blood advances

American Society of Hematology
2021 L Street NW, Suite 900,
Washington, DC 20036
Phone: 202-776-0544 | Fax 202-776-0545
bloodadvances@hematology.org

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

Tracking no: ADV-2024-012561R2

Zachariah DeFilipp (Massachusetts General Hospital, United States) Haesook Kim (Dana-Farber Cancer Inst., United States) Nikolaos Spyrou (Icahn School of Medicine at Mount Sinai, Japan) Nikolaos Katsivelos (Icahn School of Medicine at Mount Sinai, United States) Steven Kowalyk (Icahn School of Medicine at Mount Sinai, United States) Gilbert Eng (Icahn School of Medicine at Mount Sinai, United States) Stelios Kasikis (Icahn School of Medicine at Mount Sinai, United States) Rahnuma Beheshti (Icahn School of Medicine at Mount Sinai, United States) Janna Baez (Icahn School of Medicine at Mount Sinai, United States) Yu Akahoshi (Icahn School of Medicine at Mount Sinai, United States) Francis Ayuk (University Medical Center Hamburg-Eppendorf, Germany) Hannah Choe (The Ohio State University, United States) Aaron Etra (Icahn School of Medicine at Mount Sinai, United States) Stephan Grupp (The Children's Hospital of Philadelphia, United States) Elizabeth Hexner (University of Pennsylvania School of Medicine, United States) William Hogan (Mayo Medical Center, United States) Carrie Kitko (Vanderbilt University Medical Center, United States) Muna Qayed (Aflac Cancer and Blood Disorders Center, Children's Healthcare of Atlanta, and Emory University, United States) Ran Reshef (Columbia University Medical Center, United States) Ingrid Vasova (University Hospital Erlangen, Erlangen, Germany, Germany) Robert Zeiser (University Medical Center Freiburg, Germany) Rachel Young (Icahn School of Medicine at Mount Sinai, United States) Ernst Holler (University Hospital Regensburg, Germany) James Ferrara (Icahn School of Medicine at Mount Sinai, United States) Ryotaro Nakamura (City of Hope National Medical Center, United States) John Levine (Icahn School of Medicine at Mount Sinai, United States) Yi-Bin Chen (Massachusetts General Hospital, United States)

Abstract:
The significance of biomarkers at second-line treatment for acute graft-versus-host disease (GVHD) is not well characterized. We analyzed clinical data and serum samples at 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 while 102 received other systemic agents. In agreement with prospective trials, ruxolitinib resulted in higher day 28 (D28) ORR compared to non-ruxolitinib therapies (55% vs 31%, P=0.003) and patients who received ruxolitinib had significantly lower non-relapse mortality (NRM) than those who received non-ruxolitinib therapies (point estimates at 2-year: 35% vs 61%, p=0.002). 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 non-ruxolitinib therapies (point estimates at 2-year: 12% vs 41%, p=0.016). However, patients with a high MAP experienced high NRM regardless of treatment with ruxolitinib or non-ruxolitinib therapies (point estimates at 2-year: 67% vs 80%, p=0.65). A landmark analysis demonstrated that the relationship between D28 response and NRM largely depends on the MAP level at initiation of second-line therapy. In conclusion, the MAP measured at second-line systemic treatment for acute GVHD predicts treatment response and NRM. Outcomes of patients with high MAP are poor, regardless of treatment choice, and ruxolitinib appears to primarily benefit patients with low MAP.

Conflict of interest: COI declared - see note

COI notes: ZD: Research funding: Incyte, Corp., Regimmune, Corp., and Taiho Oncology, Inc; Consultancy: Sanofi, Incyte, Corp., MorphoSys AG, Inhibrx, PharmaBiome AG, and Ono Pharmaceutical. MQ: Honoraria; Novartis, Vertex. JEL; Consultancy: Equillium, Mallickrodt and Mesoblast; research funding: Equillium and Incyte. JLMF and JEL are inventors of a GVHD biomarkers patent and receive royalties from Viracor. The remaining authors declare no competing financial interests.

Preprint server: No;

Author contributions and disclosures: ZD, HTK, YBC, and JEL conceived and designed the study. ZD, NS, SK, GE, RB, JB, FA, HC, AE, SAG, EOH, WJH, CLK, MQ, RR, IV, RZ, RY, EH, RN, JEL, and YBC collected and reviewed the clinical data. HTK performed the statistical analysis. ZD wrote the report. All authors interpreted data and contributed to writing the report.

Non-author contributions and disclosures: No;

Agreement to Share Publication-Related Data and Data Sharing Statement: For original data, please contact zdefilipp@mgh.harvard.edu

Clinical trial registration information (if any):

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], 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]Dicision 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; [15]Hematology/Hematopoietic Cell Transplant, City of Hope National Medical Center, Duarte, CA:

Corresponding Author:

Zachariah DeFilipp, MD
Massachusetts General Hospital
55 Fruit Street
Boston, MA 02114
Telephone: 617-726-5765
Email: zdefilipp@mgh.harvard.edu

For original data, please contact zdefilipp@mgh.harvard.edu

Short title: MAP at 2nd line acute GVHD therapy

Abstract word count: 248
Manuscript word count: 3416
References: 21
Tables / Figures: 3 / 3
Supplemental Tables / Figures: 5 / 3

