Supplementary Information for "Immune Cell Contexture in the Bone Marrow Tumor Microenvironment Impacts Therapy Response in CML"

Oscar Brück[1], Sami Blom[2]*, Olli Dufva[1]*, Riku Turkki[2], Himanshu Chheda[2], Antonio Ribeiro[2], Panu Kovanen[3], Tero Aittokallio[2,4], Perttu Koskenvesa[1], Olli Kallioniemi[2,5], Kimmo Porkka[1], Teijo Pellinen[2], and Satu Mustjoki[1,6]

[1]Hematology Research Unit Helsinki, University of Helsinki and Department of Hematology, Helsinki University Hospital Comprehensive Cancer Center, Helsinki, Finland.

[2]Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland.

[3]Department of Pathology, University of Helsinki and HUSLAB, Helsinki University Hospital, Helsinki, Finland.

[4]Department of Mathematics and Statistics, University of Turku, Turku, Finland.

[5]Science for Life Laboratory, Karolinska Institutet, Department of Oncology and Pathology, Solna, Sweden.

[6]Department of Clinical Chemistry, University of Helsinki, Helsinki, Finland.

* These authors contributed equally to this work.

This file includes supplementary methods, reagent specification, and supplementary data supporting the original manuscript.

Supplementary Methods

Multiplexed Immunohistochemistry (mIHC)

General. TMA blocks were cut in 3.5 µm sections on Superfrost objective slides (Kindler O Gmbh, Germany) with the Microm 355S microtome (Thermo Scientific, Waltham, MA). Glass slides were dried overnight at +37°C and stored for short-term use at +4°C. All consecutive phases were performed in room temperature unless otherwise specified. Protein blocking and antibody incubations were performed in a humid chamber. Slides were washed three times with 0.1% Tween-20 (Thermo Fisher Scientific) diluted in 10 mM Tris-HCL buffered saline pH 7.4 (TBS) after peroxide block, antibody incubations, and fluorochrome reaction.

Tissue Preparation. Slides were deparaffinized in xylene and rehydrated in graded ethanol series and H_2O. Heat-induced epitope retrieval (HIER) was carried out in 10 mM Tris-HCl - 1 mM EDTA buffer (pH 9) in +99°C for 20 min (PT Module, Thermo Fisher Scientific, Waltham, MA). Peroxide activity was blocked in 0.9% H_2O_2 solution for 15 min, and subsequently 10% normal goat serum (TBS-NGS) applied for 15 min.

Fluorescence staining. Primary antibodies diluted in protein blocking solution and secondary anti-mouse or anti-rabbit horseradish peroxidase-conjugated (HRP) antibodies (Immunologic, Netherlands) diluted 1:1 in washing buffer were applied for 1h45min and 45 min, respectively. Tyramide signal amplification (TSA) Alexa Fluor 488 (PerkinElmer, Waltham, MA) diluted in TBS was applied on the slides for 10 min. Primary antibodies were denaturated and enzymatic activity of secondary antibody HRP was quenched by repeating HIER. Thereafter, peroxide and protein block were repeated, followed by application of a different primary antibody,

matching HRP-conjugated secondary antibody diluted 1:5 in washing buffer and TSA Alexa Fluor 555 (PerkinElmer). Again, HIER, peroxide block and protein block were repeated. Then, the slides were incubated with two additional primary antibodies immunized in different species overnight in +4°C. Next, AlexaFluor647 and AlexaFluor750 fluorochrome-conjugated secondary antibodies (Thermo Fisher Scientific) diluted in 1:150 in washing buffer were applied for 45 min. After washing, nuclei were counterstained with DAPI (Roche) diluted in 1:250 in TBS for 15 min. Last, we applied ProLong Gold mountant (Thermo Fisher Scientific) and a coverslip on the slides.

Chromogenic Staining. Following fluorescence imaging, tissue slides were incubated in washing buffer overnight in +4°C to detach atraumatically coverslips from the slides. Then, HIER, peroxide block and protein block were repeated similarly as before. Two primary antibodies immunized in different species and a matching pair of alkaline phosphatase (AP) and HRP-conjugated secondary antibodies (Immunologic) mixed 1:1 were applied. To detect antibody complexes red (Liquid Permanent Red; Dako) and green (VinaGreen; Biocare Medical, Concord, CA) chromogens were used according to manufacturer´s instructions. VinaGreen reaction was performed before Liquid Permanent Red. Slides were counterstained using Mayer's hematoxylin diluted 1:10 in H_2O for 1 min. Slides were washed in H_2O for 30 s after each chromogen reaction. Finally, slides were mounted with Pertex and coverslips.

Primary Antibody Validation. The primary antibody mIHC panels included a total of 30 different lymphoid, dendritic, macrophage, leukemic, immune checkpoint and immune cell activation markers selected based on their role in controlling T-cell mediated immune responses (table S3). Antibodies were used if they met antibody selection recommendations.[1,2] First, promising antibody clones were selected based

on their use in qualified publications, and their validation by distinguished antibody manufacturer, while preferring monoclonal over polyclonal antibodies. Then, the performance of potential antibodies was evaluated by their staining pattern in different immunologic and non-immunologic test tissues (lymph node, appendicitis, healthy bone marrow, different leukemic bone marrow, brain). Antibody specificity was graded based on comparison of observed intracellular (membrane, cytoplasmic, nuclear) and histological localization (cell aggregations, individual cells, extracellular) staining patterns with expected results from the literature. Finally, we selected proper dilution for each antibody to be used for mIHC panels (table S4).

Primary Antibody Denaturation Test. In order to minimize false positive signal from antibody cross-reactions during the mIHC procedure, we required that antibodies selected for mIHC must be completely denatured during the HIER step between staining rounds. Therefore, before applying in mIHC, the denaturation properties of all primary antibodies were examined by performing an additional HIER step between primary and secondary antibody incubation. Antibodies not denaturing completely were used in chromogenic staining, e.g. the last step, which does not require denaturation.

Imaging. Both 5-channel fluorescent and 3-channel brightfield images were acquired with the AxioImager.Z2 (Zeiss, Germany) microscope equipped with Zeiss Plan-Apochromat 20x objective (NA 0.8), CoolCube1 CCD camera (MetaSystems, Germany), PhotoFluor LM-75 (89 North) metal-halide light source and Zeiss EPLAX VP232-2 power supply. DAPI, FITC, Cy3, Cy5, and Cy7 filters with compatible LED light sources were used and exposure times for all fluorescence channels were optimized visually for fluorescence imaging. Scanned images were acquired and were

converted to JPEG2000 format (95% quality) for image analysis to reduce memory demand, but not significantly affecting downstream image analysis.[3]

Image Preprocessing. First, the brightfield images were deconvolved in order to separate the individual chromogen staining signals.[4] Then, using mean image of fluorescent channels and mean image of brightfield channels, spot images were registered with 2-dimensional phase correlation method.[5] For image registration, mean images were downscaled by factor of 8 and histograms of the images adjusted to match each other. The analysis was implemented in a numerical computing environment (MATLAB, MathWorks, Natick, MA, U.S.).

The staining quality was then assessed visually from gray-scale images of TMA spots. Few images were discarded due to blurred focusing or unsuccessful image registration caused mainly by air bubbles in mounting media or shattered tissue, respectively.

Statistical analysis

R packages glmnet, corrplot, survminer, survival, ggplot2, gplots, lda, factoextra, timeROC and pROC were used for statistical analysis and visualization.

Code availability. Code for statistical analyses is available upon reasonable request.

