

Original Paper

# Predictive Models for Long Term Survival of AML Patients Treated with Venetoclax and Azacitidine or 7+3 Based on Post Treatment Events and Responses: Retrospective Cohort Study

Nazmul Islam<sup>1</sup>, MS, MBA, PhD; Jamie S Reuben<sup>1</sup>, MPH, MPPA; Justin Dale<sup>1</sup>, BBA; James W Coates<sup>1</sup>, BS, MS; Karan Sapia<sup>1</sup>, MS; Frank R Markson<sup>1</sup>, BS; Craig T Jordan<sup>2</sup>, PhD; Clay Smith<sup>3</sup>, MD

<sup>1</sup>RefinedScience, Aurora, CO, United States

<sup>2</sup>Division of Hematology, University of Colorado Anschutz, Aurora, CO, United States

<sup>3</sup>Department of Medicine, University of Colorado Anschutz, Aurora, CO, United States

**Corresponding Author:**

Nazmul Islam, MS, MBA, PhD

RefinedScience

Bioscience 1, Suite 175

12635 E Montview Blvd

Aurora, CO, 80045

United States

Phone: 1 9196049705

Email: [nazmul.islam@refinedscience.com](mailto:nazmul.islam@refinedscience.com)

## Abstract

**Background:** The treatment of acute myeloid leukemia (AML) in older or unfit patients typically involves a regimen of venetoclax plus azacitidine (ven/aza). Toxicity and treatment responses are highly variable following treatment initiation and clinical decision-making continually evolves in response to these as treatment progresses. To improve clinical decision support (CDS) following treatment initiation, predictive models based on evolving and dynamic toxicities, disease responses, and other features should be developed.

**Objective:** This study aims to generate machine learning (ML)-based predictive models that incorporate individual predictors of overall survival (OS) for patients with AML, based on clinical events occurring after the initiation of ven/aza or 7+3 regimen.

**Methods:** Data from 221 patients with AML, who received either the ven/aza (n=101 patients) or 7+3 regimen (n=120 patients) as their initial induction therapy, were retrospectively analyzed. We performed stratified univariate and multivariate analyses to quantify the association between toxicities, hospital events, and short-term disease responses and OS for the 7+3 and ven/aza subgroups separately. We compared the estimates of confounders to assess potential effect modifications by treatment. 17 ML-based predictive models were developed. The optimal predictive models were selected based on their predictability and discriminability using cross-validation. Uncertainty in the estimation was assessed through bootstrapping.

**Results:** The cumulative incidence of posttreatment toxicities varies between the ven/aza and 7+3 regimen. A variety of laboratory features and clinical events during the first 30 days were differentially associated with OS for the two treatments. An initial transfer to intensive care unit (ICU) worsened OS for 7+3 patients (aHR 1.18, 95% CI 1.10-1.28), while ICU readmission adversely affected OS for those on ven/aza (aHR 1.24, 95% CI 1.12-1.37). At the initial follow-up, achieving a morphologic leukemia free state (MLFS) did not affect OS for ven/aza (aHR 0.99, 95% CI 0.94-1.05), but worsened OS following 7+3 (aHR 1.16, 95% CI 1.01-1.31) compared to that of complete remission (CR). Having blasts over 5% at the initial follow-up negatively impacted OS for both 7+3 ( $P<.001$ ) and ven/aza ( $P<.001$ ) treated patients. A best response of CR and CR with incomplete recovery (CRi) was superior to MLFS and refractory disease after ven/aza ( $P<.001$ ), whereas for 7+3, CR was superior to CRi, MLFS, and refractory disease ( $P<.001$ ), indicating unequal outcomes. Treatment-specific predictive models, trained on 120 7+3 and 101 ven/aza patients using over 114 features, achieved survival AUCs over 0.70.

**Conclusions:** Our findings indicate that toxicities, clinical events, and responses evolve differently in patients receiving ven/aza compared with that of 7+3 regimen. ML-based predictive models were shown to be a feasible strategy for CDS in both forms of AML treatment. If validated with larger and more diverse data sets, these findings could offer valuable insights for developing AML-CDS tools that leverage posttreatment clinical data.

**KEYWORDS**

Leukemia, Myeloid, Acute; Venetoclax; Azacitidine; Anthracycline; Arabinoside, Cytosine; Clinical Decision Support; Clinical Informatics; Machine Learning; Predictive Model; Overall Survival

## Introduction

Acute myeloid leukemia (AML) is an aggressive malignancy of the myeloid cells in the hematopoietic system [1]. Without treatment, patients can die within days to months due to infection, bleeding, organ damage, or other complications. The treatment approaches for AML vary significantly based on the patient's ability or willingness to tolerate intensive therapy [1,2]. For young and fit patients, a typical intensive therapy approach involves induction treatment with anthracycline and cytosine arabinoside, commonly known as 7+3 therapy. This is followed by additional consolidative chemotherapy or an allogeneic stem cell transplantation (alloSCT), depending on the genetic features of the AML at diagnosis, as well as the clinical status of the patient and the AML after induction therapy [3,4]. This intensive approach is potentially curative but is associated with high morbidity, mortality, cost, and prolonged hospital stays. For patients who are not suitable for, or choose to decline, this intensive approach due to age, fitness, or personal preference at diagnosis, the Bcl-2 inhibitor venetoclax, in combination with a hypomethylating agent such as azacitidine or decitabine, has become a new standard of care [5-7]. This strategy is typically aimed at prolonging life rather than achieving a cure and is associated with less morbidity, treatment-related mortality, and time spent in the hospital compared with intensive approaches [8].

We and others have described a variety of features of both patients and AML at diagnosis that are associated with long-term survival and other outcomes following treatment with either intensive approaches or venetoclax plus azacitidine (ven/aza)-based treatments [9-11]. However, the treatment course for patients with AML is highly variable, and factors such as "fitness" can change significantly, for better or worse, during treatment. Additionally, there is significant variability in AML responses to therapy during treatment, which are difficult to predict at diagnosis. As a result, prognosis and clinical decision-making can evolve significantly based on events and responses occurring after the initiation of treatment. Therefore, identifying key prognostic features that develop following treatment and are associated with long-term disease behavior and survival is essential for refining clinical decision-making over time. For intensive treatment approaches, events such as the achievement of a morphologic complete remission (CR), the presence or absence of minimal residual disease (MRD) detected by flow cytometry or next-generation sequencing, and other AML-related assessments that occur following the initiation of therapy are predictive of long-term outcomes [12-28]. Many of these early response indicators are useful for guiding subsequent therapeutic decisions. For example, the presence of MRD after induction therapy with 7+3 or other intensive treatments can predict the success of alloSCT, guide the choice of transplant type, and identify high-risk

patients who may benefit from post-transplant maintenance therapy [16,18-20,29-33]. In ven/aza treatment, achieving MRD negativity is associated with improved event-free survival and overall survival (OS) [34]. However, in contrast to intensive approaches, there is limited knowledge about how toxicities, early clinical events, and short-term treatment responses are associated with disease behavior and long-term patient outcomes with this therapy.

To address this gap, we evaluated clinical events, toxicities, short-term outcomes, biomarkers, and other features occurring after the initiation of treatment with either 7+3 or ven/aza to understand their association with OS. Additionally, we developed models to assess the long-term dynamic behavior of responses to 7+3 and ven/aza based on short-term disease responses. These studies reveal substantial differences in the clinical and AML features that evolve with the 2 different treatments and highlight how these differences impact prognosis and clinical decision-making.

## Methods

### Patient Populations

Adult, newly diagnosed AML patients who received initial induction therapy with either the ven/aza regimen or the 7+3 regimen at the University of Colorado Hospital (UCH) between January 1, 2013, and December 31, 2020, were included in the study. Patients with acute promyelocytic leukemia and those who voluntarily withdrew within less than 28 days of treatment were excluded. Patient baseline characteristics are summarized in Table S1 in [Multimedia Appendix 1](#). Note that this patient cohort is a subset of the analytical data set as previously described [10]. For exploratory analyses, 120 patients treated with 7+3 and 101 patients treated with ven/aza were included (Figure S1 in [Multimedia Appendix 1](#)). Best response analyses were based on 118 out of 120 (98.3%) of the 7+3 patients and all (101/101, 100%) of the ven/aza patients (including those who died before response assessment). For the multistate transition analyses, 115 out of 120 (95.8%) of the 7+3 patients and 98 out of 101 (97.0%) of the ven/aza patients had sufficient data after excluding those without at least one response assessment or who died before their first response assessment. Additionally, 111 out of 120 (92.5%) of the 7+3 patients and 91 out of 101 (90.1%) of the ven/aza patients had 30-day follow-up data adequate for developing machine learning (ML) models.

### Ethical Considerations

This study was a retrospective analysis utilizing a limited data set. A full waiver of consent and a full waiver of Health Insurance Portability and Accountability Act (HIPAA) authorization were granted by the Colorado Multiple Institutional Review Board (approval number 18-1861). The limited data set was securely stored on a HIPAA-compliant,

cloud-based data platform, and accessible only to members of the study team.

### Outcome Definitions

Treatment responses, including CR, CR with incomplete hematologic recovery (CRi), morphologic leukemia-free state (MLFS), progressive disease, and stable disease, were defined according to the standard 2017 European LeukemiaNet (ELN) criteria [35]. A patient was classified as “refractory” if the disease persisted after 90 days from the start of treatment or if the disease worsened or showed no improvement at any point during the treatment cycle. Toxicity variables were graded according to the National Comprehensive Cancer Network (NCCN) Common Terminology Criteria for Adverse Events (CTCAE) guidelines [35-37]. Ejection fraction toxicity was defined as detailed in Table S10 in [Multimedia Appendix 1](#). “Induction events” occurred during the initial treatment hospitalization. The “Day<sub>15-55</sub>” disease assessment refers to patient examinations, laboratory analyses, and bone marrow biopsy (BMB) analyses performed closest to day 30 after the initiation of treatment, but within days 15-55 to accommodate variations in assessment timing. “Day 30 readmission events” were defined as clinical events that occurred at least one day after a patient’s discharge from the initial treatment hospitalization and at least one day before the “Day<sub>15-55</sub>” BMB.

### Statistical Learning

Structured and unstructured electronic medical record data were integrated into a heme data mart on the Google Cloud Platform (Alphabet Inc.), as previously described [10]. Descriptive summary statistics of confounders were provided for both 7+3 and ven/aza treatments. Systematic differences between treatments were compared using the Mann-Whitney *U* test, Fisher exact test (for small sample sizes), chi-square test, and standardized mean differences (SMDs). Kaplan-Meier analyses were performed for OS with 95% CIs, assuming right censoring. *P* values for testing the equality of survival curves were reported using log-rank (LR), Tarone-Ware (TW), and Fleming-Harrington (FH) methods. All hypotheses were 2-sided. Cumulative incidence functions (CIFs) for hazards were reported for toxicity along with 95% CIs, LR-based *P* values, and median time to reach the worst toxicity grading from baseline. Multistate survival analyses were conducted using follow-up BMB responses. Occupation probabilities of disease states were estimated using the Aalen-Johansen estimation technique [38-40]. Transition probabilities for moving from one disease state to another over time were estimated assuming a Markov process, with standard errors reported using bootstrap methods across 300 runs [41]. Multivariable Cox proportional hazards (Cox-PH) models with a ridge penalty (ie, L2 norm penalty) were fitted to adjust for multicollinearity, and estimates of adjusted hazard ratios (aHRs) were reported [42]. Before fitting multivariable models, numeric variables were categorized based

on clinically meaningful thresholds to enhance interpretability. Noise variables were filtered out using a univariate approach based on accelerated failure time (AFT) models. Tuning parameters for the ridge penalty were selected using the 10-fold cross-validation (CV) approach. Bias-corrected 95% CIs for aHRs were constructed using the fractional random weight bootstrap method with 2000 runs, where weights were computed from a univariate Dirichlet distribution [43].

### Development and Validation of Prognosis Models

The steps for training and evaluating ML models are depicted in [Figures 1](#) and [2](#). The process consists of 2 stages. First, internal validation based on CV (steps 1-8) was conducted to select the appropriate ML model for each treatment separately. Second, subject-specific OS predictions, conditional on observed covariates, and the corresponding uncertainty quantification were performed using the selected treatment-specific ML models (steps 10-11). A total of 17 different models ranging from statistical learning-, ML-, and deep learning (DL)-based survival models were used to assess long-term outcomes. These included ensemble-based methods such as the random survival forest (RSF) [44], survival forest with bagging, and conditional inference survival forest [45], as well as Cox-PH models with boosting, penalized Cox-PH models, and parametric AFT models [46] with exponential, Weibull, and log-logistic error structures. These models generated OS probabilities by leveraging over 114 features, as highlighted in Table S12 in [Multimedia Appendix 1](#). The prognostic variable list was further enhanced by creating binary variables based on the first and fifth quintiles of numeric laboratory variables. Regularization penalties [42,47-50] were applied to reduce the risk of overfitting. The penalty terms included ridge, LASSO, elastic-net (eNet), smoothly clipped absolute deviation (SCAD), minimax concave penalty (MCP), adaptive SCAD, adaptive MCP, adaptive eNet, adaptive MCP with L2 norm penalty (mNet), and adaptive SCAD with L2 norm penalty (sNet) [47]. Adaptive models were fitted in 2 stages: in the first stage, models were fitted with ridge penalties, and in the second stage, models were refitted with covariate weights calculated as the reciprocal of parameter estimates from the first stage. Additionally, DL-based survival models (Deep-Surv [51], Deep-LogHaz [52], and Deep-Hit [53]) with 2 hidden layers exploiting neural network structures were used. Tuning parameters for the ensemble-based approaches and DL models were selected using a combination of grid search and CV. For Cox-PH models, regularization penalties were selected using 10-fold CV. Parametric AFT models were fitted with a reduced set of variables. To minimize dimensionality and avoid collinearity in AFT models, a univariate filtering approach was applied, where only variables with Bonferroni-corrected *P* values below a prespecified threshold of 0.20 were included in the final multivariable AFT models. For additional technical details, please refer to [Multimedia Appendix 1](#).

