Contribution of measurable residual disease status to prediction accuracy of relapse and survival in adults with acute myeloid leukemia undergoing allogeneic hematopoietic cell transplantation

Measurable residual disease (MRD) before allogeneic hematopoietic cell transplantation (HCT) is strongly associated with relapse risk and survival in acute myeloid leukemia (AML).[1-4] However, association is not prediction.[5,6] In association studies, inferences about a risk factor (e.g., MRD) are made at the population level, not the individual patient level.[6,7] This difference is important because, for clinical decision-making, we are most interested in predicting outcomes for a particular patient. This ability is evaluated with classification models. Good association is usually necessary but not sufficient for good classification. The degree to which MRD data improve post-HCT outcome prediction in adults with AML is unknown.

In order to address this, we studied all adults ≥18 years with AML who received a first allograft while in first or second remission between 4/2006 and 5/2021 and underwent bone marrow MRD testing by multiparameter (10-color) flow cytometry (MFC) before and approximately 1 month after HCT.[8] The MRD assay methodology has remained essentially unchanged throughout the study period, with stable assay performance over time.[8] MRD was identified using a "difference from normal" approach, with the assay detecting MRD in most cases to a level of 0.1% and in progressively smaller subsets of patients as the level of MRD decreases below that level.[8] Disease risk and treatment response was assessed via 2017 European LeukemiaNet (ELN) criteria[9] except that relapse was defined as emergence of >5% blasts by morphology or MFC in blood or marrow, emergence of cytogenetic abnormalities seen previously, or presence/emergence of any level of disease if leading to therapeutic intervention.[8] Our retrospective analysis was approved by the Fred Hutch's Institutional Review Board. Data follow-up was current as of February 10, 2022.

Overall survival (OS) and relapse-free survival (RFS) were estimated using the Kaplan-Meier method. Probabilities of relapse and non-relapse mortality (NRM) were summarized using cumulative-incidence estimates, with death without prior relapse considered a competing risk for relapse and relapse being a competing risk for NRM. Following our previous approach[8] and supported by findings from restricted cubic spline models and maximally selected rank statistics (*Online Supplementary Figure S1*), any detectable level of MRD was considered positive. We used Cox regression to assess the association between OS or RFS and covariates of interest, whereas cause-specific Cox regression models were used for relapse and NRM in the setting of competing risks. We used the C-statistic to quantify a model's ability to predict outcomes, with values of 0.6-0.7, 0.7-0.8, and 0.8-0.9 considered as poor, fair, and good. Statistical analyses were performed using R (http://www.r-project.org).

We identified 979 patients for study inclusion (Table 1). Four hundred and sixty deaths, 308 relapses, and 193 NRM events contributed to estimates for relapse, OS, RFS, and NRM with a median (range) follow-up after HCT among survivors of 62 (3-182) months. Compared to adults without MRD, those with MRD had a significantly increased relapse risk (hazard ratio [HR]: 4.28, 95% confidence interval [CI]: 3.40-5.40, P<0.001), shorter RFS (HR: 3.11, 95% CI: 2.57-3.76, P<0.001), and shorter OS (HR: 2.65, 95% CI: 2.18-3.23, P<0.001). Besides pre-HCT MRD status, several other factors were associated with relapse, RFS, OS, and/or NRM (*Online Supplementary Table S1*).

