blood advances

The MAGIC algorithm probability predicts treatment response and long-term outcomes to second-line therapy for acute GVHD

Zachariah DeFilipp,[1] Haesook T. Kim,[2] Nikolaos Spyrou,[3] Nikolaos Katsivelos,[3] Steven Kowalyk,[3] Gilbert Eng,[3] Stelios Kasikis,[3] Rahnuma Beheshti,[3] Janna Baez,[3] Yu Akahoshi,[3] Francis Ayuk,[4] Hannah Choe,[5] Aaron Etra,[3] Stephan A. Grupp,[6] Elizabeth O. Hexner,[7] William J. Hogan,[8] Carrie L. Kitko,[9] Muna Qayed,[10] Ran Reshef,[11] Ingrid Vasova,[12] Robert Zeiser,[13] Rachel Young,[3] Ernst Holler,[14] James L. M. Ferrara,[3] Ryotaro Nakamura,[15] John E. Levine,[3] and Yi-Bin Chen[1]

[1]Hematopoietic Cell Transplant and Cellular Therapy Program, Massachusetts General Hospital, Boston, MA; [2]Department of Data Science, Dana-Farber Cancer Institute and Harvard T.H. Chan School of Public Health, Boston, MA; [3]The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY; [4]Department of Stem Cell Transplantation, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; [5]Division of Hematology, James Cancer Center, The Ohio State University, Columbus, OH; [6]Division of Oncology, Children's Hospital of Philadelphia, Philadelphia, PA; [7]Blood and Marrow Transplantation Program, Abramson Cancer Center, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA; [8]Division of Hematology, Mayo Clinic, Rochester, MN; [9]Pediatric Stem Cell Transplant Program, Vanderbilt University Medical Center, Nashville, TN; [10]Aflac Cancer and Blood Disorders Center, Emory University, Atlanta, GA; [11]Blood and Marrow Transplantation Program, Columbia University Medical Center, New York, NY; [12]Department of Internal Medicine 5, Hematology/Oncology, University Hospital Erlangen, Erlangen, Germany; [13]Department of Medicine I - Medical Centre, Faculty of Medicine, University of Freiburg, Freiburg, Germany; [14]Department of Hematology and Oncology, Internal Medicine III, University of Regensburg, Regensburg, Germany; and [15]Hematology/Hematopoietic Cell Transplant, City of Hope National Medical Center, Duarte, CA

Key Points

- The MAGIC algorithm probability measured at the initiation of second-line therapy for acute GVHD predicts NRM and overall survival.

- The higher D28 response rates and survival observed with ruxolitinib than with other therapies were limited to patients with low MAP.

The significance of biomarkers in second-line treatment for acute graft-versus-host disease (GVHD) has not been well characterized. We analyzed clinical data and serum samples at the initiation of second-line systemic treatment of acute GVHD from 167 patients from 17 centers of the Mount Sinai Acute GVHD International Consortium (MAGIC) between 2016 and 2021. Sixty-two patients received ruxolitinib-based therapy, whereas 102 received other systemic agents. In agreement with prospective trials, ruxolitinib resulted in a higher day 28 (D28) overall response Frate than nonruxolitinib therapies (55% vs 31%, P = .003) and patients who received ruxolitinib had significantly lower nonrelapse mortality (NRM) than those who received nonruxolitinib therapies (point estimates at 2-year: 35% vs 61%, P = .002). Biomarker analyses demonstrated that the benefit from ruxolitinib was observed only in patients with low MAGIC algorithm probabilities (MAPs) at the start of second-line treatment. Among patients with a low MAP, those who received ruxolitinib experienced significantly lower NRM than those who received nonruxolitinib therapies (point estimates at 2-year: 12% vs 41%, P = .016). However, patients with high MAP experienced high NRM regardless of treatment with ruxolitinib or nonruxolitinib therapies (point estimates at 2-year: 67% vs 80%, P = .65). A landmark analysis demonstrated that the relationship between the D28 response and NRM largely depends on the MAP level at the initiation of second-line therapy. In conclusion, MAP measured at second-line systemic treatment for acute GVHD predicts treatment response and NRM. The outcomes of patients with high MAP are poor regardless of treatment choice, and ruxolitinib appears to primarily benefit patients with low MAP.

The full-text version of this article contains a data supplement.

Introduction

Acute graft-versus-host disease (GVHD) remains a significant barrier to the success of allogeneic hematopoietic cell transplantation (HCT).[1,2] Most patients who develop grades II to IV acute GVHD receive systemic treatment with corticosteroids. However, at least 20% of patients require additional lines of therapy, and steroid-dependent or steroid-refractory cases drive morbidity and mortality.[3] In recent years, the treatment landscape for acute GVHD has been dramatically altered following both the US Food and Drug Administration (FDA) and European Medicines Agency approval of ruxolitinib, an oral selective Janus kinase (JAK) 1/2 inhibitor, for the treatment of steroid-refractory acute GVHD.[4] The FDA approval was based on the data from REACH1, an open-label phase 2 study that demonstrated a 55% overall response rate (ORR) at day 28 (D28) with ruxolitinib in patients with steroid-refractory disease.[5] These results were further confirmed in REACH2, a randomized open-label phase 3 trial in which ruxolitinib achieved higher response rates than investigator choice of the best available therapies (D28 ORR: 62% vs 39%, $P < .001$) in the steroid-refractory setting, which led to European Medicines Agency approval.[6] In both studies, strict criteria for steroid-refractory GVHD were applied for eligibility criteria. After its approval, ruxolitinib has been used with increasing frequency in the management of acute GVHD. However, to date, characterization of ruxolitinib use outside of prospective clinical trials is limited.[7,8]