**Figure 1.** ML architecture. Notation and description of 11 steps for the development of ML models, optimum model selection, validation, prediction, and uncertainty quantification for a newly diagnosed patient with AML. AML: acute myeloid leukemia; AUC: area under the curve.

### Steps 1-8: Internal validation

- 1 Select clinical features for machine learning (ML) models; standardize and organize the analytical dataset tailored to each ML model; complete feature engineering step.
- 2 Arrange data frame with response of interest, and selected covariates; augment lab variables; initiate iterations for leave-one-out-cross-validation (LOOCV) for each type of ML models.
- 3 Create training dataset via the LOOCV technique leaving one subject out and use the left-out subject as a test set.
- 4 Use the grid search/cross-validation (CV) techniques to obtain optimum tuning parameters in the corresponding training sets; repeat this step independently for each LOOCV run.
- 5 Train respective ML model with the chosen set of tuning parameters; do it for each ML model separately.
- 6 Obtain the corresponding test set prediction using the respective trained model.
- 7 Calculate classification metrics (eg, survival AUC) based on Youden index-based cutoff.
- 8 Select the “best” performing ML model among 17 competing algorithms with respect to the stacked test results based on LOOCV runs.

### Step 9: Adversarial validation

- 9 Perform adversarial validation to detect any potential data drift in distribution between the training and independent validation set; apply selected models on the independent external cohort and re-evaluate model performances.

### Step 10: Prospective validation

- 10 Retrain the selected “best” ML model exploiting all patients’ information based on the corresponding analytical dataset; store the model artifacts/parameters; generate patient specific conditional prediction for a new patient.

### Step 11: Uncertainty quantification

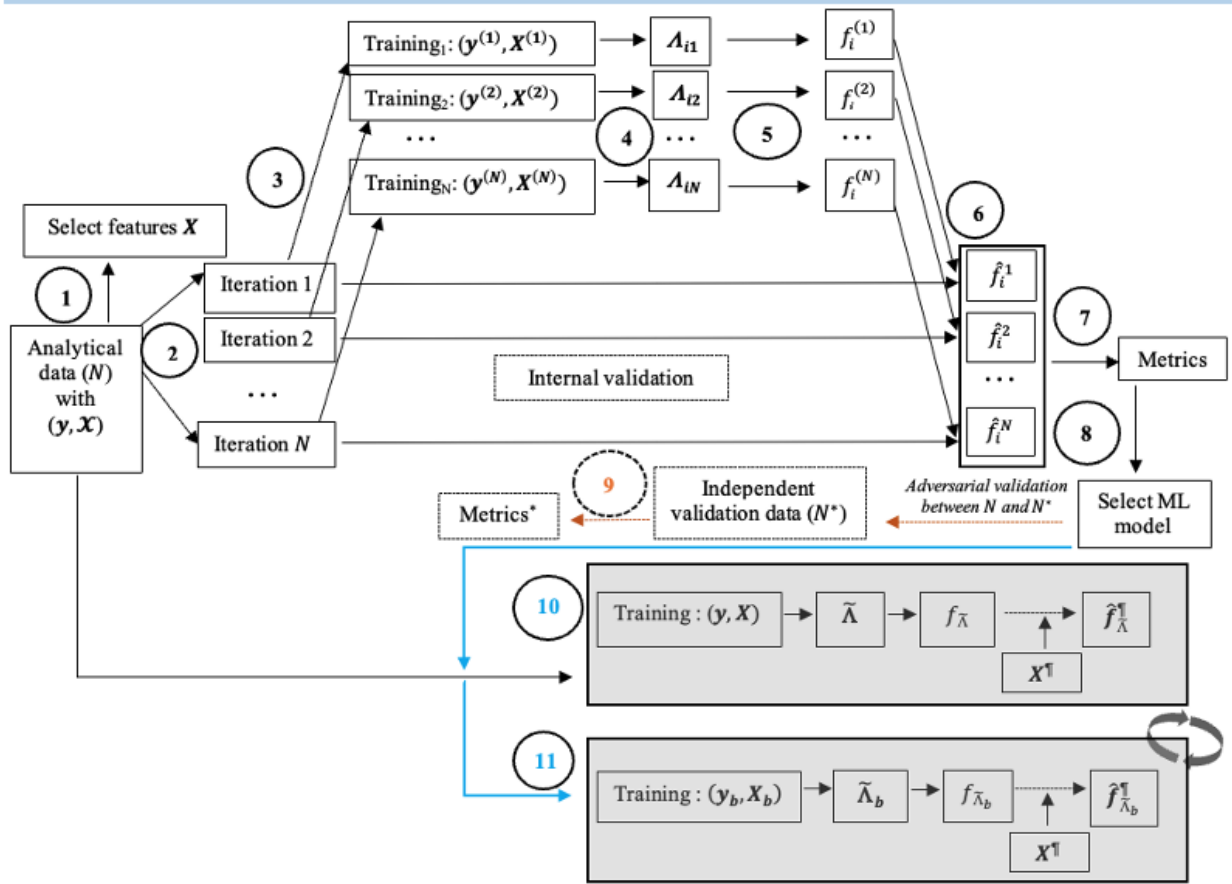
- 11 Quantify variation in prediction (eg, standard errors, inter quartile range) with respect to  $B$  (300 runs) nonparametric bootstrap samples with replacement; obtain percentile prediction intervals.

Internal validation was conducted using leave-one-out cross-validation, focusing on several metrics: dynamic area under the curve (AUC) of cumulative case dynamic control of receiver operative characteristics (ROC) curves (cAUC), incident case dynamic control ROC (iAUC) curves, integrated Brier scores, and time-dependent concordance (C) index and Brier score at 1-year survival (denoted by  $C_t$  and Brier<sub>t</sub>, respectively). The median (M) of cAUCs and iAUCs over event times within 2 years were reported. The model demonstrating the best numerical performance during the internal validation step was retrained using the full data set with appropriately selected tuning parameters. These models were then further

evaluated on 2 independent validation sets: 1 for each treatment arm (7+3,  $n_1=14$ ; ven/aza,  $n_2=30$  patients with AML) treated at the University of Colorado Hospital. Adversarial validation, utilizing a generalized linear model with a logit link function, was used to assess potential data drift between the training and validation sets. SMDs were computed, and the predictive performance of the models on the validation sets was reported. For out-of-sample patients, predicted probabilities were reported along with 95% percentile-based confidence bands, derived from 300 nonparametric bootstrap runs. As the primary aim of the study was to develop treatment-specific prognostic models, we did not apply multiple testing corrections for type I errors.

**Figure 2.** The processes for development of models, optimum model selection, validation, prediction, and uncertainty quantification for a newly diagnosed patient with AML. AML: acute myeloid leukemia; ML: machine learning.

$i$  : Subscript defines a machine learning (ML) model  
 $j$  : Superscript defines an iteration or a subject;  $j = 1, \dots, N$   
 $b$  : Subscript denotes a bootstrap sample;  $b = 1, \dots, B$   
 $(\mathbf{y}, \mathbf{X})$  : Observed data  
 $\mathbf{X} \subseteq \mathbf{X}$  : Features used in the respective ML model  
 $(\mathbf{y}^{(j)}, \mathbf{X}^{(j)})$  : Observed data leaving  $j$ th instance out  
 $\mathbf{A}_i$  : Set of tuning parameters corresponding to  $i$ th model  
 $f_i^{(j)}$  :  $i$ th trained model at  $j$ th iteration  
 $\hat{f}_i^j$  : Predicted probability for  $j$ th test case with respect to  $i$ th model  
 $(\mathbf{y}^*, \mathbf{X}^*)$  : Observed data for external validation; superscript \* denotes such independent set with size  $N^*$   
 $\mathbf{X}^\top$  : Covariates for a newly diagnosed patient with unknown response  
 $\hat{f}_{\bar{\Lambda}_b, \alpha}^\top$  :  $\alpha$ th percentile of  $B$  bootstrap-based predicted probabilities



## Results

### Statistical Learning–Based Comparison of Ven/Aza and 7+3 During the First 30 Days of Treatment

Summary statistics for the 7+3 and ven/aza cohorts are presented in Tables S1-S5 in [Multimedia Appendix 1](#). Ven/aza patients were older (median age 72 years, IQR 66-78 years; range 22-90 years) and had more comorbidities and high-risk AML features compared with 7+3 patients (median age 53 years, IQR 41-59 years; range 20-75 years), as previously described [54]. The ven/aza cohort had a higher prevalence of patients with an Eastern Cooperative Oncology Group score of 2 (15/59, 25%) compared with the 7+3 cohort (1/31, 3%). Various diagnostic criteria, including demographic features, comorbidities,

laboratory values, and AML pathology characteristics, were associated with OS for both the ven/aza cohort (Figure S2 in [Multimedia Appendix 1](#)) and the 7+3 cohort (Figure S4 in [Multimedia Appendix 1](#), “Diagnostic criteria”), consistent with findings described previously [10]. Specific covariates showing notable negative associations (aHR>1) for the ven/aza cohort included prior myelodysplastic syndrome (aHR 1.09, 95% CI 1.03-1.16), prior coagulopathy (aHR 1.12, 95% CI 1.05-1.20), abnormal white blood cell (WBC) count (aHR 1.05, 95% CI 1.00-1.10), blasts >20% (aHR 1.13, 95% CI 1.06-1.22), abnormal platelet count (aHR 1.09, 95% CI 1.04-1.14), elevated uric acid (aHR 1.10, 95% CI 1.03-1.18), high lactate dehydrogenase (aHR 1.14, 95% CI 1.09-1.19), poor cytogenetic risk (aHR 1.12, 95% CI 1.07-1.17), flow cytometry–based CD7 expression (aHR 1.21, 95% CI 1.14-1.28), CD34 expression

(aHR 1.06, 95% CI 1.01-1.11), CD38 expression (aHR 1.08, 95% CI 1.02-1.14), and CD11b expression (aHR 1.07, 95% CI 1.01-1.12). Multivariable analyses for the 7+3 cohort revealed similar effects in terms of direction for abnormal WBC, platelet count, uric acid, creatinine, lactate dehydrogenase, poor cytogenetic risk, myeloperoxidase (MPO), and isocitrate dehydrogenase 2 (IDH2). However, more pronounced adverse effects were observed for the ELN-2017–based adverse risk subgroup (aHR 1.06, 95% CI 1.01-1.12), *EGR1* mutation (aHR 1.14, 95% CI 1.07-1.22), and runt-related transcription factor (RUNX; aHR 1.07, 95% CI 1.01-1.13). For more details, see Figures S2 and S4 in [Multimedia Appendix 1](#). The direction of effects was reversed for *CBFB* and *NPM1* between the ven/aza and 7+3 treatment cohorts (Table S13 in [Multimedia Appendix 1](#)). Variables indicating genetic abnormalities are detailed in Table S11 in [Multimedia Appendix 1](#).

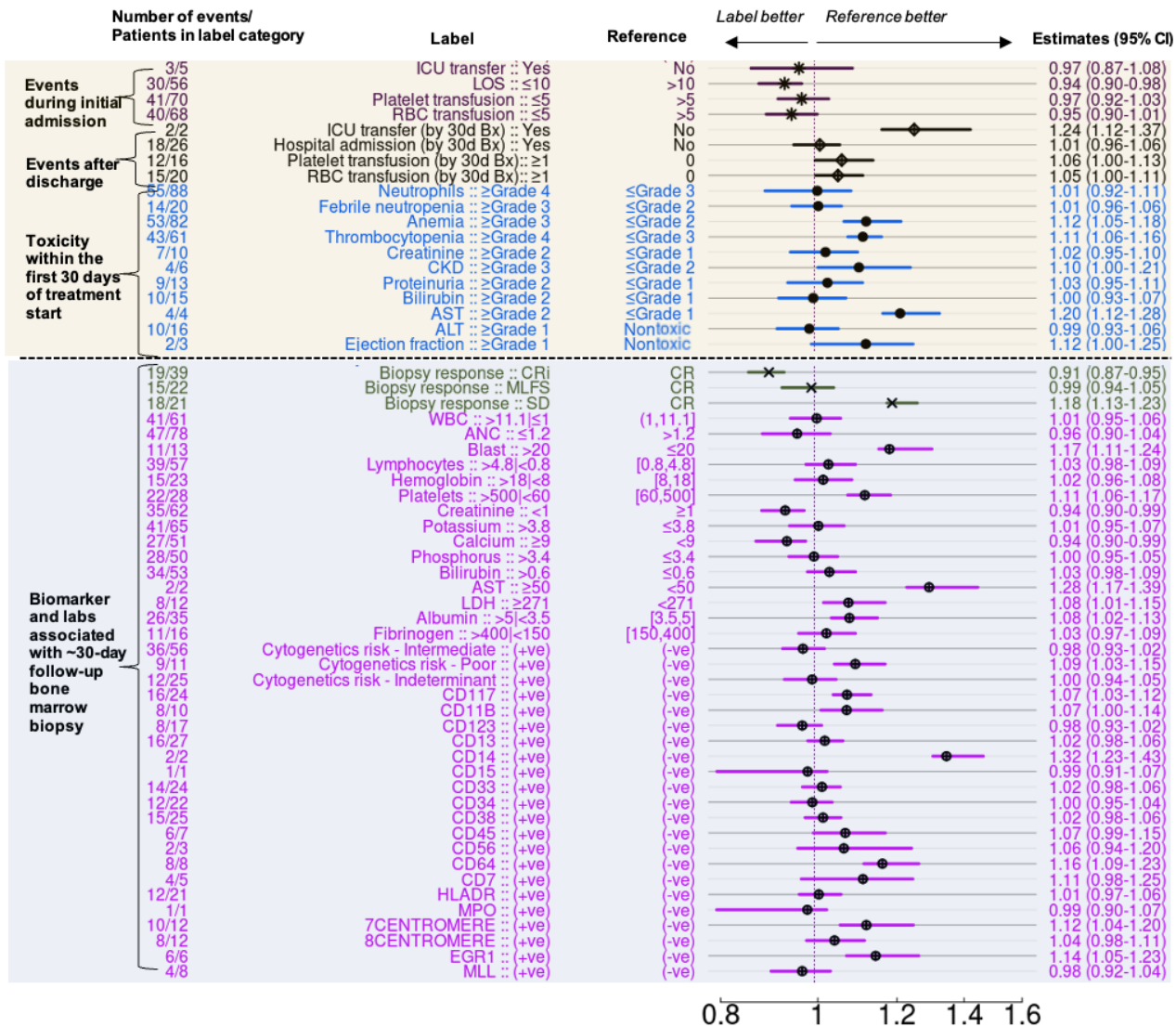
To determine whether features occurring after diagnosis and the initiation of treatment influenced long-term outcomes, we evaluated the associations between OS and factors such as toxicities, hospital events, transfusions, and short-term disease responses for both treatments separately. A summary of CTCAE toxicities, transfusions, and hospital events, including intensive care unit (ICU) transfers and readmission instances for the 2 treatment cohorts, is provided in Table S2 in [Multimedia Appendix 1](#).

