Prognostic value of European LeukemiaNet 2022 criteria and genomic clusters using machine learning in older adults with acute myeloid leukemia

Silvia Park,[1,2,3*] Tong Yoon Kim,[1,2*] Byung-Sik Cho,[1,2] Daehun Kwag,[1,2] Jong-Mi Lee,[4] Myungshin Kim,[4] Yonggoo Kim,[4] Jamin Koo,[5,6] Anjali Raman[3,7] Tae Kon Kim,[3,7] and Hee-Je Kim[1,2]

[1]Department of Hematology, Catholic Hematology Hospital, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, South Korea; [2]Leukemia Research Institute, College of Medicine, The Catholic University of Korea, Seoul, South Korea; [3]Division of Hematology/Oncology, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA; [4]Department of Laboratory Medicine, College of Medicine, The Catholic University of Korea, Seoul, South Korea; [5]Department of Chemical Engineering, Hongik University, Seoul, South Korea; [6]ImpriMedKorea Inc., Seoul, South Korea and [7]Vanderbilt University, Nashville, TN, USA

SP and TYK contributed equally as first authors.

Correspondence: B.S. Cho
cbscho@catholic.ac.kr

Received: May 25, 2023.
Accepted: September 6, 2023.
Early view: September 14, 2023.

https://doi.org/10.3324/haematol.2023.283606

Supplementary Material

Table of contents

Methods

Molecular and Cytogenetic Studies

Bone marrow karyotyping was performed on G-banded metaphase chromosomes using conventional techniques. The karyotypes were interpreted using ISCN 2016.[1] For molecular analysis, next generation sequencing was performed using a customized myeloid panel ("SM panel") as described in our previous report.[2] The SM panel contains 67 genes that are frequently mutated in patients with AML. Target-capture sequencing was performed using a customized target kit (3039061; Agilent Technologies) according to the manufacturer's instructions. DNA libraries were prepared according to the manufacturer's protocol, and sequencing was performed using the Illumina HiSeq4000 platform (San Diego, CA, USA). Variants were interpreted as mutations when there were >20 and 5% variant allele frequencies (VAFs). Sequenced reads were mapped to the human reference genome (hg19; Genome Reference Consortium, February 2009).

For the detection of the FMS-like tyrosine kinase 3-internal tandem duplication (*FLT3*-ITD) mutation, polymerase chain reaction (PCR) for fragment analysis was performed using a previously published modified protocol.[3-5] The functional domains of *FLT3* (GenBank Accession NM_004119.2) were PCR-amplified with forward primers that were 5′ end-labeled with a fluorescent dye. The PCR products were interpreted using a model 3130XL genetic analyzer (Applied Biosystems, Foster City, CA, USA), and amplicons with a size greater than that of the wild type (328 ± 1 base) were considered positive for ITD mutation. The number, area, and length of the mutant peaks on the electropherogram were analyzed using the GeneMapper analysis software (Applied Biosystems).

Unsupervised Clustering and Stratification

Unsupervised techniques were evaluated in our cohort and internally validated without testing external data. The abnormalities in karyotypes and genetic mutations that were used for risk stratification in ELN 2022 and additional mutations found in >3% of all patients in our cohort were used for the analysis. Detailed genomic variables are described in Table S1. The number of clusters explored using the parameter NbClust[6] ranged from 3 to 12, and the optimal number of clusters was chosen according to the following measures the maximum value of the index (Dunn, Krzanowski–Lai, Calinski–Harabasz, Sarle, Ratkowsky–Lance, and Milligan), the maximum difference between the hierarchy levels of the index (Hartigan and Friedman–Rubin), and the minimum value of second differences between levels of the index (Friedman–Rubin).

R package NbClust and the Ward1 algorithm with Euclidean distances were used for hierarchical agglomerative clustering and k-means. Cluster,[7] kohonen,[8] and mcclust[9] packages were used for clustering with partitioning around medoids (PAM), self-organizing maps (SOM), and the Gaussian mixture model (GMM), respectively. For internal validation, the clustering algorithms were compared by the clValid package.[10] The variation of information was analyzed using the mcclust package[11] to identify similarities between clusters. The risk stratification in the survival curve was allocated to minimize the p-value by comparing the median and the 95% confidence interval for survival. The ggraph package[12] was used for network map visualization, and the correlation network map demonstrated the weight if the correlation coefficient was greater than 0.02. The analysis was performed using R software for statistical computing (R Foundation for Statistical Computing, Vienna, Austria, version 4.0.2).

