A 29-gene and cytogenetic score for the prediction of resistance to induction treatment in acute myeloid leukemia

Tobias Herold,[1,2,3] Vindi Jurinovic,[4] Aarif M. N. Batcha,[2,3,4] Stefanos A. Bamopoulos,[1] Maja Rothenberg-Thurley,[1] Bianka Ksienzyk,[1] Luise Hartmann,[1,2,3] Philipp A. Greif,[1,2,3] Julia Phillippou-Massier,[5] Stefan Krebs,[5] Helmut Blum,[5] Susanne Amler,[3] Stephanie Schneider,[1] Nikola Konstandin,[1] Maria Cristina Sauerland,[6] Dennis Görlich,[6] Wolfgang E. Berdel,[7] Bernhard J. Wörmann,[8] Johanna Tischer,[1] Marion Subklewe,[1] Stefan K. Bohlander,[9] Jan Braess,[10] Wolfgang Hiddemann,[1,2,3] Klaus H. Metzeler,[1,2,3] Ulrich Mansmann[2,3,4*] and Karsten Spiekermann[1,2,3*]

[1]Department of Internal Medicine III, University of Munich, Germany; [2]German Cancer Consortium (DKTK), Partner Site Munich, Munich, Germany; [3]German Cancer Research Center (DKFZ), Heidelberg, Germany; [4]Institute for Medical Informatics, Biometry and Epidemiology, University of Munich, Germany; [5]Laboratory for Functional Genome Analysis (LAFUGA), Gene Center, Ludwig-Maximilians-Universität (LMU) München, Germany; [6]Institute of Biostatistics and Clinical Research, University of Munich, Germany; [7]Department of Medicine, Hematology and Oncology, University of Münster, Germany; [8]German Society of Hematology and Oncology, Berlin, Germany; [9]Department of Molecular Medicine and Pathology, University of Auckland, Auckland, New Zealand and [10]Department of Oncology and Hematology, Hospital Barmherzige Brüder, Regensburg, Germany

**UM and KS contributed equally to this work.*

©2018 Ferrata Storti Foundation. This is an open-access paper. doi:10.3324/haematol.2017.178442

Received: August 15, 2017.

Accepted: December 7, 2017.

Pre-published: December 14, 2017.

Correspondence: tobias.herold@med.uni-muenchen.de

Supplement

A 29-gene and cytogenetic score for the prediction of resistance to induction treatment in acute myeloid leukemia

Running Title: Prediction of primary resistant AML

Tobias Herold[1,2,3], Vindi Jurinovic[4], Aarif M. N. Batcha[2,3,4], Stefanos A. Bamopoulos[1], Maja Rothenberg-Thurley[1], Bianka Ksienzyk[1], Luise Hartmann[1,2,3], Philipp A. Greif[1,2,3], Julia Phillippou-Massier[5], Stefan Krebs[5], Helmut Blum[5], Susanne Amler[6], Stephanie Schneider[1], Nikola Konstandin[1], Maria Cristina Sauerland[6], Dennis Görlich[6], Wolfgang E. Berdel[7], Bernhard J. Woermann[8], Johanna Tischer[1], Marion Subklewe[1], Stefan K. Bohlander[9], Jan Braess[10], Wolfgang Hiddemann[1,2,3], Klaus H Metzeler[1,2,3], Ulrich Mansmann[2,3,4§] and Karsten Spiekermann[1,2,3§]

[1]Department of Internal Medicine III, University of Munich, Munich, Germany [2]German Cancer Consortium (DKTK), Partner Site Munich, Munich, Germany; [3]German Cancer Research Center (DKFZ), Heidelberg, Germany; [4]Institute for Medical Informatics, Biometry and Epidemiology, University of Munich, Munich, Germany; [5]Laboratory for Functional Genome Analysis (LAFUGA), Gene Center, Ludwig-Maximilians-Universität (LMU) München, München, Germany; [6]Institute of Biostatistics and Clinical Research, University of Münster, Münster, Germany; [7]Department of Medicine, Hematology and Oncology, University of Münster, Münster, Germany; [8]German Society of Hematology and Oncology, Berlin, Germany; [9]Department of Molecular Medicine and Pathology, University of Auckland, Auckland, New Zealand; [10]Department of Oncology and Hematology, Hospital Barmherzige Brüder, Regensburg, Germany

§ contributed equally

Corresponding Author:
Tobias Herold, MD; Marchioninistr. 15; 81377 München; Germany; Phone: +49 89 4400-0; FAX: +49 89 4400-74242; Email: tobias.herold@med.uni-muenchen.de

Supplement

Treatment protocols

The AMLCG-1999 trial (clinicaltrials.gov identifier NCT00266136) randomized patients <60 years to receive double induction with either one cycle of TAD-9 (thioguanine 100 mg/m² twice daily on days 3-9, cytarabine 100 mg/m²/d continuous infusion on days 1 and 2 and 100 mg/m² twice daily on days 3-8, and daunorubicin 60 mg/m² on days 3-5) followed by one cycle of HAM (cytarabine 3 g/m² twice daily on days 1-3 and mitoxantrone 10 mg/m² on days 3-5) on day 21, or two cycles of HAM 21 days apart. Older patients (≥60 years) were randomized to receive induction therapy with one cycle of either TAD-9 or HAM. A second cycle of HAM was stipulated in the protocol if on day 21 ≥ 5% residual blasts were present in the bone marrow at day 16. The trial recruited from 1999 to 2004.

The AMCG-2008 trial (clinicaltrials.gov identifier NCT01382147) randomized patients <60 years and medically fit patients ≤ 70 years to receive either double induction chemotherapy with TAD-9 and HAM (21 days apart) as stipulated in the AMLG-1999 trial, or dose-dense induction therapy (sHAM: cytarabine 3 g/m² [1 g/m² in patients ≥60 years] twice daily on days 1, 2, 8 and 9 and mitoxantrone 10 mg/m² on days 3, 4, 10 and 11). Medically unfit and older patients were randomized to receive either induction therapy according to the HAM regimen with reduced cytarabine dose (1g/m² per dose) and a second cycle of HAM if on day 21 ≥ 5% residual blasts were present in the bone marrow at day 16 or to dose-dense induction with sHAM (cytarabine, 1g/m² per dose). The trial recruited from 2009 to 2012.