For toxicities occurring after the initiation of treatment, grade  $\geq 3$  anemia (aHR 1.12, 95% CI 1.05-1.18) and grade  $\geq 4$  thrombocytopenia (aHR 1.11, 95% CI 1.06-1.16) were associated with worse OS in the ven/aza group, as observed in both multivariable and univariate analyses ([Figure 3](#), “Toxicity within the first 30 days of treatment”; [Figure S3 in Multimedia Appendix 1](#)). Elevated aspartate aminotransferase was also linked to worse OS (aHR 1.20, 95% CI 1.12-1.28) in the ven/aza group ([Figure 3](#)), but this association was not found in the 7+3 group, according to both univariate and multivariable analyses ([Figures S4 and 5A in Multimedia Appendix 1](#)). Creatinine grade  $\geq 2$  in the first 30 days of treatment was associated with worse OS in the 7+3 group, with an aHR of 1.10 (95% CI 1.01-1.20), as seen in both multivariable ([Figure S4 in Multimedia Appendix 1](#), “Toxicity within the first 30 days of treatment start”) and univariate analyses ([Figure S5B in Multimedia Appendix 1](#)). By contrast, this association appeared weaker in the ven/aza group ([Figure 3](#), “Toxicity within the first 30 days of treatment start”; [Figure S5B in Multimedia Appendix 1](#)). Despite this, worse OS was linked to chronic kidney disease (CKD) grade  $\geq 3$  in the ven/aza group, with a multivariable model-based aHR of 1.10 (95% CI 1.00-1.21; [Figure 3](#), “Toxicity within the first 30 days of treatment start”). A similar trend was observed in the 7+3 group ([Figure S4 in Multimedia Appendix 1](#), “Toxicity within the first 30 days of treatment start”). The kinetics of developing CKD differed significantly between ven/aza and 7+3 treatments ([Figure S5C in Multimedia Appendix 1](#)). In the ven/aza cohort, CKD was present at diagnosis or developed quickly, with a CIF of approximately 68% at 50 days. By contrast, CKD developed more gradually within the 7+3 cohort, showing a CIF of about 28% at 50 days. There was a trend toward worse outcomes associated with

developing ejection fraction toxicity of grade  $\geq 1$  for the ven/aza group, although the patient numbers were small ([Figure 3](#), “Toxicity within the first 30 days of treatment start”). For the 7+3 group, no significant association was found between ejection fraction grade  $\geq 1$  and OS (aHR 1.02, 95% CI 0.95-1.09; [Figure S4 in Multimedia Appendix 1](#), “Toxicity within the first 30 days of treatment start”). However, both treatment groups developed progressively higher proportions of patients with ejection fraction toxicity grade  $>1$  during the first 30 days of treatment and beyond, at roughly equal rates ([Figure S5D in Multimedia Appendix 1](#)). Febrile neutropenia, a common complication of AML therapy, did not show a clear association with OS in either treatment cohort. The aHRs were 1.01 (95% CI 0.96-1.06) for the ven/aza group and 1.02 (95% CI 0.95-1.09) for the 7+3 cohort, indicating no significant effect either by multivariable analysis ([Figure 3](#), “Toxicity within the first 30 days of treatment start”) and [Figure S4 in Multimedia Appendix 1](#), “Toxicity within the first 30 days of treatment start”), or by univariate analysis ([Figure S5E in Multimedia Appendix 1](#)). Interestingly, for the 7+3 regimen, grade  $\geq 4$  neutrophils (CIF  $\sim 100\%$  at 50 days) and grade  $\geq 3$  febrile neutropenia (CIF  $\sim 75\%$  at 50 days) occurred at high levels. By contrast, for the ven/aza cohort, there was a much lower rate of febrile neutropenia (CIF  $\sim 25\%$  at 50 days) over time, despite a nearly universal incidence of neutrophils ([Figure S5F in Multimedia Appendix 1](#)).

For transfusions occurring after the initiation of treatment, Kaplan-Meier analysis revealed that a higher number of platelet and red blood cell transfusions were associated with poorer outcomes in the 7+3 group. This association was evident in both univariate ([Figure S5G and S5H in Multimedia Appendix 1](#)) and multivariable analyses ([Figure S4 in Multimedia Appendix 1](#), “Toxicity within the first 30 days of treatment start”) and [Figure 3](#), “Toxicity within the first 30 days of treatment start”). Specifically, in the 7+3 cohort, the corresponding aHR indicated negative association with more than 5 platelet transfusions (aHR 1.11, 95% CI 1.06-1.18). This association was less pronounced in the ven/aza group. For hospital events occurring after the initiation of treatment, ICU transfer during the induction period was a particularly poor prognostic feature for patients receiving the 7+3 treatment, with an aHR of 1.18 (95% CI 1.10-1.28) indicating worse outcomes ([Figure S4 in Multimedia Appendix 1](#), “Events during initial admission,” and [Figure S5I in Multimedia Appendix 1](#)). By contrast, there was no significant association between ICU transfer following ven/aza treatment and OS during the initial admission ([Figure 3](#), “Events during initial admission”). However, if a patient was discharged and then readmitted to the hospital within the first month of treatment, ICU admission during the readmission was a poor prognostic feature for those treated with ven/aza, with an aHR of 1.24 (95% CI 1.12-1.37; [Figure 3](#), “events after discharge”). For the 7+3 group, initial admissions lasting more than 35 days were associated with worse outcomes (aHR 1.11, 95% CI 1.04-1.18; [Figure S4 in Multimedia Appendix 1](#), “Events during initial admission”). Similarly, for the ven/aza cohort, admissions lasting more than 10 days were associated with poorer outcomes (aHR 1.06, 95% CI 1.02-1.11; [Figure 3](#), “Events during initial admission”).

**Figure 3.** Adjusted hazard ratios (aHRs) for predictors of overall survival for the ven/aza cohort corresponding to events occurring during the first ~30 days of therapy. Reported are the aHRs (vertical tick) and bootstrap-based 95% CIs (horizontal line). “Reference features” correlating with a better outcome are to the right and “Label features” with a better outcome are to the left. The number of patients who died relative to the subset of patients with each feature is summarized at the far left. The table includes findings during the first 30 days and outcomes at the Day<sub>15-55</sub> bone marrow biopsy assessment at the bottom. Day<sub>15-55</sub> is defined as the day (or days) between 15 and 55 days from the initiation of treatment when bone marrow biopsy, blood test, and clinical evaluation are conducted to assess response. Different symbols for aHRs were used to differentiate the values between different types of variables. ALT: alkaline phosphatase; ANC: absolute neutrophil count; AST: aspartate transaminase; CKD: chronic kidney disease; CR: complete remission; CRi: complete remission with incomplete hematologic recovery; ICU: intensive care unit; LDH: lactate dehydrogenase; LOS: length of stay; MLFS: morphologic leukemia-free state; RBC: red blood cell; SD: stable disease; ven/aza: venetoclax plus azacitidine; WBC: white blood cell.



Next, associations between OS and patient assessments around day 30 (ie, Day<sub>15-55</sub>) after treatment initiation were analyzed. Tables S3-S5 in [Multimedia Appendix 1](#) provide a summary of follow-up patient laboratory values, biomarkers, and AML responses assessed at Day<sub>15-55</sub> for both ven/aza- and 7+3-treated patients, respectively. At the Day<sub>15-55</sub> assessment, the ven/aza cohort exhibited lower levels of alanine aminotransferase, aspartate aminotransferase, neutrophils, fibrinogen, lymphocytes, and WBC compared with that of the 7+3 cohort, with SMDs greater than 0.40 (Table S3 in [Multimedia Appendix 1](#)). Platelets and hemoglobin levels were also lower in the ven/aza cohort at the Day<sub>15-55</sub> assessment, but these differences were clinically inconsequential. Summary statistics for AML-related responses are provided in Table S5 in [Multimedia](#)

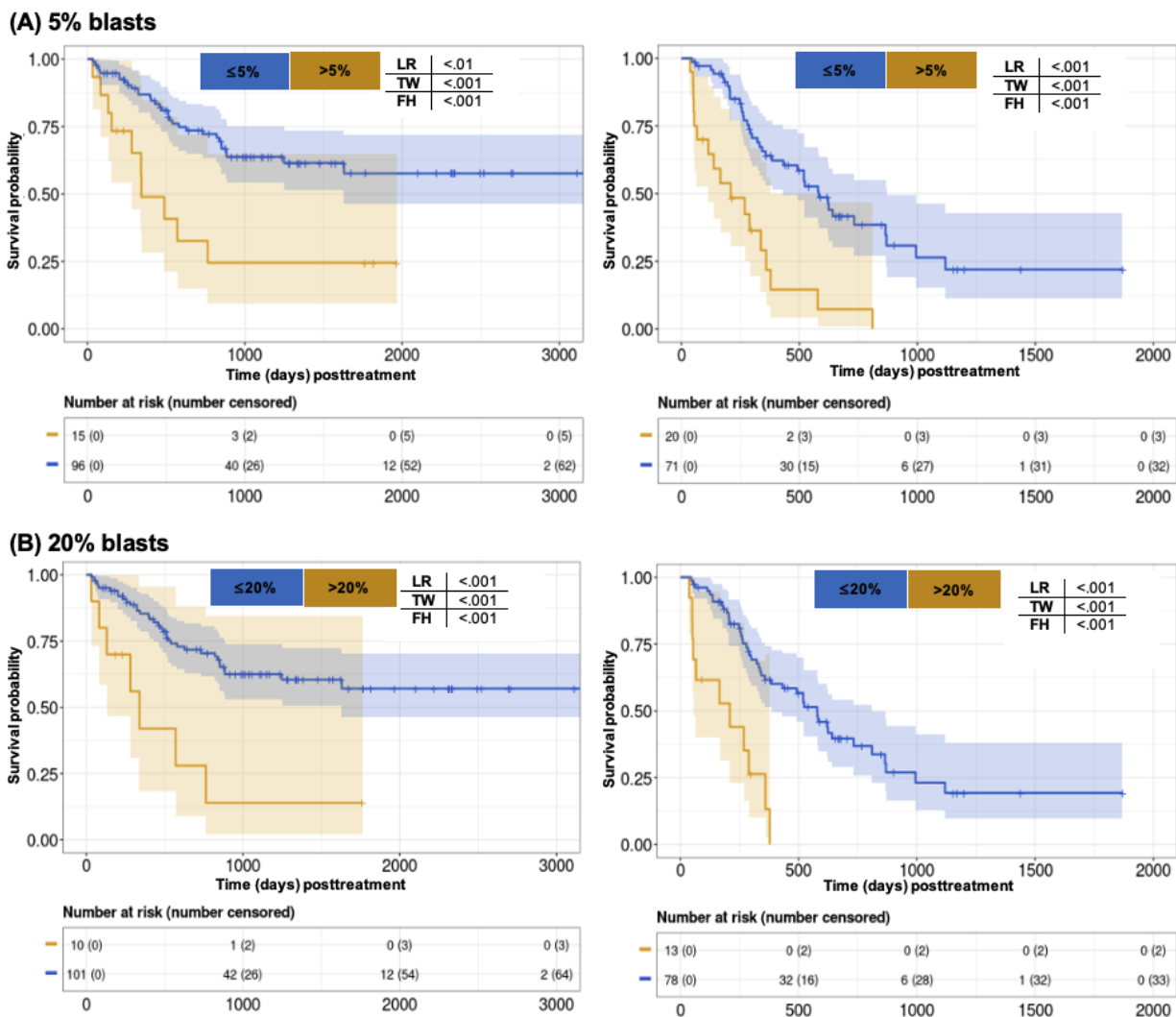
[Appendix 1](#). Notably, a lower proportion of ven/aza patients achieved CR at Day<sub>15-55</sub> (61/111, 55% for 7+3 vs 9/91, 10% for ven/aza). Conversely, a higher proportion of ven/aza patients were in CRi and MLFS compared with those treated with 7+3 (Figure S3 in [Multimedia Appendix 1](#)). Patients who achieved CR or CRi at Day<sub>15-55</sub> had better outcomes compared with those who did not, with an LR-based *P* value of <.001 (Figure S3 in [Multimedia Appendix 1](#), top panel). This was also true for patients who proceeded to receive an alloSCT, with an LR-based *P* value of <.01 (Figure S3 in [Multimedia Appendix 1](#), middle panel). Ven/aza-treated patients who achieved MLFS at Day<sub>15-55</sub> did not have worse OS compared with those who achieved CR, with an aHR of 0.99 (95% CI 0.94-1.05). By contrast, MLFS at this time point for patients treated with 7+3 was associated

with worse outcomes than CR, with an aHR of 1.16 (95% CI 1.01-1.31). This difference was observed in both univariate (Figure S3 in Multimedia Appendix 1, bottom panel) and multivariable analyses (Figure 3, “Biomarker and labs associated with ~30-day follow-up bone marrow biopsy”, and Figure S4 in Multimedia Appendix 1, “Biomarker and labs associated with ~30-day follow-up bone marrow biopsy”). Findings of persistent leukemia in the marrow as detected by flow cytometry, cytogenetics, or fluorescence in situ hybridization were associated with worse outcomes for both treatment groups according to multivariable analysis (Figure S4 in Multimedia Appendix 1, “Biomarker and labs associated with ~30-day follow-up bone marrow biopsy,” and Figure 3, “Biomarker and labs associated with ~30-day follow-up bone marrow biopsy”). Summary statistics for genetics and phenotypic features are provided in Table S4 in Multimedia Appendix 1. Specific posttreatment covariates in the ven/aza cohort demonstrated substantial negative associations, with aHRs exceeding 1, including CD117 (aHR 1.07, 95% CI 1.03-1.12), CD11B (aHR 1.07, 95% CI 1.00-1.14), CD64 (aHR 1.16, 95% CI 1.09-1.23),

7 centromere (aHR 1.12, 95% CI 1.04-1.20), and EGR1 (aHR 1.14, 95% CI 1.05-1.23; Figure 3). Similarly, for the 7+3 patients’ cohort, posttreatment covariates demonstrated inverse associations with OS, with aHRs exceeding 1. Significant negative associations were observed for cytogenetic poor risk (aHR 1.20, 95% CI 1.10-1.31), indeterminant risk (aHR 1.10, 95% CI 1.03-1.17), 8 centromere (aHR 1.22, 95% CI 1.11-1.35), EGR1 (aHR 1.14, 95% CI 1.06-1.22), and FLT3 (aHR 1.22, 95% CI 1.10-1.34; Figure S4 in Multimedia Appendix 1).

