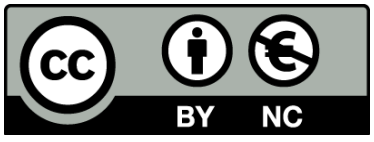
UNIVERSITAT DE BARCELONA

Automatic recognition of different types of acute leukaemia using peripheral blood cell images

Laura Boldú Nebot



AUTOMATIC RECOGNITION OF DIFFERENT TYPES OF ACUTE LEUKAEMIA USING PERIPHERAL BLOOD CELL IMAGES

WRITTEN BY:

Laura Boldú Nebot

SUPERVISED BY:

Dr Anna Merino **Prof José Rodellar**

TUTOR:

Dr Roser Sala Llonch

This thesis is presented to obtain the degree of Doctor in Biomedicine

BARCELONA, APRIL 2021

To my parents, Xavier and Cristina, and Ignasi and Oriol,
for their unconditional support and encouragement over the years.

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincerest and deepest gratitude to Dr Anna Merino and Prof José Rodellar for their trust in me and offer me the opportunity to develop this research. They have supervised, encouraged and helped me during my doctoral thesis. I am very thankful to have them as mentors and for their invaluable support and advice. My gratitude also extends to all the members of the research group CoDAlab from the Department of Mathematics of the Technical University of Catalonia (UPC), all their suggestions and contributions have been a motivation to this thesis. In particular, I want to thank to Angel Molina, Andrea Acevedo and Santiago Alférez for their friendship and support over these years.

I am deeply grateful to the Haematology and Cytology Unit of the Core Laboratory of the Hospital Clinic of Barcelona for providing me a unique working environment and allowing me the use of its facilities to obtain the cell images and the confirmed diagnoses of the patients. I greatly appreciate all the experience I have acquired there. I would like to express my special appreciation to Nayra, Laura, Esther, Rosa, Javi and Maria for their unconditional support and help every day I worked with them, and specially for giving me their friendship.

I would like to show my deepest gratitude to Dr Gaby Marquardt and all the members of her research group in Siemens Healthcare GmbH in Erlangen (Germany), for the granting of my pre-doctoral fellowship and their guidance and support during my research visit in 2021.

I am also thankful to Dr Roser Sala to be my academic tutor during my doctoral thesis.

I am deeply grateful to all my friends, who have supported me in different moments over these years, specially to my best friends Estrella, Albert, Capde, Xavi and Cris for their loyal friendship, support and advice during my life journey.

I would like to express my greatest thanks to all my family, specially to my parents Xavier and Cristina, and my brothers Ignasi and Oriol, who always have believed in me and have been my life's mainstay. Your empathy and moral support have been essential to achieve the fulfilment of this thesis. You spirit of perseverance and your philosophy of doing always your best in everything you do are reflected in these pages and in who I am now.

Finally, any omission in these brief acknowledgements does not mean lack of gratitude.

ABSTRACT

Clinical pathologists have learned to identify morphological qualitative features to characterise the different normal cells, as well as the abnormal cell types whose presence in peripheral blood is the evidence of serious haematological diseases. A drawback of visual morphological analysis is that is time consuming, requires well-trained personnel and is prone to intra-observer variability, which is particularly true when dealing with blast cells. Indeed, subtle interclass morphological differences exist for leukaemia types, which turns into low specificity scores in the routine screening. They are well-known the difficulties that clinical pathologists have in the discrimination among different blasts and the subjectivity associated with their morphological recognition.

The general objective of this thesis is the automatic recognition of different types of blast cells circulating in peripheral blood in acute leukaemia using digital image processing and machine learning techniques. In order to accomplish this objective, this thesis starts with a discrimination among normal mononuclear cells, reactive lymphocytes and three types of leukemic cells using traditional machine learning techniques and hand-crafted features obtained from cell segmentation. In the second part of the thesis, a new predictive system designed with two serially connected convolutional neural networks is developed for the diagnosis of acute leukaemia. This system was proved to distinguish neoplastic (leukaemia) and non-neoplastic (infections) diseases, as well as recognise the leukaemia lineage. Furthermore, it was evaluated for its integration in a real-clinical setting.