RNA and cDNA preparation

RNA was prepared using Trizol and the Zymo Research Direct-Zol RNA Prep Kit (Zymo Research) following the recommendations of the manufacturer. The manufactures instructions of the Sense mRNA Seq Library Prep Kit V2 (Lexogen) were applied with the following changes: The reverse transcription and ligation was performed using 2.5 µl

Starter/Stopper Mix (ST) in pre-warmed Reverse Transcription and Ligation Mix Long buffer (RTL). The following incubation time was increased to 15 min. Libraries were amplified in 21 cycles. All elutions from Purification Beads (PB) were performed at 37 degree Celsius, 1000 rpm for 30 min. Libraries were quantified using the Qubit (Invitrogen). The molarity was determined as well as their quality controlled using the Bioanalyzer (Agilent Technologies).

Processing details and sequencing metrics

Samples were aligned with STAR 2.4.0[8] to the Reference Genome (hg19).

The alignment to the reference genome was a two-step process. A genomic index was created using the reference genome and its complementary .gtf-file. All samples were aligned with the genomic index. All splice junctions (SJ.out.tab, see RNA STAR Manual 2.4.0.1) from all samples were extracted and merged into a single file. Mitochondrial splice junctions were removed by filtering. A novel genome index for 2nd pass mapping was created by integrating the extracted and processed splice junctions with the reference genome and its complementary .gtf-file. All samples were then re-aligned with the new genomic index in order to make use of the large sample count. No transcriptome reference was used, but reads mapping to the transcriptome were extracted using the --quantMode TranscriptomeSAM option. Orphaned reads, that were created during read pre-processing were aligned separately and only once (only 2nd pass mapping with the new reference genome). They were discarded in further downstream analysis due to their arbitrary count (<0.1%).

Read pairs underwent adaptor clipping and quality trimming prior to alignment by external tools using Cutadapt (Galaxy Tool Version 1.6) and Trimmomatic (Galaxy Tool Version 0.32.3). 9 bases were removed from the 5'-end of the forward reads and 6 bases were removed from the 5'-end of the reverse reads, as recommended by Lexogen. Two adaptor

sequences were clipped where necessary using default clipping options. Trimming was performed using the MAXINFO method with a target length of 40 and a strictness of 0.5. Reads were reported as valid alignments if the read pair mapped to less than 10 loci, and had less than 10 mismatches (with a ratio of mapped length/mismatches not exceeding 0.3). Splice junctions were reported (and later used for the creation of the 2nd genomic index), if they had an overhang of at least 12 (30 for non-canonical motifs) and had a minimum of 1 uniquely or multiply mapped read spanning over the splice junction (3 reads required for non-canonical motifs). A minimum overhang of 20 was required for the report of a chimeric splice junction. The minimum score difference between a chimeric alignment and the next best one was 10. Gap open penalty was 0 (-8 for non-canonical, AT/AC- or GT/AT-motifs and -4 for GC/AG and CT/GC motifs). Deletion and insertion open penalty was -2 (plus -2 per extended base).

Preprocessing of raw count data was done with the R package DESeq2.[9]

Development of the predictive classifier

10,000 bootstrap samples of both training sets were used for preselection of variables. For each bootstrap sample, univariate logistic regression was used to test all variables for a significant association with resistant disease/non-response. Variables with a Benjamini-Hochberg-adjusted p-value ≤0.05 were considered significant. Only variables that were significant in ≥50% of bootstrap samples in both training sets were considered for further analyses. The candidate models were constructed in training set 1. Different numbers of variables were included in penalized logistic regression to create multiple candidate models, ranging from 100 most frequent to 500 most frequent variables (most frequently significant variables in bootstrap samples of training set 1). The penalty parameter was chosen with 10-fold cross validation. Alternatively, we created additional models by limiting the number of variables in the model from 5 to 50. Among the numerous candidate models, the one with

the best AUC in training set 2 was chosen as the final model. The reason for this development process was that multiple signatures in training set 1 had optimal AUC´s and a selection process using training set 2 resulted in a signature that integrates patients treated with “7+3” regiments. Furthermore, this development procedure increases the robustness of the classifier by avoiding overfitting. The distribution and impact of the variables included in the final classifier are shown in Figure S8.

In its current version it is not possible to interpret the result of a single patient analyzed e.g. by microarray from a different data set. If RNA sequencing results were available a conversion of the expression results to RPKM values would be an option to make the data sets comparable. However, to allow a broader usage of the classifier the transfer to a more accessible and standardized platform like e.g. NanoString is warranted and necessary.

Calibration analysis

To test the calibration of PS29MRC in the validation set, we divided the continuous score into groups defined by its quintiles and calculated the predicted probability for resistant disease in each patient. For a well calibrated model, the distribution of these probabilities in each group should reflect this group’s true proportion of patients with resistant disease. Figure S9 shows the distribution of predicted probabilities each group and its true rate of patients with resistant disease.

Supplementary Tables

Table S1: Univariate and multivariable analysis of the prediction of resistant disease in the validation set

	Multivariable analysis, n=235		**Univariate analysis**	
Variable	OR [95%-CI]	p-value	OR [95%-CI]	p-value
PS29MRCcont	1.75 [1.26; 2.47]	**0.0011**	2.39 [1.80; 3.26]	**$8.63\ 10^{-9}$**
Age *continuous*	1.06 [1.02; 1.09]	**0.00062**	1.07 [1.04-1.10]	**$2.58\ 10^{-6}$**
NPM1mut	0.44 [0.18; 1.01]	0.059	0.23 [0.11-0.45]	**$5.21\ 10^{-5}$**
RUNX1mut	0.89 [0.38; 2.06]	0.79	2.42 [1.24-4.72]	**0.0092**
TP53mut	6.61 [1.60; 36.56]	**0.016**	11.92 [3.69-53.29]	**0.00017**

Table S2: Univariate and multivariable analysis of the prediction of resistant disease in the validation set containing only patients treated in the AMLCG 2008 trial

	Multivariable analysis, n=195		**Univariate analysis**	
Variable	OR [95%-CI]	p-value	OR [95%-CI]	p-value
PS29MRCdic	4.15 [1.42; 12.33]	**0.0091**	7.62 [3.18; 18.51]	**$5.47 \cdot 10^{-6}$**
Age *continuous*	1.06 [1.02; 1.11]	**0.0077**	1.07 [1.03; 1.11]	**0.0011**
NPM1mut	0.51 [0.15; 1.56]	0.24	0.30 [0.11; 0.72]	**0.012**
RUNX1mut	0.72 [0.21; 2.28]	0.59	2.00 [0.78; 4.83]	0.14
TP53mut	3.25 [0.56; 22.25]	0.20	12.69 [3.15; 63.13]	**0.00057**

Table S3: Univariate and multivariable analysis of the prediction of resistant disease in the validation set for resistant disease at day 60