For both treatments, the presence of >20% bone marrow blasts and >5% bone marrow blasts at the Day<sub>15-55</sub> time point were associated with very poor OS by univariate analysis (Figure 4). The corresponding aHRs of >20% blasts were 1.17 (95% CI 1.11-1.24) for the patients treated with ven/aza and 1.11 (95% CI 1.04-1.20) for the patients treated with 7+3. These findings highlight that early toxicities, treatment events, and short-term responses occurring within the first month after treatment initiation are associated with OS for both 7+3 and ven/aza. However, the impact and relevance of these features vary between the 2 treatment regimens.

**Figure 4.** Univariate analysis of blasts recorded at Day<sub>15-55</sub> response assessment and long-term outcomes (7+3 left, ven/aza right). (A) >5% blasts and outcomes and (B) >20% blasts and outcomes. As described in the "Methods" section, Day<sub>15-55</sub> is defined as a bone marrow biopsy and other clinical evaluation done within 15-55 days from the initiation of treatment and closest to day 30. P values are based on log-rank (LR), Tarone-Ware (TW), and Fleming-Harington (FH) tests. ven/aza: venetoclax plus azacitidine.





## Prospective Machine Learning Predictors of OS

To translate the statistical associations between events and responses occurring after the initiation of therapy into predictions that could be potentially applied to individual patients, we developed ML-based predictive models for OS utilizing 17 different ML algorithms based on these baseline

and early posttreatment features (Table 1 and Table S12 in Multimedia Appendix 1). The modeling steps are illustrated in Figure 1, and an example of the model development process is shown in Figure 2. Detailed information on feature engineering, model specification, optimization, and final model selection is provided in the section titled “Technical Details” in Multimedia Appendix 1.

**Table 1.** List of machine learning models.

Model abbreviation	Method definition
RSF	Ensemble survival forest—random
RSB	Ensemble survival forest—bagging
Cox-Ridge	Cox regression with ridge penalty
Cox-LASSO	Cox regression with lasso penalty
Cox-Relaxed	Cox regression with relaxed lasso penalty
Cox-Elastic	Cox regression with elastic net penalty
Cox-adElastic	Cox regression with adaptive elastic net penalty
Cox-adSCAD	Cox regression with adaptive smoothly clipped absolute deviation (SCAD) penalty
Cox-adSNET	Cox regression with adaptive SCAD coupled with L2 penalty
Cox-adMCP	Cox regression with adaptive minimax concave penalty (MCP)
Cox-adMNET	Cox regression with adaptive MCP coupled with L2 penalty
Cox-Boost	Boosted Cox regression
AFT	Accelerated failure time with exponential, Weibull, and log-logistic error
CISF	Conditional inference survival forest
Deep-Surv	Cox regression with deep neural net
Deep-LogHaz	Discrete-time survival estimates by log hazard with neural net
Deep-Hit	Deep learning–based survival analysis relaxing distributional assumptions

Among all the models, Cox-Boost (Boosted Cox regression) and RSF achieved median cAUCs of 0.85 (90% CI 0.78-0.88) and 0.80 (90% CI 0.76-0.84) for the ven/aza and 7+3 cohorts, respectively (Tables 2 and 3). In an independent validation set consisting of 16 7+3 and 30 ven/aza patients, median cAUCs of 0.71 and 0.68 were observed for the ven/aza and 7+3 cohorts, respectively (Table 4). DL models resulted in less optimal performance, primarily due to the small sample size and their susceptibility to noise variables. A comparative analysis highlighting the drift between the training and validation cohorts was conducted, with details provided in Tables S6-S9 in Multimedia Appendix 1. These tables cover laboratory values (Table S6 in Multimedia Appendix 1), phenotypic features (Table S7 in Multimedia Appendix 1), genetic biomarkers (Table S8 in Multimedia Appendix 1), and clinical events (Table S9 in Multimedia Appendix 1). For a test patient, the selected ML models were used to generate patient-specific survival probabilities. Figure 5 illustrates the features (top panel) and

predicted survival (bottom panel) probabilities for a representative patient randomly selected from the independent validation set. Similarly, subject-specific analyses were conducted for 2 additional patients randomly selected from the internal validation cohorts: 1 treated with ven/aza (Figure S6 in Multimedia Appendix 1) and 1 treated with 7+3 (Figure S7 in Multimedia Appendix 1). The selected models were retrained with 120 7+3 and 100 ven/aza patients for the ven/aza test subject and with 119 7+3 and 101 ven/aza patients for the 7+3 test subject. Although exploratory and limited by sample size, these analyses illustrate that ML predictors of OS can potentially be developed based on clinical events, early disease responses, and biomarkers for both ven/aza and 7+3 treatments. However, as with the statistical analyses, the models that perform optimally are likely to vary between ven/aza and 7+3 treatments. Therefore, they should be developed and validated on a treatment-specific basis.

**Table 2.** Machine learning models of overall survival for ven/aza-treated patients based on events occurring in the first 30 days of treatment and the Day15-55 follow-up assessment.<sup>a</sup>

Overall survival (2 years) prognostic models <sup>b</sup>	Median cAUC <sup>c</sup> (5th-95th) <sup>d</sup>	Median iAUC <sup>e</sup> (5th-95th) <sup>d</sup>	$C_t^d$	Brier <sub>t</sub> <sup>f</sup>	iBrier <sup>f</sup>
RSF	0.79 (0.71-0.86)	0.68 (0.66-0.76)	0.72	0.20	0.17
RSB	0.73 (0.64-0.82)	0.61 (0.60-0.63)	0.67	0.23	0.19
Cox-Ridge	0.82 (0.79-0.89)	0.66 (0.62-0.75)	0.72	0.20	0.17
Cox-LASSO	0.85 (0.78-0.89)	0.69 (0.60-0.72)	0.73	0.20	0.16
Cox-Relaxed	0.83 (0.77-0.88)	0.61 (0.59-0.68)	0.72	0.20	0.17
Cox-Elastic	0.83 (0.78-0.86)	0.63 (0.60-0.70)	0.72	0.20	0.17
Cox-adElastic	0.80 (0.76-0.89)	0.64 (0.60-0.72)	0.66	0.24	0.19
Cox-adSCAD	0.62 (0.58-0.72)	0.52 (0.52-0.53)	0.52	0.29	0.23
Cox-adSNET	0.78 (0.73-0.85)	0.65 (0.62-0.69)	0.66	0.23	0.19
Cox-adMCP	0.62 (0.51-0.66)	0.55 (0.54-0.56)	0.55	0.26	0.22
Cox-adMNET	0.78 (0.73-0.85)	0.65 (0.62-0.69)	0.66	0.23	0.20
Cox-Boost <sup>g</sup>	0.85 (0.78-0.88)	0.66 (0.61-0.74)	0.76	0.19	0.16
AFT-Exponential	≤0.50	0.62 (0.60-0.64)	0.73	0.19	0.16
AFT-Weibull	≤0.50	0.65 (0.64-0.66)	0.72	0.22	0.17
AFT-log-logistic	≤0.50	0.65 (0.64-0.66)	0.75	0.21	0.17
CISF	≤0.50	0.64 (0.61-0.65)	0.58	0.22	0.21
Deep-Surv	≤0.50	0.52 (0.52-0.53)	0.46	0.25	0.44
Deep-Hit	0.52 (≤0.50-0.55)	0.52 (0.51-0.52)	0.38	0.32	0.34
Deep-LogHaz	≤0.50	0.53 (0.53-0.54)	0.46	0.43	0.74

<sup>a</sup>Time-dependent AUCs (ie, median cAUC and iAUC), time-dependent concordance (C) index and Brier score at 1 year, and integrated Brier score were reported. As described in the “Methods” section, Day<sub>15-55</sub> is defined as a bone marrow biopsy and other clinical evaluation done within 15-55 days from the initiation of treatment.

<sup>b</sup>See Table 1 for models and method definitions.

<sup>c</sup>cAUC: cumulative case dynamic control of receiver operative characteristics.

<sup>d</sup>The higher value (ie, close to 1) means better numerical performance.

<sup>e</sup>iAUC: incident case dynamic control of receiver operative characteristics.

<sup>f</sup>The lower value (ie, close to 0) means better numerical performance.

<sup>g</sup>The selected final model for ven/aza.

**Table 3.** Machine learning models of overall survival for 7+3-treated patients based on events occurring in the first 30 days of treatment and the Day15-55 follow-up assessment.<sup>a</sup>

Overall survival (2 years) prognostic models <sup>b</sup>	Median cAUC <sup>c</sup> (5th-95th) <sup>d</sup>	Median iAUC <sup>e</sup> (5th-95th) <sup>d</sup>	$C_t^d$	Brier <sub>t</sub> <sup>f</sup>	iBrier <sup>f</sup>
RSF <sup>g</sup>	0.80 (0.76-0.84)	0.71 (0.70-0.73)	0.74	0.12	0.16
RSB	0.78 (0.75-0.84)	0.71 (0.71-0.72)	0.73	0.13	0.15
Cox-Ridge	0.80 (0.75-0.86)	0.70 (0.69-0.73)	0.71	0.11	0.17
Cox-LASSO	0.71 (0.67-0.83)	0.64 (0.63-0.65)	0.65	0.13	0.18
Cox-Relaxed	0.68 (0.61-0.73)	0.64 (0.63-0.65)	0.65	0.13	0.19
Cox-Elastic	0.73 (0.69-0.83)	0.65 (0.65-0.66)	0.67	0.12	0.18
Cox-adElastic	0.76 (0.70-0.79)	0.68 (0.67-0.68)	0.70	0.13	0.17
Cox-adSCAD	0.70 (0.44-0.73)	0.66 (0.66-0.67)	0.68	0.14	0.18
Cox-adSNET	0.71 (0.64-0.72)	0.66 (0.66-0.67)	0.68	0.13	0.17
Cox-adMCP	0.61 (0.47-0.65)	0.62 (0.61-0.62)	0.64	0.13	0.18
Cox-adMNET	0.71 (0.65-0.73)	0.66 (0.66-0.67)	0.68	0.13	0.17
Cox-Boost	0.70 (0.61-0.79)	0.63 (0.63-0.64)	0.64	0.13	0.18
AFT-Exponential	≤0.50	0.51 (0.50-0.51)	0.60	0.13	0.20
AFT-Weibull	≤0.50	0.53 (0.53-0.54)	0.57	0.14	0.20
AFT-log-logistic	≤0.50	0.57 (0.56-0.57)	0.60	0.14	0.20
CISF	≤0.50	0.63 (0.62-0.64)	0.61	0.13	0.23
Deep-Surv	≤0.50	0.58 (0.58-0.59)	0.39	0.15	0.44
Deep-Hit	0.55 (≤0.50-0.64)	≤0.50	0.53	0.14	0.59
Deep-LogHaz	≤0.50	0.61(0.60,0.61)	0.45	0.48	≥1.00

<sup>a</sup>Time-dependent AUCs (ie, median cAUC and iAUC), time-dependent concordance (C) index and Brier score at 1 year, and integrated Brier score were reported. As described in the “Methods” section, Day<sub>15-55</sub> is defined as a bone marrow biopsy and other clinical evaluation done within 15-55 days from the initiation of treatment.

<sup>b</sup>See Table 1 for models and method definitions.

<sup>c</sup>cAUC: cumulative case dynamic control of receiver operative characteristics.

<sup>d</sup>The higher value (ie, close to 1) means better numerical performance.

<sup>e</sup>iAUC: incident case dynamic control of receiver operative characteristics.

<sup>f</sup>The lower value (ie, close to 0) means better numerical performance.

<sup>g</sup>Selected model for 7+3.

**Table 4.** Numerical performances of the chosen machine learning models for ven/aza<sup>a</sup> and 7+3 on independent validation cohorts with respect to overall survival (≤2 years).

Treatment	Machine learning models	Median cAUC <sup>b,c</sup>	Median iAUC <sup>b,d</sup>	$C_t^b$	Brier <sub>t</sub> <sup>e</sup>	iBrier <sup>e</sup>
7+3 (n=14)	RSF <sup>f</sup>	0.71	0.69	0.66	0.19	0.21
ven/aza (n=30)	Cox-Boost <sup>f</sup>	0.68	0.64	0.66	0.29	0.23

<sup>a</sup>ven/aza: venetoclax plus azacitidine.

<sup>b</sup>The higher value (ie, close to 1) means better numerical performance.

