



## The 65th ASH Annual Meeting Abstracts

## ONLINE PUBLICATION ONLY

## 803.EMERGING TOOLS, TECHNIQUES AND ARTIFICIAL INTELLIGENCE IN HEMATOLOGY

**Machine Learning in CLL**

Mohamed Elhadary, MD<sup>1</sup>, Mervat Mattar, MD<sup>2</sup>, Khalil Al Farsi, MD<sup>3</sup>, Salem Alshemmari<sup>4</sup>, Basel ElSayed, MD<sup>5</sup>, Omar Metwalli, MD<sup>6</sup>, Amgad Elshoeibi, MD<sup>7</sup>, Ahmed Abdelrehim Badr, MD<sup>8</sup>, Awni Alshurafa, MD<sup>9</sup>, Mohamed A Yassin, MBBS, MSc<sup>10</sup>

<sup>1</sup> Qatar University, Doha, Qatar

<sup>2</sup> Cairo University, Cairo, Egypt

<sup>3</sup> Sultan Qaboos University, Muscat, OMN

<sup>4</sup> Kuwait Cancer Center, Kuwait, Kuwait

<sup>5</sup> College of Medicine, Qatar University, Doha, Qatar

<sup>6</sup> CMED, Qatar University, Doha, Qatar

<sup>7</sup> Qatar University, Doha, Qatar

<sup>8</sup> CMED Doha, Qatar University, Doha, Qatar

<sup>9</sup> Department of Hematology and Bone Marrow Transplant, National Center for Cancer Care Research, Doha, Qatar

<sup>10</sup> Hematology-BMT, National Centre For Cancer Care & Research, Doha, Qatar

**Introduction:**

Chronic lymphocytic leukemia (CLL) and small lymphocytic lymphoma (SLL) are closely related diseases with similar characteristics, classified as mature B cell neoplasms. They involve the accumulation of monoclonal B lymphocytes and are considered as one disease with different manifestations. CLL is more prevalent and is usually detected through routine blood tests, while SLL is identified by primary lymph node involvement. Although many patients with CLL/SLL are asymptomatic, some may experience lymphoma-related symptoms. The diagnosis involves blood tests and immunophenotypic analysis. CLL prognosis varies, and staging systems help in determining the appropriate treatment regimen. Machine learning (ML) is being increasingly used in hematology to aid in diagnosis, treatment, and risk-stratification. This review explores the application of ML in the classification and diagnosis of CLL, discussing the performance and limitations of existing models, with the goal of encouraging further research and integration into clinical settings to improve patient care.

**Materials and Methods:**

A literature search was conducted using the PubMed/MEDLINE and EMBASE databases. EndNote and Rayyan software were used to eliminate duplicate entries and for screening the articles. In addition, the references of the identified articles were manually screened to identify additional relevant studies. Articles with primary data pertaining to the use of ML in CLL/SLL diagnosis and classification were included without language or time restrictions.

**Results:**

A total of 169 articles were identified and duplicates were removed using Endnote® and Rayyan® software, resulting in 149 articles eligible for screening. The included articles met specific criteria: they utilized ML methods, reported conclusions on the method's reliability or accuracy, and focused on CLL diagnosis and classification. Excluded articles were non-English, animal/in vitro studies, abstracts, and review articles. After screening, 14 studies met the inclusion criteria.

Studies examining the potential of ML algorithms in enhancing the diagnosis of CLL used various data sources like blood smears, flow cytometry, genetic data, and histopathological images. The current evidence suggests that ML can effectively predict CLL diagnosis, aid in screening, identify potential biomarkers, and explore the underlying molecular mechanisms. Many of the ML models evaluated in this review demonstrate high accuracy, ranging from 83% to 100%. Implementing AI and ML in hematology can automate tasks involved in patient workup, risk assessment, and treatment, allowing hematologists to focus on critical aspects of patient care and research.