Clinical and biomarker risk stratification are being increasingly integrated in acute GVHD management as aids to predict the clinical trajectory of patients. Using initial acute GVHD organ staging, the Minnesota GVHD risk score predicts primary treatment response and survival for patients with standard- and high-risk GVHD.[9] Serum biomarkers have also emerged as an important tool in predicting clinical outcomes after developing acute GVHD. The Mount Sinai Acute GVHD International Consortium (MAGIC) has previously demonstrated that the measurement of 2 serum biomarkers, suppressor of tumorigenesis 2 (ST2) and regenerating islet-derived 3α (REG3α), at diagnosis and during the treatment of acute GVHD can generate MAGIC algorithm probabilities (MAPs) that predict nonrelapse mortality (NRM) and response to treatment for individual patients.[3,10,11] Furthermore, these biomarkers have been shown to predict NRM better than clinical assessments.[3,11]

The initiation of second-line treatment represents an important clinical event in the course of a patient with acute GVHD. However, the significance of biomarkers at this time point and the relationship between biomarker risk and the choice of therapy are not well characterized. In this study, we evaluated predictors of D28 response to second-line therapy, with special attention to the association between MAP and the use of ruxolitinib-based treatment. We also demonstrate that MAP and treatment choice can predict NRM and survival at the time of initiation of second-line therapy for acute GVHD.

Methods

Study design and patient selection

We studied patients who received second-line treatment for acute GVHD between January 2016 and December 2021 using data and serum samples from the MAGIC database and biorepository.

MAGIC collects clinical data and longitudinal serum samples from patients with HCT using the PRoBE study design.[12] The data and sample collection protocols were reviewed and approved by the Institutional Review Boards of the respective MAGIC centers and all patients gave written informed consent to participate in the study.

We included adult patients who received second-line treatment for acute GVHD after the initial treatment with corticosteroids (supplemental Figure 1). A total of 289 patients from 17 MAGIC sites met the clinical inclusion criteria. We excluded 122 patients because they did not have a collected serum sample at the initiation of second-line therapy or had missing clinical response data (patient characteristics in supplemental Table 1); these patients experienced lower 2-year NRM compared with those that were included (NRM at 6 months, 32% vs 41%, $P = .047$). The final population for the analysis consisted of 167 patients who received second-line agents for GVHD treatment.

Clinical GVHD data

The clinical severity of acute GVHD was staged according to published guidelines.[9,13] The indications for second-line therapy were categorized as either steroid-resistance (SR), steroid-dependence (SD), or steroid-sparing (SS). SR was defined as acute GVHD at the initiation of second-line therapy (compared with baseline) that met 1 of 3 conditions: (1) there was no response in any target organ; (2) there was increased involvement in any organ regardless of improvement in other organs; and (3) there was no response in gastrointestinal (GI) or liver involvement despite improvement in skin. SD was defined as GVHD for which second-line therapy was initiated for GVHD flare (increase in stage in ≥1 organ after the initial response). SS was defined as GVHD that responded to corticosteroid treatment without flare and for which second-line therapy was initiated to accelerate tapering of steroids or improve upon an ongoing response. Treatment response at D28 from the initiation of second-line therapy was assessed by ORR, which was defined as the proportion of patients who had a complete response (CR) or partial response (PR) as compared with baseline organ staging without the use of additional systemic therapies. CR was defined as the complete resolution of GVHD symptoms in all 3 target organs. PR was defined as an improvement in at least 1 organ without complete resolution in all organs or worsening in other organs.

Biomarker determination

Serial serum samples were collected from patients enrolled in the MAGIC natural history biorepository study and cryopreserved. Samples were not standardly collected at the initiation of the second-line therapy. Thus, the samples analyzed herein were obtained as (1) planned calendar-based collections or (2) event-driven collections related to the initial systemic treatment of acute GVHD. Samples collected 7 days before 3 days after initiation of second-line therapy met the criteria for analysis. Serum levels of ST2[14] and REG3α[15] were measured retrospectively using enzyme-linked immunosorbent assays as previously described.[3,11,16-19] MAP was calculated as a single value between 0.001 and 0.999 according to the formula: log[-log(1-MAP)] = −11.263 + 1.844(log$_{10}$ST2) + 0.577(log$_{10}$REG3α).[16] We used a single threshold at the initiation of second-line therapy to divide MAPs into 2 groups (high, ≥0.291 vs low, <0.291), as previously described.[11]

Statistical analyses

Baseline and transplant characteristics were reported descriptively and compared using the Fisher exact test, χ^2 test, or Wilcoxon rank-sum test, as appropriate. Univariable and multivariable logistic regression analyses were performed to assess clinical factors that were associated with D28 response to the second-line treatment. The Kaplan-Meier method was used to estimate overall survival (OS), whereas the cumulative incidence of NRM was estimated considering relapse as a competing risk. OS was defined from the initiation of second-line therapy to death from any cause or censoring at the last clinical evaluation. Log-rank and Gray tests were used to compare the OS and cumulative incidence of NRM, respectively. To assess the association between risk factors and long-term outcomes, multivariable Cox regression analysis was performed for OS and NRM. For NRM, cause-specific Cox model was performed. The association between D28 response and survival was analyzed by treating D28 response as a time-dependent variable in these models. A landmark analysis at D28, which excluded patients who died in the first 28 days of second-line acute GVHD therapy, was also performed to illustrate the association of D28 response with NRM and OS. Before modeling, the linearity assumption for continuous variables and the proportional hazards assumption were examined. All P values were 2 sided at a significance level of .05 and multiplicity was not considered. All calculations were performed using SAS version 9.4 (SAS Institute, Inc, Cary, NC) and R version 3.5.1.

The data and sample collection protocols were reviewed and approved by the Institutional Review Boards of the respective MAGIC centers and all patients gave written informed consent to participate in the study.