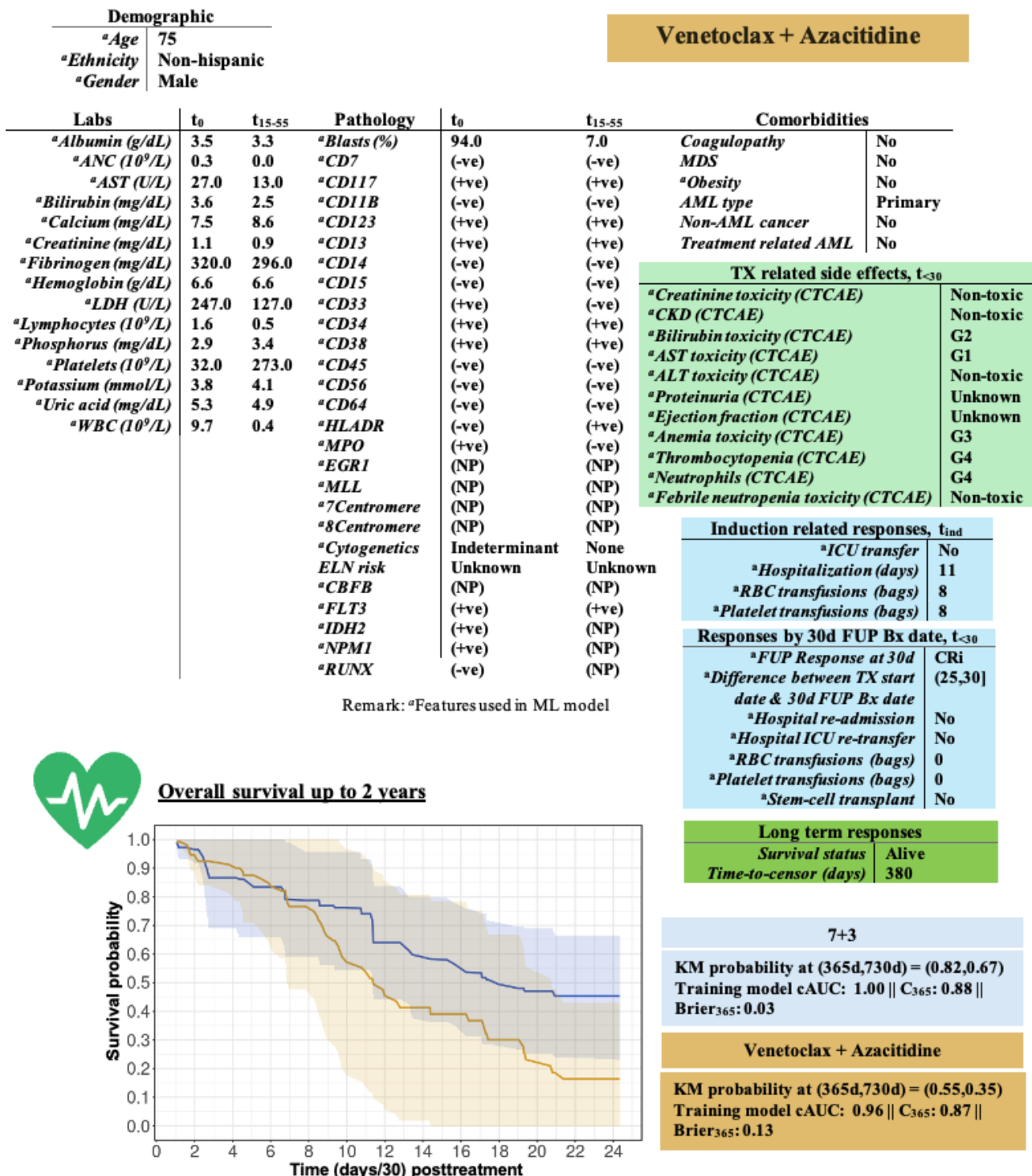
<sup>c</sup>cAUC: cumulative case dynamic control of receiver operative characteristics.

<sup>d</sup>iAUC: incident case dynamic control of receiver operative characteristics.

<sup>e</sup>The lower value (ie, close to 0) means better numerical performance.

<sup>f</sup>See Table 1 for models and method definitions.

**Figure 5.** Representative machine learning (ML)-based predictions for a patient chosen randomly from the validation cohort treated with ven/aza. Actual patient values are in the top boxes and the predicted overall survival probabilities along with 95% confidence bands based on the optimal ML models are depicted at the bottom. ALT: alkaline phosphatase; AML: acute myeloid leukemia; ANC: absolute neutrophil count; AST: aspartate transaminase; CKD: chronic kidney disease; CTCAE: Common Terminology Criteria for Adverse Events; FUP: follow-up; ICU: intensive care unit; LDH: lactate dehydrogenase; MDS: myelodysplastic syndrome; RBC: red blood cell; TX: treatment; ven/aza: venetoclax plus azacitidine; WBC: white blood cell.



### Association Between Events Occurring During the First Year of Therapy and Overall Survival

As treatment events and responses in AML can evolve beyond the first month of treatment, we investigated associations between later disease responses and OS. Initially, we examined the association between the best response after the treatment

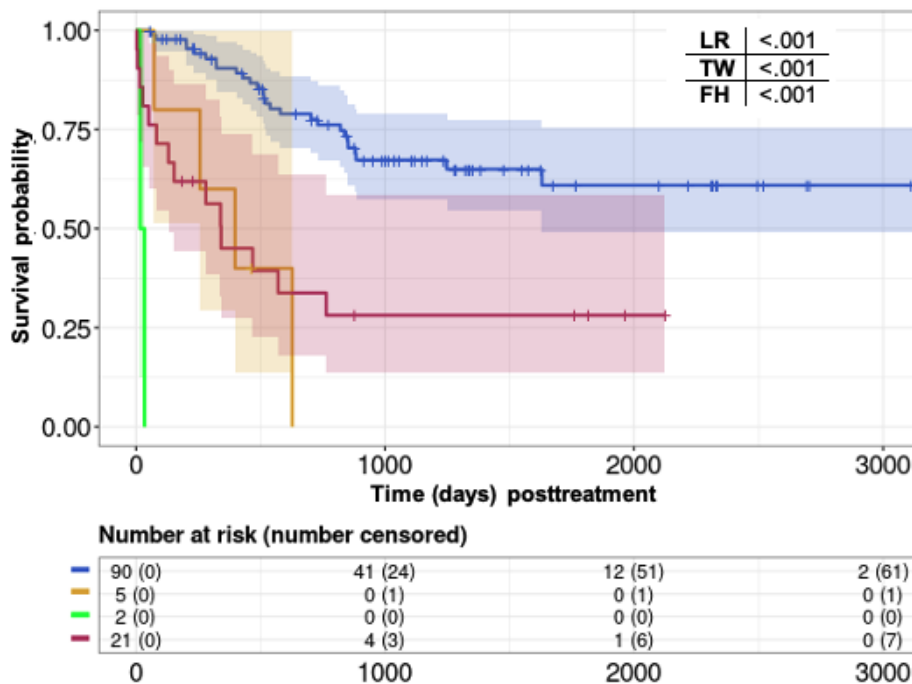
initiation and long-term OS. Among patients treated with 7+3, achieving CR as the best response correlated with a 4-year OS rate of approximately 60%. Conversely, achieving CRi, MLFS, or being nonresponsive (refractory) as the best responses correlated with an OS rate of approximately 25% or less. An LR-based *P* value <.001 indicated significant differences between survival curves (Figure 6A). In the ven/aza-treated

cohort, both CR and CRi were similarly associated with OS, whereas MLFS and nonresponses correlated with lower OS (Figure 6B). It is important to note that Kaplan-Meier survival curves might be influenced by alloSCT. For instance, out of 21

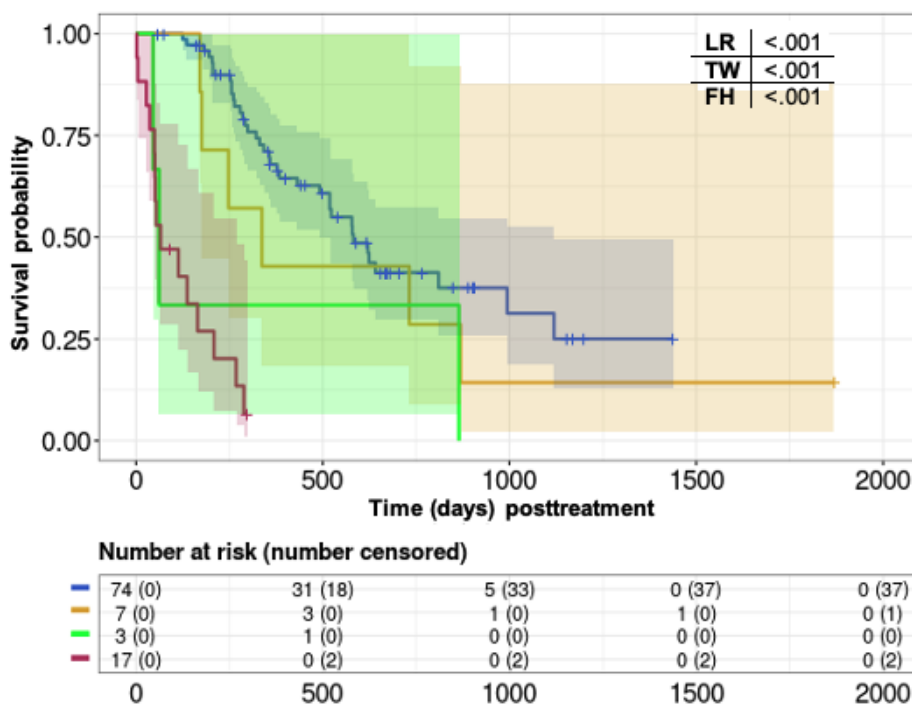
refractory 7+3 patients, 10 (48%) underwent alloSCT, potentially leading to an overestimation of the corresponding survival curve. By contrast, only 3 (18%) out of 17 refractory ven/aza patients underwent alloSCT.

**Figure 6.** Univariate Kaplan-Meier analysis of best response during the first 180 days' assessment and long-term outcomes with (A) 7+3 and (B) ven/aza. *P* values are based on log-rank (LR), Tarone-Ware (TW), and Fleming-Harington (FH) tests. CR: complete remission; CRi: complete remission with incomplete hematologic recovery; MLFS: morphologic leukemia-free state; ven/aza: venetoclax plus azacitidine.

**(A) 7+3**



**(B) ven/aza**



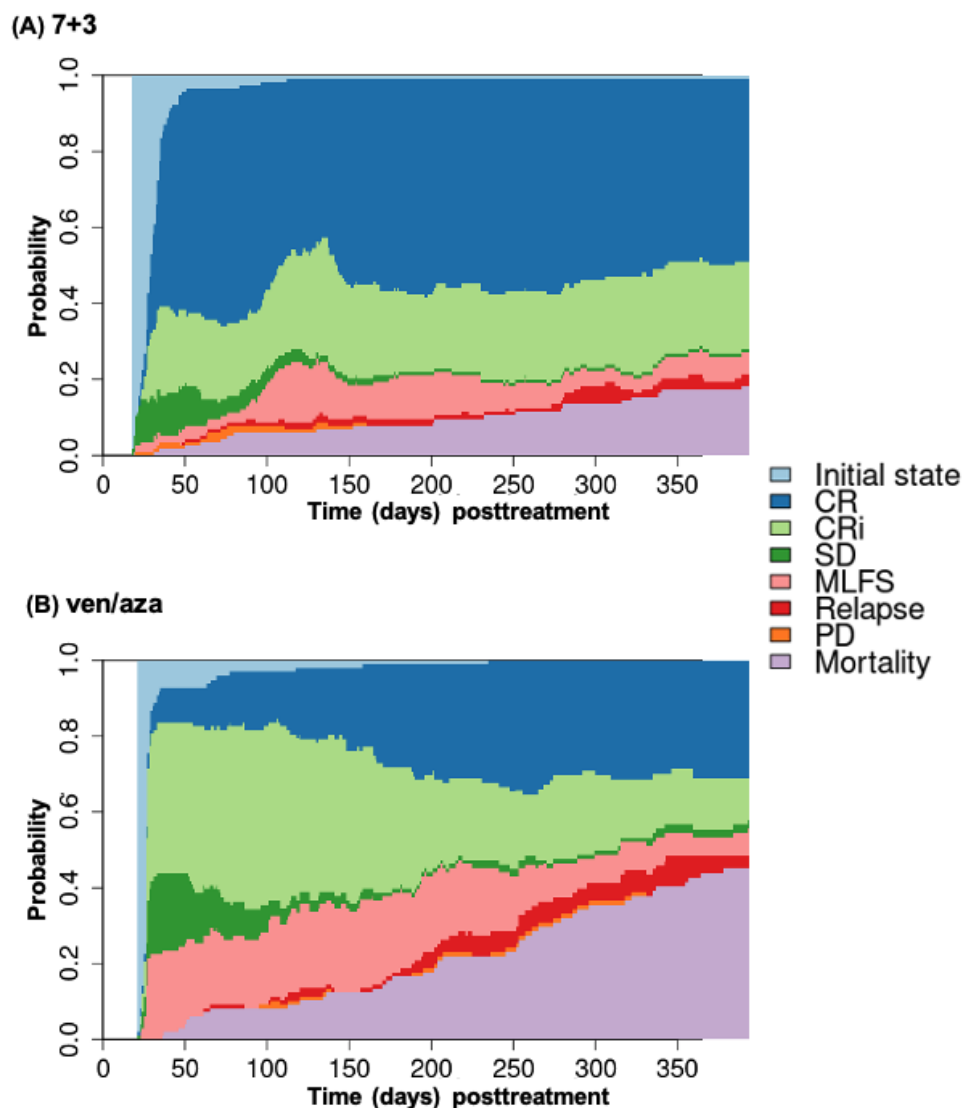
Next, we examined the kinetics of achieving the best response in the 2 treatment groups. The pattern of reaching the best responses differed between the 7+3 and ven/aza groups (Figure