Despite the promising potential, some important considerations should be acknowledged. First, many reviewed models had limited sample sizes from single centers, limiting their generalizability. It's crucial to develop models using larger, diverse datasets from multiple centers to improve generalizability. Second, there is a lack of prospective studies and research on the impact of ML models on patient outcomes. Future investigations should focus on prospectively assessing the effect of ML

models on CLL diagnosis, prognosis, and patient outcomes. Third, the integration of ML applications into direct patient care raises ethical and medico-legal concerns, including liability, data privacy, and doctor-ML application interaction. An ethical framework specific to ML applications in healthcare is needed, and doctors should receive training on ML applications to understand their capabilities and limitations.

Addressing these considerations can lead to successful implementation of ML applications in the care of CLL patients, improving diagnosis and patient outcomes.

**Conclusion:**

Our findings demonstrate the promising potential of ML algorithms in accurately predicting CLL diagnosis, conducting CLL screening, and identifying genetic biomarkers associated with CLL. Future research endeavors should prioritize the development of large, standardized datasets to effectively train ML models, conduct prospective evaluations to assess their performance, explore their impact on patient outcomes, and establish ethical frameworks to govern their utilization.

**Disclosures** No relevant conflicts of interest to declare.

<https://doi.org/10.1182/blood-2023-179388>

Reference	Outcome	Best model(s)	AUROC	ACC	SEN	SPE
Zhang, Qureshi [17]	Diagnosis and classification of tumors using targeted RNA expression profiling	GMNB	99.7% (CLL vs. normal) 98.6% (CLL vs. MCL) 98.4% (CLL vs. MZL)	NR	96.4% (CLL vs. normal) 94.6% (CLL vs. MCL) 98.7% (CLL vs. MZL)	98.8% (CLL vs. normal) 95.2% (CLL vs. MCL) 91% (CLL vs. MZL)
Xia, Leon [18]	Diagnosis and classification of SBCLs using DNA methylation profiling	SVM	NR	100%	100%	99%
Zhu, Gan [19]	Identification of diagnostic biomarkers for CLL using GEO database	LASSO	NA	NA	NA	NA
Abhishek, Jha [20]	Diagnosis and classification of leukemia using images of blood smears	VGG16 (feature extraction) – RF (classification)	NR	96.8%	100%	96%
Dese, Raj [21]	Diagnosis and classification of leukemia using images of blood smears	SVM	NR	100%	100%	100%
Mohammed, Mohamed [22]	Diagnosis of CLL using images of blood smears	MCS (SVM, KNN, and DT)	NR	85%	89%	87%
Simonson, Lee [23]	Predict whether additional antibody panel should be ordered to distinguish CLL from MCL	EnsembleCNN (CNN supplies predictions & RF integrates)	89%	94%	78%	95%
Ng and Zurowski [24]	Diagnosis and classification of B-cell malignancies using flow cytometry	RF	96.9%	96.4%	86.9%	98.3%
Zhao, Mallesh [25]	Diagnosis and classification mature B-cell neoplasms using flow cytometry	CNN	NR	83%	NR	NR
Haider, Ujjan [26]	Early diagnosis and classification of leukemia using CBC	RBFN	90.5%	NR	NR	NR
Steinbuss, Kriegsmann [27]	Diagnosis and classification of NHL using LNs histopathological images	EfficientNetB3	NR	91.2%	62%	100%
Zhang, Cui [28]	Classification of NHL subtypes using histopathological images	VGG16 (feature extraction) – Neural network (classification)	NR	99.3%	98.4%	99.8%
do Nascimento, Martins [29]	Diagnosis and classification of NHL using LNs histopathological images	PL	100% (CLL vs. FL) 100% (CLL vs. MCL)	100% (CLL vs. FL) 100% (CLL vs. MCL)	100% (CLL vs. FL) 100% (CLL vs. MCL)	100% (CLL vs. FL) 100% (CLL vs. MCL)
Féré, Gobinet [30]	Diagnose CLL using Raman data	PLS-DA	NR	NR	95%	85%

Figure 1