Flow Cytometry

'FC validation cohort'. Fresh diagnostic-phase BM aspirates (n=52) were aspirated in small volumes (1-2 mL) in EDTA tubes and visually examined to confirm the presence of BM fragments. BM samples were stained between 2006-2014 by the Department of Clinical Chemistry, HUCH. T-cell subsets were quantified from

200,000 cells using CD3-PE-Cy7 (SK7, BD Pharmingen), CD4-FITC (SK3, BD Pharmingen), CD8-APC-Cy7 (SK1, BD Pharmingen) antibodies. Fluorescence was measured with FACSCanto (BD Pharmingen). T-cell subgroups were quantified with FlowJo after cell debris was excluded based on its typical small forward scatter (FSC).

'*FC longitudinal cohort'*. Mononuclear cells separated using Ficoll-Paque gradient centrifugation from PB samples and BM aspirates and vitally frozen in 90% heat-inactivated fetal bovine serum and 10% dimethyl sulfoxide were used for FC. Cells were thawed by suspending them in phosphate-buffered saline (PBS) supplemented with 2 mM EDTA and 0.5% bovine serum albumin. After centrifugation (300g, 5 min), cells were resuspended in 1 ml PBS-EDTA, counted using Trypan Blue, and 0.5-1 million cells were distributed to 12x75 mm FACS tubes for staining. Cells were washed with 1 ml PBS-EDTA, centrifuged (300g, 5 min), and the supernatant was discarded by decanting. Surface-marker antibodies (table S6) were added in volumes listed in table S7 and cells were incubated for 15 min in RT covered from light. Cells were then washed with 1 ml PBS-EDTA, centrifuged (300g, 5 min), the supernatant was discarded, and cells were resuspended in 200 µl PBS-EDTA. Cells in tubes with only surface-marker stains ('T-cells 1', 'Progenitor cells') were resuspended in 200 µl PBS-EDTA and acquired on a FACSVerse flow cytometer (BD Biosciences) within 1 h from completing the staining. Cells in tubes with intracellular stains ('T-cells 2') were then suspended in 1 ml FoxP3 Fixation/Permeabilization working solution **(**4X Fixation/Permeabilization Concentrate (1 part) diluted in Fixation/Permeabilization Diluent (3 parts); FoxP3 staining buffer set, eBioscience 00-5523-00). Cells were incubated in +4 °C covered from light for 30 minutes and then washed with 2 ml Permeabilization Buffer (10X Permeabilization concentrate diluted in

deionized/distilled water; FoxP3 staining buffer set, eBioscience 00-5523-00) and centrifuged (300g, 5 min). The washing and centrifugation step was repeated, after which intracellular antibodies (table S6) were added according to table S7 and cells were incubated in +4 °C covered from light for 30 minutes. Cells were then washed (2 ml Permeabilization Buffer) and centrifuged (300g, 5 min) twice, resuspended in 200 µl PBS-EDTA and acquired on FACSVerse within 1 h from completing the staining. Gating was performed using FlowJo as shown in Supplementary Figure 7.

References

1. Baker M. Antibody anarchy: A call to order. *Nature*. 2015;527(7579):545–551.
2. Marx V. Finding the right antibody for the job. *Nat Meth*. 2013;10(8):703–707.
3. Konsti J, Lundin M, Linder N, et al. Effect of image compression and scaling on automated scoring of immunohistochemical stainings and segmentation of tumor epithelium. *Diagn. Pathol.* 2012;7(1):29.
4. Ruifrok AC, Johnston DA. Quantification of histochemical staining by color deconvolution. *Anal. Quant. Cytol. Histol.* 2001;23(4):291–299.
5. Reddy BS, Chatterji BN. An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Trans. Image Process.* 1996;5(8):1266–1271.

Table S1. Characteristics of control subjects* and indications for bone marrow biopsy.

Gender	Age	Indication for bone marrow biopsy
male	61	Thrombocytosis for many years
male	56	Thrombocytosis for many years
female	57	Thrombocytopenia for many months
male	13	Neutropenia for many months
male	60	Thrombocytosis and leukocytosis for many years
male	61	Thrombocytopenia for many years
male	44	Eosinofilia for many years
male	65	Anemia for many years
female	40	Thrombocytosis and leukocytosis for many months
female	25	Anemia and leukopenia for many years
female	44	Erythrocytosis and thrombocytosis for many years
female	47	Neutropenia and anemia for many months
male	54	Erythrocytosis for many years
female	43	Anemia and leukopenia for many years

*Controls subjects have not been diagnosed with any hematologic malignancy, chronic infection nor autoimmune disorder and are no longer under active surveillance at the Department of Hematology.

Table S2. Clinical factors.

Covariates	**Unit**
Time from treatment start to remission	days
Remission	1=yes, 0=no
1st generation TKI (imatinib)	1=yes, 0=no
2nd generation TKI (dasatinib or nilotinib)	1=yes, 0=no
Gender	0 = male, 1 = female
Age at diagnosis	years
Hemoglobin	g/L
Hematocrit	%
Basophils, blood, absolute count	E9/L
Basophils, blood, proportion of leukocytes	%
Blast, blood, proportion of leukocytes	%
Eosinophils, blood, absolute count	E9/L
Eosinophils, blood, proportion of leukocytes	%
Leukocytes, blood, absolute count	E9/L
Lymphocytes, blood, absolute count	E9/L
Lymphocytes, blood, proportion of leukocytes	%
Metamyelocytes, blood, proportion of leukocytes	%
Monocytes, blood, absolute count	E9/L
Monocytes, blood, proportion of leukocytes	%
Myeloblasts, blood, proportion of leukocytes	%
Neutrophils, blood, absolute count	E9/L

Neutrophils, blood, proportion of leukocytes	%
Thrombocytes, blood, absolute count	E9/L
Blast, bone marrow proportion of leukocytes	%
Bone marrow dysplasia	1 = yes, 0 = no
Bone marrow dysplasia in megakaryocytes	1 = yes, 0 = no
Bone marrow dysplasia in granulocytes	1 = yes, 0 = no
Ph+ chromosome	1 = yes, 0 = no
Ph- abnormalities	1 = yes, 0 = no
Ph+ abnormalities	1 = yes, 0 = no
Ph+ chromosomes with FISH	0-100%
Giemsa banding	0-100%
Sokal score	
Hasford score	
EUTOS score	
ELTS score	
Sokal score, risk class	1 = low, 2 = intermediate, 3 = high
Hasford score, risk class	1 = low, 2 = intermediate, 3 = high
EUTOS score, risk class	1 = low, 2 = high
ELTS score, risk class	1 = low, 2 = intermediate, 3 = high
Spleen size	cm
Spleen size	cm over 11 cm

Alkaline *phosphatase,* plasma	U/L
Alanine aminotransferase, plasma	U/L
Potassium, plasma	mmol/L
Creatinine, plasma	mmol/L
Lactate dehydrogenase, plasma	U/L
Sodium, plasma	mmol/L
Uric acid, plasma	umol/L
C-reactive protein, plasma	mg/mL
Body-mass index	kg/m2
Number of comorbidities	
Smoking	1 = smoking history, 0 = no smoking history

Table S3. Antibodies used in multiplex IHC.