7), as did the overall frequencies of various treatment response outcomes (Figure S8 in Multimedia Appendix 1). At the population level, the 7+3 cohort quickly reached a relatively

stable state by day 30 (Figure 7A). By contrast, the ven/aza cohort showed a continual evolution with conversions from CRi and MLFS to either CR or death (Figure 7B). The disease-state transition probabilities between days 30 and 365 also differed significantly between 7+3 (Figure 8A) and ven/aza (Figure 8B). Achieving CR (0.28; SE 0.05), CRi (0.37; SE 0.06), or MLFS (0.42; SE 0.07) around day 30 after ven/aza treatment showed similar probabilities of transitioning to mortality within a year. By contrast, CR (0.13; SE 0.03) and CRi (0.20; SE 0.05) had comparable transitioning rates to mortality for 7+3. Ven/aza patients with stable disease and progressive disease around day 30 had poorer OS, with 1-year mortality rates of 0.61 (SE 0.08) and 0.75 (SE 0.18), respectively. These observations contrast with that of 7+3, presumably because 7+3 patients were generally fit enough to undergo additional therapies aimed at disease control during this period. Similar observations were noted in disease-state transition probabilities between days

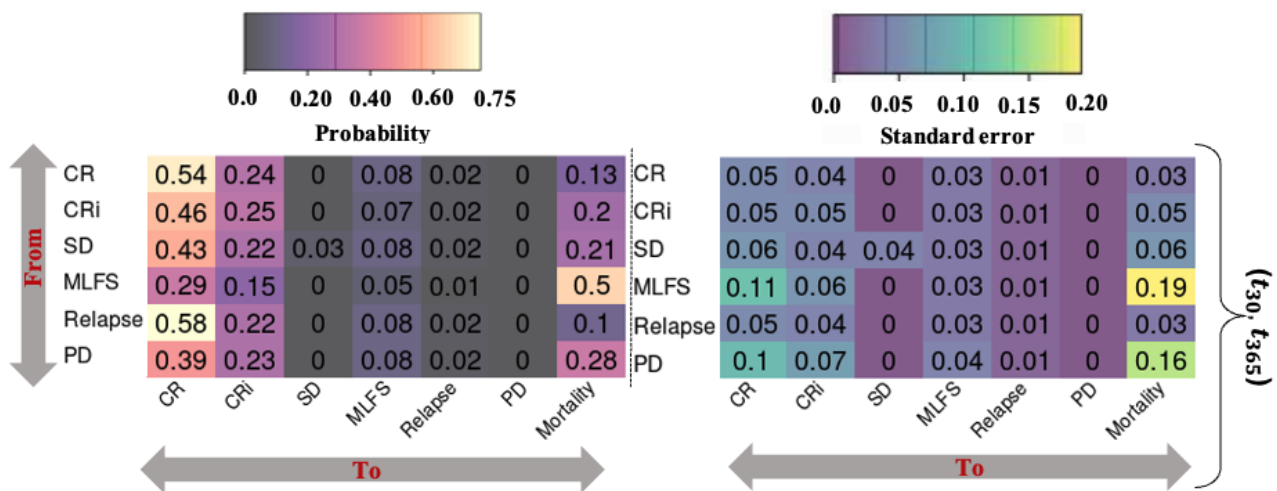
90-365 (Figure S9 in Multimedia Appendix 1) and days 180-365 (Figure S10 in Multimedia Appendix 1). Unlike the 7+3 subgroup, patients treated with ven/aza who achieved any disease state around 180 days had a higher likelihood of transitioning to mortality, with the highest probability observed for the relapse state (0.73; SE 0.16 for ven/aza and 0.54; SE 0.33 for 7+3). Specifically, patients in the MLFS disease state around 180 days transitioned more rapidly to mortality with ven/aza (0.37; SE 0.08) compared with 7+3 (0.14; SE 0.07; refer to Figure S10 in Multimedia Appendix 1). These results further confirm that associations with short- and long-term outcomes differ following ven/aza and 7+3 treatments. The kinetics of responses with ven/aza are notably more dynamic and occur over different time frames compared with those with 7+3. These observations underscore the necessity for distinct response criteria, maintenance strategies, and timing of measurements tailored to each therapy.

**Figure 7.** Overall trends in disease state changes during first year of treatment for (A) 7+3 and (B) ven/aza. CR: complete remission; CRi: complete remission with incomplete hematologic recovery; MLFS: morphologic leukemia-free state; PD: progressive disease; SD: stable disease; ven/aza: venetoclax plus azacitidine.

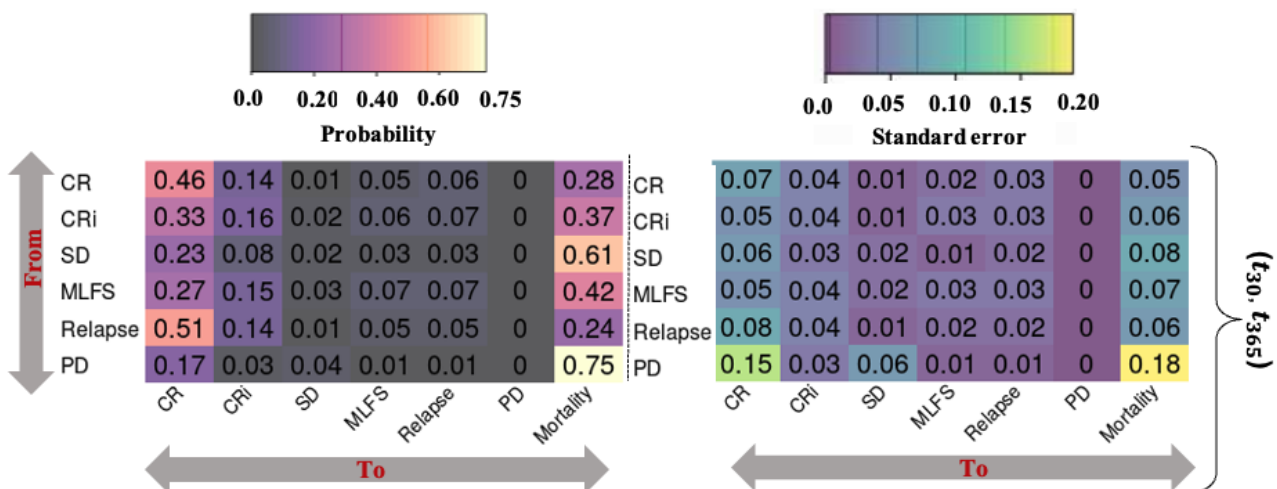


**Figure 8.** Probabilities of transitions from treatment responses (y-axis) achieved by day 30 of treatment to states (x-axis) within 365 days following treatment for (A) 7+3-treated patients and (B) ven/aza-treated patients. The state transition probabilities are on the left panels and SEs are on the right. CR: complete remission; CRi: complete remission with incomplete hematologic recovery; MLFS: morphologic leukemia-free state; PD: progressive disease; SD: stable disease; ven/aza: venetoclax plus azacitidine.

**(A) 7+3**



**(B) ven/aza**



**Discussion**

**Principal Findings**

The main findings of this study indicate that various clinical events occurring during the first month of ven/aza treatment correlate with OS, distinct from outcomes following the 7+3 treatment. Achieving CR/CRi or MLFS around day 30 (ie, Day<sub>15-55</sub>) after ven/aza treatment has a similar long-term prognostic impact, while failure to achieve MLFS around day 30 with 7+3 indicates poorer outcomes. We also identified clinical features such as bone marrow blasts >5%, flow cytometric and genetic detection of AML, and AML-related cytogenetic factors at reassessment as having negative prognostic impacts on OS. Based on these observations, detection of persistent leukemia in the bone marrow around day 30 following ven/aza treatment suggests consideration of alternative therapies. By contrast, achieving CR/CRi/MLFS around day 30 with minimal evidence of persistent leukemia

following ven/aza is associated with improved OS, indicating the benefit of continuing this treatment. However, we also found that failure to achieve CR/CRi by approximately day 180 after ven/aza initiation has negative implications for OS. This suggests that alternative therapies should be considered if the milestone of achieving CR/CRi by this time point is not met.

We also found that certain hospital events and toxicities occurring after the initiation of ven/aza treatment have prognostic implications, which differ from those seen with the 7+3 treatment. For instance, ICU admission during the initial ven/aza treatment was not associated with worse outcomes, whereas ICU transfer during the initial hospitalization for 7+3 was a poor prognostic factor. Additionally, grade  $\geq 4$  thrombocytopenia and grade  $\geq 3$  anemia were more pronounced as poor prognostic indicators for ven/aza compared with 7+3. The incidence of renal impairment was similar for both treatments; however, elevated creatinine, proteinuria, and CKD were associated with worse OS among patients treated with

7+3, but less so among those treated with ven/aza. The progression of grade  $\geq 1$  ejection fraction toxicity over time was comparable between both treatment arms. While there is a well-known association between anthracyclines in the 7+3 regimen and cardiac toxicity, such an association has not been previously described for ven/aza. These associative findings, albeit based on small sample sizes, may warrant further investigation.

### Prospective Predictive Models for AML Clinical Decision Support

To translate the statistical associations between events and outcomes following the initiation of AML therapy into patient-specific prognostic models, we developed predictive ML models independently for OS with ven/aza and 7+3 treatments. These models utilized baseline and early disease responses, biomarkers, and clinical events. Several models achieved relatively high AUCs of 0.80 and 0.85 in the internal validation step. However, in the independent validation phase, AUCs were lower at 0.71 and 0.68 for the 7+3 and ven/aza cohorts, respectively. This predictive discrepancy may stem from data drift, yet we tested the models to evaluate their performance on nonhomogeneous data. Although not ideal, we contend that such drifts are typical in real-world data sets. Nevertheless, these findings illustrate the feasibility of developing ML-based individual predictors using patient data that evolve. This capability allows clinical decision-making to adapt to individual changes in treatment side effects and responses. This effort contributes to an expanding body of research utilizing ML to predict outcomes in the treatment of AML and other hematologic malignancies [55-58]. For instance, Park et al [55] evaluated the prognostic performance of ELN genetic risk stratification models using unsupervised ML techniques and found suboptimal predictions for OS in older patients with AML, indicating a need for new risk models in this demographic. Karami et al [56] identified novel prognostic factors for survival in patients with AML, incorporating demographic and AML-specific features through ML approaches. Shaikh et al [57] developed a novel risk stratification model for patients with AML and RUNX1-RUNX1T1 using supervised machine learning models. This model assesses risk based on somatic mutations in *Flt3*, *NRAS*, and other genes. Lastly, Eckardt et al [58] conducted a review of various ML approaches for AML diagnosis, prognosis, and risk stratification, emphasizing their evolving and potentially impactful role in this specific disease area. To our knowledge, no prior predictive ML-based survival models with uncertainty quantification have been developed in the AML literature by exploiting both patient- and event-specific long-term dynamic features at this level of detail.

### Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. All statements in this report, including its findings and conclusions, are solely those of the authors and do not necessarily represent the views of the RefinedScience, its Board of directors, or the University of Colorado.

### Limitations

Our study is limited by relatively small data sets, and our results require validation on larger data sets from diverse centers ensuring heterogeneity. Although we adjusted for high collinearity among variables, missingness, and overfitting, these issues need further careful consideration in larger external data sets. Additionally, our single-center data set consists of half real-world and half clinical trial data, which may potentially bias the results compared with more diverse population-based data sets. Lastly, any comparison between the 2 primary treatments, ven/aza and 7+3, is influenced by differences in patient demographics, comorbidities, and other inherent features. Ven/aza is currently approved only for older and unfit patients, whereas 7+3 is primarily used in younger and fitter patients. Additionally, 24 out of 101 ven/aza patients (23.8%) underwent at least one alloSCT after the initiation of treatment, whereas 79 out of 120 7+3 patients (65.8%) underwent at least one alloSCT. Survival curves in AML are affected by alloSCT, which can significantly impact OS and necessitate adjustments and modifications in ML modeling, a direction we plan to explore in future studies. Because of these complexities, the extent and direction of association with OS for confounding factors vary across the treatments we evaluated. However, our primary objective was not to establish causal treatment effectiveness by treating ven/aza as the treatment group and 7+3 as the control for which a classical propensity score-based or weighted method is recommended to ensure balances in data distributions between the 2 treatment arms. Instead, our primary objective was to separately explore the variations in data to assess whether treatment modifies the effects (ie, directions) of confounders on OS. Therefore, we treated 7+3 and ven/aza as effect modifiers and conducted “stratified” multivariable analyses, creating subgroups of patients treated with 7+3 and ven/aza [59]. This stratification approach minimizes the variation attributed to treatment differences significantly, and by accounting for the same set of potential confounders in both models, it enhances the comparability of results.

### Conclusions

Despite these considerations, our results have highlighted significant clinical implications of posttreatment outcomes, clinical events, and toxicities on long-term outcomes and treatment decisions in AML, demonstrating differences between ven/aza and 7+3. Additionally, these observations suggest strong potential to develop ML-based predictive models which could ultimately offer crucial ongoing clinical decision support for patients and providers as toxicities, responses, and other events evolve dynamically throughout treatment cycles.



## Data Availability

The raw, individual patient data are protected and not available due to data privacy laws. The processed data are available at reasonable request to the corresponding author.

## Authors' Contributions

CS and NI designed the study and drafted the manuscript. JSR, JD, JWC, KS, and FRM processed and pulled the structured analytical data sets. JSR, JD, CS, and NI assessed the validity and quality of data. NI performed numerical analyses. CS, NI, CTJ, and JSR interpreted the results of the analyses. All authors reviewed, provided constructive comments, and agreed to its publication.

## Conflicts of Interest

CS is an employee of and holds equity in OncoVerity, and is a consultant to RefinedScience. All other authors declare no conflicts of interest.

## Multimedia Appendix 1

Additional pertinent results.