Results

Patient characteristics

We studied 167 patients who received second-line systemic treatment for acute GVHD and met all inclusion criteria (supplemental Methods). Sixty-two patients (37%) received ruxolitinib-based therapy and 105 (63%) received nonruxolitinib therapies as second-line therapies (supplemental Table 2). In the ruxolitinib group, 48 patients (77%) received ruxolitinib monotherapy, whereas 14 patients (23%) received ruxolitinib in combination with other agents. In the nonruxolitinib therapy group, 85 patients (81%) received monotherapy, whereas 20 patients (19%) received combination therapy as second-line treatment. The treatments administered most frequently in the nonruxolitinib therapies group were extracorporeal photopheresis (n = 30), etanercept (n = 29), and tocilizumab (n = 20). Notably, 17 patients in the nonruxolitinib therapy cohort subsequently received ruxolitinib as a third- or later-line treatment. The baseline clinical characteristics of the study population according to the treatment group are shown in Table 1. The only statistically significant different variables between the cohorts were donor type (fewer matched related donors in the ruxolitinib group, $P = .014$), time from corticosteroids to second-line treatment (longer time in the ruxolitinib group, $P = .02$), and the year of start of systemic GVHD treatment (more recent in the ruxolitinib group, $P < .001$). Since ruxolitinib was approved by the FDA for SR acute GVHD in 2019, we first assessed the year in which systemic GVHD therapy started and

found no difference in NRM and OS between 2016 to 2018 and 2019 to 2021 ($P = .7$ for NRM and 0.4 for OS, supplemental Figure 2A-B). We also tested for a possible interaction between second-line treatment and year of systemic GVHD treatment, which was not statistically significant and was therefore not considered further. In total, 81 patients had high MAPs and 86 patients had low MAPs at the start of second-line therapy, with a similar distribution in each treatment cohort ($P = .53$). When examining the key GVHD characteristics of patients at second-line treatment according to biomarker risk, patients with high MAPs (when compared with patients with low MAPs) were less likely to have skin involvement and more likely to have liver, upper GI, or lower GI involvement, which correlated with higher overall acute GVHD grades. No difference in corticosteroid sensitivity (SR vs SD vs SS) was observed among MAP groups (supplemental Table 3). Notably, median MAP scores after initiation of second-line therapy (days 1-3) were higher than those before or at second-line therapy (day 7-0) (after: 0.315, vs before/at: 0.265, $P = .01$). For the entire cohort, the median follow-up among survivors was 22 months (range, 1.5-27).

Biomarker stratification predicts response to second line treatment for acute GVHD

We first evaluated the D28 response rates for second-line acute GVHD therapy. Patients who received ruxolitinib had higher D28 ORR than patients in the nonruxolitinib group (55% vs 31%, $P = .003$), but comparable D28 CR rates (34% vs 22%, $P = .1$) (Figure 1A). When evaluating outcomes based on biomarker risk, patients with low MAP had a higher D28 ORR (49% vs 30%, $P = .012$) as compared with patients with high MAP (Figure 1B). Notably, patients with low MAP were fourfold more likely to have CR at D28 than patients with high MAP (41% vs 11%, $P < .0001$). D28 ORR did not differ based on initial corticosteroid sensitivity (SR 60%, SD 88%, SS 50%; $P = .21$). When considering the MAP (low vs high) and treatment choice (ruxolitinib vs no ruxolitinib), patients with low MAP who received ruxolitinib had higher D28 ORR compared with those receiving nonruxolitinib therapies (65% vs 39%, $P = .027$); a similar difference was observed in patients with high MAP although this did not reach statistical significance (ruxolitinib: 43%, no ruxolitinib: 23%, $P = .08$) (Figure 1C). Notably, although a trend toward higher CR rates with ruxolitinib therapy compared with nonruxolitinib therapies in patients with low MAP was observed (53% vs 33%, $P = .075$), the CR rates were low for patients with high MAP (11% vs 11%, $P = 1.0$) regardless of the treatment cohort. Low MAP, less clinically severe GVHD (grade <3), and use of ruxolitinib as second-line therapy were all significantly ($P < .05$) associated with a higher D28 ORR in multivariable analysis (Table 2).

Biomarker stratification predicts NRM and survival

We next evaluated NRM and OS of patients with acute GVHD from the initiation of second-line therapy. Patients who received ruxolitinib had lower NRM (point estimates at 2-year: 35% vs 60%, $P = .002$) and higher OS (point estimates at 2-year: 51% vs 32%, $P = .008$) compared with patients in the nonruxolitinib therapy group (Figure 2A-B). Notably, MAP was a stronger discriminator of risk with much larger differences in NRM (point estimates at 2-year: 29% vs 76%, $P < .001$) and OS (point estimates at 2-year: 63% vs 12%, $P < .0001$) (Figure 2C-D). When both biomarker risk and

Table 1. Patient characteristics

	Ruxolitinib (n = 62)	Nonruxolitinib therapies (n = 105)	P value
Median age at transplant (range)	57 (18-73)	58 (18-72)	.92
Sex, n (%)			.42
Female	30 (48)	43 (41)	
Male	32 (52)	62 (59)	
Indication for HCT, n (%)			.42
Acute leukemia	35 (56)	47 (45)	
MDS/MPN	15 (24)	45 (43)	
Lymphoma	8 (13)	9 (9)	
Other	4 (6)	4 (4)	
Donor type, n (%)			**.014**
Related	9 (15)	31 (30)	
Unrelated	46 (74)	71 (67)	
Haploidentical	7 (11)	3 (3)	
HLA match, n (%)			.08
Matched	47 (76)	89 (85)	
Mismatched	8 (13)	13 (12)	
Haploidentical	7 (11)	3 (3)	
Stem cell source, n (%)			.55
Peripheral blood	51 (82)	79 (75)	
BM	9 (15)	20 (19)	
Cord blood	2 (3)	6 (6)	
Conditioning regimen intensity, n (%)			.26
Myeloablative	35 (56)	49 (47)	
Reduced intensity	27 (44)	56 (53)	
GVHD prophylaxis, n (%)			.08
CNI-based	48 (77%)	94 (90%)	
PTCy-based	12 (19%)	7 (7%)	
Other	2 (3%)	4 (3%)	
Median days from steroids to second line therapy (range)	18 (1-45)	12 (2-80)	**.02**
Reason for second line treatment, n (%)			.65
Steroid resistance	53 (84)	93 (89)	
Steroid dependence	4 (6)	4 (4)	
Steroid sparing	5 (8)	8 (7)	
Organ involvement at second line therapy, n (%)			
Skin	32 (52)	48 (46)	.52
Liver	4 (10)	16 (15)	.35
Upper GI	11 (18)	24 (23)	.56
Lower GI	38 (61)	70 (67)	.51
GVHD grade at second line therapy, n (%)			.14
0	4 (7)	1 (1)	
1	7 (11)	12 (11)	
2	13 (21)	30 (29)	
3	19 (31)	40 (38)	
4	19 (31)	22 (21)	

