94P - A retrospective study comparing the results of leukocyte differentiation between two laboratories and an AI application

Hauspurg H[1], Elsner F[2], Kunzweiler F[2], Becker M[2], Pfisterer, F[2], Reufer Y[1], Plümer P[1], Fuhrmann S[3], Schabath R[3]

Introduction

Several concepts use AI to support flow cytometry in lymphoma and leukemia diagnostics. Many applications focus on identifying specific cell populations or detecting minimal residual disease (MRD). Often, samples are analyzed with little clinical information; a screening tube is used first to determine leukocyte subpopulations, followed by additional tubes as needed. AI-supported automatic gating and leukocyte differentiation could simplify laboratory work.

This study aimed to show that fully automated AI-supported leukocyte differentiation produces results comparable to conventional laboratory reports.

Materials and Methods

FACS data (FCS files) from peripheral blood analyses between January 1, 2024, and June 30, 2025, from two laboratories were evaluated using the hema.to platform. Included were cases where at least the LYS tube (for lymphocyte subpopulations with 10 colours and 11 antibodies from Lab 1 or 8 colours and 9 antibodies from Lab 2), NHL1 (for clonal B cells and aberrant marker expression of B cells), and NHL2 (for further aberrant marker expression on B cells) were analyzed. For this study, only the LYS tube was used, requiring at least 5,000 events, with leukocyte count and lymphocyte differentiation as structured data. Leukocyte counts were determined by a hematology device (Sysmex XN-350), and both machine-generated and, if needed, manual differential counts were prepared. FACS analysis confirmed the distribution of leukocyte and lymphocyte subpopulations.

On hema.to, AI automatically classified lymphocytes (B cells, NK cells, T cells—further divided into αβ T cells [CD4+, CD8+] and γδ T cells) and myeloid cells (eosinophils, basophils, neutrophils, monocytes).

Relative frequencies from the labs were statistically compared to hema.to results.

A total of 6,564 records met criteria: 67 from Laboratory 1 and 6,497 from Laboratory 2. Structured data included leukocytes, neutrophils, eosinophils, basophils, monocytes, lymphocytes, B cells, NK cells, T cells, CD4+ T cells, and CD8+ T cells. The study assessed consistency between lab and AI-generated results.

Results

Agreement between the hema.to AI model and labs was high for granulocytes, lymphocytes, B cells, NK cells, T cells, CD4+ T cells, CD8+ T cells, and the CD4/CD8 ratio. This held true across pathological and non-pathological cases and various disease entities.

Lower agreement was seen for monocytes, mainly due to a few cases with divergent values. In these, monocyte counts from the lab matched the hematology device, which can be falsely elevated in B-cell neoplasms like hairy cell leukemia. In one case, the model overestimated monocytes due to monoblasts in acute monocytic leukemia; the AI model is not yet trained for acute leukemias.

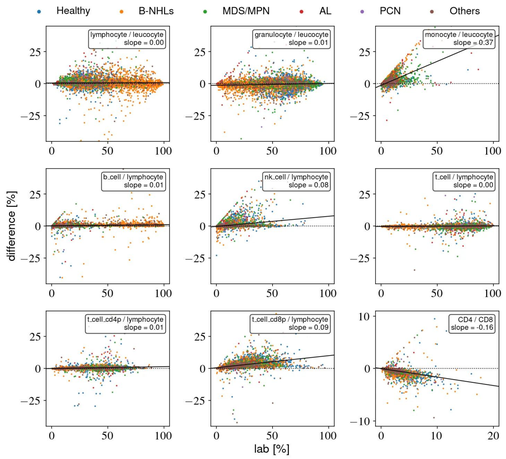
A systematic difference was found for CD8+ T cells between Laboratory 2 and the model, due to different gating strategies. Laboratory 1 showed very good agreement.