	Multivariable analysis, n=211		**Univariate analysis**	
Variable	OR [95%-CI]	p-value	OR [95%-CI]	p-value
PS29MRCdic	4.56 [1.95; 11.05]	**0.00058**	8.09 [3.88; 17.46]	**4.29×10^{-8}**
Age *continuous*	1.06 [1.02; 1.10]	**0.0011**	1.07 [1.04; 1.10]	**3.64×10^{-5}**
NPM1mut	0.42 [0.15; 1.09]	0.080	0.21 [0.09-0.44]	**0.00014**
RUNX1mut	1.30 [0.52; 3.19]	0.57	3.03 [1.46; 6.27]	**0.0026**
TP53mut	8.45 [1.87; 47.77]	**0.0082**	9.75 [2.78; 45.27]	**0.00091**

Table S4: Multivariable analysis of PS29MRCdic for overall survival (not censored for SCT)

	Multivariable analysis, n=250		**Univariate analysis**	
Variable	OR [95%-CI]	p-value	OR [95%-CI]	p-value
PS29MRCdic	2.15 [1.39; 3.31]	**0.00052**	2.81 [1.98; 3.99]	**$7.73\ 10^{-9}$**
Age *continuous*	1.03 [1.02; 1.05]	**$1.21\ 10^{-5}$**	1.04 [1.02; 1.05]	**$5.16\ 10^{-7}$**
***de novo*-AML**	0.65 [0.43; 0.99]	**0.043**	0.61 [0.42; 0.90]	**0.012**
ASXL1	0.80 [0.47; 1.38]	0.42	1.63 [1.08; 2.46]	**0.019**
RUNX1mut	1.21 [0.78; 1.88]	0.40	1.69 [1.17; 2.46]	**0.0057**
TP53mut	2.39 [1.32; 4.31]	**0.0040**	3.55 [2.14; 5.90]	**$1.04\ 10^{-6}$**
U2AF1	1.12 [0.54; 2.33]	0.76	2.34 [1.23; 4.46]	**0.0096**

Table S5: Univariate and multivariable analysis of the prediction of resistant disease of PS29MRCcont and alternative models in the validation set

	Multivariable analysis, n=225§		**Univariate analysis***		
Variable	OR [95%-CI]	p-value	OR [95%-CI]	p-value	AUC
PS29MRCcont	2.32 [1.72; 3.23]	**$1.23{\cdot}10^{-7}$**	2.46 [1.84; 3.40]	**$8.29{\cdot}10^{-9}$**	**0.76**
AML-score by Walter et al.	1.18 [1.02; 1.36]	**0.023**	1.28 [1.12; 1.48]	**0.00068**	**0.71**
	Multivariable analysis, n=225§		**Univariate analysis***		
Variable	OR [95%-CI]	p-value	OR [95%-CI]	p-value	AUC
PS29MRCcont	2.35 [1.74; 3.27]	**$1.13{\cdot}10^{-7}$**	2.46 [1.84; 3.40]	**$8.29{\cdot}10^{-9}$**	**0.76**
Molecular Version of the AML-score by Walter et al.	1.15 [0.91; 1.45]	0.25	1.37 [1.11; 1.69]	**0.0032**	**0.63**
	Multivariable analysis, n=235		**Univariate analysis**		
Variable	OR [95%-CI]	p-value	OR [95%-CI]	p-value	
PS29MRCcont	2.16 [1.59; 3.00]	**$1.90{\cdot}10^{-6}$**	2.39 [1.80; 3.26]	**$8.63{\cdot}10^{-9}$**	**0.76**
LSC17	7.11 [0.75; 70.04]	0.088	51.36 [7.80; 388.15]	**$7.29{\cdot}10^{-5}$**	**0.66**
	Multivariable analysis, n=235		**Univariate analysis**		
Variable	OR [95%-CI]	p-value	OR [95%-CI]	p-value	
PS29MRCcont	2.38 [1.76; 3.32]	**$8.44{\cdot}10^{-8}$**	2.39 [1.80; 3.26]	**$8.63{\cdot}10^{-9}$**	**0.76**
Retrained response LSC17	1.02 [0.54; 1.92]	0.95	1.97 [1.16; 3.39]	**0.013**	**0.61**

§ n=10 patients had to be excluded due to missing variables to calculate the AML-score by Walter et al.

***** To allow a fair comparison, univariate analyses were performed on the subset of patients with available information on all compared variables

Supplementary References

1. Buchner T, Krug UO, Peter Gale R, et al: Age, not therapy intensity, determines outcomes of adults with acute myeloid leukemia. Leukemia 30:1781-4, 2016
2. Buchner T, Berdel WE, Schoch C, et al: Double induction containing either two courses or one course of high-dose cytarabine plus mitoxantrone and postremission therapy by either autologous stem-cell transplantation or by prolonged maintenance for acute myeloid leukemia. J Clin Oncol 24:2480-9, 2006
3. Herold T, Metzeler KH, Vosberg S, et al: Isolated trisomy 13 defines a homogeneous AML subgroup with high frequency of mutations in spliceosome genes and poor prognosis. Blood 124:1304-11, 2014
4. Li Z, Herold T, He C, et al: Identification of a 24-gene prognostic signature that improves the European LeukemiaNet risk classification of acute myeloid leukemia: an international collaborative study. J Clin Oncol 31:1172-81, 2013
5. Wouters BJ, Lowenberg B, Erpelinck-Verschueren CA, et al: Double CEBPA mutations, but not single CEBPA mutations, define a subgroup of acute myeloid leukemia with a distinctive gene expression profile that is uniquely associated with a favorable outcome. Blood 113:3088-91, 2009
6. Taskesen E, Bullinger L, Corbacioglu A, et al: Prognostic impact, concurrent genetic mutations, and gene expression features of AML with CEBPA mutations in a cohort of 1182 cytogenetically normal AML patients: further evidence for CEBPA double mutant AML as a distinctive disease entity. Blood 117:2469-75, 2011
7. Braess J, Kreuzer K-A, Spiekermann K, et al: High Efficacy and Significantly Shortened Neutropenia Of Dose-Dense S-HAM As Compared To Standard Double Induction: First Results Of a Prospective Randomized Trial (AML-CG 2008). Blood 122:619-619, 2013
8. Dobin A, Davis CA, Schlesinger F, et al: STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15-21, 2013
9. Love MI, Huber W, Anders S: Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550, 2014