We then assessed the ability of covariates to predict relapse, RFS, and OS in individual patients. In univariate analyses, pre-HCT MRD status was the strongest (albeit poor) individual predictor for relapse, RFS, and OS (C-statistics: 0.64 [relapse], 0.60 [RFS], 0.59 [OS]; Table 2). Comparable results were obtained with the ELN-recommended ≥0.1% MRD positivity threshold (C-statistics: 0.61, 0.58, 0.57). A basic multivariable model that included age, cytogenetic risk at diagnosis, remission number, time between most recent remission and HCT, Karnofsky score, neutrophil recovery before HCT, conditioning intensity, and secondary AML status yielded C-statistics of 0.64, 0.62, and 0.62 for relapse, RFS, and OS (Table 2). Inclusion of pre-HCT cytogenetic data (normalized *vs*. not normalized for patients presenting with abnormal karyotypes), which we very recently found to provide complementary information to flow cytometric MRD data despite its low sensitivity,[10] improved models only minimally, yielding C-statistics of 0.66, 0.63, and 0.63 for relapse, RFS, and OS. The prediction accuracy could be further improved when information from pre-HCT MRD testing was included (C-statistics of 0.70, 0.66, and 0.65 for relapse, RFS, and OS; C-statistics using ELN cut-off: 0.70, 0.67, 0.65). In contrast, once the pre-HCT MRD status was included, adding pre-HCT cytogenetic data did not further improve the models' predictive ability (Table 2).

In our models for relapse, RFS, and OS, we identified a significant interaction between pre-HCT MRD status and con-

Table 1. Demographic and clinical characteristics of the study cohorts.

	Pre-HCT MRD cohort (N=979)	Peri-HCT MRD dynamics cohort (N=926)
Median age at HCT in years (range)	55 (18-81)	55 (18-81)
Male sex, N (%)	525 (54)	495 (53)
Median WBC at diagnosis, x 10^9/L (range)	8 (0-348)	8 (0-348)
2017 ELN cytogenetic risk*, N (%)		
Favorable	72 (7)	68 (7)
Intermediate	618 (63)	590 (64)
Adverse	251 (26)	231 (25)
Missing/Not reported	38 (4)	37 (4)
2017 ELN cytogenetic/molecular risk, N (%)		
Favorable	171 (17)	164 (18)
Intermediate	87 (9)	83 (9)
Adverse	128 (13)	121 (13)
Missing/Not reported	593 (61)	558 (60)
NPM1 mutational status at diagnosis, N (%)		
Positive	142 (15)	134 (14)
Negative	441 (45)	422 (46)
Missing/Not reported	396 (40)	370 (40)
FLT3-ITD mutational status at diagnosis, N (%)		
Positive	172 (18)	162 (17)
Negative	451 (46)	429 (46)
Missing/Not reported	356 (36)	335 (36)
Secondary AML**, N (%)	257 (26)	239 (26)
Disease status at HCT, N (%)		
First remission	747 (76)	708 (76)
Second remission	232 (24)	218 (24)
Median remission duration before HCT, days (range)	98 (7-788)	98 (7-788)
Recovered ANC before HCT***, N (%)	885 (90)	840 (91)
Recovered platelet count before HCT***, N (%)	691 (71)	656 (71)
Recovered peripheral blood counts before HCT***, N (%)	680 (69)	646 (70)
Pre-HCT Cytogenetics HCT, N (%)		
Normalized karyotype	387 (40)	363 (39)
Abnormal karyotype	161 (16)	151 (16)
Non-informative karyotype****	431 (44)	412 (44)
Pre-HCT MRD status, N (%)		
MRD^{pos}	191 (20)	186 (20)
MRD^{neg}	788 (80)	740 (80)
Post-HCT MRD status day +20-40, N (%)		
MRD^{pos}	71 (7)	50 (5)
MRD^{neg}	876 (89)	876 (95)
Not available	32 (3)	-
Peri-HCT MRD dynamics, N (%)		
MRD^{pos}/MRD^{pos}	50 (5)	50 (5)
MRD^{pos}/MRD^{neg}	136 (14)	136 (15)
MRD^{neg}/MRD^{neg}	740 (76)	740 (80)
MRD^{neg}/MRD^{pos}	21 (2)	-
Not available	32 (3)	-
Karnofsky score, % (range)	90 (40-100)	90 (50-100)
HCT Comorbidity Index, N (%)		
0-1	339 (35)	324 (35)
2-3	347 (35)	334 (36)
≥4	293 (30)	268 (29)
Conditioning intensity, N (%)		
MAC	583 (60)	558 (60)
Non-MAC	396 (40)	368 (40)

Continued on following page.