Marker	**Manufacturer**	**Clone**	**Host species**
CD3	Thermo Scientific	EP449E	Rabbit
CD4	Abcam	EPR6855	Rabbit
CD8	BioSB	C8/144B	Mouse
CD11b	BioSB	EP45	Rabbit
CD11c	Abcam	EP1347Y	Rabbit
BDCA-1	Abcam	2F4	Mouse
BDCA-3	Abcam	EPR4051	Rabbit
CD20	BioSB	L26	Mouse
CD34	Dako	QBEnd 10	Mouse
CD56	Cell Marque	MRQ-46	Rabbit
FOXP3	Abcam	236A/E7	Mouse
CD25	Abcam	EPR6452	Rabbit
CD27	Sigma	Polyclonal	Rabbit
CD57	Sigma	VC1.1	Mouse
PD1	LsBio	PDCD1	Mouse
TIM3	Cell Signaling	D5D5R	Rabbit
OX40	Biolegend	ACT35	Mouse
HLA-DR	Abcam	TAL 1B5	Mouse
CD45RO	Abcam	UCH-L1	Mouse
Granzyme B	Novocastra	11F1	Mouse
PDL1	Cell Signaling	E1L3N	Rabbit
PDL2	Sigma	Polyclonal	Rabbit

HLA ABC	MBL International Corporation	EMR8-5	Mouse
HLA G	Santa Cruz	4H84	Mouse
CD68	Abcam	KP1	Mouse
pSTAT1 (S727)	Cell Signaling	D3B7	Rabbit
cMAF	Abcam	EPR16484	Rabbit
CD33	LsBio	LS-C338084	Mouse
LAG3	Abcam	EPR4329	Rabbit
CTLA4	Santa Cruz	F-8	Mouse

Table S4. Multiplex IHC antibody panels and antibody dilutions.

LPR, liquid permanent red. VG, vina green. GrB, granzyme B. MDSCs, myeloid-derived suppressor cells.

Panels/Detection probes	**GFP**	**Cy3**	**Cy5**	**Cy7**	**LPR**	**VG**
T cell activity	GrB 1:100	CD57 1:400	CD8 1:25	CD4 1:25		CD3 1:250
Memory T-cells	CD27 1:500	CD25 1:25	CD8 1:25	CD4 1:25		CD3 1:250
Immune checkpoints 1	PD-1 1:1500	TIM-3 1:2500	CD8 1:25	CD4 1:25	FOXP3 1:150	CD3 1:250
Immune checkpoints 2	LAG-3 1:150	CTLA-4 1:150	CD8 1:25	CD4 1:25		CD3 1:250
Immune checkpoints 3	PD-1 1:1500	OX40 1:25	CD8 1:25	CD4 1:25	CD45RO 1:250	CD3 1:250
Cancer cell ligands	HLA G 1:25	PD-L1 1:50	TIM-3 1:100	HLA-ABC 1:100	PD-L2 1:250	CD34 1:100
B-cells, NK-cells and macrophages	CD56 1:300	pSTAT1 1:100	CD3 1:25	CD20 1:25	CMAF 1:150	CD68 1:250
Dendritic cells and MDSCs	CD11b 1:250	CD33 1:200	BDCA-3 1:25	HLA-DR 1:200	CD11c 1:250	BDCA-1 1:100

Table S5. Covariates used in the development of the CMLi model

Covariates	**HR (95% CI for HR)**	**P-value**
2nd generation TKI (dasatinib or nilotinib) [1=yes, 0=no]	1.65 (0.792-3.43)	0.177
Giemsa banding [0-100%]	0.991 (0.979-1)	0.0888
Age at diagnosis [years]	1.01 (0.993-1.04)	0.204
Basophils, blood, absolute count [E9/L]	1 (0.951-1.06)	0.892
Basophils, blood, proportion of leukocytes [%]	1.09 (1.02-1.17)	0.00658
Blast, bone marrow proportion of leukocytes [%]	1.13 (0.966-1.32)	0.126
Bone marrow dysplasia [1 = yes, 0 = no]	0.543 (0.193-1.53)	0.241
Bone marrow dysplasia in megakaryocytes [1 = yes, 0 = no]	0.562 (0.173-1.83)	0.333
Bone marrow dysplasia in granulocytes [1 = yes, 0 = no]	0.553 (0.0754-4.06)	0.555
Ph+ chromosome [1 = yes, 0 = no]	0.664 (0.236-1.87)	0.434
Ph- abnormalities [1 = yes, 0 = no]	0.612 (0.259-1.45)	0.26
Ph+ abnormalities [1 = yes, 0 = no]	0.634 (0.268-1.5)	0.297
Ph+ chromosomes with FISH [0-100%]	0.927 (0.887-0.968)	5.99E-05
Spleen size [cm over 11 cm]	0.955 (0.895-1.02)	0.156
Sokal score	0.947 (0.733-1.22)	0.679
Hasford score	1 (1-1)	0.39
EUTOS score	1.01 (0.996-1.01)	0.297
ELTS score	0.704 (0.433-1.15)	0.158
Sokal score, risk class [1 = low, 2 = intermediate, 3 =	1.21 (0.83-1.77)	0.317

high]		
Hasford score, risk class [1 = low, 2 = intermediate, 3 = high]	1.19 (0.802-1.78)	0.38
EUTOS score, risk class [1 = low, 2 = high]	2.08 (0.863-5.03)	0.0951
ELTS score, risk class [1 = low, 2 = intermediate, 3 = high]	0.872 (0.561-1.36)	0.543
Spleen size [cm]	0.95 (0.89-1.01)	0.114
Blast, blood, proportion of leukocytes [%]	0.953 (0.846-1.07)	0.418
Eosinophils, blood, absolute count [E9/L]	0.986 (0.927-1.05)	0.654
Eosinophils, blood, proportion of leukocytes [%]	1.11 (0.998-1.22)	0.0537
Blood hemoglobin [g/L]	1.01 (0.991-1.02)	0.476
Blood hematocrit [%]	1.01 (0.965-1.06)	0.608
Leukocytes, blood, absolute count [E9/L]	0.996 (0.992-1)	0.0239
Lymphocytes, blood, absolute count [E9/L]	0.921 (0.841-1.01)	0.0697
Lymphocytes, blood, proportion of leukocytes [%]	1.03 (0.967-1.09)	0.4
Metamyelocytes, blood, proportion of leukocytes [%]	0.956 (0.886-1.03)	0.238
Monocytes, blood, absolute count [E9/L]	0.904 (0.814-1)	0.0562
Monocytes, blood, proportion of leukocytes [%]	0.953 (0.846-1.07)	0.422
Myeloblasts, blood, proportion of leukocytes [%]	0.984 (0.949-1.02)	0.385
Neutrophils, blood, absolute count [E9/L]	0.992 (0.986-0.999)	0.0161
Neutrophils, blood, proportion of leukocytes [%]	1 (0.973-1.03)	0.968
Thrombocytes, blood, absolute count [E9/L]	1 (1-1)	0.0024
Alkaline phosphatase, plasma [U/L]	0.999 (0.993-1.01)	0.838
Alanine aminotransferase, plasma [U/L]	0.991 (0.969-1.01)	0.402
Potassium, plasma [mmol/L]	2.44 (1.04-5.73)	0.0407

