

LETTER OPEN



MINIMAL RESIDUAL DISEASE

Computational measurable residual disease assessment in acute myeloid leukemia: a retrospective validation in the HOVON-SAKK-132 trial

Tim R. Mocking^{1,2}, Lukas H. Haaksma^{1,2}, Tom Reuvekamp^{1,2}, Angèle Kelder^{1,2}, Willemijn J. Scholten^{1,2}, Lok Lam Ngai^{1,2}, Dimitri A. Breems³, Thomas Fischer⁴, Bjørn T. Gjertsen⁵, Laimonas Griškevičius⁶, Gunnar Juliusson⁷, Johan A. Maertens⁸, Markus G. Manz^{9,10}, Thomas Pabst^{10,11}, Jakob R. Passweg^{10,12}, Kimmo Porkka¹³, Peter J. M. Valk¹⁴, Patrycja Gradowska^{14,15}, Bob Löwenberg¹⁴, David C. de Leeuw^{1,2}, Jeroen J. W. M. Janssen^{1,2,16}, Gert J. Ossenkoppele^{1,2}, Arjan A. van de Loosdrecht^{1,2}, Jacqueline Cloos^{1,2} and Costa Bachas^{1,2}✉

© The Author(s) 2025

Leukemia (2025) 39:2559–2562; <https://doi.org/10.1038/s41375-025-02747-8>

TO THE EDITOR:

European Leukemia Net (ELN) guidelines recommend flow cytometry for widely applicable (~90% of patients) measurable residual disease (MRD) assessment in acute myeloid leukemia (AML) [1]. This immunophenotypic MRD is assessed by manual gating (mgMRD), which identifies leukemic cells based on manual inspection of two-dimensional plots. However, this process is time-consuming (>30 min per sample) and requires extensive standardization for reproducibility, particularly outside specialized laboratories [2, 3]. The growing number of markers that can be evaluated together with spectral cytometry platforms exacerbates the analytical complexity, rendering mgMRD increasingly impractical [3].

To address these issues, we previously developed a fully automated (~3 s) computational MRD (cMRD) pipeline that can identify leukemic blasts in flow cytometry data using interpretable machine learning [4]. This algorithm first automatically detects healthy and leukemic blasts, after which a statistical model detects and enumerates the cells with aberrant marker expression (Fig. 1a). To assess the prognostic relevance of cMRD, we retrospectively analyzed the HOVON-SAKK-132 trial [5], which prospectively evaluated MRD-guided consolidation therapy using a standardized four-tube eight-color assay [6]. We included all flow MRD measurements on bone marrow (BM) from AML patients in remission after two cycles of induction chemotherapy, with measurements carried out at the central laboratory at Amsterdam UMC according to ELN guidelines [7]. If multiple MRD measurements were available from the same patient, the first measurement after completion of induction therapy was used.

We included 399 patients (Table S1) and determined the cMRD % for each patient using the pipeline. Leukemic burden (defined

as % of white blood cells, WBCs) assessed by cMRD and mgMRD was significantly correlated ($r_s = 0.55$, $p < 0.001$) (Fig. 1b). To investigate the independent prognostic value of cMRD% (independent of cut-offs), we performed a multivariable analysis adjusted for AML type, WBC count at diagnosis, ELN risk group (2017 and 2022), age, and a time-dependent covariable for consolidation treatment (Table S2–S5). This analysis confirmed cMRD as an independent prognostic factor for overall survival (OS) and relapse-free survival (RFS).

To define cMRD-positivity, the maximally selected rank statistic was used to select a cut-off based on patient outcomes (RFS). We consistently identified two robust prognostic cMRD cut-offs (0.1% and 0.56%) in 1,000 different permutations of the cohort (Fig. S1, Table S6). Although the conventional 0.1% mgMRD cut-off was prognostic for cMRD (Fig. S2) the higher number of predicted leukemic cells in cMRD compared to mgMRD (Fig. 1b) led to a high number of cMRD+ patients (63.7%, 254/399) and lower concordance with mgMRD status (52.4%, 209/399). Elevated cMRD levels (Fig. 1b) may result from false positive misclassifications at the cell level. However, they can also reflect a technical difference: while the cMRD pipeline estimates the total leukemic burden, manual gating generally reports the percentage of the dominant leukemic population using two-dimensional analysis, which may underestimate the overall leukemic burden. Based on these clinical and technical considerations, we decided to classify patients as cMRD+ using the 0.56% cut-off. With this cut-off, 12.3% (49/399) of patients were cMRD+, compared to 17.5% (70/399) mgMRD+ patients using the 0.1% cut-off. mgMRD and cMRD status were concordant for 85.2% (340/399) of patients. OS, RFS, and cumulative incidence of relapse (CIR) were comparable between mgMRD and cMRD negative groups (Fig. 1c–e). cMRD-