This thesis also contributes in advancing the state of the art of the automatic recognition of acute leukaemia by providing a more realistic approach which reflects the real-life complexity of acute leukaemia diagnosis.

Keywords: *Leukaemia*; *Morphological analysis*; *Blood cell automatic recognition*; *Image analysis*; *Convolutional neural networks*; *Diagnostic support*.

TABLE OF CONTENTS

LIST OF PUBLICATIONS

This thesis is based on a compendium of the following publications:

1. **Boldú L**, Merino A, Alférez S, Molina A, Acevedo A, Rodellar J. Automatic recognition of different types of acute leukaemia in peripheral blood by image analysis. *Journal of Clinical Pathology*, **2019**;72(11):755-761. doi: 10.1136/jclinpath-2019-205949. PMID: 31256009.

2. **Boldú L**, Merino A, Acevedo A, Molina A, Rodellar J. A deep learning model (ALNet) for the diagnosis of acute leukaemia lineage using peripheral blood cell images. *Computer Methods and Programs in Biomedicine*, **2021**;202:105999. doi: 10.1016/j.cmpb.2021.105999. PMID: 33618145.

The research work leading to this thesis resulted in additional journal and conference publications, which are listed below and sorted by publication date.

Journals

3. Merino A, Vlagea A, Molina A, Egri N, Laguna J, Barrera K, **Boldú L**, Acevedo A, Díaz-Pavón M, Sibina F, Bascón F, Sibila O, Juan M, Rodellar J. Atypical lymphoid cells circulating in blood in COVID-19 infection: morphology, immunophenotype and prognosis value. *Journal of Clinical Pathology*, **2020**:1-8. doi: 10.1136/jclinpath-2020-207087. PMID: 33310786.

4. Acevedo A, Merino A, Alférez S, Molina A, **Boldú L**, Rodellar J. A dataset of microscopic peripheral blood cell images for development of automatic recognition systems. *Data in Brief*, **2020**;30:105474. doi: 10.1016/j.dib.2020.105474. PMID: 32346559.

5. Molina A, Alférez S, **Boldú L**, Acevedo A, Rodellar J, Merino A. Sequential classification system for recognition of malaria infection using peripheral blood cell images. *Journal of Clinical Pathology*, **2020**; 73(10):665-670. doi: 10.1136/jclinpath-2019-206419. PMID: 32179558.

6. Puigví L, Merino A, Alférez S, **Boldú L**, Acevedo A, Rodellar J. Quantitative cytologic descriptors to differentiate CLL, Sézary, Granular, and Villous Lymphocytes through image analysis. *American Journal of Clinical Pathology*, **2019**;152(1):74-85. doi: 10.1093/ajcp/aqz025. PMID: 30989170.

7. Merino A, **Boldú L**, Ermens A. Acute myeloid leukaemia: How to combine multiple tools. *International Journal of Laboratory Hematology*, **2018**;40(Suppl.1):109-119. doi: https://doi.org/10.1111/ijlh.12831. PMID: 29741260.

8. Merino A, Puigví L, **Boldú L**, Alférez S, Rodellar J. Optimizing morphology through blood cell image analysis. *International Journal of Laboratory Hematology*, **2018**;40(Suppl.1):54-61. doi: https://doi.org/10.1111/ijlh.12832. PMID: 29741256.

Conferences

9. **Boldú L**, Molina A, Laguna J, Rodríguez M, Bascón F, García S, Carballo C, Merino A. Evaluation of flagging performance of the Mindray BC-6800Plus haematology analyser. In XIV National Congress of Clinical Laboratory, Virtual Meeting, **2020**. (abstract)

10. Rodríguez M, Molina A, **Boldú L**, Laguna J, Merino A. Algorithm to predict the diagnosis of hereditary spherocytosis and autoimmune haemolytic anaemia using haematological parameters. In XIV National Congress of Clinical Laboratoy, Virtual Meeting, **2020**. (abstract)