Creatinine, plasma [mmol/L]	1 (0.982-1.02)	0.806
Lactate dehydrogenase, plasma [U/L]	0.999 (0.998-1)	0.113
Sodium, plasma [mmol/L]	1.01 (0.914-1.12)	0.842
Uric acid, plasma [umol/L]	1 (0.997-1)	0.747
C-reactive protein, plasma [mg/mL]	0.988 (0.96-1.02)	0.393
Body-mass index [kg/m2]	1.02 (0.955-1.09)	0.586
Number of comorbidities	1.25 (0.954-1.63)	0.103
Deletion of 9q (cytogenetics)	0.768 (0.185-3.19)	0.715
Smoking [1 = smoking history, 0 = no smoking history]	0.994 (0.556-1.78)	0.983
CD3+CD4+/CD27+CD25+ [%]	1 (0.888-1.13)	0.958
CD3+CD4+/CD27+CD25- [%]	0.859 (0.625-1.18)	0.345
CD3+CD4+/CD27-CD25+ [%]	1.01 (0.971-1.06)	0.555
CD3+CD4+/CD27-CD25- [%]	0.993 (0.958-1.03)	0.685
CD3+CD8+/CD27+CD25+ [%]	1.01 (0.943-1.09)	0.696
CD3+CD8+/CD27+CD25- [%]	0.974 (0.771-1.23)	0.825
CD3+CD8+/CD27-CD25+ [%]	1 (0.976-1.03)	0.988
CD3+CD8+/CD27-CD25- [%]	0.997 (0.976-1.02)	0.745
CD3+CD4+/CD27+ [%]	0.972 (0.794-1.19)	0.782
CD3+CD4+/CD25+ [%]	1.02 (0.949-1.09)	0.604
CD3+CD8+/CD27+ [%]	1.02 (0.889-1.17)	0.773
CD3+CD8+/CD25+ [%]	1 (0.959-1.05)	0.899
CD11c+ [%]	0.995 (0.948-1.04)	0.825
CD11c+BDCA1+ [%]	4.64 (0.0148-1450)	0.601
CD11c+BDCA3+ [%]	463 (0.26-827000)	0.109
CD11c+BDCA1+ MHCIIhigh [%]	0.774 (0.121-4.96)	0.787

CD11c+BDCA3+ MHCIIhigh [%]	0.756 (0.122-4.69)	0.764
CD11c+BDCA1+/CD11b+ [%]	0.712 (0.116-4.38)	0.714
CD11c+BDCA3+/CD11b+ [%]	1.74 (0.17-17.7)	0.641
CD11b+CD33+HLADR- [%]	0.832 (0.671-1.03)	0.0928
CD3+CD4+/GrB+CD57+ [%]	0.973 (0.915-1.03)	0.376
CD3+CD4+/GrB+CD57- [%]	1 (0.943-1.06)	0.967
CD3+CD4+/GrB-CD57+ [%]	0.979 (0.948-1.01)	0.201
CD3+CD4+/GrB-CD57- [%]	1.02 (0.994-1.05)	0.135
CD3+CD8+/GrB+CD57+ [%]	1 (0.922-1.09)	0.922
CD3+CD8+/GrB+CD57- [%]	0.978 (0.931-1.03)	0.368
CD3+CD8+/GrB-CD57+ [%]	0.979 (0.933-1.03)	0.389
CD3+CD8+/GrB-CD57- [%]	1.02 (0.988-1.05)	0.255
CD3+CD4+/GrB+ [%]	0.989 (0.952-1.03)	0.585
CD3+CD4+/CD57+ [%]	0.98 (0.955-1.01)	0.133
CD3+CD8+/GrB+ [%]	0.989 (0.954-1.02)	0.538
CD3+CD8+/CD57+ [%]	0.987 (0.949-1.03)	0.502
CD3+CD4+/CTLA4+LAG3+ [%]	0.998 (0.955-1.04)	0.917
CD3+CD4+/CTLA4+LAG3- [%]	0.99 (0.965-1.02)	0.456
CD3+CD4+/CTLA4-LAG3+ [%]	1.03 (0.998-1.07)	0.0577
CD3+CD4+/CTLA4-LAG3- [%]	0.999 (0.972-1.03)	0.92
CD3+CD8+/CTLA4+LAG3+ [%]	1 (0.967-1.03)	0.992
CD3+CD8+/CTLA4+LAG3- [%]	0.983 (0.959-1.01)	0.174
CD3+CD8+/CTLA4-LAG3+ [%]	1.04 (0.994-1.1)	0.073
CD3+CD8+/CTLA4-LAG3- [%]	1.01 (0.983-1.04)	0.408
CD3+CD4+/CTLA4+ [%]	0.989 (0.964-1.02)	0.414

CD3+CD8+/CTLA4+ [%]	0.977 (0.95-1.01)	0.114
CD3+CD4+/LAG3+ [%]	1.01 (0.987-1.04)	0.335
CD3+CD8+/LAG3+ [%]	1.01 (0.984-1.03)	0.517
CD3-CD56+ [%]	0.98 (0.922-1.04)	0.522
CD3-CD56+/pSTAT1+ [%]	1.03 (1.01-1.06)	0.0136
CD68+CD3- [%]	0.975 (0.95-1)	0.0568
CD68+CD3-/pSTAT1-cMAF+ [%]	0.957 (0.886-1.03)	0.262
CD68+CD3-/pSTAT1+cMAF+ [%]	1.15 (0.948-1.39)	0.156
CD68+CD3-/pSTAT1+cMAF- [%]	1.04 (0.994-1.09)	0.0832
CD68+CD3-/pSTAT1-cMAF- [%]	0.982 (0.945-1.02)	0.372
CD68+CD3-/pSTAT1+ [%]	1.03 (0.995-1.07)	0.0842
CD68+CD3-/cMAF+ [%]	0.98 (0.917-1.05)	0.561
HLAABC+/All cells [%]	1.01 (0.973-1.04)	0.651
HLAG+/All cells [%]	1.01 (0.983-1.05)	0.371
PDL1+/All cells [%]	1.01 (0.903-1.13)	0.836
PDL2+/All cells [%]	1.15 (0.943-1.39)	0.172
HLAABC+/CD34+ cells [%]	1.03 (1-1.06)	0.0384
HLAG+/CD34+ cells [%]	1.1 (0.899-1.34)	0.355
PDL1+/CD34+ cells [%]	2.73 (0.0753-98.9)	0.581
PDL2+/CD34+ cells [%]	0.714 (0.231-2.21)	0.559
CD3+CD4+/PD1+TIM3+ [%]	1.02 (0.826-1.25)	0.879
CD3+CD4+/PD1+TIM3- [%]	0.951 (0.909-0.995)	0.0284
CD3+CD4+/PD1-TIM3+ [%]	1.03 (0.986-1.07)	0.203
CD3+CD4+/PD1-TIM3- [%]	1.01 (0.985-1.04)	0.348
CD3+CD8+/PD1+TIM3+ [%]	0.999 (0.838-1.19)	0.992