Key Points

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

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

Abstract

The significance of biomarkers at second-line treatment for acute graft-versus-host disease (GVHD) is not well characterized. We analyzed clinical data and serum samples at 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 while 102 received other systemic agents. In agreement with prospective trials, ruxolitinib resulted in higher day 28 (D28) ORR compared to non-ruxolitinib therapies (55% vs 31%, P=0.003) and patients who received ruxolitinib had significantly lower non-relapse mortality (NRM) than those who received non-ruxolitinib therapies (point estimates at 2-year: 35% vs 61%, p=0.002). 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 non-ruxolitinib therapies (point estimates at 2-year: 12% vs 41%, p=0.016). However, patients with a high MAP experienced high NRM regardless of treatment with ruxolitinib or non-ruxolitinib therapies (point estimates at 2-year: 67% vs 80%, p=0.65). A landmark analysis demonstrated that the relationship between D28 response and NRM largely depends on the MAP level at initiation of second-line therapy. In conclusion, the MAP measured at second-line systemic treatment for acute GVHD predicts treatment response and NRM. Outcomes of patients with high MAP are poor, regardless of treatment choice, and ruxolitinib appears to primarily benefit patients with low MAP.

Keywords: graft-versus-host disease; allogeneic hematopoietic cell transplantation; biomarker

Manuscript Text

Introduction

Acute graft-versus-host disease (GVHD) remains a significant barrier to the success of allogeneic hematopoietic cell transplantation (HCT).[1, 2] The majority of patients who develop grades II-IV acute GVHD will receive systemic treatment with corticosteroids. However, at least 20% of patients will require additional lines of therapy, and these 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 US Food and Drug Administration (FDA) and European Medicines Agency (EMA) 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 which 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 as compared to investigator choice of best available therapies (D28 ORR: 62% vs 39%, $p<0.001$) in the steroid-refractory setting, which led to EMA approval.[6] In both studies, strict criteria for steroid-refractory GVHD were applied for eligibility criteria. Following 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-risk 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 two serum biomarkers, suppressor of tumorigenesis 2 (ST2) and regenerating islet-derived 3α (REG3α), at the diagnosis and during the treatment of acute GVHD can generate MAGIC algorithm probabilities (MAPs) that predict non-relapse 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 timepoint and the relationship between biomarker risk and 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 the MAP and the use of ruxolitinib-based treatment. We also demonstrate that the 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 with data and serum samples in the MAGIC database and biorepository. MAGIC collects clinical data and longitudinal serum samples from HCT patients using a PRoBE study design.[12] 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 following initial treatment with corticosteroids (**Supplemental Figure 1**). In total, 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 were missing clinical response data (patient characteristics in **Supplemental Table 1**); these patients experienced lower 2-year non-relapse mortality (NRM) compared to those that were included (NRM at 6 months, 32% vs 41%, p=0.047).. The final population for analysis consisted of 167 patients who received second line agents for GVHD treatment.

Clinical GVHD data

Clinical severity of acute GVHD was staged according to published guidelines.[9, 13] The indication for second-line therapy was 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 to baseline) which met one of three conditions: 1) there was no response in any target organ; 2) there was increased involvement in any organ regardless of improvement in other organs; 3) there was no response in gastrointestinal 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 initial response). SS was defined as GVHD that had 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 initiation of second-line therapy was assessed by overall response rate (ORR), which was defined as the proportion of patients who had a complete response (CR) or partial response (PR) as compared to baseline organ staging without use of additional systemic therapies. CR was defined as complete resolution of GVHD symptoms in all 3 target organs. PR was defined as an improvement in at least one organ without complete resolution in all organs or worsening in other organs.

Biomarker determination

Serial serum samples were collected prospectively and cryopreserved for patients enrolled on the MAGIC natural history biorepository study. Samples are not standardly collected at the initiation of 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 from 7 days prior to 3 days after initiation of second-line therapy met criteria for analysis. Serum levels of ST2[14] and REG3α[15] were measured retrospectively by enzyme-linked immunosorbent assays, as previously described.[3, 11, 16-19] The MAP was calculated as a single value between 0.001 and 0.999 according to the formula: $\log[-\log(1 - \text{MAP})] = -11.263 + 1.844(\log_{10}\text{ST2}) + 0.577(\log_{10}\text{REG3}\alpha)$.[16] We utilized a single threshold at initiation of second-line therapy to divide MAPs into two groups (high, ≥0.291 vs low, <0.291), as previously described.[11]

Statistical analyses

Baseline and transplant characteristics were reported descriptively and compared using Fisher's exact test, χ^2 test or Wilcoxon rank-sum test, as appropriate. Univariable and multivariable logistic regression analysis 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 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. The log-rank and Gray tests were used to compare OS and cumulative

incidence of NRM, respectively. To assess the association of risk factors with long term outcomes, multivariable Cox regression analysis was performed for OS and NRM. For NRM, cause-specific Cox model was performed. The association of D28 response with survival was analyzed 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 from second-line acute GVHD therapy, was also performed to illustrate the association of D28 response on NRM and OS. Prior to modeling, linearity assumption for continuous variables and the proportional hazards assumption were examined. All P-values were two-sided at a significance level of 0.05 and multiplicity was not considered. All calculations were performed using SAS 9.4 (SAS Institute, Inc., Cary, NC) and R version 3.5.1.