Figure S1: Flow Chart Patient Cohorts

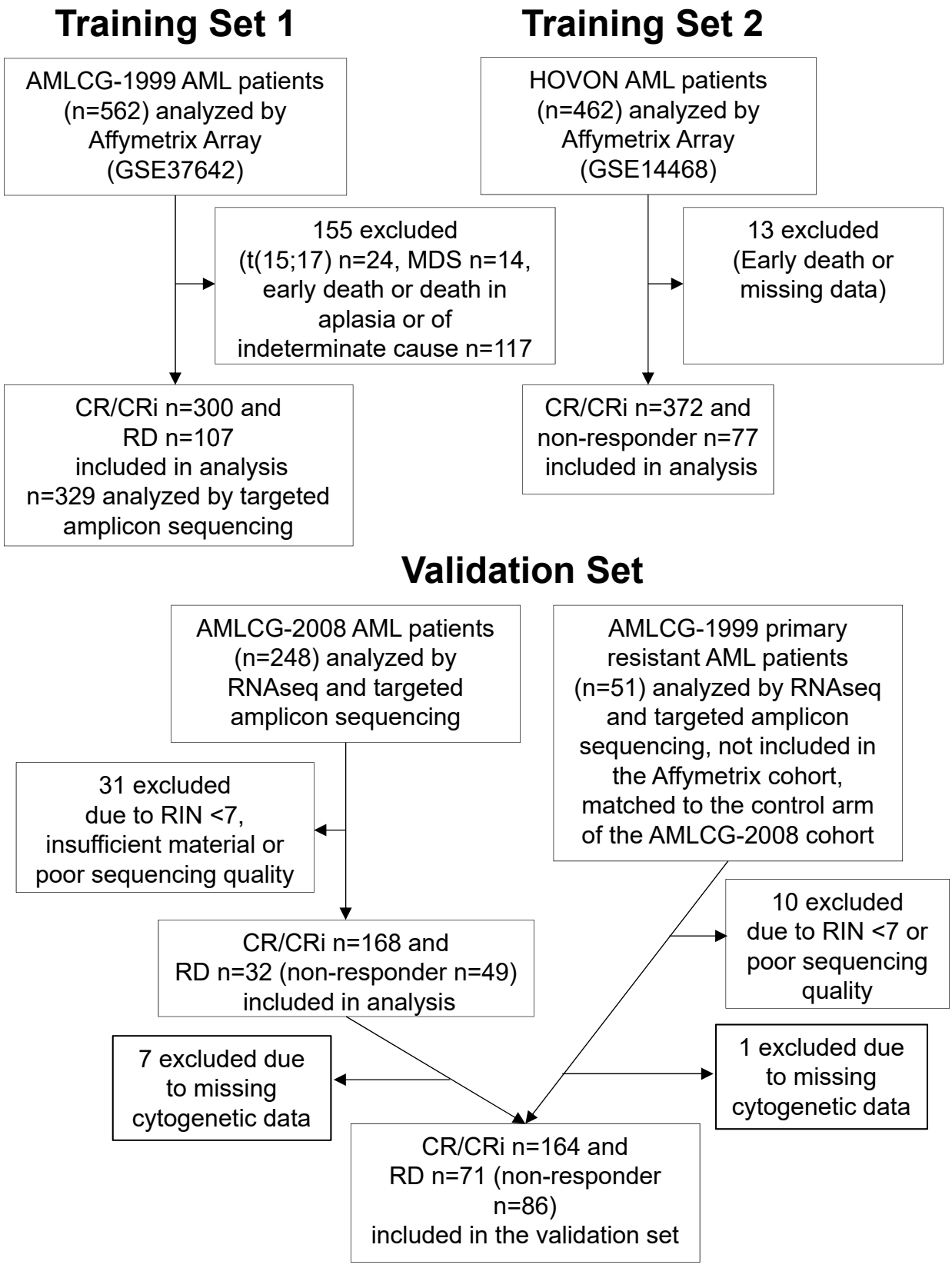


Figure S1: Flow chart describing the patient cohorts and selection process. In training set 2 only the information of responder (CR/CRi) and non-responder was available (see definition or response Figure S3). CR: Complete remission; CRi: CR with incomplete recovery; RD: Resistant disease; RIN: RNA integrity number

Figure S2: Flow Chart Signature Generation

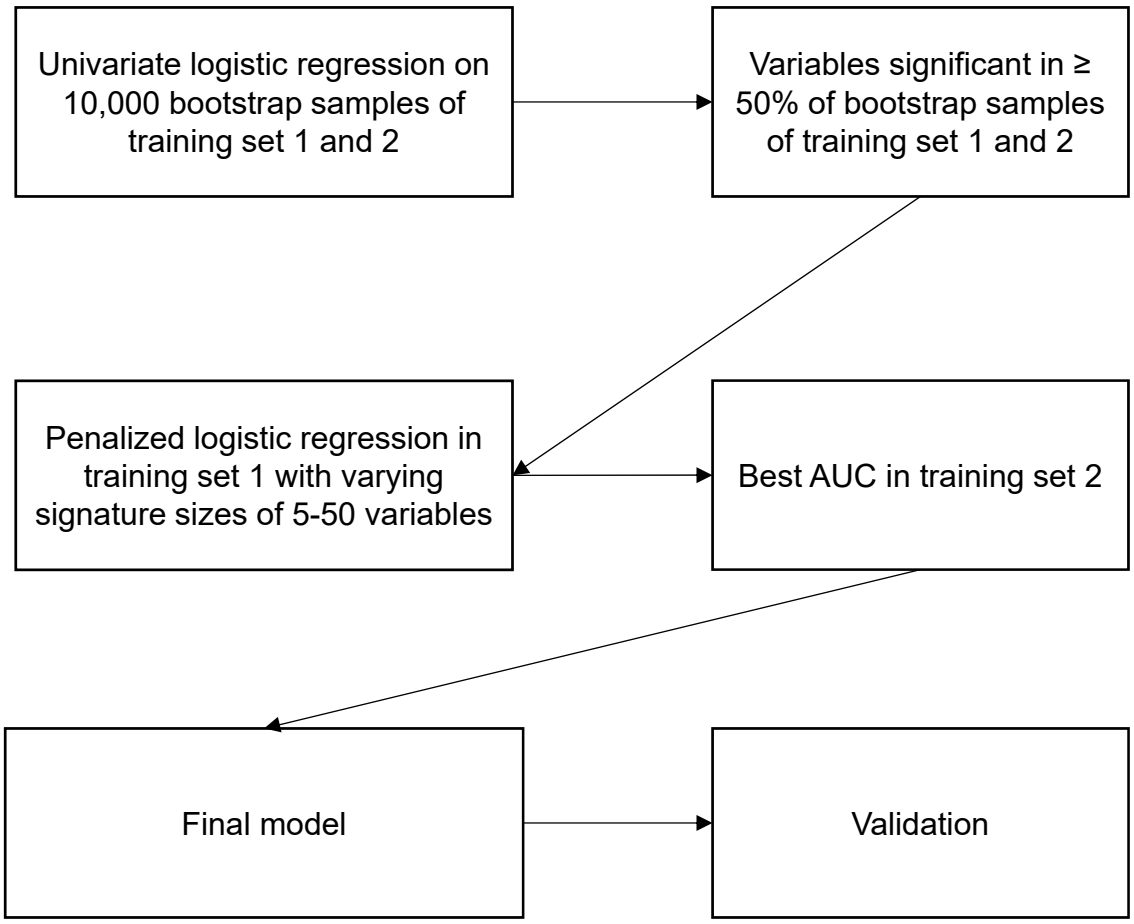


Figure S2: Flow chart describing the signature development process.