¹Amsterdam UMC, Amsterdam, The Netherlands. ²Cancer Center Amsterdam, Amsterdam, The Netherlands. ³Ziekenhuis aan de Stroom, Antwerpen, Belgium. ⁴Otto von Guericke University Hospital Magdeburg, Magdeburg, Germany. ⁵Haukeland University Hospital, Bergen, Norway. ⁶Vilnius University Hospital Santaros Klinikos and Vilnius University, Vilnius, Lithuania. ⁷Skanes University Hospital, Lund, Sweden. ⁸University Hospital Gasthuisberg, Leuven, Belgium. ⁹University Hospital, Zurich, Switzerland. ¹⁰Swiss Group for Clinical Cancer Research, Bern, Switzerland. ¹¹University Hospital Inselspital, Bern, Switzerland. ¹²University Hospital, Basel, Switzerland. ¹³Helsinki University Hospital Cancer Center, Helsinki, Finland. ¹⁴Erasmus Medical Center Cancer Institute, Rotterdam, The Netherlands. ¹⁵HOVON Foundation, Rotterdam, The Netherlands. ¹⁶Radboud University Medical Center, Nijmegen, The Netherlands. ✉email: c.bachas@amsterdamumc.nl

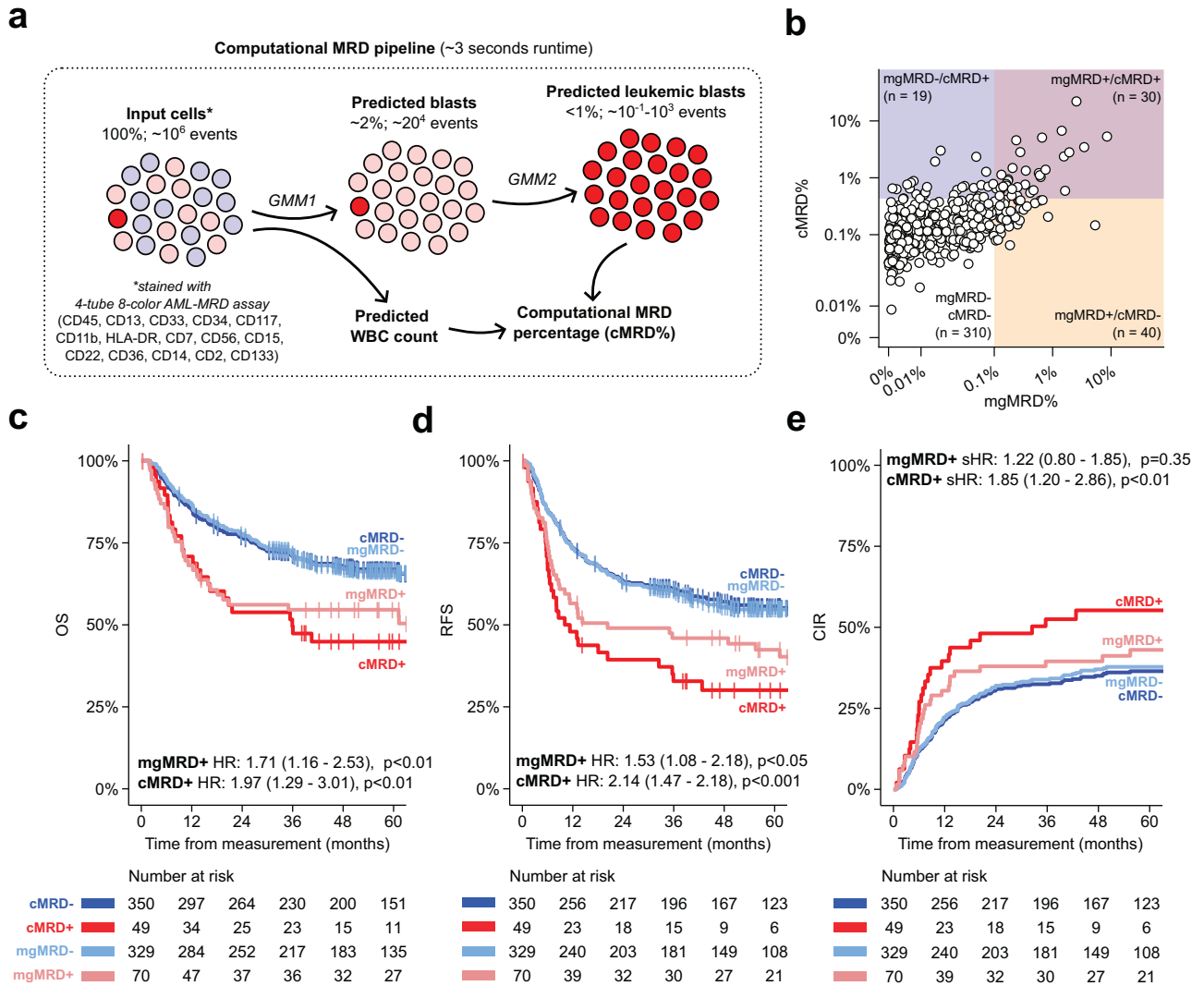


Fig. 1 Prognostic value of computational MRD assessment (cMRD) in AML compared with manual gating MRD (mgMRD). **a** Schematic overview of cMRD pipeline. Each cell is sequentially processed by two Gaussian mixture models (GMM1, GMM2) that subset blasts and leukemic cells respectively. **b** Association between MRD% evaluated by mgMRD vs. cMRD. Kaplan Meier curves for OS (**c**) RFS (**d**), and the CIR curve (**e**). Hazard ratios were calculated by univariate Cox proportional hazards (OS, RFS) or by univariate proportional subdistribution hazards (CIR). P-values were derived from Wald tests based on regression coefficients. Confidence intervals are based on 2.5% and 97.5% percentiles. MRD measurable residual disease, AML acute myeloid leukemia, HR hazard ratio, OS overall survival, RFS relapse-free survival, CIR cumulative incidence of relapse.