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 (see Methods). Sixty-two patients (37%) received ruxolitinib-based therapy and 105 (63%) received non-ruxolitinib therapies as second-

line therapies (**Supplemental Table 2**). In the ruxolitinib group, 48 patients (77%) received ruxolitinib monotherapy, while 14 patients (23%) received ruxolitinib in combination with other agents. In the non-ruxolitinib therapies group, 85 patients (81%) received monotherapy, while 20 patients (19%) received combination therapy for second-line treatment. The treatments administered most frequently in the non-ruxolitinib therapies group were extracorporeal photopheresis (n=30), etanercept (n=29), and tocilizumab (n=20). Of note, 17 patients in the non-ruxolitinib therapies cohort subsequently received ruxolitinib as a third or later line of treatment. Baseline clinical characteristics of the study population according to 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=0.014), time from corticosteroids to second-line treatment (longer time in the ruxolitinib group, P=0.02) and the year of start of systemic GVHD treatment (more recent in the ruxolitinib group, P<0.001). Since ruxolitinib was approved by the FDA for SR acute GVHD in 2019, we first assessed the year of systemic GVHD therapy started and found no difference in NRM and OS between 2016-2018 and 2019-2021 (p=0.7 for NRM and 0.4 for OS, **Supplemental Figures 2A-B**). We also tested for a possible interaction between the second-line treatment and year of systemic GVHD treatment, which was not statistically significant and 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 similar distribution in each treatment cohort (P=0.53). When examining key GVHD characteristics of patients at second-line treatment according to biomarker risk, patients with high MAPs (when compared to patients with low MAPs) were less likely to have skin involvement and

more likely to have liver, upper gastrointestinal (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**). Of note, median MAP scores after initiation of second-line therapy (Days 1-3) were higher than those from before or at second line therapy (Day -7 to 0) (after: 0.315, vs. before/at: 0.265, p=0.01). For the entire cohort, the median follow-up among survivors was 22 months (range, 1.5-27).

Biomarker stratification predicts response to 2nd line treatment for acute GVHD

We first evaluated D28 response rates for second-line acute GVHD therapy. Patients who received ruxolitinib had higher D28 ORR as compared patients in the non-ruxolitinib group (55% vs 31%, P=0.003), but comparable D28 CR rates (34% vs 22%, p=0.1) (**Figure 1A**). When evaluating outcomes based on biomarker risk, patients with low MAP had higher D28 ORR (49% vs 30%, P=0.012) as compared to patients with high MAP (**Figure 1B**). Notably, patients with low MAP were four-fold more likely to have a CR at D28 compared to patients with high MAP (41% vs 11%. P<0.0001). D28 ORR did not differ based on initial corticosteroid sensitivity (SR 60%, SD 88%, SS 50%; p=0.21). 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 to those receiving non-ruxolitinib therapies (65% vs 39%, P=0.027); a similar difference was observed in patients with high MAP although this did not reach statistical significance (ruxolitinib: 43%, no ruxolitinib: 23%, P=0.08) (**Figure 1C**). Of note, while a

trend towards higher CR rates with ruxolitinib therapy compared to non-ruxolitinib therapies in patients with low MAP was observed (53% vs 33%, P=0.075), the CR rates were low for patients with high MAP (11% vs 11%, P=1.0) regardless of treatment cohort. Low MAP, less clinically severe GVHD (grade <3), and use of ruxolitinib as second-line therapy were all significantly (P<0.05) associated with higher D28 ORR in multivariable analysis (**Table 2**).

Biomarker stratification predicts non-relapse mortality 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=0.002) and higher OS (point estimates at 2-year: 51% vs 32%, P=0.008) as compared to patients in the non-ruxolitinib therapies group (**Figures 2A-B**). Notably, the MAP was a stronger discriminator of risk with much larger differences in NRM (point estimates at 2-year: 29% vs 76%, P<0.001) and OS (point estimates at 2-year: 63% vs 12%, P<0.0001) (**Figures 2C-D**). When both biomarker risk and 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 non-ruxolitinib therapies (P<0.0001; **Figures 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%; **Figures 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<0.05) associated with lower NRM and higher OS, respectively (**Table 3**). When an interaction between the second line treatment and MAP level was included in multivariable analysis, the hazard ratios (HRs) of non-ruxolitinib therapies over ruxolitinib were 3.3 (p=0.04) for NRM and 2.51 (p=0.049) for OS within the low MAP group. For patients with a high MAP, 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) remain significantly associated with 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 depth of response (CR, PR, or non-response [NR]) may be associated with NRM and OS. Thus, we performed a landmark analysis at D28 (excluding the 18 patients who died in the first 28 days after second-line treatment for acute GVHD). Patients who achieved a D28 CR had significantly less NRM than patients with PR or NR (16% vs 55% and 60%, respectively, P<0.0001) and better OS (70% vs 28% and 32%, P<0.0001) (**Figures 3A-B**). As expected, acute GVHD was the primary cause of death for 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) (**Supplemental Table 5**). The 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=0.12), PR (71% vs 29%, P=0.13), or NR (80% vs 38%, P<0.0001), although not all differences were statistically significant. Similarly, 2-year OS was lower for patients with high MAP compared low MAP, regardless if D28 response was CR (33% vs 79%, P=0.01), PR (10% vs 69%, P=0.06), or NR (14% vs 53%, P<0.0001) (**Supplemental Figures 3A-F**).

Discussion

We investigated the association between the MAP measured at the initiation of second-line acute GVHD therapy with D28 clinical responses and long-term outcomes. The MAP separated patients into 2 groups (low vs high) with significantly different overall response rates and survival, an association which remained significant in multivariable analysis. Patients who received ruxolitinib had better outcomes when compared to those not receiving ruxolitinib, but this difference was limited to patients with low MAP, as the outcomes of patients with high MAP was dismal regardless of second-line treatment choice. Of note, a recent exploratory analysis of samples obtained from patients who participated in a phase 3 randomized trial that compared ruxolitinib to 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 the MAP and evaluates the potential utilization of the MAP in relationship to second-line therapy for acute GVHD.