11. Merino A, Vlagea A, Molina A, Egri N, Laguna J, **Boldú L**, Bascón F, Rodellar J. Células linfoides atípicas circulantes en la sangre periférica de pacientes con infección por COVID-19: Morfología, Inmunofenotipo y Valor pronóstico. In XIV National Congress of Clinical Laboratory, Virtual Meeting, **2020**. (abstract)

12. **Boldú L**, Acevedo A, Molina A, Alférez S, Rodellar J, Merino A. A deep learning model for the diagnosis of acute leukaemia subtypes using peripheral blood cell images. In: XXXIII International Symposium on Technical Innovations in Laboratory Hematology, Virtual Meeting, **2020**. (abstract)

13. Merino A, **Boldú L**, Acevedo A, Alférez S, Molina A, Serrando M, Rodellar J. Automatic differentiation of acute leukaemia, lymphoma and reactive lymphocytes in peripheral blood using a novel convolutional network. In: XXXIII International Symposium on Technical Innovations in Laboratory Hematology, Virtual Meeting, **2020**. (Top scoring abstract)

14. Molina A, Rodellar J, **Boldú L**, Acevedo A, Alférez S, Merino A. Automatic classification of red blood cell inclusions using deep learning model. In: XXXIII International Symposium on Technical Innovations in Laboratory Hematology, Virtual Meeting, **2020**. (Awardee abstract)

15. Acevedo A, Merino A, **Boldú L**, Molina A, Alférez S, Rodellar J. A new model for the automatic detection of dysplastic cells in peripheral blood: DysplasiaNet. In: XXXIII International Symposium on Technical Innovations in Laboratory Hematology, Virtual Meeting, **2020**. (Awardee abstract)

16. Barrera K, Pérez A, Acevedo A, Alférez S, **Boldú L**, Molina A, Rodellar J, Merino A. Automatic generation of artificial images of peripheral blood cells using generative artificial networks (GANs). In: XXXIII International Symposium on Technical Innovations in Laboratory Hematology, Virtual Meeting, **2020**. (Awardee abstract with oral session)

17. **Boldú L**, Molina A, García S, Carballo C, Rodríguez M, Laguna J, Merino A. Performance evaluation of the Mindray BC-6800Plus haematology analyser on comparison with the Advia®2120 and manual review using the CellaVision®DM96. In: XXXIII International Symposium on Technical Innovations in Laboratory Hematology, Virtual Meeting, **2020**. (abstract)

18. Fabregat A, **Boldú L**, Merino A. Detection of neoplastic cells in body fluid samples on the Mindray BC-6800Plus analyser. In: XXXIII International Symposium on Technical Innovations in Laboratory Hematology, Virtual Meeting, **2020**. (oral session)

19. **Boldú L**, Bedini JL, Merino A. Detection of neoplastic cells in body fluid samples on the Sysmex UF-5000 analyser. In *Clinica Chimica Acta* **2019**;493:S379–S433. https://doi.org/10.1016/j.cca.2019.03.888. (abstract)

20. **Boldú L**, Alférez S, Molina A, Acevedo A, Rodellar J, Merino A. Automatic recognition of different types of acute leukaemias in peripheral blood by image analysis. In: XXXII International Symposium on Technical Innovations in Laboratory Hematology, Vancouver – Canada, **2019**. (Awardee abstract)

21. Merino A, Alférez S, **Boldú L**, Molina A, Puigví L, Acevedo A, Rodellar J. A new image-based machine-learning system (CellsiMatic) for the automatic recognition of hematologic neoplasia versus infections in peripheral blood. In: XXXII International Symposium on Technical Innovations in Laboratory Hematology, Vancouver – Canada, **2019**. (Top scoring abstract with oral session)

22. Molina A, Alférez S, **Boldú L**, Acevedo A, Rodellar J, Merino A. Automatic classification of red blood cells infected with malaria by means of digital image analysis. In: XXXII International Symposium on Technical Innovations in Laboratory Hematology, Vancouver – Canada, **2019**. (oral session)

23. Alférez S, Merino A, **Boldú L**, Acevedo A, Molina A, Rodellar J. A deep learning approach to automatically classify pathological cell images in peripheral blood. In: XXXII International Symposium on Technical Innovations in Laboratory Hematology, Vancouver – Canada, **2019**. (abstract)