Supplementary Table S1. Distribution of genetic alteration in the nine subgroups by hierarchical clustering methods.

Variables (N, %)	**C1, N=29**	**C2, N=30**	**C3, N=32**	**C4, N=26**	**C5, N=26**	**C6, N=67**	**C7, N=19**	**C8, N=30**	**C9, N=20**	***p*-value**
Age at diagnosis (years),	65.0 [62.0; 67.0]	67.0 [63.0; 76.0]	69.0 [64.5; 74.5]	68.0 [63.0; 70.0]	66.0 [63.0; 71.0]	68.0 [64.0; 74.5]	66.0 [64.5; 75.5]	72.0 [68.0; 74.0]	68.5 [63.0; 74.5]	0.054
Sex, female, n	13 (44.8)	14 (46.7)	6 (18.8)	14 (53.8)	16 (61.5)	29 (43.3)	7 (36.8)	10 (34.5)	10 (50.0)	0.657
Disease type										0.45
De novo	26 (89.7)	28 (93.3)	25 (78.1)	21 (80.8)	24 (92.3)	55 (82.1)	12 (63.2)	29 (96.7)	16 (80.0)	
Secondary	3 (10.3)	2 (6.7)	7 (21.9)	5 (19.2)	2 (7.7)	12 (17.9)	7 (36.8)	1 (3.3)	4 (20.0)	
Treatment										0.018
IC	22 (75.9)	16 (53.3)	11 (34.4)	15 (57.7)	17 (65.4)	24 (35.8)	5 (26.3)	13 (43.3)	8 (40.0)	
HMA	1 (3.4)	12 (40.0)	12 (37.5)	5 (19.2)	3 (11.5)	23 (34.3)	4 (21.1)	10 (33.3)	6 (30.0)	
HMA/VEN	6 (20.7)	2 (6.7)	9 (28.1)	6 (23.1)	6 (23.1)	20 (29.9)	10 (52.6)	7 (23.3)	6 (30.0)	
Complex	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	3 (4.5)	0 (0.0)	0 (0.0)	20 (100)	<0.001
-5 or del(5q);-7;17/abn(17p)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (3.8)	0 (0.0)	19 (100)	0 (0.0)	0 (0.0)	<0.001
Monosomal	0 (0.0)	0 (0.0)	1 (3.1)	0 (0.0)	0 (0.0)	0 (0.0)	12 (63.2)	0 (0.0)	7 (35.0)	<0.001
RUNX1-RUNX1T1	0 (0.0)	1 (3.3)	0 (0.0)	15 (57.7)	0 (0.0)	1 (1.5)	0 (0.0)	0 (0.0)	0 (0.0)	<0.001
CBFB-MYH11	0 (0.0)	0 (0.0)	0 (0.0)	10 (38.5)	0 (0.0)	1 (1.5)	0 (0.0)	0 (0.0)	0 (0.0)	<0.001
NPM1mut without FLT3-ITD	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (7.7)	21 (31.3)	0 (0.0)	9 (30.0)	0 (0.0)	<0.001
DNMT3A	0 (0.0)	9 (30.0)	6 (18.8)	2 (7.7)	14 (53.8)	21 (31.3)	4 (21.1)	0 (0.0)	4 (20.0)	<0.001
TET2	0 (0.0)	10 (33.3)	5 (15.6)	0 (0.0)	1 (3.8)	8 (11.9)	4 (21.1)	25 (83.3)	3 (15.0)	<0.001
FLT3-ITD	0 (0.0)	29 (96.7)	5 (15.6)	0 (0.0)	2 (7.7)	2 (3.0)	0 (0.0)	0 (0.0)	0 (0.0)	<0.001
FLT3-TKD	0 (0.0)	3 (10.0)	2 (6.2)	2 (7.7)	0 (0.0)	4 (6.0)	1 (5.3)	0 (0.0)	0 (0.0)	>0.999
IDH1	0 (0.0)	1 (3.3)	3 (9.4)	0 (0.0)	2 (7.7)	11 (16.4)	1 (5.3)	0 (0.0)	0 (0.0)	0.117
IDH2	0 (0.0)	5 (16.7)	1 (3.1)	0 (0.0)	23 (88.5)	3 (4.5)	0 (0.0)	0 (0.0)	2 (10.0)	<0.001
RUNX1	0 (0.0)	0 (0.0)	25 (78.1)	0 (0.0)	2 (7.7)	3 (4.5)	0 (0.0)	1 (3.3)	1 (5.0)	<0.001
ASXL1	0 (0.0)	0 (0.0)	18 (56.2)	0 (0.0)	3 (11.5)	4 (6.0)	0 (0.0)	4 (13.3)	0 (0.0)	<0.001
NRAS	0 (0.0)	0 (0.0)	4 (12.5)	0 (0.0)	0 (0.0)	13 (19.4)	1 (5.3)	1 (3.3)	1 (5.0)	0.009
PTPN11	0 (0.0)	0 (0.0)	1 (3.1)	0 (0.0)	0 (0.0)	17 (25.4)	1 (5.3)	0 (0.0)	0 (0.0)	<0.001
SRSF2	0 (0.0)	0 (0.0)	1 (3.1)	0 (0.0)	2 (7.7)	1 (1.5)	1 (5.3)	14 (46.7)	0 (0.0)	<0.001
BCOR	0 (0.0)	0 (0.0)	7 (21.9)	0 (0.0)	8 (30.8)	1 (1.5)	0 (0.0)	0 (0.0)	1 (5.0)	<0.001
TP53 VAF ≥ 10%	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (1.5)	2 (10.5)	1 (3.3)	7 (35.0)	<0.001
KIT	0 (0.0)	0 (0.0)	0 (0.0)	12 (46.2)	0 (0.0%)	0 (0.0)	1 (5.3)	0 (0.0)	1 (5.0)	<0.001
U2AF1	0	0	4	0	1	8	1	0	0	0.009