Previous studies have established that the MAP can predict NRM at multiple timepoints in relation to upfront systemic therapy with corticosteroids, including treatment initiation, 1 week and 4 weeks into treatment.[3, 10, 11] The start of second-line therapy marks an important clinical event, mainly that 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 a steroid-sparing approach.

The divergent survival of patients according to the 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 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 initiation of 2nd line therapy and should be considered the current standard of care in this lower-risk population. However, in patients with high MAP measured when 2nd 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 the MAP into clinical trial design, either as part of eligibility criteria or as a key secondary analysis, can provide 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 the MAP is also of particular importance for trials that focus on the treatment of lower GI GVHD. 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 when compared to those with low MAP in the current analysis.

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

Both the MAP and choice of second-line treatment separated patients into groups with different D28 ORR. However, 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 a low MAP who did not receive ruxolitinib, but D28 ORR for both groups was approximately 40%. Furthermore, the landmark analysis demonstrated that CR and PR at D28 have different long-term survivals, which is not the case for 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, with the MAP stratifying NRM and OS according to D28 response. Thus, for patients not participating in a clinical trial, knowledge of the MAP at the initiation of second line treatment may significantly impact discussions of potential therapeutic strategies and expected long-term outcomes.

The current study has several limitations, mainly related to the size of the study population and the retrospective nature of the analysis. While the database reflects real-world practice, we are unable to comprehensively characterize use of ruxolitinib in the real-world. In addition, the analysis was limited to patients with a biomarker evaluation collected on the MAGIC natural history study. Since the initiation of second-line therapy

is a not a standard timepoint for serum specimen collection, many patients were lacking an available sample for analysis. Although differences in MAP scores based on timing around second-line therapy were observed, larger sample sizes would be 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 with strict eligibility criteria to define SR and SD disease.[6] While 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 comparison to REACH2 is the utilization 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 select 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, the 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 low MAP patients drives higher survival rates. Incorporation of MAP into clinical trials studying 2nd line acute GVHD therapy warrants investigation. The results also support the current use of ruxolitinib as standard second-line treatment for acute GVHD, particularly in patients with low MAP.

Acknowledgements

The authors thank the patients, their families, and the research staff for their participation. This work was supported by the National Institutes of Health, National Cancer Institute (grant PO1CA03942), the Pediatric Cancer Foundation, and the German Jose Carreras Leukemia Foundation (grants DJCLS 01 GVHD 2016 and DJCLS 01 GVHD 2020).
YA is a recipient of the Japan Society for the Promotion of Science Postdoctoral Fellowship for Research Abroad.

Author Contributions

ZD, HTK, YBC, and JEL conceived and designed the study. ZD, NS, NK, SK, GE, RB, JB, YA, FA, HC, AE, SAG, EOH, WJH, CLK, MQ, RR, IV, RZ, RY, EH, RN, JEL, and YBC collected and reviewed the clinical data. HTK performed the statistical analysis. ZD wrote the report. All authors interpreted data and contributed to writing the report.

Conflicts of Interest Statement

ZD: Research funding: Incyte, Corp., Regimmune, Corp., and Taiho Oncology, Inc; Consultancy: Sanofi, Incyte, Corp., MorphoSys AG, Inhibrx, PharmaBiome AG, and

Ono Pharmaceutical. FA: Research funding: Mallinckrodt/Therakos, Honoraria: Mallinckrodt/Therakos, Novartis, BMS/Celgene, Kite/Gilead, Janssen, Medac, Miltenyi. CLK: Honoraria – Incyte, Horizon Therapeutics. MQ: Honoraria; Novartis, Vertex. RZ: Honoraria: Mallinckrodt/Therakos, Novartis, Neovii, Medac, Incyte. JEL: Consultancy: Equillium, Mallickrodt and Mesoblast; Research funding: Equillium and Incyte. JLMF and JEL are inventors of a GVHD biomarkers patent and receive royalties from Viracor. YBC: Consultancy: Incyte, Takeda, Pharmacosmos, Editas, Novo Nordisk, Equilium, Vor. The remaining authors declare no competing financial interests.

References

1. Ferrara JL, Levine JE, Reddy P, Holler E. Graft-versus-host disease. *Lancet* 2009; **373**(9674)**:** 1550-1561. doi: 10.1016/S0140-6736(09)60237-3

2. Zeiser R, Blazar BR. Acute Graft-versus-Host Disease - Biologic Process, Prevention, and Therapy. *N Engl J Med* 2017; **377**(22)**:** 2167-2179. doi: 10.1056/NEJMra1609337

3. Srinagesh HK, Ozbek U, Kapoor U, Ayuk F, Aziz M, Ben-David K *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. doi: 10.1182/bloodadvances.2019000791

4. Przepiorka D, Luo L, Subramaniam S, Qiu J, Gudi R, Cunningham LC *et al.* FDA Approval Summary: Ruxolitinib for Treatment of Steroid-Refractory Acute Graft-Versus-Host Disease. *Oncologist* 2020; **25**(2)**:** e328-e334. e-pub ahead of print 2020/02/12; doi: 10.1634/theoncologist.2019-0627

5. Jagasia M, Perales MA, Schroeder MA, Ali H, Shah NN, Chen YB *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. e-pub ahead of print 2020/03/12; doi: 10.1182/blood.2020004823

6. Zeiser R, von Bubnoff N, Butler J, Mohty M, Niederwieser D, Or R *et al.* Ruxolitinib for Glucocorticoid-Refractory Acute Graft-versus-Host Disease. *N Engl J Med* 2020; **382**(19)**:** 1800-1810. e-pub ahead of print 2020/04/23; doi: 10.1056/NEJMoa1917635