Figure S3: Barplots showing the specificity and sensitivity of the classifier

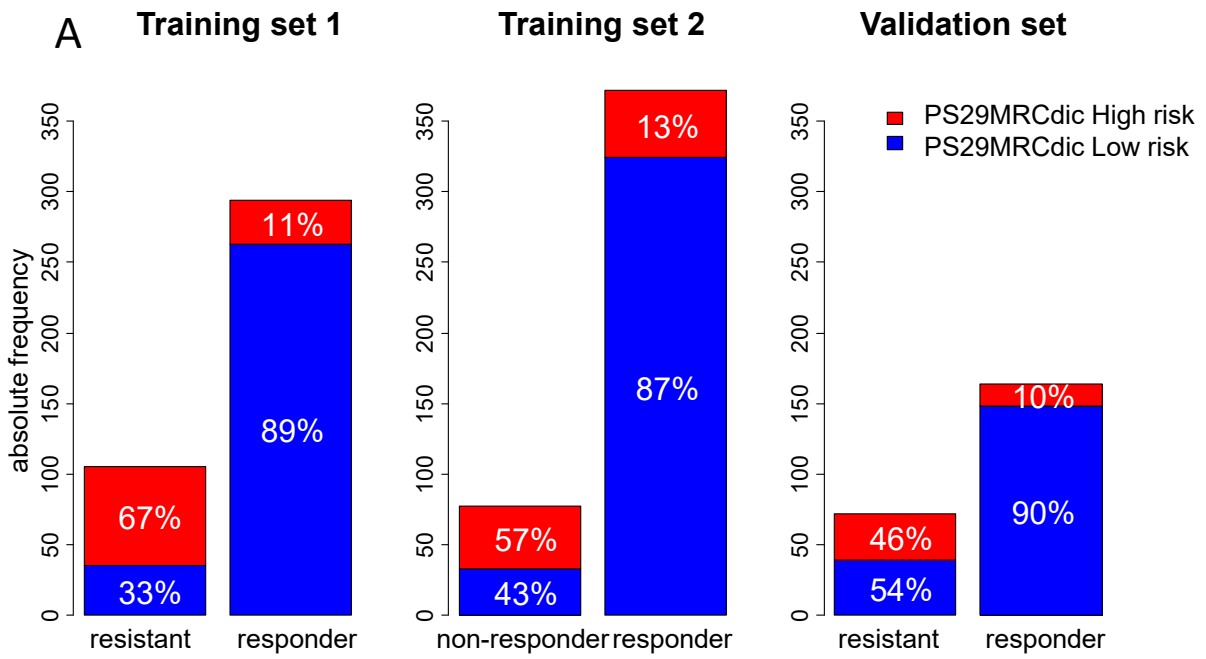


Figure S3: Barplots showing the predictive ability of the classifier PS29MRCdic in the training sets 1 and 2 and the independent validation set. The y-axis shows the absolute number of patients included. Patients in blue were predicted to respond to treatment (Low risk). Patients in red were predicted as non-responders (training set 2) or resistant disease (training set 1 and validation set) – High risk. The sensitivity of the predictor is shown as percentage in the red part of the barplot in non-responders/resistant disease and the specificity is shown in the blue part of the barplot in responders.

Figure S4: Predictive ability of PS29MRC in cytogenetic risk groups in the validation set

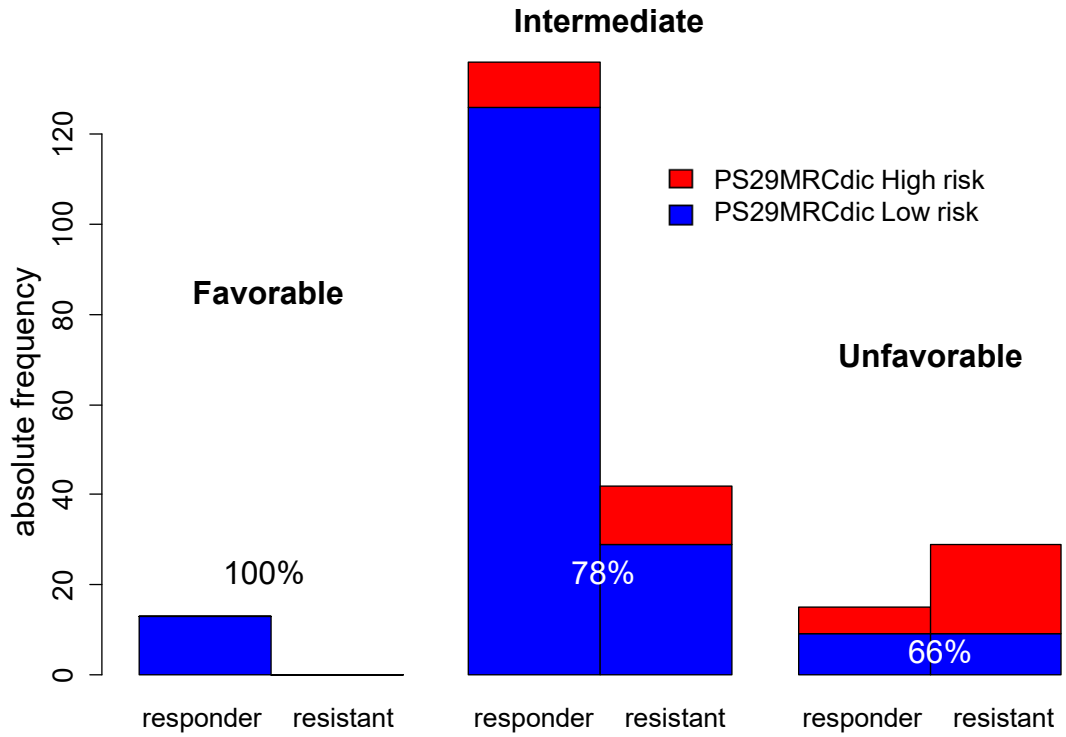


Figure S4: Barplots showing the predictive ability of PS29MRC in cytogenetic subgroups. The y-axis shows the absolute number of patients included. Patients in blue were predicted to respond to treatment (Low risk). Patients in red were predicted as resistant (High risk). The accuracy is given as percentage.