Discussion

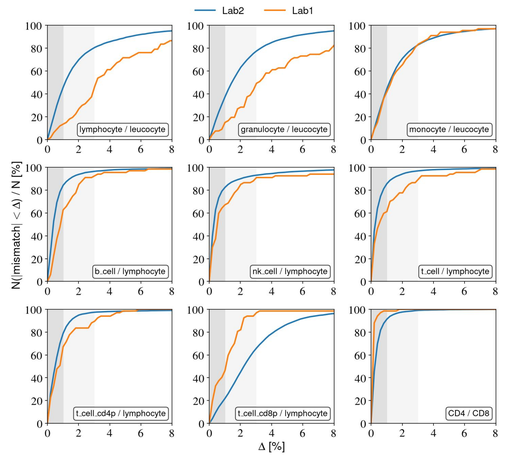
This study compared conventional lab results and the hema.to AI model for leukocyte differentiation, focusing on major subpopulations. High concordance was found for most parameters, regardless of tube composition (LYS10 or LYS8), supporting the robustness of the AI model across validated FACS panels and workflows.

Agreement Across Major Leukocyte Subpopulations:

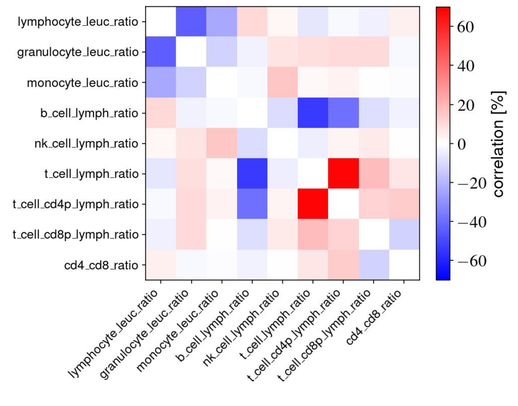
Strong agreement was observed for granulocytes, lymphocytes, B cells, NK cells, T cells, CD4+ T cells, CD8+ T cells, and the CD4/CD8 ratio, across disease entities. This aligns with recent literature showing AI can match or surpass manual methods in leukocyte differentiation. The AI model worked well with different FACS panels, relevant for real-world labs.



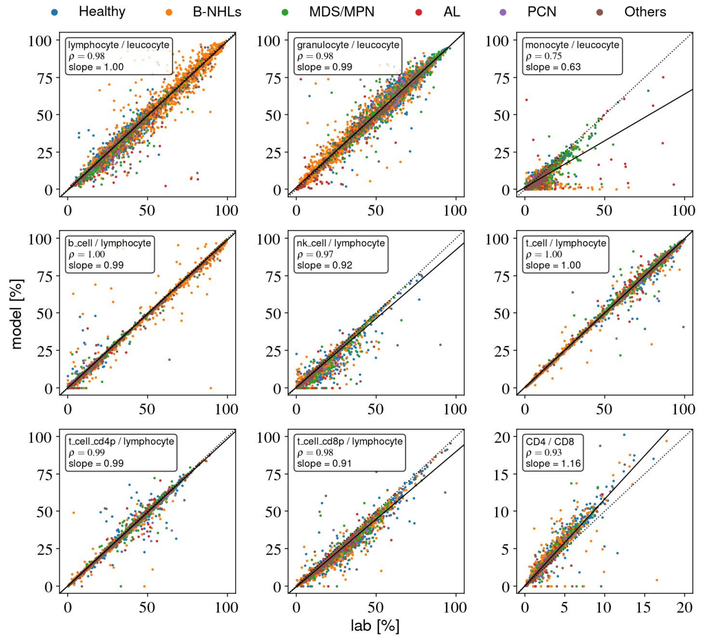
Difference plots showing the lab-result on the x-axis and the mismatch between lab and model results on the y-axis by disease label (6 disease categories: healthy, B-cell Non-Hodgkin-Lymphoma (B-NHL), myelodysplastic and/or myeloproliferative diseases (MDS/MPN), Acute Leukemia (AL), Plasma Cell Neoplasia (PCN), Others)



Above, we show the distribution of simple differences between Lab 1 (yellow), Lab 2 (blue), and model results (without any normalization), allowing to estimate what fraction of data sets have a mismatch of at most X %. Regions, where the discrepancy is below 1% and 3% are highlighted in gray.



Differences in measured cell ratios show statistical correlations between individual cell populations. Above, the correlation matrix for all cell types is shown. We find noticeable correlation pattern, both positive and negative, between certain classes.



Global agreement between population sizes by disease label (6 disease categories: healthy, B-cell Non-Hodgkin-Lymphoma (B-NHL), myelodysplastic and/or myeloproliferative diseases (MDS/MPN), Acute Leukemia (AL), Plasma Cell Neoplasia (PCN), Others). Each subpanel lists the Pearson correlation coefficient (ρ, with a value of 1.0 corresponding to perfect correlation) as well as the slope of a linear regression model (with a value of 1 representing a perfect correspondence between lab results and model predictions).