HLA matching, N (%)		
HLA-identical related donor	227 (23)	219 (24)
HLA-matched unrelated donor	482 (49)	456 (49)
1-2 allele/antigen mismatched unrelated donor	101 (10)	95 (10)
HLA-haploidentical donor	37 (4)	34 (4)
UCB	132 (13)	122 (13)
Source of stem cells, N (%)		
PB	766 (78)	728 (79)
BM	81 (8)	76 (8)
UCB	132 (13)	122 (13)
GvHD prophylaxis, N (%)		
CNI + MMF ± sirolimus	473 (48)	444 (48)
CNI + MTX ± other	370 (38)	352 (38)
PTCy	122 (12)	116 (13)
Other	14 (1)	14 (2)

*Risk stratification according to the presence of cytogenetic abnormalities only. **Secondary AML was defined as disease following an antecedent hematologic disorder or treatment with systemic chemotherapy, radiotherapy, or both for a different disorder. ***ANC ≥1,000/μL and platelets ≥100,000/μL. ****Normal cytogenetics in cytogenetically normal AML or missing cytogenetics at diagnosis. AML: acute myeloid leukemia; ANC: absolute neutrophil count; BM: bone marrow; CNI: calcineurin inhibitor; ELN: European LeukemiaNet; HCT: hematopoietic cell transplantation; MAC: myeloablative conditioning; MMF: mycophenolate mofetil; MRD: measurable residual disease; MTX: methotrexate; PB: peripheral blood; PTCy: post transplantation cyclophosphamide; UCB: umbilical cord blood; WBC: total white blood cell count. GvHD: graft-*versus*-host disease.

ditioning intensity, with similar outcomes for MRD-positive patients across regimens but better outcomes for MRD-negative patients with myeloablative conditioning. However, including this interaction term in models only minimally improved the predictive performance for relapse (C-statistic 0.71 *vs*. 0.70) but not for RFS or OS.

Currently, there is focus on pre-HCT testing because of the perceived value in using data for decision-making regarding use/not use of allogeneic HCT and specifics of the allografting approach. However, there is interest in MRD-directed post-HCT interventions, for which early post-HCT MRD data could be useful. That is because our recent data indicated that, across conditioning intensities, combined use of pre-HCT MRD data and early post-HCT MRD data, obtained 20-40 days after HCT ("peri-HCT MRD dynamics"), improves risk assessment over isolated pre-HCT MRD assessments and identifies four groups of patients with distinct clinical outcomes.[8] We therefore performed a second set of analyses, restricting the dataset to 926 patients with pre- and early post-HCT MRD data available (Table 1). Patients in the MRDneg/MRDpos group (n=21) were excluded from these models due to unstable hazard ratio estimates caused by most patients in the subgroup having relapsed by time 0 of the relapse and RFS measurements. In day +40 landmark analyses, peri-HCT MRD dynamics were strongly associated with relapse, RFS, and OS and, weakly, with NRM; as in the entire cohort, several other factors were associated with relapse, RFS, OS, and/or NRM in this patient subset (*Online Supplementary Table S2*). In univariate models, peri-HCT MRD dynamics were equally accurate as pre-HCT MRD data for the prediction of relapse (C-statistic=0.70), RFS (C-statistic=0.63), and OS (C-statistic=0.61; Table 2). Likewise, C-statistic values in multivariable models remained unchanged when data on peri-HCT MRD dynamics rather than pre-HCT MRD status were included (0.71, 0.68, and 0.67 for the prediction of relapse, RFS, and OS).