positivity was associated with shorter OS (HR (95% CI): 1.97 (1.16–2.53), $p < 0.01$; Fig. 1c) and RFS (HR (95% CI): 2.14 (1.29–3.01), $p < 0.001$; Fig. 1d) compared to cMRD-negativity. CIR was significantly higher for cMRD+ patients compared to cMRD-patients (sHR (95% CI): 1.85 (1.20–2.86), $p < 0.01$; Fig. 1e), whereas this effect was absent for mgMRD (sHR (95% CI): 1.22 (0.80–1.85), $p = 0.35$; Fig. 1e). Prognostic differences for cMRD in RFS were found in both intermediate and adverse ELN2017 (Fig. S3) and ELN2022 (Fig. S4) risk groups.

To understand the differences between manual and computational MRD assessment, we investigated the cases with discordant MRD status. All discordant cases were independently re-examined by two manual gating experts in two rounds. In the first round, leukemic cells were gated according to current standards and checked by the second operator. Both operators were blinded to previous analyses and MRD status. In the second round, cMRD

output was added, allowing for a side-by-side comparison of manual and computational cell classifications.

The largest discordant-group (mgMRD+/cMRD-, $n = 40$) had a 5-year RFS of 54%, which was not significantly different from mgMRD-/cMRD- patients (5-year RFS: 56%; log-rank: $p = 0.6$; Fig. 2a–c). However, their RFS was significantly longer than mgMRD+/cMRD+ patients (5-year RFS: 27%; log-rank: $p < 0.05$), suggesting potential false-positive mgMRD results. This was confirmed by manual re-analysis, in which 10 out of 40 samples initially classified as mgMRD+ were now considered mgMRD- with current gating procedures. This difference originated from specific expression patterns (e.g., CD15, CD22) now recognized as transient phenotypes associated with BM regeneration. As the computational pipeline identified and modeled clusters with these cells in our regenerating BM reference cohort [4], the cMRD algorithm did not classify these cells as aberrant (Fig. 2d, e). Other