24. Acevedo A, Merino A, Alférez S, **Boldú L**, Molina A, Rodellar J. Automatic detection of dysplastic cells using deep learning. In: XXXII International Symposium on Technical Innovations in Laboratory Hematology, Vancouver – Canada, **2019**. (Awardee abstract)

25. Macias L, **Boldú L**, Molina A, Bascon F, Bedini JL, Merino A. Blast detection on the Advia 2120i analyser. In: XXXII International Symposium on Technical Innovations in Laboratory Hematology, Vancouver – Canada, **2019**. (abstract)

26. **Boldú L**, Alférez S, Puigví L, Acevedo A, Rodellar J, Merino A. Descriptores cuantitativos para el reconocimiento automático de células blásticas en sangre periférica mediante análisis digital de imagen. In XII National Congress of Clinical Laboratory, Bilbao – Spain, **2018**. (abstract)

27. Puigví L, Merino A, Alférez S, **Boldú L**, Acevedo A, Rodellar J. Evaluación de características cuantitativas para la detección de granulación citoplasmática y textura de la cromatina en células linfoides anormales de sangre periférica. In XII National Congress of Clinical Laboratory, Bilbao – Spain, **2018**. (abstract)

28. Merino A, Alférez S, Puigví L, **Boldú L**, Acevedo A, Molina A, Rodellar J. Reconocimiento automático de células linfoides reactivas, anormales y blásticas mediante análisis de imágenes. In XII National Congress of Clinical Laboratory, Bilbao – Spain, **2018**. (abstract)

29. **Boldú L**, Alférez S, Puigví L, Acevedo A, Rodellar J, Merino A. Definition of new quantitative features for the automatic classification of blast cells in peripheral blood by image analysis. In: XXXIst International Symposium on Technical Innovations in Laboratory Hematology, Brussels – Belgium, **2018**. (Awardee abstract with oral session)

30. Alférez S, Merino A, Puigví L, **Boldú L**, Acevedo A, Rodellar J. Automatic recognition of acute leukemia from peripheral blood smears. In: XXXIst International Symposium on Technical Innovations in Laboratory Hematology, Brussels – Belgium, **2018**. (Awardee abstract)

31. Puigví L, Merino A, González C, Redin M.E, Alférez S, **Boldú L**, Acevedo A, Rodellar J. Quantitative features in peripheral blood cell images obtained by diferent hospitals and acquisition methods: utility for reactive and neoplàstic lymphoid cell discrimination. In: XXXIst International Symposium on Technical Innovations in Laboratory Hematology, Brussels – Belgium, **2018**. (abstract)

32. Acevedo A, Merino A, Alférez S, Puigví L, **Boldú L**, Rodellar J. Training a convolutional neural network for automatic classification of peripheral blood cells. In: XXXIst International Symposium on Technical Innovations in Laboratory Hematology, Brussels – Belgium, **2018**. (Awardee abstract with oral session)

33. Macías L, Rico N, **Boldú L**, Filella X, Augé JM, Parra M, Deulofeu R, Bedini JL. Estudio del impacto de un sistema de refrigeración automatizado en la estabilidad de magnitudes de inmunoanálisis en muestras séricas. In XIII Catalan Congress of Clinical Laboratory Sciences, Reus – Spain, **2018**. (abstract)

34. **Boldú L**, Alférez S, Puigví L, Acevedo A, Rodellar J, Merino A. Obtención de nuevos descriptores cuantitativos de células blásticas en sangre periférica mediante técnicas de análisis de imagen. In XI National Congress of Clinical Laboratory, Málaga – Spain, **2017**. (abstract)

35. Merino A, Alférez S, Puigví L, Acevedo A, **Boldú L**, Rodellar J. Detección de la presencia de blastos, promielocitos anormales, linfocitos reactivos y anormales en sangre periférica utilizando análisis de imágenes. In XI National Congress of Clinical Laboratory, Málaga – Spain, **2017**. (abstract)