Figure S5: Refinement of the ELN 2017 genetic risk stratification by PS29MRC (not censored for SCT)

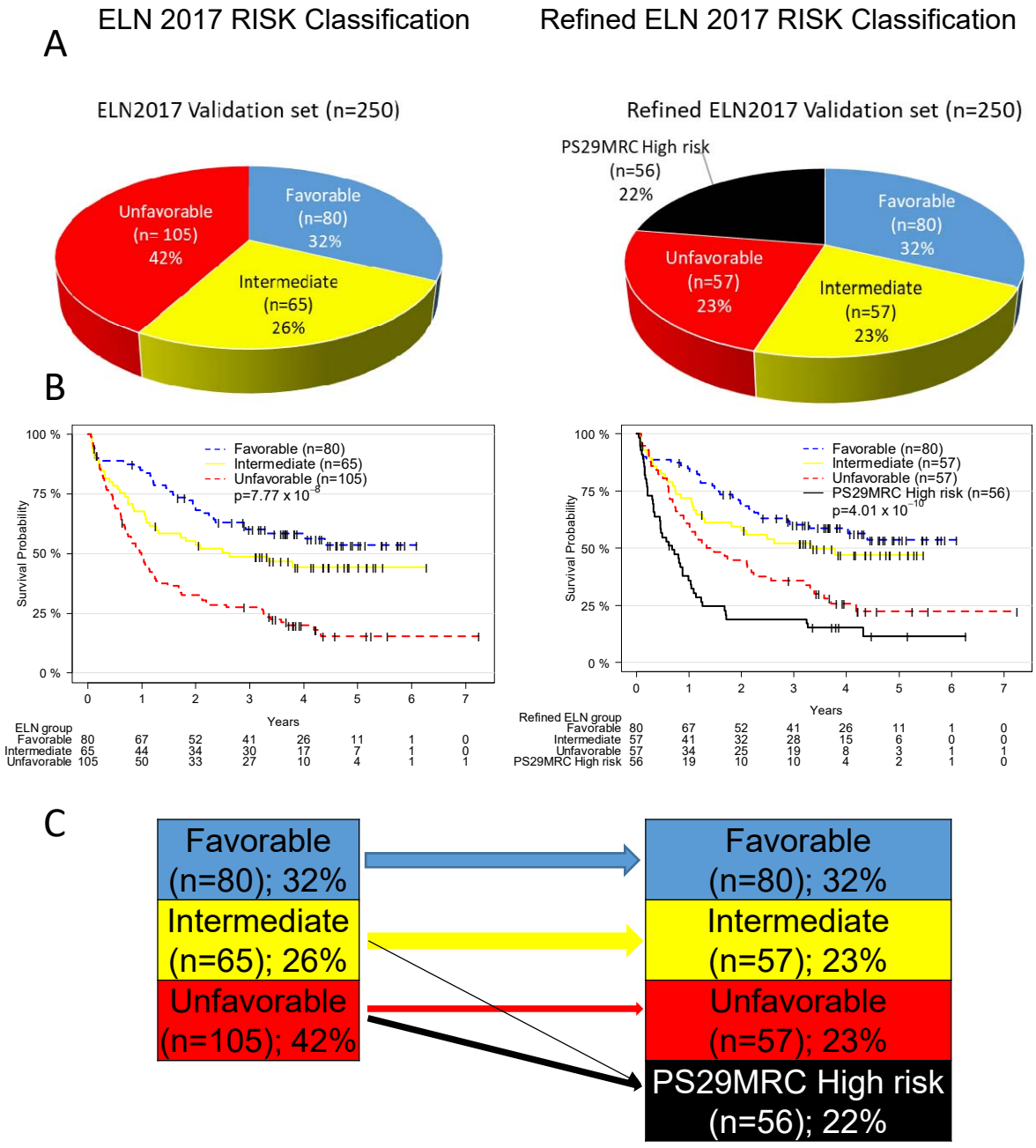


Figure S5: (A) Pie charts showing the distribution of patients according to ELN2017 and refined risk criteria. **(B)** Kaplan–Meier estimates of AML patients in the validation set according to ELN2017 and the refined ELN2017 classification. **(C)** Scheme of reclassification of the three ELN2017 risk groups into four groups by integrating PS29MRCdic (high risk) with the ELN2017 risk classification.

Median survival in months [95%-CI]: 8 [6-13], 18 [12-39], 41 [16- Not reached], not reached). 24 months survival probability: 19%, 45%, 60%, 71%.