We previously reported that detailed molecular data modestly improve the prediction of therapeutic resistance or survival in adults receiving intensive AML chemotherapy.[11] We therefore considered that refined cytogenetic/molecular risk categorization could improve post-HCT outcome prediction. As a limitation of our dataset, molecular data sufficient for risk classification based on ELN 2017 criteria were only available for 260 patients in the more recent period since 2016 when extended molecular profiling became routine at our institution. In this latter subset, the addition of ELN 2017 cytogenetic/molecular disease risk to a multivariable model including pre-HCT MRD status yielded C-statistics of 0.70, 0.69 and 0.68 for relapse, RFS and OS.

Decision-making in most areas of medicine, including HCT, entails important uncertainties.[12] While MRD data increased the accuracy of outcome prediction in our cohort, our finding of C-statistics not exceeding 0.71 for relapse, 0.68 for RFS, and 0.67 for OS highlights that our ability to predict important outcomes in adults with AML undergoing allogeneic HCT is limited even with MRD information available. This is reminiscent of findings we obtained when building multivariable models to predict outcomes after intensive AML chemotherapy.[13] With this, our data caution against over-reliance on MRD data (or any other routine clinical information) to guide decision-making and prognostication in individual AML patients considered for allografting or post-HCT interventions.

Several reasons may underlie the modest contribution of MRD data to post-HCT outcome prediction. First, there are inherent limitations to MRD testing that can result in patient misclassification.[14] Second, it is conceivable that MRD used as a binary readout rather than continuous biomarker could

Table 2. C-statistics for univariate and multivariate Cox regression analyses.

	Pre-HCT MRD cohort (N=979)			Peri-HCT MRD dynamics cohort (N=926)		
Parameter	**Relapse**	**RFS**	**OS**	**Relapse**	**RFS**	**OS**
Univariate analyses						
Pre-HCT MRD status	0.64	0.60	0.59	0.66	0.62	0.60
Post-HCT MRD status day +20-40	-	-	-	0.57	0.55	0.54
Peri-HCT MRD dynamics	-	-	-	0.67	0.63	0.61
2017 ELN cytogenetic risk at diagnosis*	0.59	0.56	0.55	0.59	0.56	0.55
Pre-HCT cytogenetics**	0.57	0.56	0.56	0.57	0.57	0.57
Age at HCT	0.55	0.57	0.57	0.53	0.57	0.57
Time between CR and HCT	0.55	0.52	0.52	0.55	0.53	0.52
Conditioning (MAC *vs.* non-MAC)	0.54	0.56	0.55	0.52	0.55	0.55
Karnofsky score	0.54	0.55	0.55	0.54	0.55	0.55
NPM1 mutational status at diagnosis	0.54	0.53	0.53	0.55	0.52	0.53
FLT3-ITD mutational status at diagnosis	0.54	0.52	0.51	0.54	0.51	0.50
Disease status at HCT (CR2 *vs.* CR1)	0.53	0.53	0.53	0.53	0.53	0.53
ANC recovery before HCT	0.52	0.52	0.52	0.52	0.52	0.52
Secondary AML*** (*vs. de novo*)	0.52	0.52	0.52	0.52	0.52	0.52
Multivariate analyses						
Basic covariates****	0.64	0.62	0.62	0.64	0.62	0.63
Basic covariates + *NPM1/FLT3*-ITD mutational status at diagnosis	0.64	0.62	0.62	0.64	0.62	0.63
Basic covariates + pre-HCT cytogenetics	0.66	0.63	0.63	0.65	0.64	0.64
Basic covariates + pre-HCT MRD status	0.69	0.66	0.65	0.70	0.68	0.66
Basic covariates + pre-HCT cytogenetics + pre-HCT MRD status	0.70	0.66	0.65	0.71	0.68	0.67
Basic covariates + pre-HCT cytogenetics + post-HCT MRD status day +20-40	-	-	-	0.67	0.66	0.65
Basic covariates + pre-HCT cytogenetics** + peri-HCT MRD dynamics	-	-	-	0.71	0.68	0.67