Table 1 (continued)

	Ruxolitinib (n = 62)	Nonruxolitinib therapies (n = 105)	P value
Total bilirubin concentration (mg/dL) at second line therapy, n (%)			.31
<2	55 (89)	88 (84)	
2 to 3.9	2 (3)	10 (10)	
≥4	5 (8)	7 (7)	
MAP at second line therapy, n (%)			.53
High	28 (45)	53 (51)	
Low	34 (55)	52 (49)	
Year of systemic GVHD therapy			<.001
2016 to 2018	23 (37)	74 (70)	
2019 to 2021	39 (63)	31 (30)	

Bold values indicate statistical significance P < 0.05.

CNI, calcineurin inhibitor; HLA, human leukocyte antigen; MDS, myelodysplastic syndrome; MPN, myeloproliferative neoplasm; PTCy, posttransplant cyclophosphamide.

treatment choice were evaluated together, patients with low MAP who received ruxolitinib had significantly lower NRM (point estimates at 2-year: 12% vs 41%) and better survival (point estimates at 2-year: 79% vs 52%) than patients who received nonruxolitinib therapies (P < .0001; Figure 2E-F). In contrast, patients with a high MAP had very poor outcomes regardless of treatment with or without ruxolitinib (point estimates at 2-year: NRM, 67% vs 80%; OS, 16% vs 10%; Figure 2E-F). In multivariable analysis, low MAP, less clinically severe GVHD (grade <3), treatment with ruxolitinib, low bilirubin concentration (<2 mg/dL), and D28 ORR were all (P < .05) associated with lower NRM and higher OS, respectively (Table 3). When an interaction between second-line treatment and MAP level was included in the multivariable analysis, the hazard ratios (HRs) of nonruxolitinib therapies over ruxolitinib were 3.3 (P = .04) for NRM and 2.51 (P = .049) for OS within the low MAP group. For patients with high MAP, the HRs were not significant (supplemental Table 4). When the model evaluated D28 CR instead of D28 ORR, the same variables (low MAP, acute GVHD grade <3, treatment with ruxolitinib, and D28 CR) remained significantly associated with a lower NRM and higher OS (data not shown).

Biomarker risk stratifies long-term outcomes according to day 28 response

We observed markedly poor survival for patients with high MAP despite D28 clinical responses and hypothesized that the depth of response (CR, PR, or nonresponse [NR]) may be associated with NRM and OS. Thus, we performed a landmark analysis at D28 (excluding 18 patients who died within the first 28 days after second-line treatment for acute GVHD). Patients who achieved a D28 CR had significantly lower NRM than patients with PR or NR (16% vs 55% and 60%, respectively, P < .0001) and better OS (70% vs 28% and 32%, respectively, P < .0001) (Figure 3A-B). As expected, acute GVHD was the primary cause of death in patients with D28 NR (n = 60; 82% of deaths). For patients with D28 CR or PR, the primary cause of death was most commonly acute GVHD (n = 12) or disease relapse (n = 7)
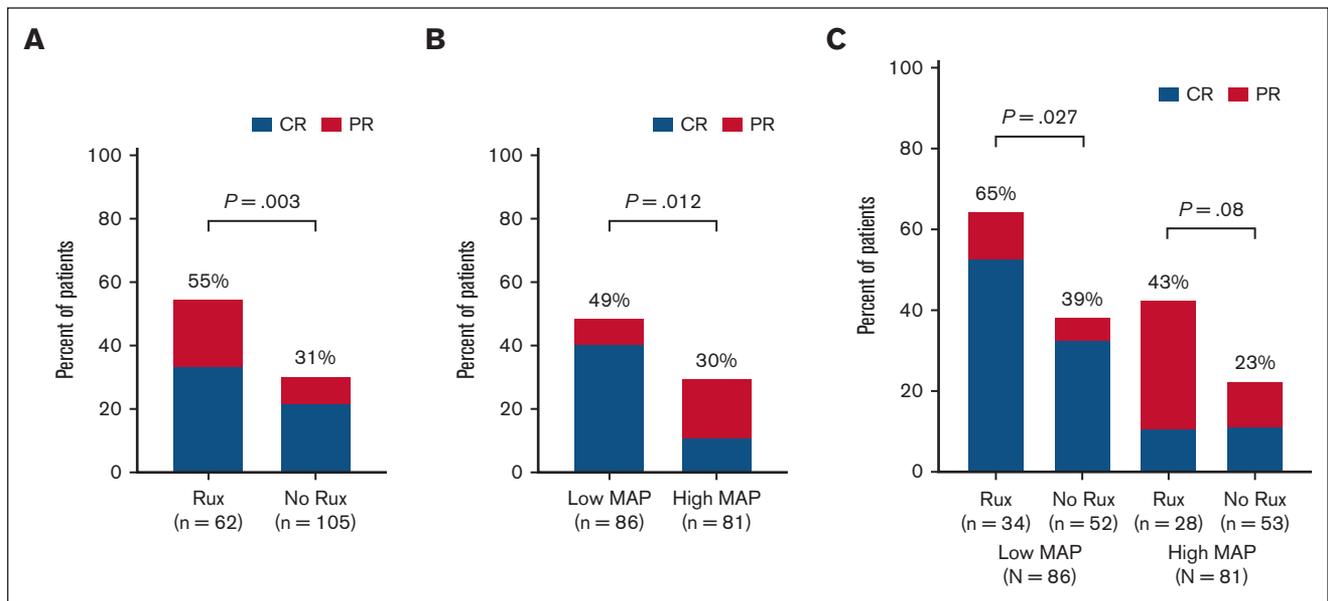
Figure 1. Overall response at day 28 after the initiation of second line therapy for GVHD. ORR (CR or PR) shown for patients according to (A) MAP (low vs high), (B) choice of second-line therapy (ruxolitinib vs no-ruxolitinib), and (C) both MAP (low vs high) and second line therapy group (ruxolitinib vs no-ruxolitinib). Rux, ruxolitinib.