7. Zeiser R, Burchert A, Lengerke C, Verbeek M, Maas-Bauer K, Metzelder SK *et al.* Ruxolitinib in corticosteroid-refractory graft-versus-host disease after allogeneic stem cell transplantation: a multicenter survey. *Leukemia* 2015; **29**(10)**:** 2062-2068. doi: 10.1038/leu.2015.212

8. Escamilla Gomez V, Garcia-Gutierrez V, Lopez Corral L, Garcia Cadenas I, Perez Martinez A, Marquez Malaver FJ *et al.* Ruxolitinib in refractory acute and chronic graft-versus-host disease: a multicenter survey study. *Bone Marrow Transplant* 2020; **55**(3)**:** 641-648. e-pub ahead of print 20191107; doi: 10.1038/s41409-019-0731-x

9. MacMillan ML, Robin M, Harris AC, DeFor TE, Martin PJ, Alousi A *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. e-pub ahead of print 20150110; doi: 10.1016/j.bbmt.2015.01.001

10. Levine JE, Braun TM, Harris AC, Holler E, Taylor A, Miller H *et al.* A prognostic score for acute graft-versus-host disease based on biomarkers: a multicentre study. *Lancet Haematol* 2015; **2**(1)**:** e21-29. doi: 10.1016/S2352-3026(14)00035-0

11. Major-Monfried H, Renteria AS, Pawarode A, Reddy P, Ayuk F, Holler E *et al.* MAGIC biomarkers predict long-term outcomes for steroid-resistant acute GVHD. *Blood* 2018; **131**(25)**:** 2846-2855. doi: 10.1182/blood-2018-01-822957

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. e-pub ahead of print 20081007; doi: 10.1093/jnci/djn326

13. Harris AC, Young R, Devine S, Hogan WJ, Ayuk F, Bunworasate U *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. doi: 10.1016/j.bbmt.2015.09.001

14. Zhang J, Ramadan AM, Griesenauer B, Li W, Turner MJ, Liu C *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. doi: 10.1126/scitranslmed.aab0166

15. Zhao D, Kim YH, Jeong S, Greenson JK, Chaudhry MS, Hoepting M *et al.* Survival signal REG3alpha prevents crypt apoptosis to control acute gastrointestinal graft-versus-host disease. *J Clin Invest* 2018; **128**(11)**:** 4970-4979. e-pub ahead of print 20180924; doi: 10.1172/JCI99261

16. Hartwell MJ, Ozbek U, Holler E, Renteria AS, Major-Monfried H, Reddy P *et al.* An early-biomarker algorithm predicts lethal graft-versus-host disease and survival. *JCI Insight* 2017; **2**(3)**:** e89798. doi: 10.1172/jci.insight.89798

17. Etra A, Gergoudis S, Morales G, Spyrou N, Shah J, Kowalyk S *et al.* Assessment of systemic and gastrointestinal tissue damage biomarkers for GVHD risk stratification. *Blood Adv* 2022; **6**(12)**:** 3707-3715. doi: 10.1182/bloodadvances.2022007296

18. Spyrou N, Akahoshi Y, Ayuk FA, Holler E, Choe HK, Etra AM *et al.* The utility of biomarkers in acute GVHD prognostication. *Blood Adv* 2023. e-pub ahead of print 20230504; doi: 10.1182/bloodadvances.2023009929

19. Akahoshi Y, Spyrou N, Hogan WJ, Ayuk FA, DeFilipp Z, Weber D *et al.* Incidence, clinical presentation, risk factors, outcomes, and biomarkers in de

novo late acute GVHD. *Blood Adv* 2023. e-pub ahead of print 20230614; doi: 10.1182/bloodadvances.2023009885

20. Socie G, Niederwieser D, von Bubnoff N, Mohty M, Szer J, Or R *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. doi: 10.1182/blood.2022018579

21. Reichenbach DK, Schwarze V, Matta BM, Tkachev V, Lieberknecht E, Liu Q *et al.* The IL-33/ST2 axis augments effector T-cell responses during acute GVHD. *Blood* 2015; **125**(20)**:** 3183-3192. e-pub ahead of print 20150326; doi: 10.1182/blood-2014-10-606830

Figure Legend

Figure 1. Overall Response at Day 28 after initiation of 2nd line therapy for GVHD. Overall response rate (complete response or partial response) 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 2nd line therapy group (ruxolitinib vs no-ruxolitinib).

Abbreviations: CR: complete response; MAP: MAGIC algorithm probability; PR: partial response; Rux: ruxolitinib;

Figure 2. Long term outcomes following initiation of 2nd 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 use of ruxolitinib-based therapies (yes or no) for 2nd line therapy of GVHD. (C) NRM and (D) OS for patients according to high or low MAP. (E) NRM and (F) OS for patients following initiation of second-line treatment for GVHD, stratified by choice of 2nd line therapy and MAP. Log-rank test was used for the comparison of OS and Gray test was used for the comparison of cumulative incidence of NRM.

Abbreviations: MAP: MAGIC algorithm probability; NRM: non-relapse mortality; OS: overall survival; Rux: ruxolitinib;

Figure 3. Landmark analysis at D28: long term outcomes according to Day 28 response to 2nd line therapy.