	(0.0)	(0.0)	(12.5)	(0.0)	(3.8)	(11.9)	(5.3)	(0.0)	(0.0)	
non-bZIP *CEBPA*	0	6	2	0	1	0	0	2	0	<0.001
	(0.0)	(20.0)	(6.2)	(0.0)	(3.8)	(0.0)	(0.0)	(6.7)	(0.0)	
bZIP in-frame CEBPA	0	0	0	0	0	6	0	0	0	0.117
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(9.0)	(0.0)	(0.0)	(0.0)	
JAK2	0	0	1	0	0	8	1	0	0	0.108
	(0.0)	(0.0)	(3.1)	(0.0)	(0.0)	(11.9)	(5.3)	(0.0)	(0.0)	
SETBP1	0	0	2	0	0	3	2	2	0	>0.999
	(0.0)	(0.0)	(6.2)	(0.0)	(0.0)	(4.5)	(10.5)	(6.7)	(0.0)	
KRAS	0	1	0	0	2	4	1	1	0	>0.999
	(0.0)	(3.3)	(0.0)	(0.0)	(7.7)	(6.0)	(5.3)	(3.3)	(0.0)	
SF3B1	0	2	2	0	1	2	0	0	1	>0.999
	(0.0)	(6.7)	(6.2)	(0.0)	(3.8)	(3.0)	(0.0)	(0.0)	(5.0)	
WT1	0	0	1	1	1	3	0	2	0	>0.999
	(0.0)	(0.0)	(3.1)	(3.8)	(3.8)	(4.5)	(0.0)	(6.7)	(0.0)	
RAD21	0	0	2	1	1	2	0	1	0	>0.999
	(0.0)	(0.0)	(6.2)	(3.8)	(3.8)	(3.0)	(0.0)	(3.3)	(0.0)	
PHF6	2	0	2	0	0	3	0	0	0	>0.999
	(6.9)	(0.0)	(6.2)	(0.0)	(0.0)	(4.5)	(0.0)	(0.0)	(0.0)	
CSF3R	1	1	0	1	1	2	0	0	0	>0.999
	(3.4)	(3.3)	(0.0)	(3.8)	(3.8)	(3.0)	(0.0)	(0.0)	(0.0)	
NF1	0	1	1	0	0	1	1	2	0	>0.999
	(0.0)	(3.3)	(3.1)	(0.0)	(0.0)	(1.5)	(5.3)	(6.7)	(0.0)	
KMT2A rearranged	1	0	0	0	1	1	2	0	0	>0.999
	(3.4)	(0.0)	(0.0)	(0.0)	(3.8)	(1.5)	(10.5)	(0.0)	(0.0)	
MLLT3-KMT2A	2	0	0	0	0	2	0	0	0	>0.999
	(6.9)	(0.0)	(0.0)	(0.0)	(0.0)	(3.0)	(0.0)	(0.0)	(0.0)	
DEK-NUP214	0	0	0	0	1	1	0	0	0	>0.999
	(0.0)	(0.0)	(0.0)	(0.0)	(3.8)	(1.5)	(0.0)	(0.0)	(0.0)	
BCR-ABL1	1	0	0	0	0	1	0	0	0	>0.999
	(3.4)	(0.0)	(0.0)	(0.0)	(0.0)	(1.5)	(0.0)	(0.0)	(0.0)	
MECOM	0	0	0	0	0	0	1	0	1	>0.999
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(5.3)	(0.0)	(5.0)	

C, cluster; N, number; IC, intensive chemotherapy; HMA, hypomethylation agent; HMA/VEN, hypomethylation plus venetoclax; VAF, variant allele frequency.