(supplemental Table 5). Landmark analysis also confirmed the utility of MAP measurement at the time of second-line treatment. Patients with high MAP experienced higher 2-year NRM than patients with low MAP, regardless if the D28 response was CR (33% vs 11%, $P = .12$), PR (71% vs 29%, $P = .13$), or NR (80% vs 38%, $P < .0001$), although not all differences were statistically significant. Similarly, 2-year OS was lower for patients with a high MAP than for those with MAP, regardless of whether the D28 response was CR (33% vs 79%, $P = .01$), PR (10% vs 69%, $P = .06$), or NR (14% vs 53%, $P < .0001$) (supplemental Figure 3A-F).

Discussion

We investigated the association between MAP measured at the initiation of second-line acute GVHD therapy and D28 clinical responses and long-term outcomes. MAP separated patients into 2 groups (low vs high) with significantly different ORRs and survival, an association that remained significant in the multivariable analysis. Patients who received ruxolitinib had better outcomes when compared with those not receiving ruxolitinib, but this difference was limited to patients with low MAP, as the outcomes of patients with high MAP were dismal regardless of second-line treatment

Table 2. Multivariable logistic regression analysis for day 28 clinical response after initiation of second line therapy for acute GVHD

		Day 28 response		
		OR	95% CI	P value
Age (y)	≤60 vs >60	0.86	(0.39-1.88)	.71
Sex	Male vs female	1.88	(0.87-4.03)	.11
Indication for HCT	AML/MDS vs other	0.87	(0.43-1.76)	.69
Donor type	Unrelated vs related	0.87	(0.39-1.95)	.74
HLA match	Mismatched vs matched	3.22	(0.95-10.90)	.06
Stem cell source	Peripheral blood vs BM	1.16	(0.45-3.01)	.75
Stem cell source	Cord blood vs BM	0.59	(0.07-4.76)	.62
Conditioning regimen intensity	Reduced-intensity vs myeloablative	0.84	(0.39-1.81)	.66
GVHD prophylaxis	Non-CNI-based vs CNI-based	2.88	(0.87-9.58)	.08
Median days from steroids to second line therapy	<14 d vs ≥14 d	1.70	(0.83-3.47)	.15
GVHD Grade at second line therapy	0 to 2 vs 3 to 4	**3.22**	**(1.55-6.70)**	**.0018**
MAP at second line therapy	Low vs high	**2.09**	**(1.02-4.28)**	**.043**
Second line treatment	Ruxolitinib vs no ruxolitinib	**2.75**	**(1.26-6.00)**	**.011**
Year of systemic GVHD therapy	2019 to 2021 vs 2016 to 2018	1.17	(0.54-2.54)	.68

Bold values indicate statistical significance $P < 0.05$.
CI, confidence interval; OR, odds ratio.

Figure 2. Long-term outcomes after initiation of second line therapy for GVHD, stratified by choice of treatment and biomarker risk. (A) Cumulative incidence of NRM and (B) Kaplan-Meier estimates of OS for patients according to the use of ruxolitinib-based therapies (yes or no) for second line therapy of GVHD. (C) NRM and (D) OS for patients according to high or low MAP. (E) NRM and (F) OS for patients after initiation of second-line treatment for GVHD, stratified by the choice of second line therapy and MAP. The log-rank test was used for the comparison of OS and the Gray test was used for the comparison of the cumulative incidence of NRM. OS, overall survival; Rux, ruxolitinib.
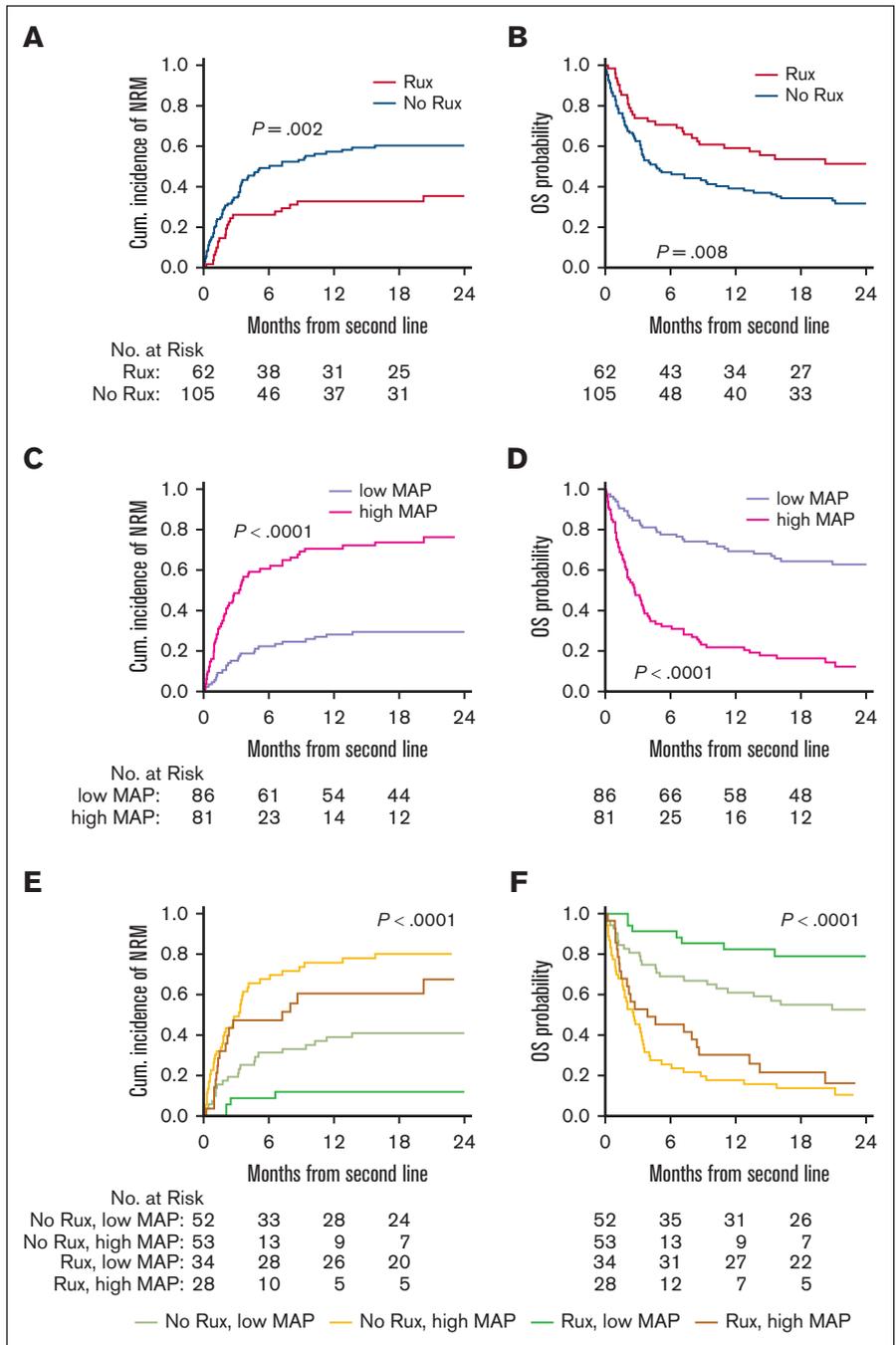
choice. Notably, a recent exploratory analysis of samples obtained from patients who participated in phase 3 randomized trial that compared ruxolitinib with other systemic therapies in SR/SD acute GVHD (REACH2) identified both ST2 and REG3α as predictors of response.[20]