(A) Cumulative incidence of NRM and (B) Kaplan-Meier estimates of OS for patients according to D28 response (CR, PR or NR). Gray test was used for the group comparison of NRM and log-rank test was used for the group comparison of OS

Abbreviations: CR: complete response; NR: no response; NRM: non-relapse mortality; OS: overall survival; PR: partial response;

Table 1. Patient characteristics.

		Ruxolitinib (n=62)	Non-ruxolitinib therapies (n=105)	P-value
Median age at transplant (range)		57 (18-73)	58 (18-72)	0.92
Sex, n (%)	Female	30 (48)	43 (41)	0.42
	Male	32 (52)	62 (59)	
Indication for HCT, n (%)	Acute leukemia	35 (56)	47 (45)	0.42
	MDS/MPN	15 (24)	45 (43)	
	Lymphoma	8 (13)	9 (9)	
	Other	4 (6)	4 (4)	
Donor type, n (%)	Related	9 (15)	31 (30)	**0.014**
	Unrelated	46 (74)	71 (67)	
	Haploidentical	7 (11)	3 (3)	
HLA match, n (%)	Matched	47 (76)	89 (85)	0.08
	Mismatched	8 (13)	13 (12)	
	Haploidentical	7 (11)	3 (3)	
Stem cell source, n (%)	Peripheral blood	51 (82)	79 (75)	0.55
	Bone marrow	9 (15)	20 (19)	
	Cord blood	2 (3)	6 (6)	
Conditioning regimen intensity, n (%)	Myeloablative	35 (56)	49 (47)	0.26
	Reduced intensity	27 (44)	56 (53)	
GVHD prophylaxis, n (%)	CNI-based	48 (77%)	94 (90%)	0.08
	PTCy-based	12 (19%)	7 (7%)	
	Other	2 (3%)	4 (3%)	
Median days from steroids to 2nd line therapy (range)		18 (1, 45)	12 (2, 80)	**0.02**
Reason for 2nd line treatment, n (%)	Steroid resistance	53 (84)	93 (89)	0.65
	Steroid dependence	4 (6)	4 (4)	
	Steroid sparing	5 (8)	8 (7)	
Organ Involvement at 2nd line therapy, n (%)	Skin	32 (52)	48 (46)	0.52
	Liver	4 (10)	16 (15)	0.35
	Upper GI	11 (18)	24 (23)	0.56
	Lower GI	38 (61)	70 (67)	0.51
GVHD Grade at 2nd line therapy, n (%)	0	4 (7)	1 (1)	0.14
	1	7 (11)	12 (11)	
	2	13 (21)	30 (29)	
	3	19 (31)	40 (38)	
	4	19 (31)	22 (21)	
Total bilirubin concentration (mg/dL) at 2nd line therapy, n (%)	<2	55 (89)	88 (84)	0.31
	2-3.9	2 (3)	10 (10)	
	≥4	5 (8)	7 (7)	
MAP at 2nd line therapy, n (%)	High	28 (45)	53 (51)	0.53
	Low	34 (55)	52 (49)	
Year of systemic GVHD therapy	2016-2018	23 (37)	74 (70)	<0.001
	2019-2021	39 (63)	31 (30)	

Abbreviations:
CNI: calcineurin inhibitor; HCT: hematopoietic cell transplantation; HLA: human leukocyte antigen; GI: gastrointestinal; GVHD: graft-versus-host disease; MAP: MAGIC

algorithm probability; MDS: myelodysplastic syndrome; MPN: myeloproliferative neoplasm; PTCy: post-transplant cyclophosphamide

Table 2. Multivariable logistic regression analysis for Day 28 clinical response following the initiation of 2nd line therapy for acute GVHD.

		Day 28 Response		
		OR	95%CI	P-value
Age (years)	≤60 vs >60	0.86	(0.39, 1.88)	0.71
Sex	Male vs Female	1.88	(0.87, 4.03)	0.11
Indication for HCT	AML/MDS vs Other	0.87	(0.43, 1.76)	0.69
Donor type	Unrelated vs Related	0.87	(0.39, 1.95)	0.74
HLA match	Mismatched vs Matched	3.22	(0.95, 10.90)	0.06
Stem cell source	Peripheral blood vs Bone marrow	1.16	(0.45, 3.01)	0.75
Stem cell source	Cord blood vs Bone marrow	0.59	(0.07, 4.76)	0.62
Conditioning regimen intensity	Reduced-intensity vs Myeloablative	0.84	(0.39, 1.81)	0.66
GVHD prophylaxis	Non-CNI-based vs CNI-based	2.88	(0.87, 9.58)	0.08
Median days from steroids to 2nd line therapy	<14 days vs ≥14 days	1.70	(0.83, 3.47)	0.15
GVHD Grade at 2nd line therapy	0-2 vs 3-4	**3.22**	(1.55, 6.70)	**0.0018**
MAP at 2nd line therapy	Low vs High	**2.09**	(1.02, 4.28)	**0.043**
2nd line treatment	Ruxolitinib vs no ruxolitinib	**2.75**	(1.26, 6.00)	**0.011**
Year of systemic GVHD therapy	2019-2021 vs 2016-2018	1.17	(0.54, 2.54)	0.68

Abbreviations:
AML: acute myeloid leukemia; CI: confidence interval; CNI: calcineurin inhibitor; HCT: hematopoietic cell transplantation; HLA: human leukocyte antigen; GVHD: graft-versus-host disease; MAP: MAGIC algorithm probability; MDS: myelodysplastic syndrome; OR: odds ratio

Table 3. Multivariable regression analysis for overall survival and non-relapse mortality following the initiation of 2nd line therapy for acute GVHD.