CD3+CD8+/PD1+TIM3- [%]	0.941 (0.896-0.988)	0.0152
CD3+CD8+/PD1-TIM3+ [%]	1.07 (0.993-1.14)	0.0775
CD3+CD8+/PD1-TIM3- [%]	1.03 (0.99-1.07)	0.148
CD3+CD4+/PD1+ [%]	0.961 (0.924-1)	0.0488
CD3+CD8+/PD1+ [%]	0.947 (0.905-0.992)	0.0215
CD3+CD4+/TIM3+ [%]	1.02 (0.985-1.06)	0.239
CD3+CD8+/TIM3+ [%]	1.05 (0.985-1.11)	0.144
CD3-CD20+ [%]	1.01 (0.947-1.07)	0.799
CD3-CD20+/pSTAT1+ [%]	1.01 (0.992-1.03)	0.249
CD3+CD4+/PD1+OX40+ [%]	0.869 (0.774-0.975)	0.0157
CD3+CD4+/PD1+OX40- [%]	0.969 (0.928-1.01)	0.154
CD3+CD4+/PD1-OX40+ [%]	1.01 (0.878-1.15)	0.926
CD3+CD4+/PD1-OX40- [%]	1.04 (0.999-1.07)	0.0539
CD3+CD8+/PD1+OX40+ [%]	0.92 (0.823-1.03)	0.141
CD3+CD8+/PD1+OX40- [%]	0.961 (0.917-1.01)	0.0941
CD3+CD8+/PD1-OX40+ [%]	1.09 (0.958-1.24)	0.188
CD3+CD8+/PD1-OX40- [%]	1.03 (0.992-1.07)	0.134
CD3+CD4+/PD1+ [%]	0.966 (0.932-1)	0.0594
CD3+CD8+/PD1+ [%]	0.966 (0.931-1)	0.0716
CD3+CD4+/OX40+ [%]	0.928 (0.852-1.01)	0.0835
CD3+CD8+/OX40+ [%]	0.985 (0.91-1.07)	0.709
CD3+CD4+/CD45RO+ [%]	1.03 (1-1.05)	0.022
CD3+CD8+/CD45RO+ [%]	1.03 (1-1.06)	0.0264
CD3+ [%]	1.12 (1.01-1.23)	0.0265
CD3+CD4+ [%]	1.89 (1.24-2.87)	0.00244

CD3+CD8+ [%]	1.25 (0.95-1.65)	0.109
CD3+CD4+FOXP3+ [%]	2860 (0.809-10100000)	0.0542

HR = hazard ratio, CI = confidence interval

Table S6. Antibodies used in flow cytometry in '*FC longitudinal cohort*'.

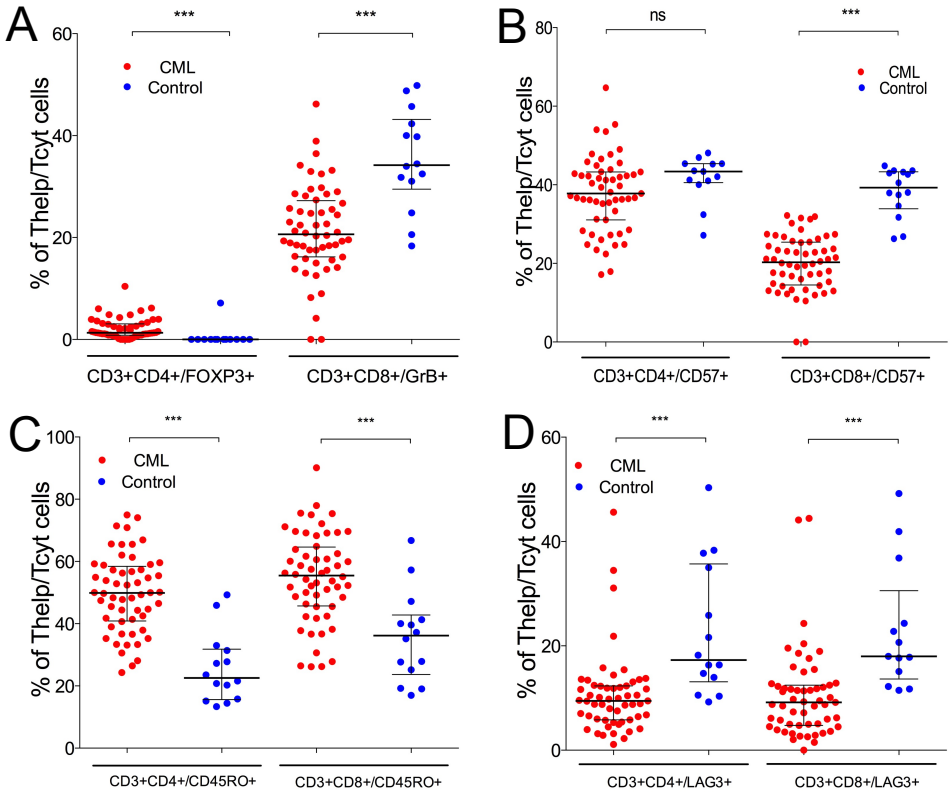
Marker	**Marker type**	**Manufacturer**	**Clone**	**Host species**
CCR7	Surface	RnD	FAB197P	Mouse
CD3	Surface	BD Biosciences	SK7 (PerCP-Cy5.5)	Mouse
CD3	Surface	BD Biosciences	UCHT-1 (AF700)	Mouse
CD4	Surface	BD Biosciences	SK3	Mouse
CD8	Surface	BD Biosciences	SK1	Mouse
CD14	Surface	Invitrogen	MHCD1428	Mouse
CD25	Surface	BD Biosciences	2A3	Mouse
CD34	Surface	BD Biosciences	8G12	Mouse
CD38	Surface	BD Biosciences	HB-7	Mouse
CD45	Surface	BD Biosciences	2D1	Mouse
CD80	Surface	BD Biosciences	L307.4	Mouse
CD86	Surface	BD Biosciences	2331(FUN-1)	Mouse
CD45RA	Surface	BD Biosciences	HI100	Mouse
CXCR4	Surface	BD Biosciences	12G5	Mouse
FOXP3	Intracellular	eBioscience	2362/E7	Mouse
Ki-67	Intracellular	BD Biosciences	B56	Mouse
LAG3	Surface	eBioscience	3DS223H	Mouse
PD1	Surface	BD Biosciences	EH12.1	Mouse
PDL1	Surface	BD Biosciences	MIH1	Mouse
PDL2	Surface	BD Biosciences	MIH18	Mouse
TIM3	Surface	BioLegend	F38-2E2	Mouse