Figure S6: Overall survival in the TCGA set using the predictive signature

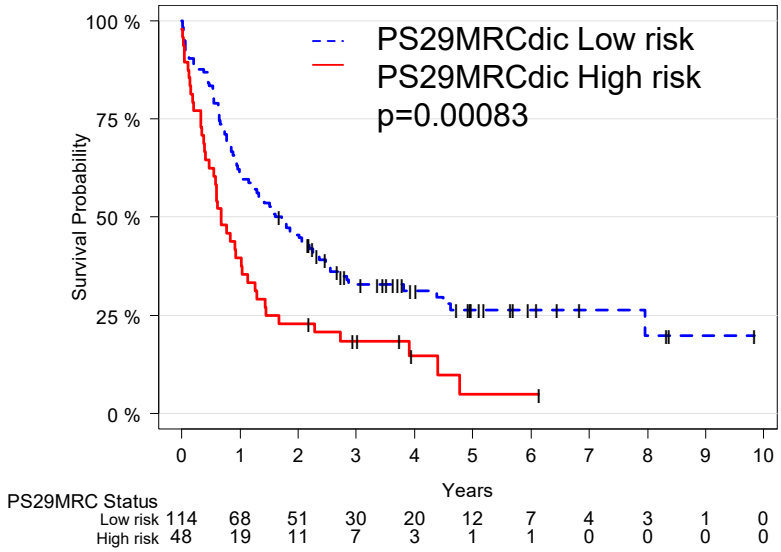
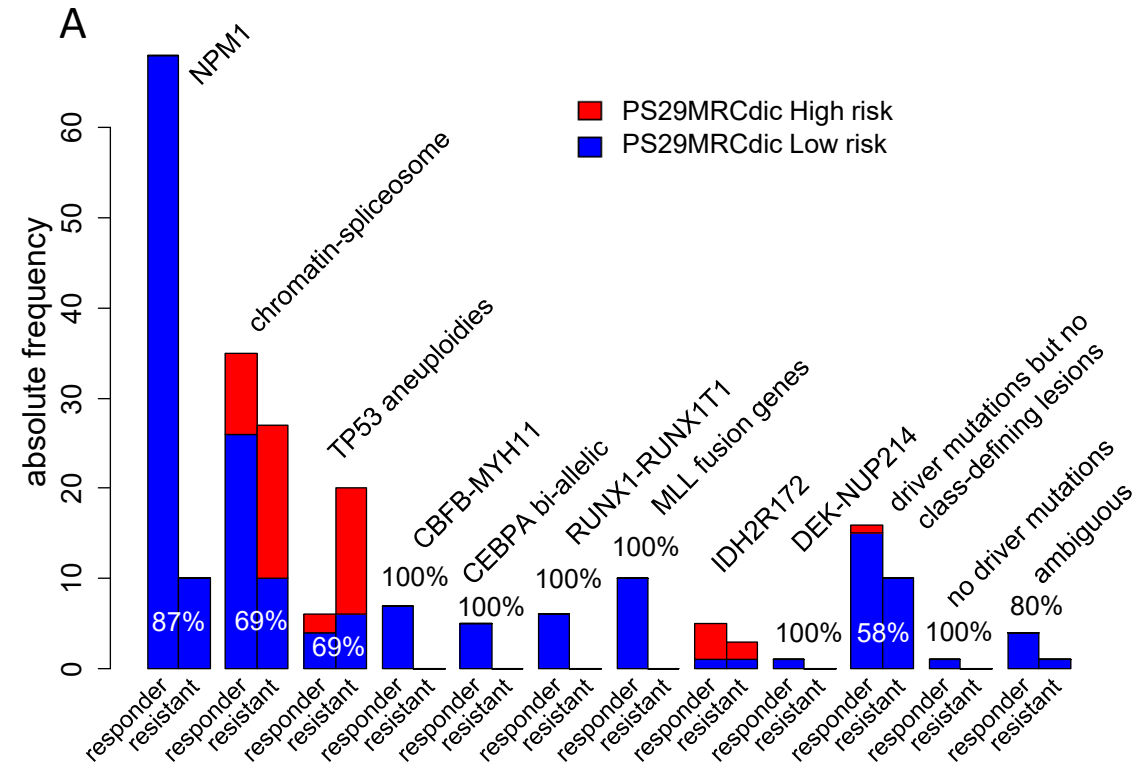


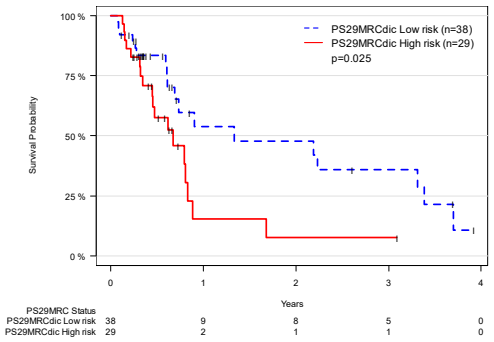
Figure S6: Overall survival of AML patients in the TCGA data set. Kaplan–Meier estimates of AML patients classified according to PS29MRC in low or high risk. The presented data is not censored for SCT.

Figure S7: Predictive ability of PS29MRC in genetic subgroups of AML according to Papaemmanuil et al.



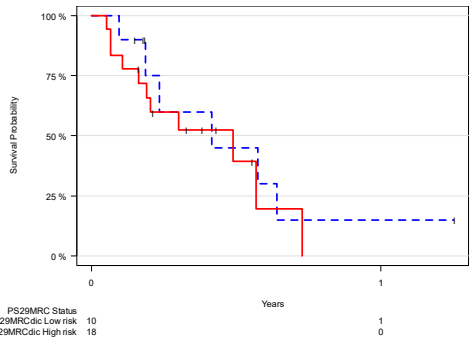
B

AML with mutated chromatin, RNA-splicing genes, or both



C

AML with TP53 mutations, chromosomal aneuploidy, or both



D

AML with driver mutations but no detected class-defining lesions

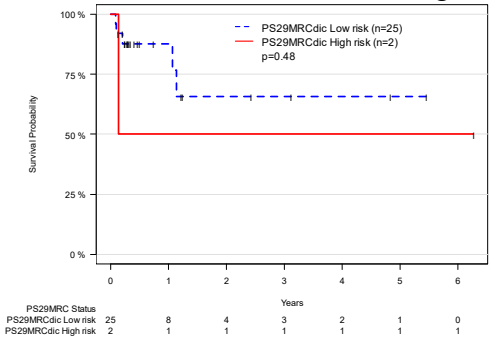


Figure S7: A: Barplots showing the predictive ability of PS29MRC in various genetic subgroups. The y-axis shows the absolute number of patients included. Patients in blue were predicted to respond to treatment. Patients in red were predicted as resistant. The accuracy is given as percentage. B-D: Overall survival of AML patients in selected genetic subgroups. Kaplan–Meier estimates of AML patients classified according to PS29MRC in Low risk and High risk.

Figure S8: Heatmap and barplot showing the distribution and impact of variables included in PS29MRC