The current analysis expands upon previous evaluations of MAP and evaluates the potential use of MAP in relation to second-line therapy for acute GVHD. Previous studies have established that the MAP can predict NRM at multiple time points in relation to upfront systemic therapy with corticosteroids, including treatment

initiation and 1 week and 4 weeks into treatment.[3,10,11] The start of second-line therapy marks an important clinical event, mainly because treatment with corticosteroids alone is inadequate. In this study, which reflects real-world clinical practice, the cause of this inadequacy was mostly steroid-refractory disease (>80%), rather than steroid-dependent disease or the need for an SS approach.

The divergent survival of patients according to MAP is pertinent to clinical trial design for SR acute GVHD. Ruxolitinib has become widely adopted in the treatment of acute GVHD, as reflected by its more common use in recent years in the current analysis. Our

Table 3. Multivariable regression analysis for OS and NRM after the initiation of second line therapy for acute GVHD

		OS			NRM		
		HR	95% CI	P value	HR	95% CI	P value
Age (y)	≤60 vs >60	1.28	(0.79-2.06)	.32	1.21	(0.71-2.06)	.48
Sex	Male vs female	1.17	(0.75-1.83)	.49	1.09	(0.67-1.77)	.74
Indication for HCT	AML/MDS vs other	0.95	(0.61-1.48)	.82	1.07	(0.66-1.74)	.79
Donor type	Unrelated vs related	1.32	(0.81-2.16)	.26	1.19	(0.70-2.02)	.53
HLA match	Mismatched vs matched	0.77	(0.33-1.77)	.54	0.82	(0.33-2.05)	.68
Stem cell source	Peripheral blood vs BM	1.16	(0.65-2.08)	.62	1.14	(0.62-2.11)	.67
Stem cell source	Cord blood vs BM	2.12	(0.61-7.37)	.24	2.46	(0.68-8.95)	.17
Conditioning regimen intensity	Reduced-intensity vs myeloablative	1.02	(0.63-1.64)	.94	1.09	(0.66-1.82)	.74
GVHD prophylaxis	Non-CNI-based vs CNI-based	0.90	(0.42-1.93)	.78	0.89	(0.40-2.00)	.78
Median days from steroids to second line therapy	<14 d vs ≥14 d	1.33	(0.86-2.05)	.19	1.46	(0.90-2.36)	.13
GVHD Grade at second line therapy	3 to 4 vs 0 to 2	**2.02**	**(1.2-3.37)**	**.0078**	**2.23**	**(1.24-3.99)**	**.007**
MAP at second line therapy	High vs low	**3.36**	**(2.05-5.50)**	**<.0001**	**3.00**	**(1.74-5.16)**	**<.0001**
Second line treatment	Ruxolitinib vs no ruxolitinib	**0.56**	**(0.34-0.90)**	**.017**	**0.54**	**(0.31-0.94)**	**.03**
Bilirubin level (mg/dL) at second line therapy*	2 to 3.9 vs <2	**2.31**	**(1.05-5.08)**	**.038**	2.11	(0.93-4.82)	.08
	≥4 vs <2	**2.48**	**(1.19-5.14)**	**.015**	**2.30**	**(1.08-4.9)**	**.03**
Year of systemic GVHD therapy	2019 to 2021 vs 2016 to 2018	1.53	(0.95-2.45)	.08	1.38	(0.82-2.31)	.22
Day 28 response†	CR/PR vs no response	**0.55**	**(0.34-0.91)**	**.02**	**0.38**	**(0.21-0.66)**	**.0007**

Bold values indicate statistical significance $P < 0.05$.
BM, bone marrow; CI, confidence interval; CNI, calcineurin inhibitor; HLA, human leukocyte antigen; HR, hazard ratio; MDS, myelodysplastic syndrome.
*Due to collinearity between MAP and bilirubin levels, 2 multivariable models were performed: 1 includes MAP and the other includes bilirubin level.
†D28 response was treated as a time-dependent variable.

