicts relapse in acute myeloid leukemia. *Bone Marrow Transplant*. 2007;39:285-291.
8. Lee HC, Saliba RM, Rondon G, et al. Mixed T lymphocyte chimerism after allogeneic hematopoietic transplantation is predictive for relapse of acute myeloid leukemia and myelodysplastic syndromes. *Biol Blood Marrow Transplant*. 2015;21:1948-1954.
9. Ware JH. The limitations of risk factors as prognostic tools. *N Engl J Med*. 2006;355:2615-2617.
10. Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol*. 2004;159:882-890.
11. Armand P, Kim HT, Logan BR, et al. Validation and refinement of the Disease Risk Index for allogeneic stem cell transplantation. *Blood*. 2014;123:3664-3671.
12. Grimwade D, Hills RK, Moorman AV, et al. Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials. *Blood*. 2010;116:354-365.
13. Koreth J, Kim HT, Nikiforow S, et al. Donor chimerism early after reduced-intensity conditioning hematopoietic stem cell transplantation predicts relapse and survival. *Biol Blood Marrow Transplant*. 2014;20:1516-1521.
14. Reshef R, Hexner EO, Loren AW, et al. Early donor chimerism levels predict relapse and survival after allogeneic stem cell transplantation with reduced-intensity conditioning. *Biol Blood Marrow Transplant*. 2014;20:1758-1766.
15. Lim AB, Roberts AW, Mason K, Bajel A, Szer J, Ritchie DS. Validating the allogeneic stem cell transplantation disease risk index: sample size, follow-up, and local data are important. *Transplantation*. 2015;99:128-132.
16. Walter RB, Gyurkocza B, Storer BE, et al. Comparison of minimal residual disease as outcome predictor for AML patients in first complete remission undergoing myeloablative or nonmyeloablative allogeneic hematopoietic cell transplantation. *Leukemia*. 2015;29:137-144.
17. Zhou Y, Othus M, Araki D, et al. Pre- and post-transplant quantification of measurable ("minimal") residual disease via multiparameter flow cytometry in adult acute myeloid leukemia. *Leukemia*. 2016;30:1456-1464.
18. Kwon M, Martinez-Laperche C, Infante M, et al. Evaluation of minimal residual disease by real-time quantitative PCR of Wilms' tumor 1 expression in patients with acute myelogenous leukemia after allogeneic stem cell transplantation: correlation with flow cytometry and chimerism. *Biol Blood Marrow Transplant*. 2012;18:1235-1242.