Table S7. Flow cytometry antibody panels and added antibody volumes in

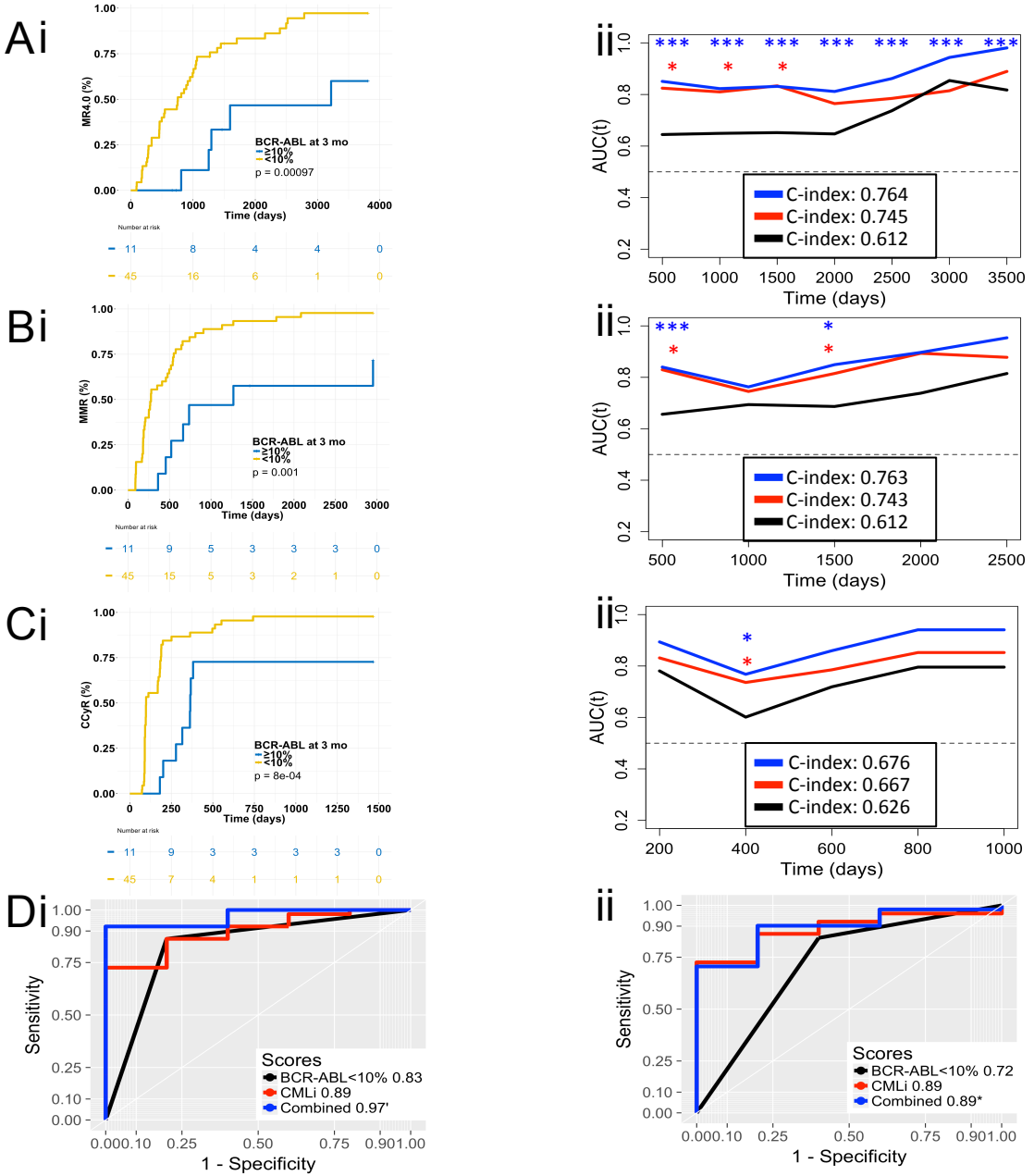
Panels/Channels	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
T-cells 1	CD8 FITC 4 µl	CCR7 PE 10 µl	CD3 PerCP-Cy5.5 5 µl	TIM3 PE-Cy7 3 µl	PD1 AF647 0.5 µl	CD45RA AF700 5 µl	CD4 BV510 3 µl	CXCR4 BV421 5 µl
T-cells 2	CD8 FITC 4 µl	CD25 PE 10 µl	LAG3 PerCP-eF710 5 µl	FOXP3 PE-Cy7 5 µl	PD1 AF647 0.5 µl	CD3 AF700 2 µl	CD4 BV510 3 µl	Ki-67 BV421 3 µl
Progenitor cells	CD34 FITC 10 µl	CD38 PE 5 µl	CD45 PerCP 5 µl	PDL1 PE-Cy7 10 µl	PDL2 APC 10 µl	CD86 AF700 3 µl	CD80 BV510 5 µl	CD14 PB 2 µl

'FC longitudinal cohort'.

AF = Alexa Fluor, BV = Brilliant Violet, PB = Pacific Blue

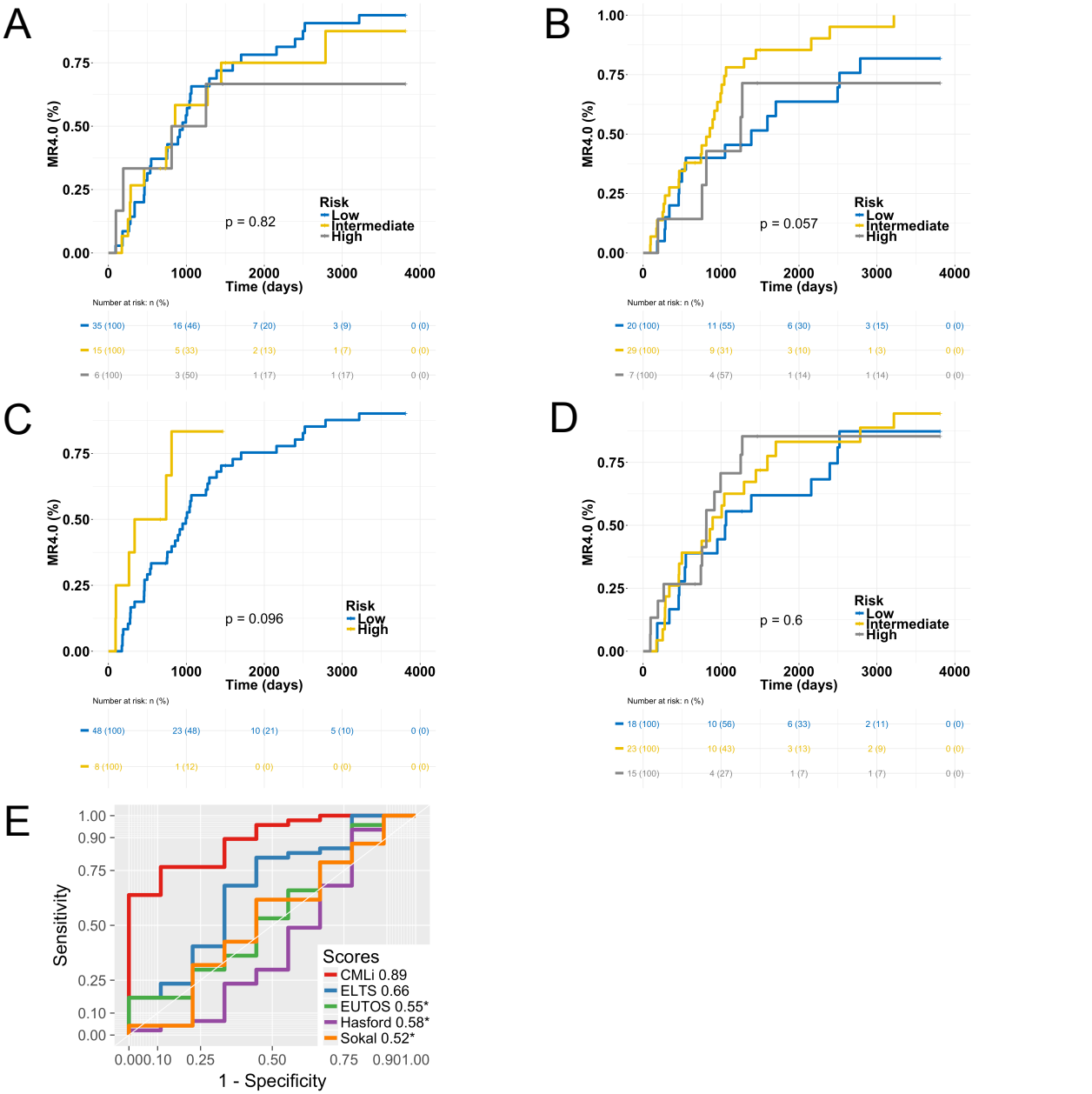


Supplementary Fig 1. T cell phenotype in chronic myeloid leukemia (CML) vs. control bone marrow (BM). Levels of (A) CD3+CD4+/ FOXP3+ regulatory T cells and granzyme expression in CD8+ T cells, (B) expression of CD57, (C) CD45RO, and (D) LAG3 in both T cell subsets were compared with Mann-Whitney U test after outlier exclusion (Grubb's test, alpha=0.20), and p-values adjusted with Benjamini-Hochberg procedure to reduce false discovery rate (q- values), and results displayed with scatter plots. Bars show median with interquartile range. * <0.05, ** <0.01, ***<0.001