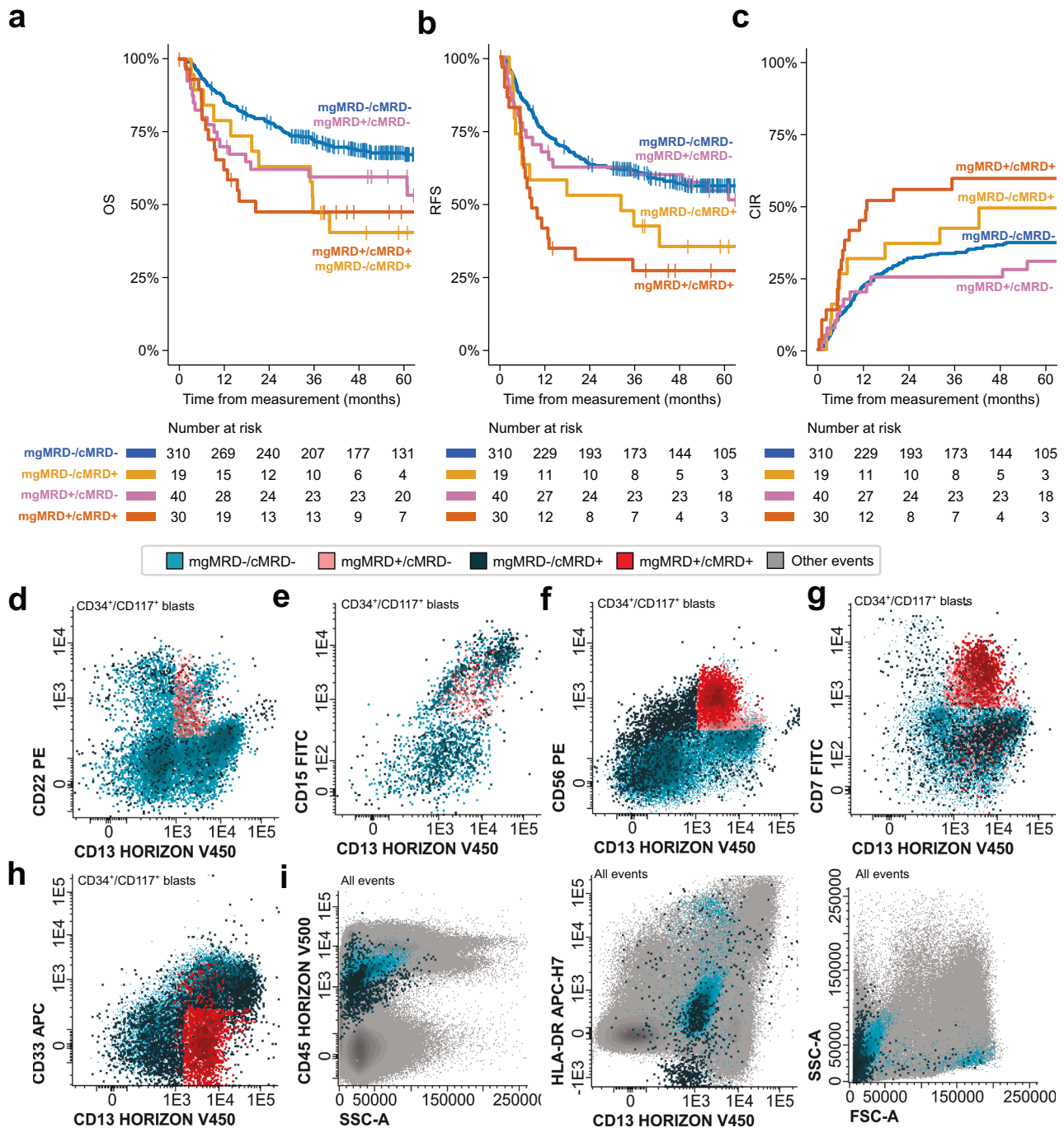


Fig. 2 Side-by-side comparison of computational MRD (cMRD) and manual gating MRD (mgMRD) analysis in AML. Kaplan Meier curves for OS (a) and RFS (b), and the CIR curve (c). Example of CD13⁺CD22⁺ (d) and CD13⁺CD15⁺ (e) populations associated with regenerating bone marrow that are ignored by cMRD. f–h Leukemic populations identified by both mgMRD and cMRD. i Example of aberrant blasts identified by cMRD with debris scatter characteristics. MRD measurable residual disease, AML acute myeloid leukemia, HR hazard ratio, OS overall survival, RFS relapse-free survival, CIR cumulative incidence of relapse.

discrepant cases resulting in cMRD-negativity originated from patients with mature phenotypes (CD34⁺CD117⁺) that were removed due to the initial selection of blasts in the cMRD pipeline (Fig. 1a) or corresponded with manual gating but did not exceed the 0.56% cMRD cut-off.