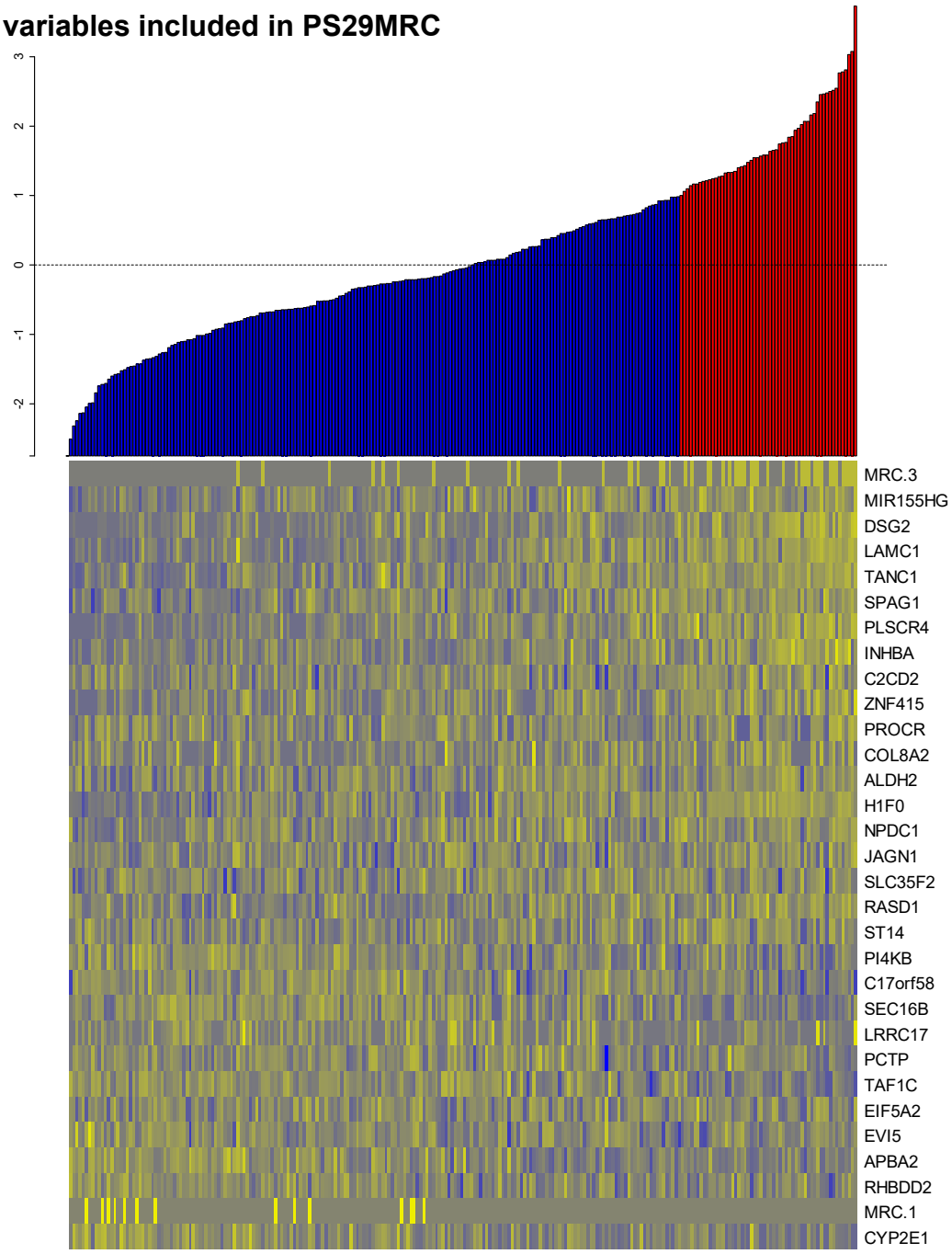


Figure S8: PS29MRC score for all patients of the validation set, along with the MRC cytogenetic groups and gene expression variables. In the heatmap yellow indicate higher and blue lower gene expression, respectively.

Figure S9: Calibration analysis

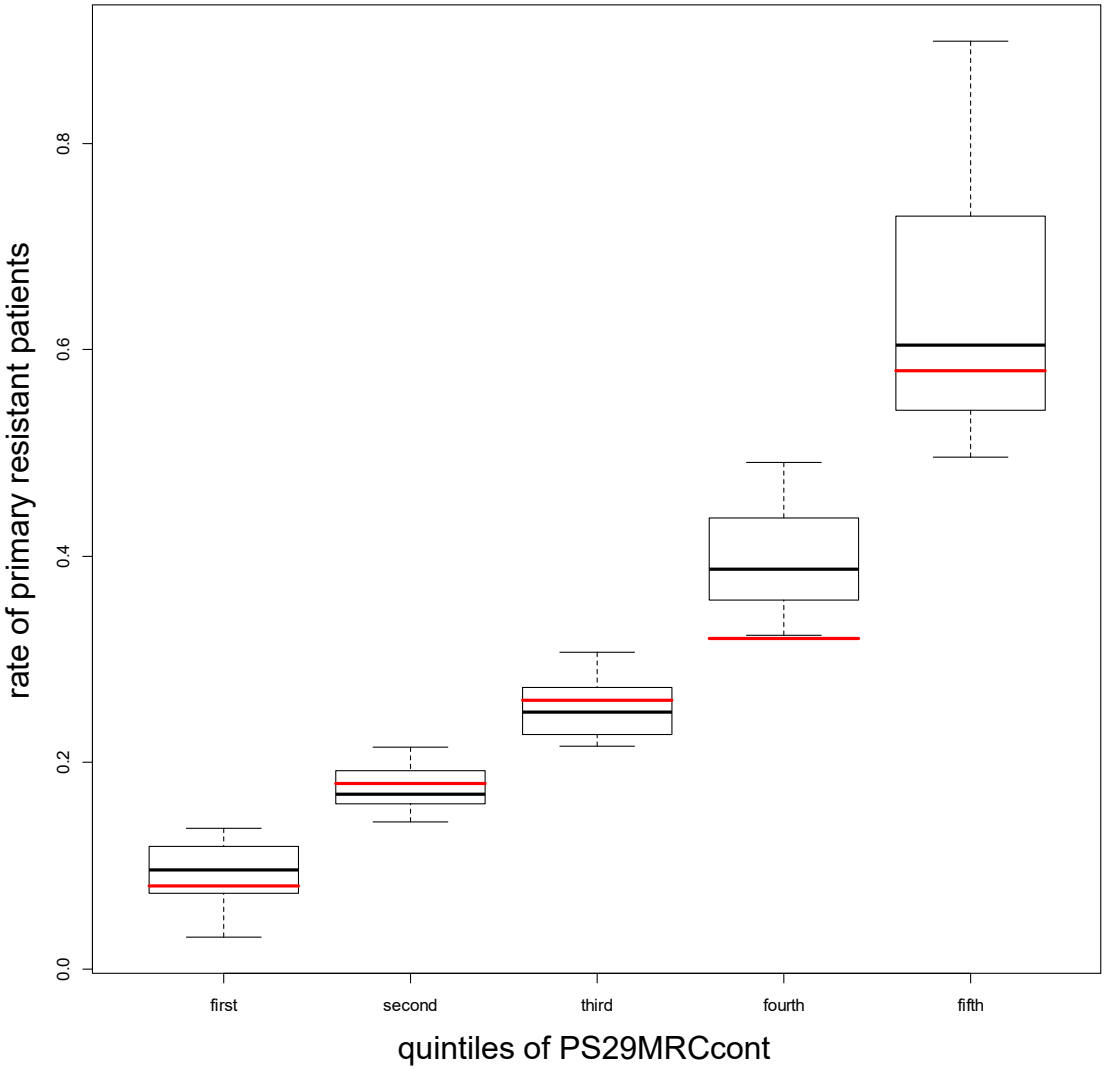


Figure S9: Predicted versus actual rates of resistant diseases of PS29MRCcont. Patients of the validation set were divided into quintiles according to their individual predicted probability of resistant disease.