Results

Supplementary Table S2. The optimal number of clusters chosen by each index.

Number of clusters	Number of criteria	Detailed criteria in NbClust package	Detailed criteria reference
0	2	Hubert, Dindex	Graphical method (Hubert and Arabie 1985, Lebart et al. 2000)
2	1	Frey	the cluster level before that index value < 1.00 (Frey and Van Groenewoud 1972)
3	4	Duda, PseudoT2, Beale, McClain	Smallest n_c index>criticalValue (Duda and Hart 1973), Smallest $n_c$index<criticalValue (Duda and Hart 1973), n_c such that critical value of the index >= alpha (Beale 1969), Minimum value of the index (McClain and Rao 1975)
4	2	TrCovW, Ball	Maximum difference between (Milligan and Cooper 1985, Ball and Hall 1965) hierarchy levels of the index
5	1	Marriot	Maximum value of second differences (Marriot 1971)
8	2	Scott, Friedman	Maximum difference between (Scott and Symons 1971, Friedman and Rubin 1967) hierarchy levels of the index
9	8	KL, CH, Hartigan, TraceW, Rubin, Ratkowsky, PtBiserial, Dunn	Maximum value of the index (Krzanowski and Lai 1988, Calinski and Harabasz 1974, Calinski and Harabasz 1974, Milligan 1980, 1981, Dunn 1974), Maximum difference between (Hartigan 1975) hierarchy levels of the index, Maximum value of absolute second (Milligan and Cooper 1985) differences between levels of the index, Minimum value of second differences (Friedman and Rubin 1967) between levels of the index
11	6	CCC, Cindex,DB, Silhouette, Sdindex, SDbw	Maximum value of the index (Sarle 1983), Minimum value of the index (Hubert and Levin 1976, Davies and Bouldin 1979, Halkidi et al. 2000, Halkidi and Vazirgiannis 2001)

Supplementary Table S3. Comparison between the clustering algorithms.

	Optimal Algorithm	Clusters in HAC	Clusters in K-means	Clusters In GMM	Clusters In PAM	Clusters In SOM
Connectivity	HAC	66.616	159.531	130.276	151.425	172.900
Dunn	K-means	0.333	0.378	0.354	0.333	0.378
Silhouette	PAM	0.134	0.172	0.159	0.188	0.171
APN	HAC	0.043	0.352	0.140	0.147	0.286
AD	PAM	1.697	1.556	1.405	1.388	1.464
ADM	HAC	0.162	0.528	0.206	0.227	0.419
FOM	GMM	0.256	0.252	0.248	0.248	0.253

HAC, hierarchical agglomerative clustering; GMM, gaussian mixture model; PAM, partitioning around medoids; SOM, self-organizing maps; APN, average proportion of non-overlap; AD, the average distance; ADM, the average distance between means; FOM, the figure of merit.

Supplementary Table S4. Overall survival in nine subgroups by hierarchical clustering methods.

	C1 (N=29)	C2 (N=30)	C3 (N=32)	C4 (N=26)	C5 (N=26)	C6 (N=67)	C7 (N=19)	C8 (N=30)	C9 (N=20)
Total Median OS, 95% CI, Months	NR [18.2-NR]	8.9 [6.8-12.9]	24.1 [10.4-NR]	23.6 [11.6-NR]	31.3 [15.1-NR]	9.0 [7.7-14.1]	6.7 [5.6-NR]	11.1 [7.4-17.5]	10.3 [6.1-17.8]
IC	NR [15.7-NR]	8.9 [6.9-12.9]	24.1 [10.4-NR]	23.6 [10.4-NR]	31.3 [15.1-NR]	8.6 [7.7-14.1]	6.7 [5.6-18.0]	11.1 [7.4-17.5]	10.3 [6.1-17.8]
HMA	15.7 [NR-NR]	4.1 [2.8-NR]	12.3 [8.2-NR]	11.6 [6.0-NR]	15.1 [4.7-NR]	8.0 [3.6-14.2]	11.8 [3.6-NR]	11.1 [3.9-NR]	8.6 [5.3-NR]
HMA/VEN	NR [NR-NR]	3.7 [3.7-NR]	31.3 [5.2-NR]	12.0 [2.4-NR]	NA [5.1-NR]	6.2 [5.5-17.2]	7.9 [3.5-NR]	7.4 [3.7-NR]	13.6 [2.6-NR]