In the mgMRD-/cMRD+ group ($n = 19$), manual re-analysis identified novel populations absent at diagnosis, resulting in mgMRD-positivity in 9 out of 19 patients. In Fig. 2f–h, we provide three examples (CD13⁺CD56⁺, CD13⁺CD7⁺, CD13⁺CD33⁺ blasts) identified both by manual and computational analysis. Although

differences in RFS between mgMRD-/cMRD- (5-year RFS: 56%) and mgMRD-/cMRD+ (5-year RFS: 35%) did not exceed statistical significance (log-rank: $p = 0.06$), we previously reported the complementary value of emerging cell populations in MRD assessment [8]. In the remaining cases, we often observed non-malignant cells with low abundance in regenerating bone marrow that were incorrectly labeled as aberrant by cMRD, such as CD34⁺CD117^{dim} lymphocyte precursors and CD117⁺⁺ mast cells. However, some cMRD+ cases could also be explained based on technical limitations. In Fig. 2i we show an extreme case, in which

the CD45^{dim} blast compartment contained cells with aberrant expression (CD13⁺HLA-DR⁺) but were identified as cell debris based on scatter (SSC^{low}/FSC^{low}). Because scatter characteristics are not well standardized in flow cytometry data, we could not properly include these parameters in the modeling. For such samples, pre-analytical standardization and quality control is key for evaluating whether samples are fit for computational analysis.

Overall, our results show that cMRD delivers a fast (~3 s) and accurate MRD assessment with clinically relevant relapse associations. Using a quantitative approach to define leukemic phenotypes not only allows for eliminating inter-operator and inter-center variability but also provides utility in re-evaluating AML-MRD gating strategies. Although integration of cMRD into routine diagnostics requires future external, multi-center, and prospective validation to conform to regulatory requirements, the cMRD pipeline we developed was designed with clinical use in mind by avoiding “black box” methodology through robust statistical modeling. Moreover, given its minimal requirements of a small training set ($n = 18$), we envision that this approach is relatively easy to implement in other centers compared to previously proposed computational methods [3], avoiding common regulatory (e.g., data-sharing) and technical (e.g., batch-effect) difficulties. Consequently, the hurdles of implementing AML-MRD in clinical practice can be reduced.

REFERENCES

1. Döhner H, Wei AH, Appelbaum FR, Craddock C, DiNardo CD, Dombret H, et al. Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. *Blood*. 2022;140:1345–77.
2. Tetterto JM, Freeman S, Buecklein V, Venditti A, Maurillo L, Kern W, et al. Technical aspects of flow cytometry-based measurable residual disease quantification in acute myeloid leukemia: experience of the European LeukemiaNet MRD Working Party. *Hemasphere*. 2022;6:e676.
3. Mocking TR, van de Loosdrecht AA, Cloos J, Bachas C. Applications of machine learning for immunophenotypic measurable residual disease assessment in acute myeloid leukemia. *HemaSphere*. 2025;9:e70138.
4. Mocking TR, Kelder A, Reuvekamp T, Ngai LL, Rutten P, Gradowska P, et al. Computational assessment of measurable residual disease in acute myeloid leukemia using mixture models. *Commun Med*. 2024;4:271.
5. Löwenberg B, Pabst T, Maertens J, Gradowska P, Biemond BJ, Spertini O, et al. Addition of lenalidomide to intensive treatment in younger and middle-aged adults with newly diagnosed AML: the HOVON-SAKK-132 trial. *Blood Adv*. 2021;5:1110–21.
6. Zeijlemaker W, Kelder A, Cloos J, Schuurhuis GJ. Immunophenotypic detection of measurable residual (stem cell) disease using LAIP approach in acute myeloid leukemia. *Curr Protoc Cytom*. 2019;91:e66.
7. Heuser M, Freeman SD, Ossenkoppele GJ, Buccisano F, Hourigan CS, Ngai LL, et al. 2021 Update on MRD in acute myeloid leukemia: a consensus document from the European LeukemiaNet MRD Working Party. *Blood*. 2021;138:2753–67.
8. Ngai LL, Hanekamp D, Kelder A, Scholten W, Carbaat-Ham J, Fayed MM, et al. The Laip-based-Dfn approach is superior in terms of useful MRD results as compared to the Laip approach after cycle II in acute myeloid leukemia. *Blood*. 2023;142:1572.