*Risk stratification according to the presence of cytogenetic abnormalities only. **Abnormal *vs.* normalized *vs.* missing/non-informative. ***Secondary AML was defined as disease following an antecedent hematologic disorder or treatment with systemic chemotherapy, radiotherapy, or both for a different disorder. ****Age at HCT, 2017 ELN cytogenetic risk at diagnosis (favorable *vs.* intermediate *vs.* adverse *vs.* unknown), remission (first *vs.* second), time between most recent CR and HCT, Karnofsky score, recovery of ANC before HCT to 1,000/μL (yes *vs.* no), conditioning intensity (MAC *vs.* non-MAC), secondary AML (*vs. de novo*). AML: acute myeloid leukemia; ANC: absolute neutrophil count; CR: complete remission; ELN: European LeukemiaNet; HCT: hematopoietic cell transplantation; MAC: myeloablative conditioning; MRD: measurable residual disease; OS: overall survival; RFS: relapse-free survival.

decrease prediction accuracy. However, we explored this through non-linear modeling of relationships between MRD and post-HCT endpoints with cubic spline functions and did not find noticeable differences in relapse or survival risks across MRD levels. This suggests the consideration of any detectable level of MRD as positive did not negatively impact models' predictive accuracies. Third, since different MRD test modalities provide complementary rather than congruent prognostic information,[15] the incorporation of molecular MRD data may improve outcome prediction. Unfortunately, molecular MRD testing was not routine part of the pre-HCT evaluation in the study period. Fourth, because of the lack of detailed molecular data, we could not assess a possible interaction between MRD and genetic/molecular disease risk. Fifth, NRM represents a major contribution to mortality after allogeneic HCT. Imperfectly accounted for in competing risks analyses, NRM events may interfere with the ability to predict relapse and survival outcomes with MRD data. Lastly, sustained AML remission after allogeneic HCT largely depends on graft-*versus*-leukemia effects. Pre-HCT and early post-HCT MRD primarily measure AML sensitivity to prior chemotherapy and conditioning regimens and may not predict immune-mediated eradication of leukemic cells. The identification of robust immune biomarkers predicting graft-*versus*-leukemia effects after allogeneic HCT could help refine relapse risk stratification.

Authors

Eduardo Rodríguez-Arbolí,[1,2] Megan Othus,[3] Corentin Orvain,[1] Lucas C.

Zarling,[4] Brenda M. Sandmaier,[1,5] Filippo Milano,[1,5] Gary Schoch,[1] Chris Davis,[1] H. Joachim Deeg,[1,5] Frederick R. Appelbaum,[1,5] Rainer Storb[1,5] and Roland B. Walter[1,6,7,8]

[1]Clinical Research Division, Fred Hutchinson Cancer Center, Seattle, WA, USA; [2]Department of Hematology, Hospital Universitario Virgen del Rocío, Instituto de Biomedicina de Sevilla (IBIS/CSIC/CIBERONC), University of Seville, Seville, Spain; [3]Public Health Science Division, Fred Hutchinson Cancer Center, Seattle, WA, USA; [4]Department of Medicine, Residency Program, University of Washington, Seattle, WA, USA; [5]Department of Medicine, Division of Medical Oncology, University of Washington, Seattle, WA, USA; [6]Department of Medicine, Division of Hematology, University of Washington, Seattle, WA, USA; [7]Department of Laboratory Medicine and Pathology, University of Washington, Seattle, WA, USA and [8]Department of Epidemiology, University of Washington, Seattle, WA, USA.

Correspondence:
R.B. WALTER - rwalter@fredhutch.org

https://doi.org/10.3324/haematol.2022.281631

Received: June 26, 2022.
Accepted: September 14, 2022.
Prepublished: September 22, 2022.

Disclosures

No conflicts of interest to disclose.