Discrepancies in Monocyte Enumeration: Monocyte discrepancies were due to (1) **B-cell Lymphomas:** Malignant B cells misclassified as monocytes by the hematology device, inflating monocyte counts compared to the AI model. (2) **Acute Leukemias:** Blasts misclassified as monocytes by the AI model or included in lab counts, causing overestimation. The AI model is not yet specifically trained for acute leukemias. Accurate identification of rare or aberrant populations remains challenging for both manual and automated methods. Data limitations, especially from Laboratory 1, highlight the need for standardized, comprehensive data for robust analysis.

Systematic Differences in CD8+ T-Cell Enumeration: CD8+ T cell discrepancies, mainly in Laboratory 2, were due to different gating strategies, a known source of variability. Both labs passed external proficiency tests, supporting overall reliability.

Implications and Future Directions: AI-based leukocyte differentiation offers speed, reproducibility, and scalability, but discrepancies for monocytes and CD8+ T cells require ongoing validation and harmonization. Future work should expand training datasets, improve data structures, and harmonize protocols. Retrospective and prospective comparisons will further clarify the strengths and limitations of AI-assisted analysis.

Limitations

Limitations include retrospective design, lack of AI training for rare diseases, and small numbers of discordant cases for some subpopulations.

Conclusion

The hema.to AI model showed strong agreement with conventional methods for most leukocyte subpopulations, supporting its use for automated differentiation in clinical practice. Discrepancies for monocytes and CD8+ T cells highlight the need for ongoing validation and protocol harmonization.

Literature

1. Ng DP, Simonson PD, Tarnok A, et al. Recommendations for using artificial intelligence in clinical flow cytometry. *Cytometry B Clin Cytom.* 2024;106(4):228-238. doi:10.1002/cyto.b.22166
2. Spies NC, Rangel A, English P, Morrison M, O'Fallon B, Ng DP. Machine Learning Methods in Clinical Flow Cytometry. *Cancers.* 2025;17(3):483. doi:10.3390/cancers17030483
3. Cheng FM, Lo SC, Lin CC, et al. Deep learning assists in acute leukemia detection and cell classification via flow cytometry using the acute leukemia orientation tube. *Sci Rep.* 2024;14(1):8350. doi:10.1038/s41598-024-58580-z
4. Irvine A, Moustafa MM, Patel S, et al. Automation of flow cytometry data analysis with elastic image registration. *Sci Rep.* 2025;15(1):16949. doi:10.1038/s41598-025-99118-1
5. Chen J, Ionita M, Feng Y, et al. Automated cytometric gating with human-level performance using bivariate segmentation. *Nat Commun.* 2025;16(1):1576. doi:10.1038/s41467-025-56622-2
6. Ahmad N, Asif H, Burhan M, Ahmad AH, Tariq Mahmood M. Coexistence of Acute Myeloid Leukemia and B-cell Non-Hodgkin Lymphoma Diagnosed on a Bone Marrow Trephine Biopsy. *Cureus.* 2025;17(3):e81107. doi:10.7759/cureus.81107
7. Gdańsk MU of. European Journal of Translational and Clinical Medicine. Accessed October 19, 2025. *https://ejtcm.gumed.edu.pl/articles/190793*
8. Cheung M, Campbell JJ, Whitby L, Thomas RJ, Braybrook J, Petzing J. Current trends in flow cytometry automated data analysis software. *Cytometry A.* 2021;99(10):1007-1021. doi:10.1002/cyto.a.24320
9. Cheung M, Campbell JJ, Thomas RJ, Braybrook J, Petzing J. Assessment of Automated Flow Cytometry Data Analysis Tools within Cell and Gene Therapy Manufacturing. *Int J Mol Sci.* 2022;23(6):3224. doi:10.3390/ijms23063224

HH has no conflicts of interest to declare. Corresponding email address hauspurg@hp-hamburg.de
1 Institute of Hematopathology, Hamburg, Germany; 2 hema.to GmbH, Munich, Germany; 3 Pathodiagnostik Berlin MVZ GmbH, Germany

HEMATO

Institut für Hämatopathologie Hamburg