IC, intensive chemotherapy; HMA, hypomethylation agent; HMA/VEN, hypomethylation plus venetoclax; C, cluster; N, number; OS, overall survival; CI, confidence interval.

References

1. International Standing Committee on Human Cytogenomic Nomenclature. ISCN: an international system for human cytogenomic nomenclature (2016). Basel: Karger; 2016.
2. Lee J-M, Kim Y-J, Park S-S, Han E, Kim M, Kim Y. Simultaneous monitoring of mutation and chimerism using next-generation sequencing in myelodysplastic syndrome. J Clin Med. 2019;8:2077.
3. Kim Y, Lee GD, Park J, et al. Quantitative fragment analysis of *FLT3*-ITD efficiently identifying poor prognostic group with high mutant allele burden or long ITD length. Blood Cancer J. 2015;5:e336.
4. Yoon J-H, Kim H-J, Park S-S, et al. Long-term clinical outcomes of hematopoietic cell transplantation for intermediate-to-poor-risk acute myeloid leukemia during first remission according to available donor types. Oncotarget. 2017;8:41590-41604.
5. Lee GD, Kim J, Lee SY, et al. Fragment analysis for detection of the *FLT3*-internal tandem duplication: comparison with conventional PCR and Sanger sequencing. Lab Med Online. 2017;7:13-19.
6. Charrad M, Ghazzali N, Boiteau V, Niknafs A. NbClust: An R package for determining the relevant number of clusters in a data set. J Stat Softw. 2014;61:1-36.
7. Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K. Cluster: Cluster analysis basics and extensions. https://www.researchgate.net/publication/272176869_Cluster_Cluster_Analysis_Basics_and_Extensions Accessed May 17, 2023.
8. Wehrens R, Kruisselbrink J. Flexible self-organizing maps in kohonen 3.0. J Stat Softw. 2018;87:1-18.
9. Fraley C, Raftery A. Model-based methods of classification: using the mclust software in

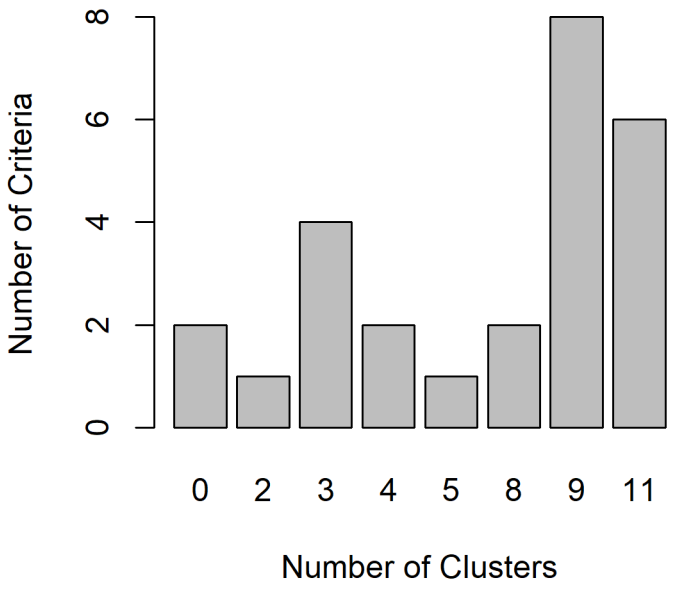
chemometrics. J Stat Softw. 2007;18:1-13.