Contributions

ERA contributed to the collection and assembly of data, data analyses and interpretation, and drafting of the manuscript. MO conducted all statistical analyses and participated in data interpretation and drafting of the manuscript. CO, LCZ, GS, and CD contributed to the collection and assembly of data. BMS, FM, HJD, FRA and RS contributed to the provision of study material, patient recruitment, and acquisition of data. RBW conceptualized and designed this study and participated in data analysis and interpretation and drafting of the manuscript. All authors revised the manuscript critically and gave final approval to submit for publication.

Acknowledgments

The authors acknowledge the excellent care provided by the physicians and nurses of the HCT teams, the staff in the Long-Term Follow-up office at the Fred Hutchinson Cancer Center, and the patients for participating in our research protocols.

Funding

Research reported in this publication was supported by grants P01-CA078902, P01-CA018029, and P30-CA015704 from the National Cancer Institute/National Institutes of Health (NCI/NIH), Bethesda, MD, USA.

Data-sharing statement

Original data can be made available upon request to the corresponding author.

References

1. Hourigan CS, Gale RP, Gormley NJ, Ossenkoppele GJ, Walter RB. Measurable residual disease testing in acute myeloid leukaemia. Leukemia. 2017;31(7):1482-1490.
2. Short NJ, Zhou S, Fu C, et al. Impact of measurable residual disease on survival outcomes in patients with acute myeloid leukemia: a meta-analysis. JAMA Oncol. 2020;6(12):1890-1899.
3. Heuser M, Freeman SD, Ossenkoppele GJ, et al. 2021 Update on MRD in acute myeloid leukemia: a consensus document from the European LeukemiaNet MRD Working Party. Blood. 2021;138(26):2753-2767.
4. Hourigan CS, Dillon LW, Gui G, et al. Pre-MEASURE: multicenter evaluation of the prognostic significance of measurable residual disease testing prior to allogeneic transplantation for adult patients with AML in first remission. J Clin Oncol. 2022;40(16 Suppl):7006.
5. Kattan MW. Judging new markers by their ability to improve predictive accuracy. J Natl Cancer Inst. 2003;95(9):634-635.
6. Feng Z. Classification versus association models: should the same methods apply? Scand J Clin Lab Invest Suppl. 2010;242:53-58.
7. Walter RB, Estey EH. Selection of initial therapy for newly-diagnosed adult acute myeloid leukemia: limitations of predictive models. Blood Rev. 2020;44:100679.
8. Paras G, Morsink LM, Othus M, et al. Conditioning intensity and peritransplant flow cytometric MRD dynamics in adult AML. Blood. 2022;139(11):1694-1706.
9. Döhner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood. 2017;129(4):424-447.
10. Orvain C, Wilson JA, Fang M, et al. Relative impact of residual cytogenetic abnormalities and flow cytometric measurable residual disease on outcome after allogeneic hematopoeitic cell transplantation in adult acute myeloid leukemia. Haematologica. 2022 Aug 4. doi: 10.3324/haematol.2022.281585. [Epub ahead of print]
11. Walter RB, Othus M, Paietta EM, et al. Effect of genetic profiling on prediction of therapeutic resistance and survival in adult acute myeloid leukemia. Leukemia. 2015;29(10):2104-2107.
12. Dhawale T, Steuten LM, Deeg HJ. Uncertainty of physicians and patients in medical decision making. Biol Blood Marrow Transplant. 2017;23(6):865-869.
13. Othus M, Wood BL, Stirewalt DL, et al. Effect of measurable ('minimal') residual disease (MRD) information on prediction of relapse and survival in adult acute myeloid leukemia. Leukemia. 2016;30(10):2080-2083.
14. Othus M, Gale RP, Hourigan CS, Walter RB. Statistics and measurable residual disease (MRD) testing: uses and abuses in hematopoietic cell transplantation. Bone Marrow Transplant. 2020;55(5):843-850.
15. Jongen-Lavrencic M, Grob T, Hanekamp D, et al. Molecular minimal residual disease in acute myeloid leukemia. N Engl J Med. 2018;378(13):1189-1199.