36. **Boldú L**, Puigví L, Merino A. Standardization and quantification of teardrop cells in primary myelofibrosis and other hematological or oncological diseases using digital morphology. In: XXXth International Symposium on Technical Innovations in Laboratory Hematology, Honolulu – Hawaii, **2017**. (abstract)

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

PB: peripheral blood

CD: clusters of differentiation

AML: acute myeloid leukaemia

ALL: acute lymphoid leukaemia

MPAL: mixed-phenotype acute leukaemia

FAB: French-American-British group

WHO: World Health Organization

MPO: myeloperoxidase

HLA-DR: human leukocyte antigen

APL: acute promyelocytic leukaemia

AML-MRC: acute myeloid leukaemia with myelodysplasia-related changes

MDS: myelodysplastic syndrome

MPN: myeloproliferative neoplasm

t-MNs: therapy-related myeloid neoplasms

AML-NOS: acute myeloid leukaemia not otherwise specified

TAM: transient abnormal myelopoiesis

ROI: region of interest

GLCM: grey level co-ocurrence matrix

ROC: Receiver Operating Characteristic

TPR: true positive rate

TNR: true negative rate

AUC: area under the curve

CNN: convolutional neural network

RGB: red, green, blue

CMYK: cyan, magenta, yellow, black

HSV: hue, saturation, brightness

Lab: lightness and chromaticities in the red-green colour (a) and blue gradient (b)

ReLU: rectifier linear unit

LDA: linear discriminant analysis

GAN: generative adversarial network

CHAPTER 1

Introduction

1.1. Motivation

Severe haematological diseases, especially leukaemia and lymphoma, are common in all the stages of life. The early detection of leukemic cells in peripheral blood and the possibility of a subsequent quick treatment are essential for patients' survival. The World Health Organization considers morphology, along with other complementary tests such as immunophenotype, cytogenetic and molecular biology, essential for the integrated diagnosis of haematological diseases. Nevertheless, smear review is time consuming, requires well-trained personnel and is prone to intra-observer variability. Most of the morphological descriptions are given in qualitative terms and there is a lack of quantitative measures. This is particularly true when dealing with blast cells in acute leukaemia. Indeed, subtle interclass morphological differences exist for leukaemia types, which turns into low specificity scores in the routine screening.

The haematological diagnosis starts with the morphological analysis and continues with other more complex procedures such as immunophenotyping, chromosome analysis and molecular-orientated disciplines, which are available only in highly specialized clinical laboratories because of the equipment's costs and the required human skills. In this context, a methodology able to automatically analyse objective morphological features of blast cells from images obtained through optical microscope, could be a practical support tool for the morphological diagnosis in clinical laboratories.

Nowadays, there are automated analysers able to perform an automatic pre-classification of blood cells based on digital image processing. They show a high efficiency in the recognition of normal blood cells such as neutrophils, eosinophils, basophils, lymphocytes and monocytes. Nevertheless, the automatic morphological discrimination of the abnormal cells by these devices is still an unresolved problem. Moreover, these analysers are not able to discriminate between myeloid or lymphoid blast cell lineages, which is crucial since the prognostic and immediate therapeutic consequences drastically depend on this differentiation.

Image analysis, quantitative morphological features and machine learning approaches have been the main technological tools adopted in the last decade to overcome such drawbacks. The late explosion of deep learning has shifted the focus towards new classification models using convolutional neural networks. This thesis is motivated by the urgent need, as demanded by clinical practise and featured in the recent literature research, to contribute with new developments, combining medical, engineering and mathematical backgrounds, to assist the clinical pathologist reach a faster, more objective and accurate diagnosis of acute leukaemia.

1.2. Objectives

The ultimate goal of this thesis is to develop a complete automatic classification system to predict the diagnosis of acute leukaemia using digital images of peripheral blood cells. The system input is a set of cell images of an individual smear, and the output is the prediction of one of the following diagnoses: myeloid leukaemia, promyelocytic leukaemia, lymphoid leukaemia or infection. We turn this goal into the problem of recognising the following cell types:

- Malignant leukemic cells: blast cells from myeloid and lymphoid origin and abnormal promyelocytes.
- Non-malignant mononuclear cells: lymphocytes, monocytes and reactive lymphocytes.