		Overall survival			Non-relapse mortality		
		HR	95%CI	P-value	HR	95%CI	P-value
Age (years)	≤60 vs >60	1.28	(0.79, 2.06)	0.32	1.21	(0.71, 2.06)	0.48
Sex	Male vs Female	1.17	(0.75, 1.83)	0.49	1.09	(0.67, 1.77)	0.74
Indication for HCT	AML/MDS vs Other	0.95	(0.61, 1.48)	0.82	1.07	(0.66, 1.74)	0.79
Donor type	Unrelated vs Related	1.32	(0.81, 2.16)	0.26	1.19	(0.70, 2.02)	0.53
HLA match	Mismatched vs Matched	0.77	(0.33, 1.77)	0.54	0.82	(0.33, 2.05)	0.68
Stem cell source	Peripheral blood vs Bone marrow	1.16	(0.65, 2.08)	0.62	1.14	(0.62, 2.11)	0.67
Stem cell source	Cord blood vs Bone marrow	2.12	(0.61, 7.37)	0.24	2.46	(0.68, 8.95)	0.17
Conditioning regimen intensity	Reduced-intensity vs Myeloablative	1.02	(0.63, 1.64)	0.94	1.09	(0.66, 1.82)	0.74
GVHD prophylaxis	Non-CNI-based vs CNI-based	0.90	(0.42, 1.93)	0.78	0.89	(0.40, 2.00)	0.78
Median days from steroids to 2nd line therapy	<14 days vs ≥14 days	1.33	(0.86, 2.05)	0.19	1.46	(0.90, 2.36)	0.13
GVHD Grade at 2nd line therapy	3-4 vs 0-2	**2.02**	(1.2, 3.37)	**0.0078**	**2.23**	(1.24, 3.99)	**0.007**
MAP at 2nd line therapy	High vs Low	**3.36**	(2.05, 5.50)	**<.0001**	**3.00**	(1.74, 5.16)	**<0.0001**
2nd line treatment	Ruxolitinib vs no ruxolitinib	**0.56**	(0.34, 0.90)	**0.017**	**0.54**	(0.31, 0.94)	**0.03**
Bilirubin level (mg/dL) at 2nd line therapy*	2-3.9 vs <2	**2.31**	(1.05, 5.08)	**0.038**	**2.11**	(0.93, 4.82)	0.08
	≥4 vs <2	**2.48**	(1.19, 5.14)	**0.015**	**2.30**	(1.08, 4.9)	**0.03**
Year of systemic GVHD therapy	2019-2021 vs 2016-2018	1.53	(0.95, 2.45)	0.08	1.38	(0.82, 2.31)	0.22
Day 28 Response**	CR/PR vs no response	**0.55**	(0.34, 0.91)	**0.02**	**0.38**	(0.21, 0.66)	**0.0007**

Note:

* due to collinearity between MAP and bilirubin level, two multivariable models were performed: one includes MAP and the other includes bilirubin level

** D28 response was treated as a time dependent variable

Abbreviations:

AML: acute myeloid leukemia; CI: confidence interval; CNI: calcineurin inhibitor; HCT: hematopoietic cell transplantation; HLA: human leukocyte antigen; HR: hazard ratio; GVHD: graft-versus-host disease; MAP: MAGIC algorithm probability; MDS: myelodysplastic syndrome;