10. Brock G, Pihur V, Datta S, Datta S. clValid: An R package for cluster validation. J Stat Softw. 2008;25:1-22.

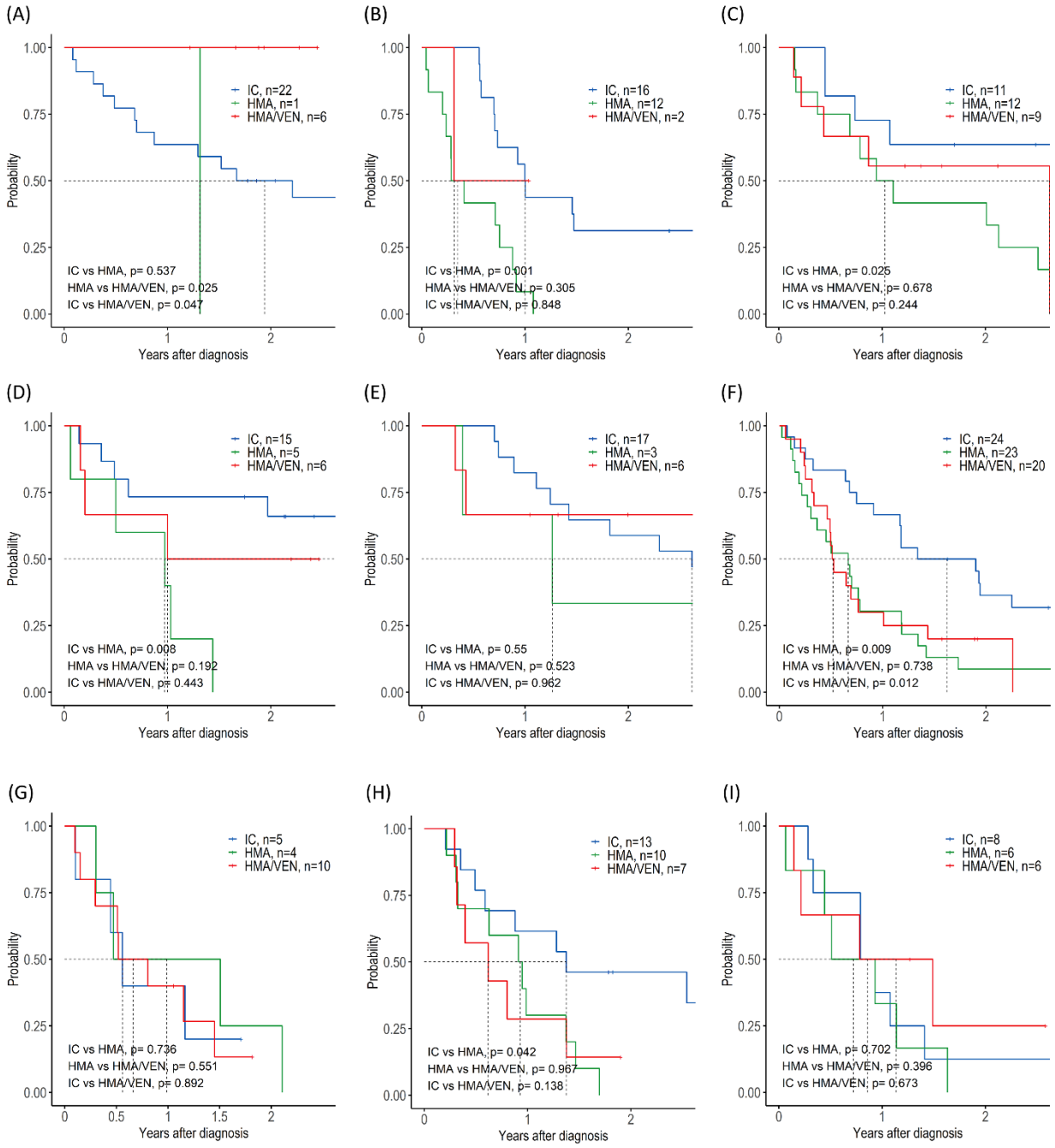
11. Fritsch A, Ickstadt K. Improved criteria for clustering based on the posterior similarity matrix. Bayesian Anal. 2009;4:367-391.

12. Pedersen TL. ggraph: an implementation of grammar of graphics for graphs and networks. R package version 2.0.3. 2020.

Supplementary Figure S1: The optimal number of clusters, nine in this case, was determined through a voting process involving multiple indices.



Supplementary Figure S2: Impact of each cluster on overall survival by treatment groups. Cluster 1 (A), Cluster 2 (B), Cluster 3 (C), Cluster 4 (D), Cluster 5 (E), Cluster 6 (F), Cluster 7 (G), Cluster 8 (H), and Cluster 9 (I).



Supplementary Figure S3: Comparison of the overall survival by hierarchical clustering according to treatment arms. Intensive chemotherapy (A), hypomethylating agent (B), and hypomethylating agent plus venetoclax (C).