[\[DOCX File , 7134 KB-Multimedia Appendix 1\]](#)

## References

1. Premnath N, Madanat YF. Paradigm shift in the management of acute myeloid leukemia-approved options in 2023. *Cancers (Basel)*. May 31, 2023;15(11):3002. [FREE Full text] [doi: [10.3390/cancers15113002](https://doi.org/10.3390/cancers15113002)] [Medline: [37296964](https://pubmed.ncbi.nlm.nih.gov/37296964/)]
2. Zeidan AM, Pollyea DA, Borate U, Vasconcelos A, Potluri R, Rotter D, et al. Venetoclax plus azacitidine compared with intensive chemotherapy as induction for patients with acute myeloid leukemia: retrospective analysis of an electronic medical record database in the United States. *Ann Hematol*. Apr 2023;102(4):749-754. [FREE Full text] [doi: [10.1007/s00277-023-05109-5](https://doi.org/10.1007/s00277-023-05109-5)] [Medline: [36732419](https://pubmed.ncbi.nlm.nih.gov/36732419/)]
3. Bittencourt MCB, Ciurea SO. Recent advances in allogeneic hematopoietic stem cell transplantation for acute myeloid leukemia. *Biol Blood Marrow Transplant*. Sep 2020;26(9):e215-e221. [FREE Full text] [doi: [10.1016/j.bbmt.2020.06.007](https://doi.org/10.1016/j.bbmt.2020.06.007)] [Medline: [32561339](https://pubmed.ncbi.nlm.nih.gov/32561339/)]
4. Blum WG, Mims AS. Treating acute myeloid leukemia in the modern era: a primer. *Cancer*. Nov 01, 2020;126(21):4668-4677. [FREE Full text] [doi: [10.1002/cncr.32904](https://doi.org/10.1002/cncr.32904)] [Medline: [32767757](https://pubmed.ncbi.nlm.nih.gov/32767757/)]
5. Chua CC, Roberts AW, Reynolds J, Fong CY, Ting SB, Salmon JM, et al. Chemotherapy and Venetoclax in Elderly Acute Myeloid Leukemia Trial (CAVEAT): a phase Ib dose-escalation study of venetoclax combined with modified intensive chemotherapy. *J Clin Oncol*. Oct 20, 2020;38(30):3506-3517. [doi: [10.1200/JCO.20.00572](https://doi.org/10.1200/JCO.20.00572)] [Medline: [32687450](https://pubmed.ncbi.nlm.nih.gov/32687450/)]
6. Pollyea DA, Amaya M, Strati P, Konopleva MY. Venetoclax for AML: changing the treatment paradigm. *Blood Adv*. 2019;3(24):4326-4335. *Blood Adv*. Mar 24, 2020;4(6):1020. [FREE Full text] [doi: [10.1182/bloodadvances.2020001769](https://doi.org/10.1182/bloodadvances.2020001769)] [Medline: [32187375](https://pubmed.ncbi.nlm.nih.gov/32187375/)]
7. DiNardo CD, Jonas BA, Pullarkat V, Thirman MJ, Garcia JS, Wei AH, et al. Azacitidine and venetoclax in previously untreated acute myeloid leukemia. *N Engl J Med*. Aug 13, 2020;383(7):617-629. [doi: [10.1056/NEJMoa2012971](https://doi.org/10.1056/NEJMoa2012971)] [Medline: [32786187](https://pubmed.ncbi.nlm.nih.gov/32786187/)]
8. Zimmer M, Kadia T. Approach to the older patient with acute myeloid leukemia. *Curr Oncol Rep*. Nov 2023;25(11):1203-1211. [doi: [10.1007/s11912-023-01450-0](https://doi.org/10.1007/s11912-023-01450-0)] [Medline: [37688738](https://pubmed.ncbi.nlm.nih.gov/37688738/)]
9. DiNardo CD, Pratz K, Pullarkat V, Jonas BA, Arellano M, Becker PS, et al. Venetoclax combined with decitabine or azacitidine in treatment-naïve, elderly patients with acute myeloid leukemia. *Blood*. Jan 03, 2019;133(1):7-17. [FREE Full text] [doi: [10.1182/blood-2018-08-868752](https://doi.org/10.1182/blood-2018-08-868752)] [Medline: [30361262](https://pubmed.ncbi.nlm.nih.gov/30361262/)]
10. Islam N, Reuben JS, Dale J, Gutman J, McMahon CM, Amaya M, et al. Machine learning-based exploratory clinical decision support for newly diagnosed patients with acute myeloid leukemia treated with 7 + 3 type chemotherapy or venetoclax/azacitidine. *JCO Clin Cancer Inform*. Sep 2022;6:e2200030. [FREE Full text] [doi: [10.1200/CCI.22.00030](https://doi.org/10.1200/CCI.22.00030)] [Medline: [36194842](https://pubmed.ncbi.nlm.nih.gov/36194842/)]
11. Pollyea DA. Venetoclax in AML: where we are and where we are headed. *Clin Lymphoma Myeloma Leuk*. Sep 2020;20 Suppl 1:S25-S26. [doi: [10.1016/S2152-2650\(20\)30450-X](https://doi.org/10.1016/S2152-2650(20)30450-X)] [Medline: [32862856](https://pubmed.ncbi.nlm.nih.gov/32862856/)]
12. Aitken MJL, Ravandi F, Patel KP, Short NJ. Prognostic and therapeutic implications of measurable residual disease in acute myeloid leukemia. *J Hematol Oncol*. Sep 03, 2021;14(1):137. [FREE Full text] [doi: [10.1186/s13045-021-01148-5](https://doi.org/10.1186/s13045-021-01148-5)] [Medline: [34479626](https://pubmed.ncbi.nlm.nih.gov/34479626/)]
13. Baron F, Labopin M, Ruggeri A, Sierra J, Robinson S, Labussière-Wallet H, et al. Impact of detectable measurable residual disease on umbilical cord blood transplantation. *Am J Hematol*. Sep 2020;95(9):1057-1065. [FREE Full text] [doi: [10.1002/ajh.25879](https://doi.org/10.1002/ajh.25879)] [Medline: [32449964](https://pubmed.ncbi.nlm.nih.gov/32449964/)]
14. Bataller A, Oñate G, Diaz-Beyá M, Guijarro F, Garrido A, Vives S, et al. Grupo Cooperativo Para el Estudio y Tratamiento de las Leucemias Agudas y Mielodisplasias (CETLAM). Acute myeloid leukemia with NPM1 mutation and favorable

- European LeukemiaNet category: outcome after preemptive intervention based on measurable residual disease. *Br J Haematol.* Oct 2020;191(1):52-61. [FREE Full text] [doi: [10.1111/bjh.16857](https://doi.org/10.1111/bjh.16857)] [Medline: [32510599](https://pubmed.ncbi.nlm.nih.gov/32510599/)]
15. Bernasconi P, Borsani O. Eradication of measurable residual disease in AML: a challenging clinical goal. *Cancers (Basel).* Jun 25, 2021;13(13):13. [FREE Full text] [doi: [10.3390/cancers13133170](https://doi.org/10.3390/cancers13133170)] [Medline: [34202000](https://pubmed.ncbi.nlm.nih.gov/34202000/)]
  16. Buccisano F, Palmieri R, Piciocchi A, Maurillo L, Del Principe MI, Paterno G, et al. Use of Measurable Residual Disease to Evolve Transplant Policy in Acute Myeloid Leukemia: A 20-Year Monocentric Observation. *Cancers (Basel).* Mar 03, 2021;13(5):1083. [FREE Full text] [doi: [10.3390/cancers13051083](https://doi.org/10.3390/cancers13051083)] [Medline: [33802502](https://pubmed.ncbi.nlm.nih.gov/33802502/)]
  17. Gilleece MH, Shimoni A, Labopin M, Robinson S, Beelen D, Socié G, et al. Measurable residual disease status and outcome of transplant in acute myeloid leukemia in second complete remission: a study by the acute leukemia working party of the EBMT. *Blood Cancer J.* May 12, 2021;11(5):88. [FREE Full text] [doi: [10.1038/s41408-021-00479-3](https://doi.org/10.1038/s41408-021-00479-3)] [Medline: [33980810](https://pubmed.ncbi.nlm.nih.gov/33980810/)]
  18. Guolo F, Di Grazia C, Minetto P, Raiola AM, Clavio M, Miglino M, et al. Pre-transplant minimal residual disease assessment and transplant-related factors predict the outcome of acute myeloid leukemia patients undergoing allogeneic stem cell transplantation. *Eur J Haematol.* Nov 2021;107(5):573-582. [doi: [10.1111/ejh.13694](https://doi.org/10.1111/ejh.13694)] [Medline: [34297437](https://pubmed.ncbi.nlm.nih.gov/34297437/)]
  19. Kim H, Kim Y, Kang D, Kim HS, Lee J, Kim M, et al. Prognostic value of measurable residual disease monitoring by next-generation sequencing before and after allogeneic hematopoietic cell transplantation in acute myeloid leukemia. *Blood Cancer J.* Jun 04, 2021;11(6):109. [FREE Full text] [doi: [10.1038/s41408-021-00500-9](https://doi.org/10.1038/s41408-021-00500-9)] [Medline: [34088902](https://pubmed.ncbi.nlm.nih.gov/34088902/)]
  20. Konuma T, Kondo T, Masuko M, Shimizu H, Shiratori S, Fukuda T, et al. Adult Acute Myeloid Leukemia Working Group of the Japan Society for Hematopoietic Cell Transplantation. Prognostic value of measurable residual disease at allogeneic transplantation for adults with core binding factor acute myeloid leukemia in complete remission. *Bone Marrow Transplant.* Nov 2021;56(11):2779-2787. [doi: [10.1038/s41409-021-01409-4](https://doi.org/10.1038/s41409-021-01409-4)] [Medline: [34272486](https://pubmed.ncbi.nlm.nih.gov/34272486/)]
  21. Kövy P, Órfi Z, Bors A, Kozma A, Gopcsa L, Dolgos J, et al. Nucleophosmin1 and isocitrate dehydrogenase 1 and 2 as measurable residual disease markers in acute myeloid leukemia. *PLoS One.* 2021;16(6):e0253386. [FREE Full text] [doi: [10.1371/journal.pone.0253386](https://doi.org/10.1371/journal.pone.0253386)] [Medline: [34153064](https://pubmed.ncbi.nlm.nih.gov/34153064/)]
  22. Percival M, Wang H, Zhang M, Saber W, de Lima M, Litzow M, et al. Impact of depth of clinical response on outcomes of acute myeloid leukemia patients in first complete remission who undergo allogeneic hematopoietic cell transplantation. *Bone Marrow Transplant.* Sep 2021;56(9):2108-2117. [doi: [10.1038/s41409-021-01261-6](https://doi.org/10.1038/s41409-021-01261-6)] [Medline: [33864019](https://pubmed.ncbi.nlm.nih.gov/33864019/)]
  23. Rautenberg C, Lauseker M, Kaivers J, Jäger P, Fischermanns C, Pechtel S, et al. Prognostic impact of pretransplant measurable residual disease assessed by peripheral blood WT1-mRNA expression in patients with AML and MDS. *Eur J Haematol.* Aug 2021;107(2):283-292. [doi: [10.1111/ejh.13664](https://doi.org/10.1111/ejh.13664)] [Medline: [33987857](https://pubmed.ncbi.nlm.nih.gov/33987857/)]
  24. Rossi G, Giambra V, de Waure C, Giacchetta I, Minervini MM, Abbenante MC, et al. Log reduction of leukemic cells and minimal residual disease by flow cytometry represent effective predictors of clinical outcome in elderly patients with acute myeloid leukemia. *Cytometry B Clin Cytom.* Jan 2022;102(1):26-33. [FREE Full text] [doi: [10.1002/cyto.b.22010](https://doi.org/10.1002/cyto.b.22010)] [Medline: [33983682](https://pubmed.ncbi.nlm.nih.gov/33983682/)]
  25. Shang L, Cai X, Sun W, Cheng Q, Mi Y. Time point-dependent concordance and prognostic significance of flow cytometry and real time quantitative PCR for measurable/minimal residual disease detection in acute myeloid leukemia with t(8;21)(q22;q22.1). *Cytometry B Clin Cytom.* Jan 2022;102(1):34-43. [doi: [10.1002/cyto.b.22028](https://doi.org/10.1002/cyto.b.22028)] [Medline: [34232569](https://pubmed.ncbi.nlm.nih.gov/34232569/)]
  26. Tiong IS, Dillon R, Ivey A, Teh T, Nguyen P, Cummings N, et al. Venetoclax induces rapid elimination of NPM1 mutant measurable residual disease in combination with low-intensity chemotherapy in acute myeloid leukaemia. *Br J Haematol.* Mar 2021;192(6):1026-1030. [FREE Full text] [doi: [10.1111/bjh.16722](https://doi.org/10.1111/bjh.16722)] [Medline: [32458446](https://pubmed.ncbi.nlm.nih.gov/32458446/)]
  27. Yu S, Fan Z, Ma L, Wang Y, Huang F, Zhang Q, et al. Association between measurable residual disease in patients with intermediate-risk acute myeloid leukemia and first remission, treatment, and outcomes. *JAMA Netw Open.* Jul 01, 2021;4(7):e2115991. [FREE Full text] [doi: [10.1001/jamanetworkopen.2021.15991](https://doi.org/10.1001/jamanetworkopen.2021.15991)] [Medline: [34232303](https://pubmed.ncbi.nlm.nih.gov/34232303/)]
  28. Zheng W, Hu Y, Guan L, Peng B, Wang S. The effect of the detection of minimal residual disease for the prognosis and the choice of post-remission therapy of intermediate-risk acute myeloid leukemia without FLT3-ITD, NPM1 and biallelic CEBPA mutations. *Hematology.* Dec 2021;26(1):179-185. [doi: [10.1080/16078454.2021.1880753](https://doi.org/10.1080/16078454.2021.1880753)] [Medline: [33594943](https://pubmed.ncbi.nlm.nih.gov/33594943/)]
  29. Byrne M, Danielson N, Sengsayadeth S, Rasche A, Culos K, Gatwood K, et al. The use of venetoclax-based salvage therapy for post-hematopoietic cell transplantation relapse of acute myeloid leukemia. *Am J Hematol.* Sep 2020;95(9):1006-1014. [FREE Full text] [doi: [10.1002/ajh.25859](https://doi.org/10.1002/ajh.25859)] [Medline: [32390196](https://pubmed.ncbi.nlm.nih.gov/32390196/)]
  30. Daneshbod Y, Kohan L, Taghadosi V, Weinberg OK, Arber DA. Prognostic significance of complex karyotypes in acute myeloid leukemia. *Curr Treat Options Oncol.* Feb 11, 2019;20(2):15. [doi: [10.1007/s11864-019-0612-y](https://doi.org/10.1007/s11864-019-0612-y)] [Medline: [30741367](https://pubmed.ncbi.nlm.nih.gov/30741367/)]
  31. Kongtim P, Hasan O, Perez JMR, Varma A, Wang SA, Patel KP, et al. Novel disease risk model for patients with acute myeloid leukemia receiving allogeneic hematopoietic cell transplantation. *Biol Blood Marrow Transplant.* Jan 2020;26(1):197-203. [FREE Full text] [doi: [10.1016/j.bbmt.2019.09.006](https://doi.org/10.1016/j.bbmt.2019.09.006)] [Medline: [31518645](https://pubmed.ncbi.nlm.nih.gov/31518645/)]
  32. Kongtim P, Parmar S, Milton DR, Perez JMR, Rondon G, Chen J, et al. Impact of a novel prognostic model, hematopoietic cell transplant-composite risk (HCT-CR), on allogeneic transplant outcomes in patients with acute myeloid leukemia and myelodysplastic syndrome. *Bone Marrow Transplant.* Jun 2019;54(6):839-848. [FREE Full text] [doi: [10.1038/s41409-018-0344-9](https://doi.org/10.1038/s41409-018-0344-9)] [Medline: [30258129](https://pubmed.ncbi.nlm.nih.gov/30258129/)]