ACKNOWLEDGEMENTS

The authors thank all participating patients and centers of the Dutch-Belgian Cooperative Trial Group for Hematology-Oncology (HOVON) and the Swiss Group for Clinical Cancer Research (SAKK) 132 trial for their contribution to the study.

AUTHOR CONTRIBUTIONS

T.R.M., J.C. and C.B. designed the current study; Sample collection was done by D.A.B., T.F., B.T.G., L.G., G.J., J.A.M., M.G.M., T.P., J.R.P., K.P., B.L., D.d.L., A.A.v.d.L., J.J.W.M.J. and G.J.O. in the HOVON-SAKK-132 trial; Immunophenotypic and molecular assays and

analysis were performed by A.K., W.J.S. and P.J.M.V.; Statistical analysis was performed by T.R.M., L.H.H. and T.R.; The manuscript was written by T.R.M. and revised by L.H.H., T.R., L.N., D.A.B., P.G., P.J.M.V., B.L., J.J.W.M.J., D.d.L., A.A.v.d.L., J.C. and C.B. Results were reviewed and the manuscript was approved by all authors.

COMPETING INTERESTS

Gjertsen: BerGenBio: Consultancy; GreinDX: Consultancy; Immedica: Consultancy; InCyte: Consultancy; Mendus AB: Consultancy, Research Funding; Novartis: Consultancy, Research Funding; Otsuka: Consultancy; Pfizer: Consultancy, Research Funding; Sanofi: Consultancy; in Alden Cancer Therapy AS: Current holder of stock options in a privately-held company; KinN Therapeutics AS: Current holder of stock options in a privately-held company; Coegin: Consultancy. Griškevičius: Miltenyi Biomedicine: Membership on an entity's Board of Directors or advisory committees. Juliusson: AbbVie: Honoraria; Jazz: Honoraria; Laboratoire Delbert: Other: Research cooperation; Novartis: Honoraria; Servier: Honoraria. Löwenberg: Servier Advisory Board; Syndax Pharmaceuticals Advisory Board; Ryvu Pharmaceuticals Consultant. de Leeuw: Takeda: Membership on an entity's Board of Directors or advisory committees; Abbvie: Consultancy; Roche: Consultancy; Servier: Consultancy, Membership on an entity's Board of Directors or advisory committees; Ellipses Pharma: Research Funding. van de Loosdrecht: BMS: Membership on an entity's Board of Directors or advisory committees, Research Funding; Celgene: Membership on an entity's Board of Directors or advisory committees, Research Funding; Roche: Research Funding. Ossenkoppele: Servier: Consultancy; Abbvie: Consultancy; Roche: Consultancy, Membership on an entity's Board of Directors or advisory committees; Astellas: Consultancy, Honoraria; Gilead: Consultancy; Amgen: Consultancy; AGIOS: Consultancy, Honoraria; Janssen: Novartis and Bristol Myers Squibb: research funding; Incyte: speaker's fee; AbbVie, Novartis, Pfizer, and Incyte: honoraria (all to institute); president of nonprofit Apps for Care and Science foundation which received unrestricted educational grants from AbbVie, Alexion, Amgen, Astellas, Astra Zeneca, Bristol Myers Squibb, Daiichi Sankyo, Janssen-Cilag, Novartis, Novo-Nordisk, Incyte, Sanofi Genzyme, Servier, Sobi, Jazz, and Takeda; Cloos: Navigate: Consultancy, Patents & Royalties: Royalties MRD assay; BD Biosciences: Patents & Royalties: Royalties LSC tube; Takeda: Research Funding; Novartis: Consultancy, Research Funding.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41375-025-02747-8>.

Correspondence and requests for materials should be addressed to Costa Bachas.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025