Figure 1

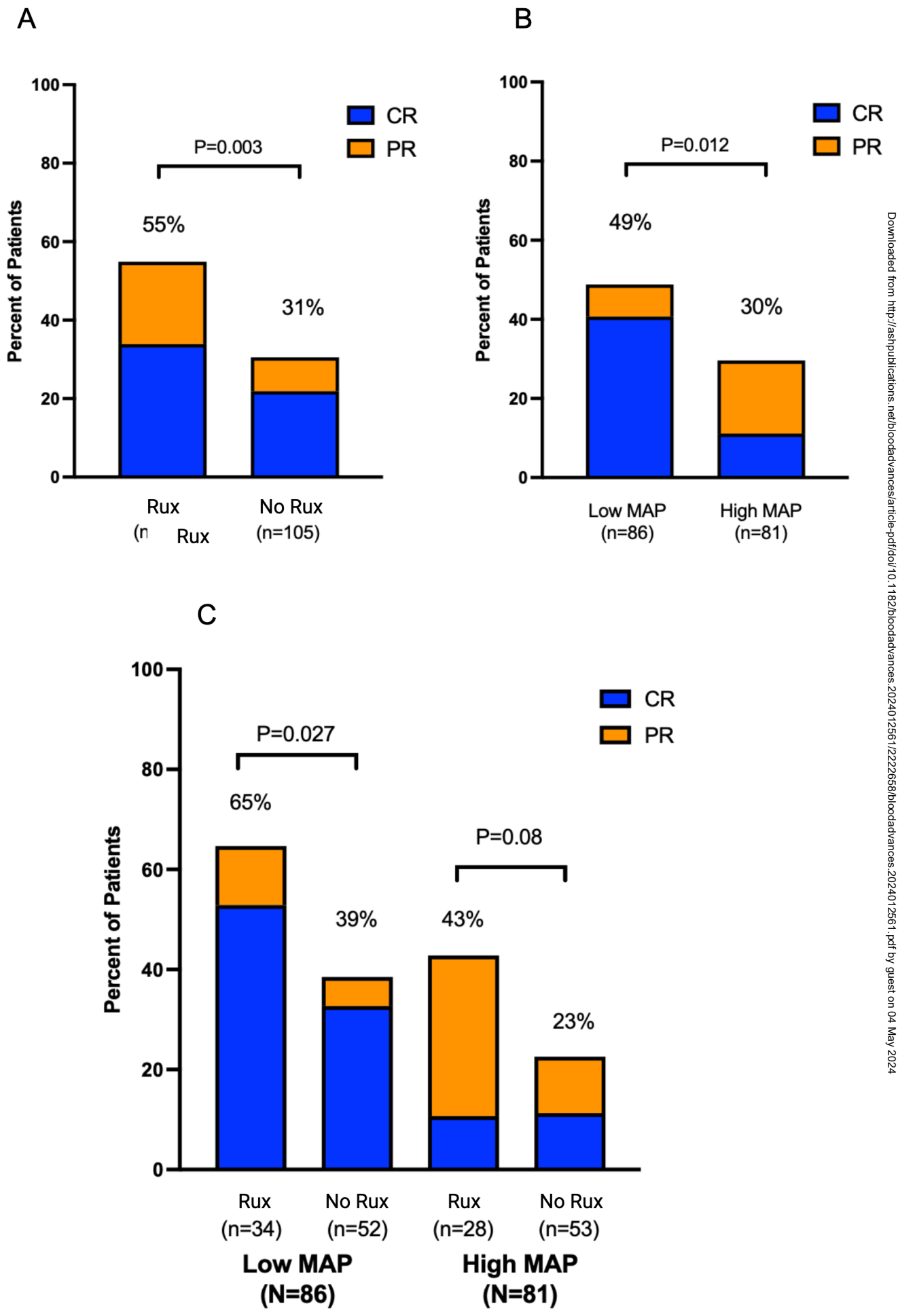

Figure 2

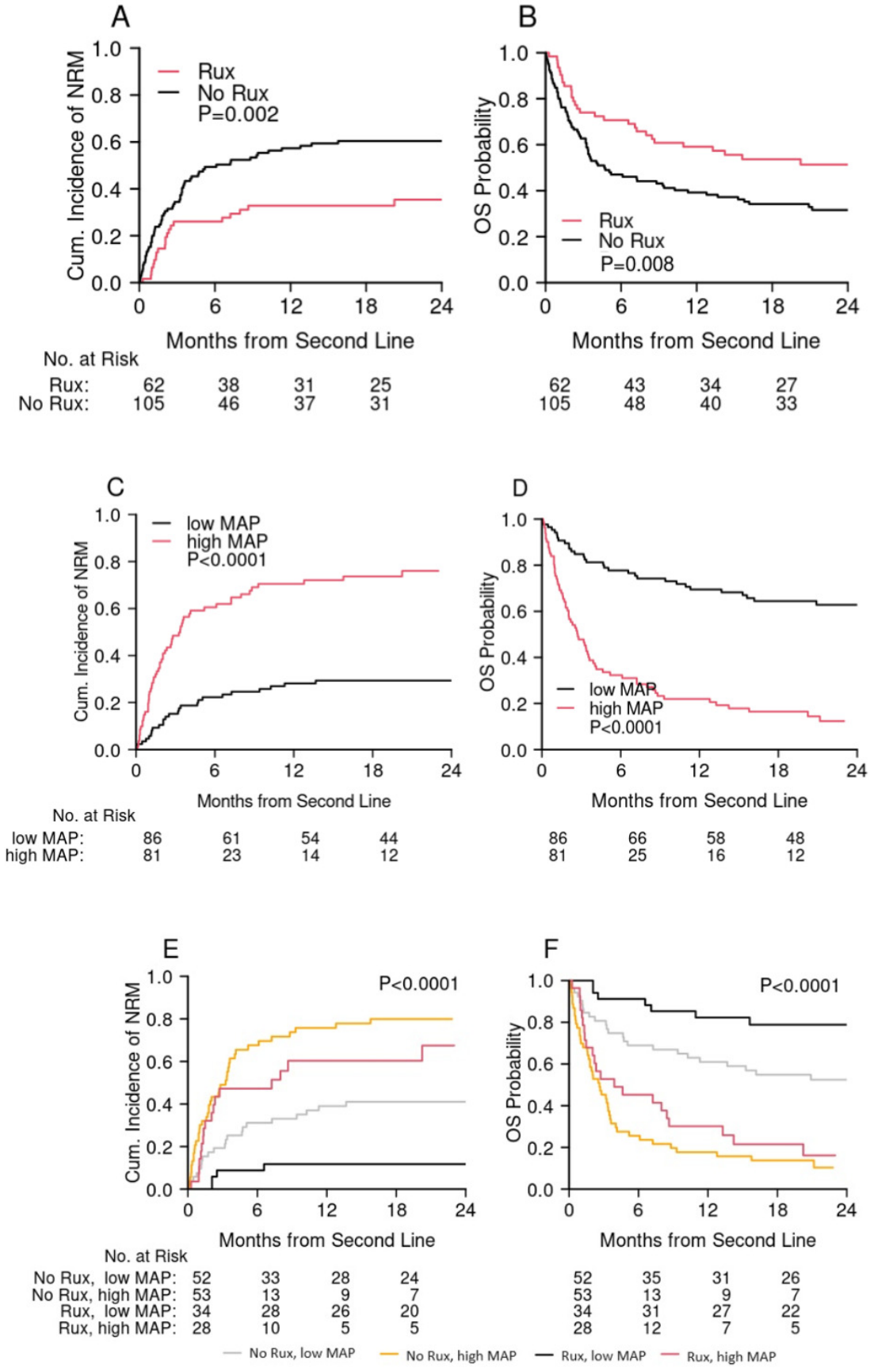

Figure 3

Figure 3