results suggest that the benefit of ruxolitinib in terms of both D28 response and long-term survival is mainly for patients with low MAP measured at the initiation of second line therapy and should be considered the current standard of care in this lower-risk population. However, in patients with high MAP measured when second line therapy is needed, clinical trials investigating novel therapies, alone or in combination with ruxolitinib, are clearly warranted to try to improve the poor outcomes for this high-risk population. Incorporating MAP into clinical trial design, either as part of eligibility criteria or as a key secondary analysis, can provide a proper clinical context to benchmark the outcomes of SR acute GVHD trials. This is especially important for single-arm trials, in which a randomized control arm is lacking. Evaluation of MAP is also of particular importance for trials that focus on the treatment of lower GI GVHD. The concentrations of ST2 and REG3α reflect the extent of GI crypt damage,[15,21] and as expected, patients with high MAP were more likely to have lower GI involvement than those with low MAP in the current analysis.
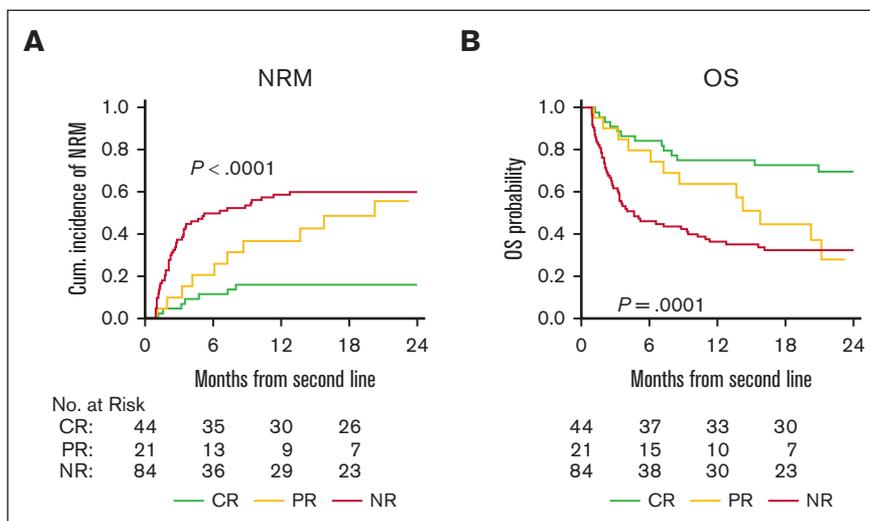

Figure 3. Landmark analysis at D28: long-term outcomes according to day 28 response to second line therapy. (A) Cumulative incidence of NRM and (B) Kaplan-Meier estimates of OS for patients according to D28 response (CR, PR, or NR). The gray test was used for the group comparison of NRM and the log-rank test was used for the group comparison of OS. NR, no response.

Our findings also emphasize the importance of D28 response evaluations in acute GVHD.

Both the MAP and choice of second-line treatment separated the patients into groups with different D28 ORR. However, the D28 ORR may not be the best predictor of long-term outcomes in a MAP-stratified high-risk patient population. Two-year survival was 70% lower for patients with high MAP who received ruxolitinib than for patients with low MAP who did not receive ruxolitinib but the D28 ORR for both groups was ~40%. Furthermore, landmark analysis demonstrated that CR and PR at D28 have different long-term survivals, which is not the case for the primary treatment of acute GVHD.[10] In addition, the association of D28 response with survival largely depends on the MAP level at the initiation of second-line treatment, with MAP stratifying NRM and OS according to D28 response. Thus, for patients not participating in a clinical trial, knowledge of MAP at the initiation of second-line treatment may significantly impact discussions of potential therapeutic strategies and expected long-term outcomes.

This study has several limitations, mainly related to the size of the study population and the retrospective nature of the analysis. Although the database reflects real-world practice, we are unable to comprehensively characterize the use of ruxolitinib in the real-world. In addition, the analysis was limited to patients with biomarker evaluation collected from the MAGIC natural history study. Because the initiation of second-line therapy is not a standard time point for serum specimen collection, many patients lacked an available sample for analysis. Although differences in MAP scores based on timing around second-line therapy were observed, larger sample sizes are needed to investigate this finding while accounting for other clinical factors. Other factors limit the ability to directly compare our results to the results of the prospective randomized REACH2 trial. For example, REACH2 was conducted using strict eligibility criteria to define SR and SD disease.[6] Although the MAGIC database applies standardized criteria to define responsiveness to first-line corticosteroid therapy, the heterogeneity of clinical practice cannot always be captured, such as the clinical threshold or motivation to initiate second-line therapy, especially in patients with SD or SS disease. Another reflection of real-world practice that limits the comparison with REACH2 is the use of combination therapies in our analysis. For purposes of the analysis, we identified treatment group according to the use of ruxolitinib, but we acknowledge that the selected use of multiple agents may influence study outcomes. Finally, the analysis was limited to the adult population and may not be applicable to pediatric patients.

In conclusion, MAP measured at the initiation of second-line systemic treatment for acute GVHD predicts treatment response as well as long-term NRM and OS. The outcomes of patients with high MAP are poor regardless of second-line treatment choice, and the higher CR rate in patients with low MAP drives higher survival rates.

Incorporation of MAP into clinical trials studying second line acute GVHD therapy warrants further investigation. The results also support the current use of ruxolitinib as a standard second-line treatment for acute GVHD, particularly in patients with a low MAP.

Authorship

Contribution: Z.D., H.T.K., Y-B.C., and J.E.L. conceived and designed the study; Z.D., N.S., N.K., S. Kowalyk, G.E., S. Kasikis, R.B., J.B., Y.A., F.A., H.C., A.E., S.A.G., E.O.H., W.J.H., C.L.K., M.Q., R.R., I.V., R.Z., R.Y., E.H., R.N., J.E.L., and Y-B.C. collected and reviewed the clinical data; H.T.K. performed the statistical analysis; Z.D. wrote the article; and all authors interpreted the data and contributed to writing the article.