To achieve this goal, we follow two alternative directions. The first one is based on machine learning approaches using hand-crafted features obtained from cell segmentation. The second approach is based on convolutional neural networks, which autonomously extract their own features and perform the classification.

The following specific objectives are proposed for the first approach:

1. Obtain quantitative features from each of the segmented regions of interest (nucleus, cytoplasm, whole cell and peripheral zone around the cell) to characterise the different cell types under study.

2. Explore several machine learning methods to obtain the best effective classifier along with the selection of the most relevant quantitative features.

3. Perform a detailed analysis and interpretation of the most relevant features using statistical tools to understand their importance in the classification.

4. Evaluate the diagnostic performance by means of a proof of concept using smears from patients with acute leukaemia, infection and healthy controls.

The following specific objectives are proposed for the second approach:

5. Develop and evaluate a convolutional neural network-based model to automatically classify images of the six cell groups of interest.

6. Evaluate the diagnostic performance of the convolutional neural network-based model by means of a proof of concept using smears from patients with acute leukaemia, infections and healthy controls.

1.3. Research Context

In 2012, a collaboration started between members of the research group CoDAlab (Control, Modelling, Identification and Applications), in the Department of Mathematics of the Technical University of Catalonia (UPC), led by Prof. José Rodellar and Dr Anna Merino, who develops her healthcare, research and teaching tasks at the Biomedical Diagnostic Centre of the Hospital Clinic of Barcelona. This teamwork arose from the experience of the daily needs in the reference haematology laboratory of the Hospital Clinic and the incorporation of new laboratory equipment with software tools, which presented serious limitations for the recognition of malignant cells. A new research line was initiated to explore the application of digital image processing and pattern recognition techniques for the automatic classification of lymphoid cells circulating in peripheral blood, which led to the publication of the first two doctoral thesis by Alférez (1) and Puigví (2). Recently, Acevedo (3) presented her doctoral thesis related to the automatic classification of normal and dysplastic blood cells. In the last years, the research group has identified other key scenarios where it would be urgent to have computational assistance during the morphological assessment, such as in the recognition of acute leukaemia, lymphoma, malaria and acquired haemolytic anaemias.

This thesis has been carried out in the Haematology and Cytology Unit of the Core Laboratory of the Hospital Clinic of Barcelona, with strong interaction and use of the facilities of CoDAlab at the Department of Mathematics in the Barcelona East School of Engineering (EEBE); and it has been conducted within the context of two research projects:

- Characterization and morphologic classification of leukemic cells by means of digital image processing and pattern recognition to support diagnosis. Project funded by the Ministry of Economy and Competitiveness of Spain (ref. DPI2015-64493-R). The ultimate objective of this project was the development of an automatic recognition system

able to supply an accurate classification among the abnormal leukocytes circulating in peripheral blood in specific lymphoid neoplasms and acute leukaemia.

- Computational hematopathology: deep learning solutions for diagnosis of haematological diseases from peripheral blood cell images (CellsiMaticDeep). Project funded by the Ministry of Science and Innovation of Spain (ref. PID2019-104087RB-I00). The main objective of this project was to develop computational methodologies to assist the clinical pathologist to reach a reliable diagnosis in acute leukaemia, lymphoma, myelodysplastic syndromes, malaria and acquired haemolytic anaemias using deep learning techniques.

This work has also been possible thanks to the pre-doctoral scholarship granted by Siemens Healthineers S.L.U (Erlangen, Germany) to Laura Boldú, as a part of a collaborative agreement with the Hospital Clinic of Barcelona. This made possible to conclude the thesis with an international internship in the multidisciplinary research group from the Technical University of Munich and the Department of Technologies for Precision Medicine of Siemens Healthcare GmbH, located in the Siemens Campus Erlangen and supervised by Dr Gaby Marquardt.