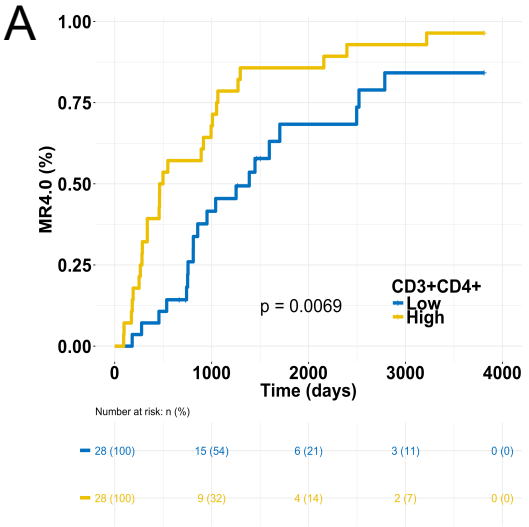
Supplementary Fig 2. Survival analysis of the '*IHC cohort*' stratified by the BCR-ABL level <10% international scale (IS) at 3 months (A) according to MR4.0, (B) MMR and (C) CCyR. Figures (i) display Cox regression analysis (log-rank test), and figures (ii) their time-dependent ROC curves (IPCW approach) and C-statistic values in relation to the CMLi model and the combination of both models. (D) ROC curves for all three models (Bootstrap method, n=4000). The ROC curve for MR4.0 is showed in Fig. 4D. '<0.10, *<0.05, **<0.01, ***<0.001

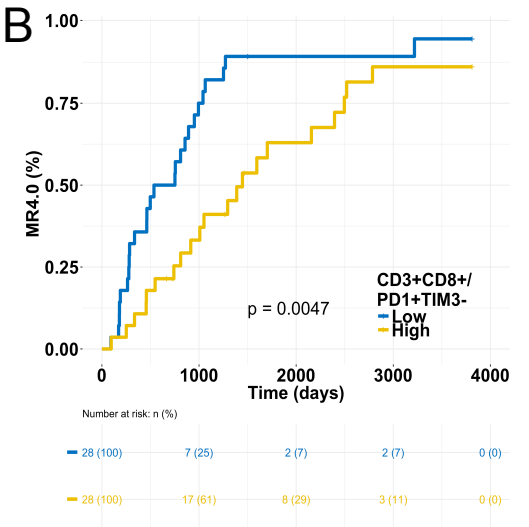
MR4.0, Molecular response 4.0. MMR, Major molecular response. CCyR, Complete cytogenic response. IPCW, inverse probability of censoring weighting.

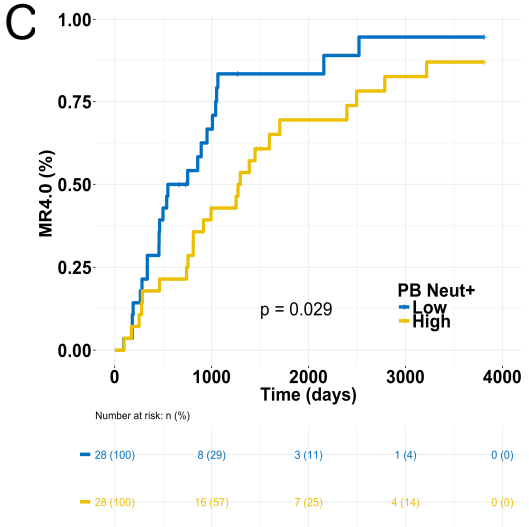


Supplementary Fig 3. Univariate Cox regression analysis (log-rank test) of the '*IHC cohort*' stratified by the (A) Sokal, (B) Hasford, (C) EUTOS, and (D) ELTS risk scores. The clinically-used risk scores are used in current treatment considerations. They have been developed to assess patient prognosis, but not to predict MR4.0. (E) The MR4.0 prediction power of the CMLi model to clinically-used risk scores was measured with area under the receiver operating characteristic curve (AUROC; bootstrap method). (*) CMLi outperform compared risk score ($p<0.05$).

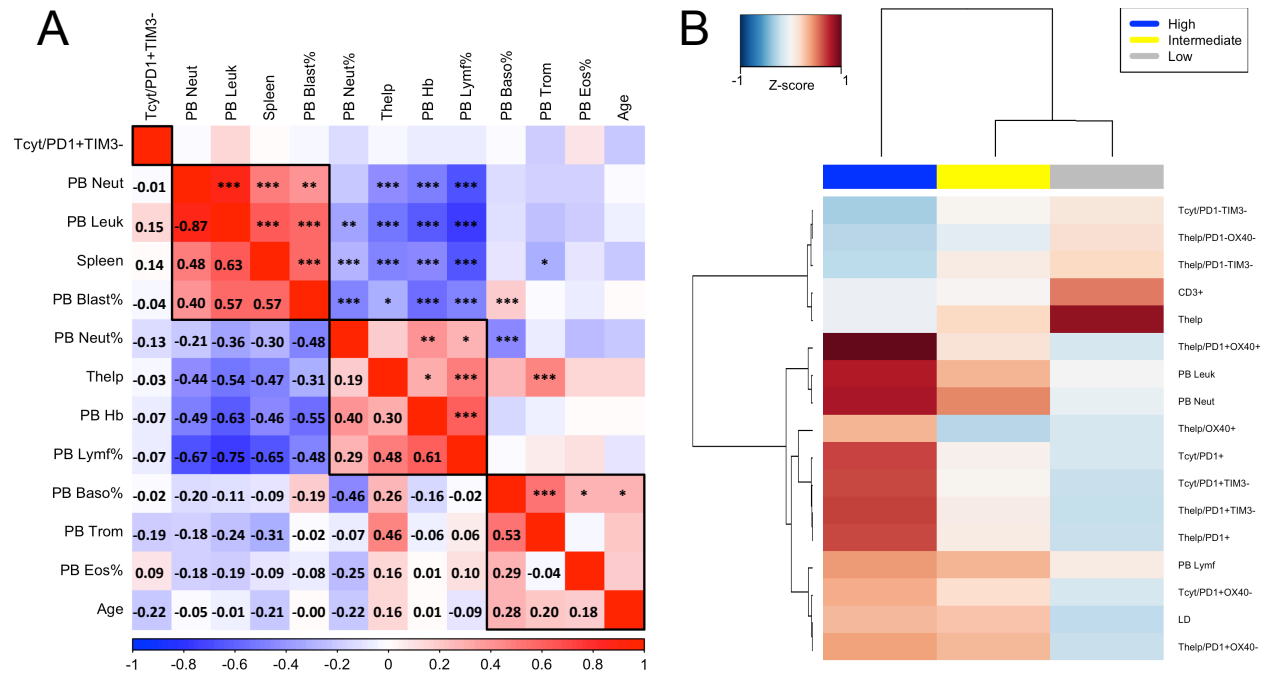
MR4.0, Molecular response 4.0.