33. Moors I, Vandepoele K, Philippé J, Deeren D, Selleslag D, Breems D, et al. Clinical implications of measurable residual disease in AML: review of current evidence. *Crit Rev Oncol Hematol*. Jan 2019;133:142-148. [doi: [10.1016/j.critrevonc.2018.11.010](https://doi.org/10.1016/j.critrevonc.2018.11.010)] [Medline: [30661650](https://pubmed.ncbi.nlm.nih.gov/30661650/)]
34. Maiti A, DiNardo CD, Wang SA, Jorgensen J, Kadia TM, Daver NG, et al. Prognostic value of measurable residual disease after venetoclax and decitabine in acute myeloid leukemia. *Blood Adv*. Apr 13, 2021;5(7):1876-1883. [FREE Full text] [doi: [10.1182/bloodadvances.2020003717](https://doi.org/10.1182/bloodadvances.2020003717)] [Medline: [33792630](https://pubmed.ncbi.nlm.nih.gov/33792630/)]
35. Döhner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Büchner T, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*. Jan 26, 2017;129(4):424-447. [FREE Full text] [doi: [10.1182/blood-2016-08-733196](https://doi.org/10.1182/blood-2016-08-733196)] [Medline: [27895058](https://pubmed.ncbi.nlm.nih.gov/27895058/)]
36. Division of cancer treatment and diagnosis. National Cancer Institute Division of Cancer Treatment and Diagnosis. URL: [https://ctep.cancer.gov/protocoldevelopment/electronic\\_applications/ctc.htm](https://ctep.cancer.gov/protocoldevelopment/electronic_applications/ctc.htm) [accessed 2023-10-18]
37. Benicio MTDL, Ribeiro AFT, Américo AD, Furtado FM, Glória AB, Lima AS, et al. Evaluation of the European LeukemiaNet recommendations for predicting outcomes of patients with acute myeloid leukemia treated in low- and middle-income countries (LMIC): a Brazilian experience. *Leuk Res*. Sep 2017;60:109-114. [doi: [10.1016/j.leukres.2017.07.005](https://doi.org/10.1016/j.leukres.2017.07.005)] [Medline: [28777950](https://pubmed.ncbi.nlm.nih.gov/28777950/)]
38. Datta S, Satten GA. Estimating future stage entry and occupation probabilities in a multistage model based on randomly right-censored data. *Statistics & Probability Letters*. Oct 2000;50(1):89-95. [doi: [10.1016/s0167-7152\(00\)00086-9](https://doi.org/10.1016/s0167-7152(00)00086-9)]
39. Datta S, Satten GA. Validity of the Aalen–Johansen estimators of stage occupation probabilities and Nelson–Aalen estimators of integrated transition hazards for non-Markov models. *Statistics & Probability Letters*. Dec 15, 2001;55(4):403-411. [doi: [10.1016/S0167-7152\(01\)00155-9](https://doi.org/10.1016/S0167-7152(01)00155-9)]
40. Ferguson N, Brock G, Datta S. msSurv: an R Package for nonparametric estimation of multistate models. *J Stat Softw*. 2012;50(14):1-24. [doi: [10.18637/jss.v050.i14](https://doi.org/10.18637/jss.v050.i14)]
41. Andersen P, Borgan O, Gill R, Keiding N. *Statistical Models Based on Counting Processes*. New York City, NY: Springer; 1995.
42. Simon N, Friedman J, Hastie T, Tibshirani R. Regularization paths for Cox's proportional hazards model via coordinate descent. *J Stat Softw*. Mar 2011;39(5):1-13. [doi: [10.18637/jss.v039.i05](https://doi.org/10.18637/jss.v039.i05)] [Medline: [27065756](https://pubmed.ncbi.nlm.nih.gov/27065756/)]
43. Xu L, Gotwalt C, Hong Y, King CB, Meeker WQ. Applications of the fractional-random-weight bootstrap. *The American Statistician*. Apr 17, 2020;74(4):345-358. [doi: [10.1080/00031305.2020.1731599](https://doi.org/10.1080/00031305.2020.1731599)]
44. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. *Ann Appl Stat*. Sep 1, 2008;2(3):841-860. [doi: [10.1214/08-aos169](https://doi.org/10.1214/08-aos169)]
45. Hothorn T, Hornik K, Zeileis A. Unbiased recursive partitioning: a conditional inference framework. *Journal of Computational and Graphical Statistics*. Sep 2006;15(3):651-674. [doi: [10.1198/106186006X133933](https://doi.org/10.1198/106186006X133933)]
46. Sayehmiri K, Eshraghian MR, Mohammad K, Alimoghaddam K, Foroushani AR, Zeraati H, et al. Prognostic factors of survival time after hematopoietic stem cell transplant in acute lymphoblastic leukemia patients: Cox proportional hazard versus accelerated failure time models. *J Exp Clin Cancer Res*. Nov 23, 2008;27(1):74. [FREE Full text] [doi: [10.1186/1756-9966-27-74](https://doi.org/10.1186/1756-9966-27-74)] [Medline: [19025609](https://pubmed.ncbi.nlm.nih.gov/19025609/)]
47. Benner A, Zucknick M, Hielscher T, Itrich C, Mansmann U. High-dimensional Cox models: the choice of penalty as part of the model building process. *Biom J*. Feb 2010;52(1):50-69. [doi: [10.1002/bimj.200900064](https://doi.org/10.1002/bimj.200900064)] [Medline: [20166132](https://pubmed.ncbi.nlm.nih.gov/20166132/)]
48. Goeman JJ. L1 penalized estimation in the Cox proportional hazards model. *Biom J*. Feb 2010;52(1):70-84. [doi: [10.1002/bimj.200900028](https://doi.org/10.1002/bimj.200900028)] [Medline: [19937997](https://pubmed.ncbi.nlm.nih.gov/19937997/)]
49. Song J, Chen Z, Huang D, Wu Y, Lin Z, Chi P, et al. Nomogram predicting overall survival of resected locally advanced rectal cancer patients with neoadjuvant chemoradiotherapy. *Cancer Manag Res*. 2020;12:7375-7382. [FREE Full text] [doi: [10.2147/CMAR.S255981](https://doi.org/10.2147/CMAR.S255981)] [Medline: [32884350](https://pubmed.ncbi.nlm.nih.gov/32884350/)]
50. Xiao N, Qing-Song X, Miao-Zhu L, Harrell F, Potapov S, Adler W. Benchmarking and visualization toolkit for penalized Cox models. GitHub. Apr 24, 2023. URL: <https://github.com/nanxstats/hdnom/blob/HEAD/R/hdnom.R> [accessed 2024-07-08]
51. Katzman JL, Shaham U, Cloninger A, Bates J, Jiang T, Kluger Y. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Med Res Methodol*. Feb 26, 2018;18(1):24. [FREE Full text] [doi: [10.1186/s12874-018-0482-1](https://doi.org/10.1186/s12874-018-0482-1)] [Medline: [29482517](https://pubmed.ncbi.nlm.nih.gov/29482517/)]
52. Kvamme H, Borgan. Continuous and discrete-time survival prediction with neural networks. *Lifetime Data Anal*. Oct 2021;27(4):710-736. [FREE Full text] [doi: [10.1007/s10985-021-09532-6](https://doi.org/10.1007/s10985-021-09532-6)] [Medline: [34618267](https://pubmed.ncbi.nlm.nih.gov/34618267/)]
53. Lee C, Yoon J, Schaar MVD. Dynamic-DeepHit: a deep learning approach for dynamic survival analysis with competing risks based on longitudinal data. *IEEE Trans Biomed Eng*. Jan 2020;67(1):122-133. [doi: [10.1109/TBME.2019.2909027](https://doi.org/10.1109/TBME.2019.2909027)] [Medline: [30951460](https://pubmed.ncbi.nlm.nih.gov/30951460/)]
54. Abbott D, Cherry E, Amaya M, McMahan C, Schwartz M, Winters A, et al. The propriety of upgrading responses to venetoclax + azacitidine in newly diagnosed patients with acute myeloid leukemia. *Leuk Lymphoma*. Jun 2021;62(6):1466-1473. [doi: [10.1080/10428194.2020.1864358](https://doi.org/10.1080/10428194.2020.1864358)] [Medline: [33375853](https://pubmed.ncbi.nlm.nih.gov/33375853/)]
55. Park S, Kim TY, Cho B, Kwag D, Lee J, Kim M, et al. Prognostic value of European LeukemiaNet 2022 criteria and genomic clusters using machine learning in older adults with acute myeloid leukemia. *Haematologica*. Apr 01, 2024;109(4):1095-1106. [doi: [10.3324/haematol.2023.283606](https://doi.org/10.3324/haematol.2023.283606)] [Medline: [37706344](https://pubmed.ncbi.nlm.nih.gov/37706344/)]

56. Karami K, Akbari M, Moradi M, Soleymani B, Fallahi H. Survival prognostic factors in patients with acute myeloid leukemia using machine learning techniques. *PLoS One*. 2021;16(7):e0254976. [[FREE Full text](#)] [doi: [10.1371/journal.pone.0254976](https://doi.org/10.1371/journal.pone.0254976)] [Medline: [34288963](#)]
57. Shaikh AF, Kakirde C, Dhamne C, Bhanshe P, Joshi S, Chaudhary S, et al. Machine learning derived genomics driven prognostication for acute myeloid leukemia with RUNX1-RUNX1T1. *Leuk Lymphoma*. Dec 2020;61(13):3154-3160. [[FREE Full text](#)] [doi: [10.1080/10428194.2020.1798951](https://doi.org/10.1080/10428194.2020.1798951)] [Medline: [32757686](#)]
58. Eckardt J, Bornhäuser M, Wendt K, Middeke JM. Application of machine learning in the management of acute myeloid leukemia: current practice and future prospects. *Blood Adv*. Dec 08, 2020;4(23):6077-6085. [[FREE Full text](#)] [doi: [10.1182/bloodadvances.2020002997](https://doi.org/10.1182/bloodadvances.2020002997)] [Medline: [33290546](#)]
59. Corraini P, Olsen M, Pedersen L, Dekkers OM, Vandenbroucke JP. Effect modification, interaction and mediation: an overview of theoretical insights for clinical investigators. *Clin Epidemiol*. 2017;9:331-338. [[FREE Full text](#)] [doi: [10.2147/CLEP.S129728](https://doi.org/10.2147/CLEP.S129728)] [Medline: [28652815](#)]

---

## Abbreviations

**AFT:** accelerated failure time  
**aHR:** adjusted hazard ratio  
**alloSCT:** allogeneic stem cell transplantation  
**AML:** acute myeloid leukemia  
**AUC:** area under the curve  
**BMB:** bone marrow biopsy  
**C:** concordance  
**cAUC:** cumulative case dynamic control of receiver operative characteristics  
**CDS:** clinical decision support  
**CIF:** cumulative incidence function  
**CKD:** chronic kidney disease  
**CR:** complete remission  
**CRi:** complete remission with incomplete hematologic recovery  
**CTCAE:** Common Terminology Criteria for Adverse Events  
**CV:** cross-validation  
**DL:** deep learning  
**ELN:** European Leukemia Net  
**eNet:** elastic-net  
**iAUC:** incident case dynamic control of receiver operative characteristics  
**ICU:** intensive care unit  
**LOOCV:** leave-one-out-cross-validation  
**M:** median  
**MCP:** minimax concave penalty  
**ML:** machine learning  
**MLFS:** morphologic leukemia-free state  
**MRD:** minimal residual disease  
**NCCN:** National Comprehensive Cancer Network  
**OS:** overall survival  
**PD:** progressive disease  
**PH:** proportional hazard  
**RBC:** red blood cell  
**ROC:** receiver operative characteristic  
**RSF:** random survival forest  
**RUNX:** runt-related transcription factor  
**SCAD:** smoothly clipped absolute deviation  
**ven/aza:** venetoclax plus azacitidine  
**WBC:** white blood cell

*Edited by N Cahill; submitted 20.11.23; peer-reviewed by Z Su, W Wei; comments to author 06.02.24; revised version received 12.03.24; accepted 08.07.24; published 21.08.24*

*Please cite as:*

*Islam N, Reuben JS, Dale J, Coates JW, Sapia K, Markson FR, Jordan CT, Smith C*

*Predictive Models for Long Term Survival of AML Patients Treated with Venetoclax and Azacitidine or 7+3 Based on Post Treatment Events and Responses: Retrospective Cohort Study*

*JMIR Cancer 2024;10:e54740*

*URL: <https://cancer.jmir.org/2024/1/e54740>*

*doi: [10.2196/54740](https://doi.org/10.2196/54740)*

*PMID:*

©Nazmul Islam, Jamie S Reuben, Justin Dale, James W Coates, Karan Sapia K, Frank R Markson, Craig T Jordan, Clay Smith. Originally published in JMIR Cancer (<https://cancer.jmir.org>), 21.08.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Cancer, is properly cited. The complete bibliographic information, a link to the original publication on <https://cancer.jmir.org/>, as well as this copyright and license information must be included.