1.4. Thesis Structure

The present thesis consists of seven chapters starting with the introduction where the motivation, the objectives and the research context are described. Chapter 2 describes acute leukaemia and also includes their latest classification regarding the World Health Organization. Chapter 3 focuses on the first step of acute leukaemia diagnosis, which is the morphological analysis of peripheral blood cells. It is the stage where we have applied the computer vision techniques. This chapter also includes a brief revision of the most important steps for the design and development of the methodology of this thesis. Chapter 4 is constituted by the first paper derived from this thesis, which develops a methodology based on traditional machine learning algorithms to predict the diagnosis of acute leukaemia using peripheral blood cell images. Chapter 5 presents the second paper in which a deep learning-based system is designed, developed and evaluated for its integration in a real clinical setting. Following these two chapters is the directors' report that supports the presentation of this thesis as a compendium of publications. Chapter 6 summarises the main results of the articles, and presents a brief discussion on their main contributions. Finally, Chapter 7 summarizes the conclusions of the thesis and some perspective on future works.

CHAPTER 2

Acute Leukaemia and their Classification

This chapter consists of two sections. Section 2.1 introduces the concept of haematopoiesis. Section 2.2 describes acute leukaemia and presents their latest classification regarding the 2016 revision of the World Health Organization, which aims that the diagnosis of acute leukaemia is accomplished through the integration of clinical findings, cytomorphology, cytochemistry, immunophenotyping, cytogenetics and molecular biology.

2.1. Haematopoiesis

Haematopoiesis is defined as the production, development, differentiation, and maturation of all blood cells (4) (Figure 2.1). It takes place in the bone marrow and all cellular blood components are derived from the maturation of myeloid or lymphoid lineages of stem cells, which have two main basic functions: self-renewal, and differentiation and maturation. The lymphoid stem cell produces lymphoblasts, which can differentiate in B and T lymphoid cells and natural killer cells. The myeloid stem cell originates the myeloid, monocytic, megakaryocytic and erythroid lineages.

Once cells have matured, they pass through bone marrow's sinuses to enter the peripheral circulation. Peripheral blood (PB) carries the following cell types suspended in plasma, all essential for immunity and life; erythrocytes (red blood cells), leukocytes (white blood cells) and thrombocytes (platelets). White blood cells include neutrophils, eosinophils, basophils, lymphocytes and monocytes. For lymphocytes, there is an extra maturation in the lymph nodes, where cells learn their defensive role in the immune system.

In normal conditions, there is a self-regulated balance between the quiescence and the ability for self-renewal of haematopoietic stem cells, which controls the proliferation, apoptosis and differentiation of the progenitors to mature cells. This results in blood carrying normal cells in the adequate proportions. Moreover, cell differentiation implies the progressive development of biochemical, functional and structural characteristics for a specific cell type. Leukocytes express distinct assortment of molecules on their cell surfaces, cytoplasm and nucleus, many of which

reflect either different stages of their lineage-specific differentiation or different states of activation or inactivation. The detection of these clusters of antigens, named clusters of differentiation (CD), by means of monoclonal antibodies has made possible the characterization of the different leukocyte subpopulations and their maturation process, which is required for the diagnosis and classification of the haematological neoplastic diseases (5).