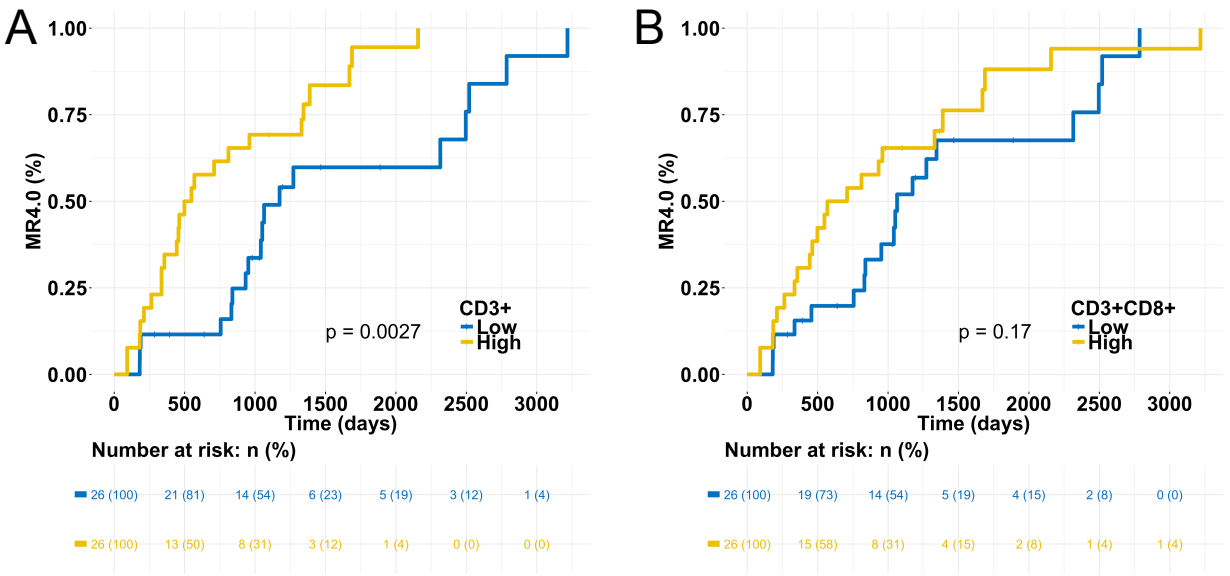




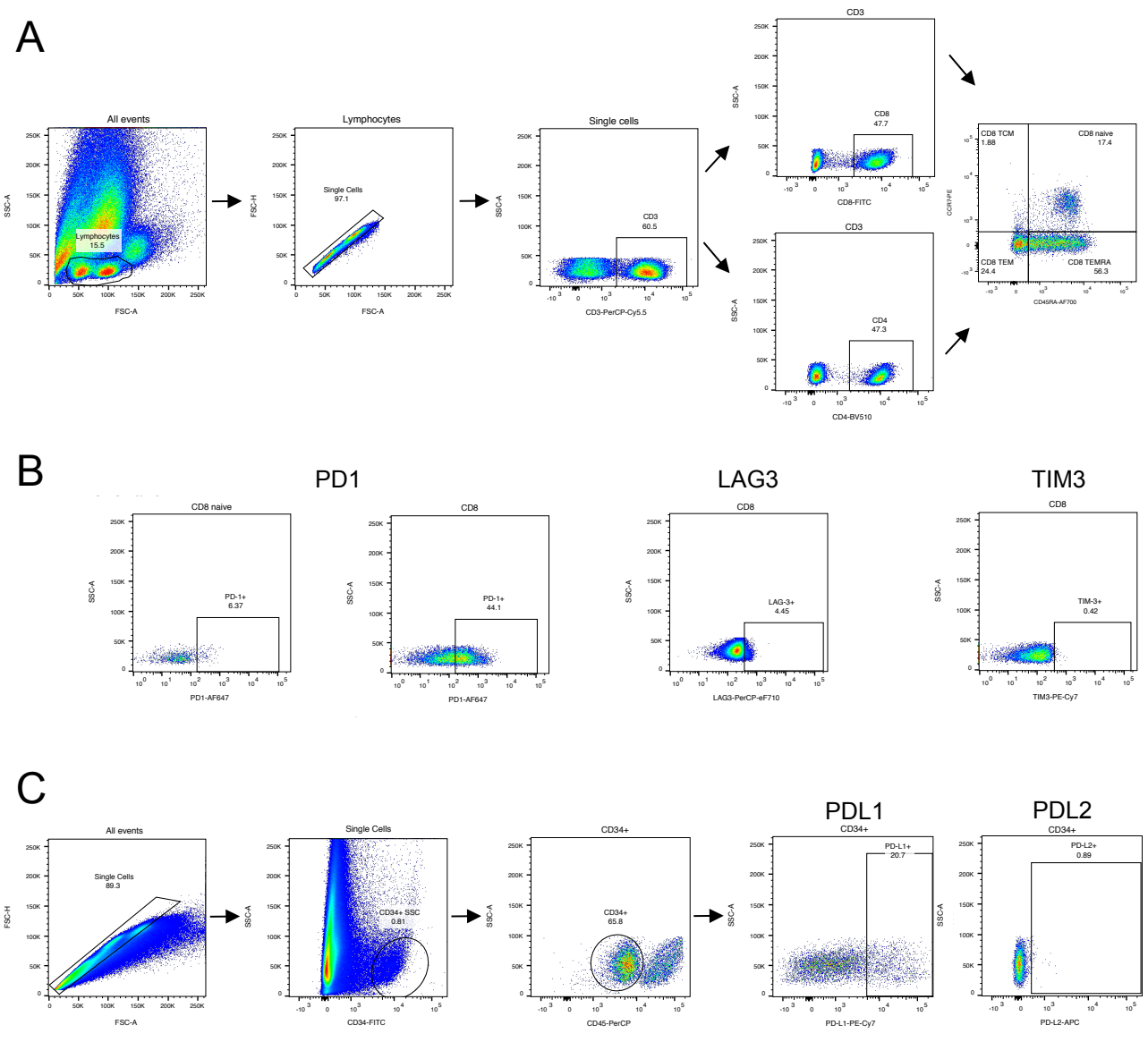
Supplementary Fig 4. Prediction of MR4.0 by CMLi covariates. Univariate Cox regression analysis (log-rank test) of molecular response 4.0 (MR4.0) predicted by (A) bone marrow (BM) CD4+ T cells, (B) proportion of BM PD1+TIM3- CD8+ T cells, and (C) peripheral blood (PB) neutrophils categorized by median of the '*IHC cohort*'.



Supplementary Fig 5. Correlation plot and clustering analysis of the CMLi score. (A) Correlation matrix (Spearman correlation) of covariates present in the CMLi model, Sokal, Hasford, EUTOS and ELTS risk scores and components of the complete blood count. Correlation factors are presented in the lower panel and significant correlations marked with asterix. (B) Continuous clinicoimmunological variables were compared in each risk groups (high, intermediate, low) of the CMLi model with Kruskal-Wallis and p-values adjusted with Benjamini&Hochberg procedure. Significant features (q<0.05) were z-scored and organized by hierarchical clustering using Spearman correlation distance and Ward linkage (ward.D2) method. * <0.05, ** <0.01, ***<0.001. LD, lactate dehydrogenase. PB, peripheral blood.



Supplementary Fig 6. Prediction of MR4.0 by BM T cells and CD8+ T cells of the '*FC validation cohort*'. Univariate Cox regression analysis (log-rank test) of molecular response 4.0 (MR4.0) predicted by the proportion of (A) total T cell and (B) CD8+ T cell count from all BM cells and categorized by median of the '*FC validation cohort*' (n=52).



Supplementary Fig 7. Flow cytometry gating strategies in the *'FC longitudinal cohort*'. (A) Gating of T-cell subgroups in tubes 'T-cells 1' and 'T-cells 2' (Supplementary Information). (B) Gating of cells positive for immune checkpoint receptors on T-cell subgroups in tubes 'T-cells 1' and 'T-cells 2' (Supplementary Methods). PD1+ gate was set at the edge of the negative population in naïve (CD45RA+CCR7+) T-cells expressing low levels of PD1. (C) Gating of CD34+ progenitor cells positive for PD1 ligands.