Conflict-of-interest disclosure: Z.D. received research funding from Incyte, Corp, Regimmune, Corp, and Taiho Oncology, Inc, and consultancy from Sanofi, Incyte, Corp, MorphoSys AG, Inhibrx, PharmaBiome AG, and Ono Pharmaceutical. F.A. received research funding from Mallinckrodt/Therakos and honoraria from Mallinckrodt/Therakos, Novartis, Bristol Myers Squibb/Celgene, Kite/Gilead, Janssen, Medac, and Miltenyi. C.L.K. received honoraria from Incyte and Horizon Therapeutics. M.Q. received honoraria from Novartis and Vertex. R.Z. received honoraria from Mallinckrodt/Therakos, Novartis, Neovii, Medac, and Incyte. J.E.L. reports consultancy from Equillium, Mallinckrodt, and Mesoblast, and research funding from Equillium and Incyte. J.L.M.F. and J.E.L. are inventors of GVHD biomarkers patent and received royalties from Viracor. Y.B.C. reports consultancy from Incyte, Takeda, Pharmacosmos, Editas, Novo Nordisk, Equilium, and Vor. The remaining authors declare no competing financial interests.

ORCID profiles: Z.D., 0000-0002-7994-8974; N.K., 0009-0004-7955-1119; Y.A., 0000-0001-6825-9340; E.O.H., 0000-0002-1125-4060; W.J.H., 0000-0002-5841-4105; M.Q., 0000-0001-7689-343X; R.R., 0000-0003-2185-9546; R.N., 0000-0002-9082-0680; J.E.L., 0000-0002-5611-7828; Y.-B.C., 0000-0002-9554-1058.

Correspondence: Zachariah DeFilipp, Massachusetts General Hospital, 55 Fruit St, Boston, MA 02114; email: zdefilipp@mgh.harvard.edu.

References

1. Ferrara JL, Levine JE, Reddy P, Holler E. Graft-versus-host disease. *Lancet*. 2009;373(9674):1550-1561.

2. Zeiser R, Blazar BR. Acute graft-versus-host disease - biologic process, prevention, and therapy. *N Engl J Med*. 2017;377(22):2167-2179.

3. Srinagesh HK, Ozbek U, Kapoor U, et al. The MAGIC algorithm probability is a validated response biomarker of treatment of acute graft-versus-host disease. *Blood Adv*. 2019;3(23):4034-4042.

4. Przepiorka D, Luo L, Subramaniam S, et al. FDA approval summary: ruxolitinib for treatment of steroid-refractory acute graft-versus-host disease. *Oncologist*. 2020;25(2):e328-e334.

5. Jagasia M, Perales MA, Schroeder MA, et al. Ruxolitinib for the treatment of steroid-refractory acute GVHD (REACH1): a multicenter, open-label phase 2 trial. *Blood*. 2020;135(20):1739-1749.

6. Zeiser R, von Bubnoff N, Butler J, et al. Ruxolitinib for glucocorticoid-refractory acute graft-versus-host disease. *N Engl J Med*. 2020;382(19): 1800-1810.

7. Zeiser R, Burchert A, Lengerke C, et al. Ruxolitinib in corticosteroid-refractory graft-versus-host disease after allogeneic stem cell transplantation: a multicenter survey. *Leukemia*. 2015;29(10):2062-2068.

8. Escamilla Gomez V, Garcia-Gutierrez V, Lopez Corral L, et al. Ruxolitinib in refractory acute and chronic graft-versus-host disease: a multicenter survey study. *Bone Marrow Transplant*. 2020;55(3):641-648.

9. MacMillan ML, Robin M, Harris AC, et al. A refined risk score for acute graft-versus-host disease that predicts response to initial therapy, survival, and transplant-related mortality. *Biol Blood Marrow Transplant*. 2015;21(4):761-767.

10. Levine JE, Braun TM, Harris AC, et al. A prognostic score for acute graft-versus-host disease based on biomarkers: a multicentre study. *Lancet Haematol*. 2015;2(1):e21-e29.

11. Major-Monfried H, Renteria AS, Pawarode A, et al. MAGIC biomarkers predict long-term outcomes for steroid-resistant acute GVHD. *Blood*. 2018; 131(25):2846-2855.

12. Pepe MS, Feng Z, Janes H, Bossuyt PM, Potter JD. Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design. *J Natl Cancer Inst*. 2008;100(20):1432-1438.

13. Harris AC, Young R, Devine S, et al. International, multicenter standardization of acute graft-versus-host disease clinical data collection: a report from the mount sinai acute GVHD international consortium. *Biol Blood Marrow Transplant*. 2016;22(1):4-10.

14. Zhang J, Ramadan AM, Griesenauer B, et al. ST2 blockade reduces sST2-producing T cells while maintaining protective mST2-expressing T cells during graft-versus-host disease. *Sci Transl Med*. 2015;7(308):308ra160.

15. Zhao D, Kim YH, Jeong S, et al. Survival signal REG3alpha prevents crypt apoptosis to control acute gastrointestinal graft-versus-host disease. *J Clin Invest*. 2018;128(11):4970-4979.

16. Hartwell MJ, Ozbek U, Holler E, et al. An early-biomarker algorithm predicts lethal graft-versus-host disease and survival. *JCI Insight*. 2017;2(3):e89798.

17. Etra A, Gergoudis S, Morales G, et al. Assessment of systemic and gastrointestinal tissue damage biomarkers for GVHD risk stratification. *Blood Adv*. 2022;6(12):3707-3715.

18. Spyrou N, Akahoshi Y, Ayuk FA, et al. The utility of biomarkers in acute GVHD prognostication. *Blood Adv*. 2023;7(17):5152-5155.

19. Akahoshi Y, Spyrou N, Hogan WJ, et al. Incidence, clinical presentation, risk factors, outcomes, and biomarkers in de novo late acute GVHD. *Blood Adv*. 2023;7(16):4479-4491.

20. Socie G, Niederwieser D, von Bubnoff N, et al. Prognostic value of blood biomarkers in steroid-refractory or steroid-dependent acute graft-versus-host disease: a REACH2 analysis. *Blood*. 2023;141(22):2771-2779.

21. Reichenbach DK, Schwarze V, Matta BM, et al. The IL-33/ST2 axis augments effector T-cell responses during acute GVHD. *Blood*. 2015;125(20): 3183-3192.