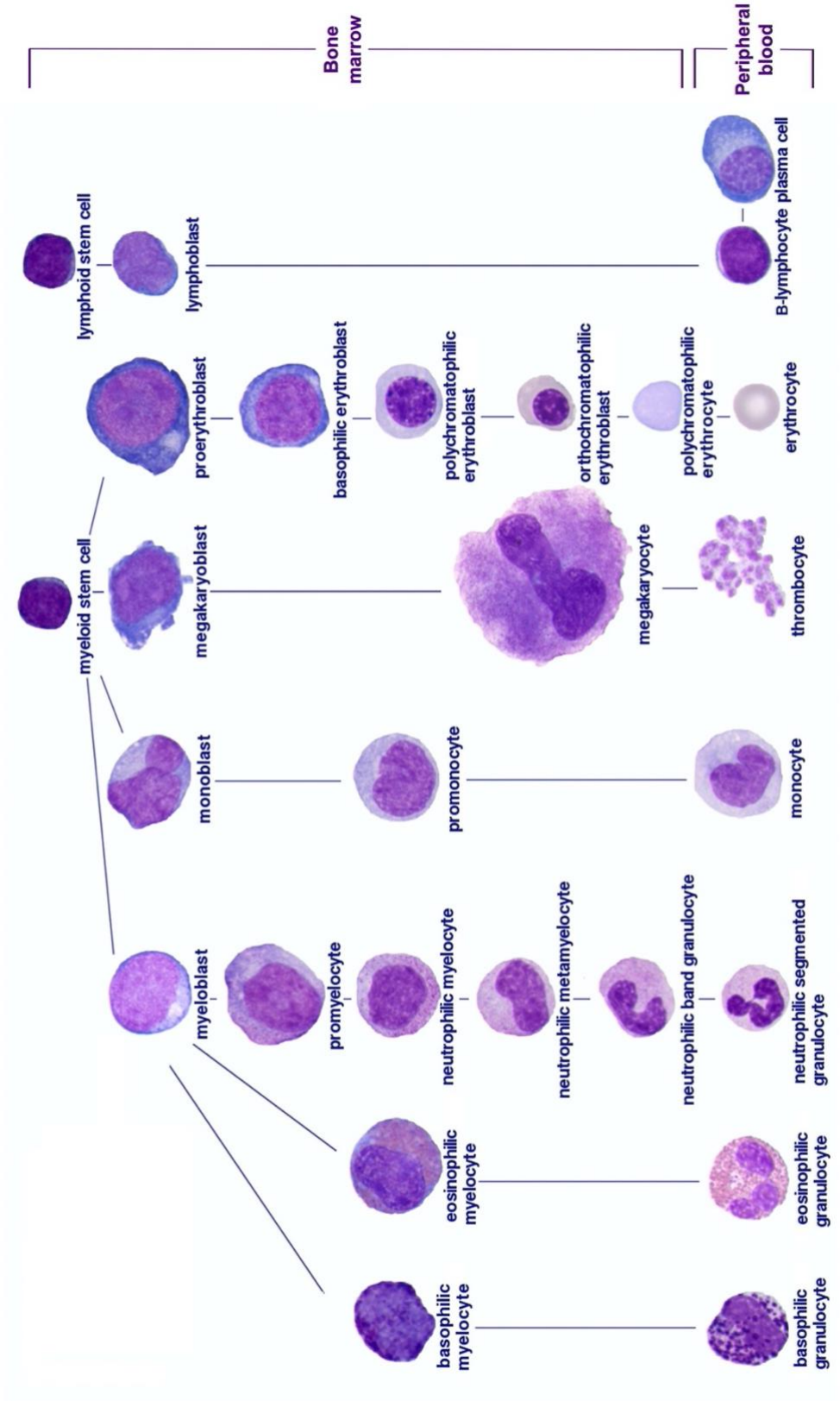


Figure 2.1. Scheme of the haematopoiesis process, where the different lineages of mature cells originate from a pluripotent haematopoietic stem cell. Source: Erasmus MC Rotterdam.

2.2. Acute Leukaemia Classification

Acute leukaemia are neoplastic diseases of the blood system which are the result of a somatic mutation in a haematopoietic stem cell which causes a clonal proliferation of immature leukemic cells. These diseases are characterised by the abnormal proliferation of blast cells in bone marrow, where causes a replacement of normal cells and a decrease of the three haematopoietic lines in PB, resulting in anaemia, neutropenia and thrombocytopenia. As a consequence, acute leukaemia frequently go with infection and/or haemorrhage, and the proliferation of blast cells in other organs causes hepatomegaly and splenomegaly (6).

Genetic alterations responsible for this leukemic transformation are usually acquired chromosomic alterations which include the inappropriate expression of oncogenes and function loss of tumour suppressor genes (those responsible for cellular circle and apoptosis). Oncogenes may be either normal cellular genes (protooncogenes) which have mutated or are dysregulated, or novel hybrid genes resulting from the fusion of two genes. In acute leukaemia there is a defect in maturation, since cells of the leukemic clone continue to proliferate without maturing to end cells and dying. Moreover, leukemic cells are genetically unstable and further mutations can occur in these clone's cells. If one of these new mutations gives the progenitor cell a growth or survival advantage it tends to replace the parent clone and lead to transformation into a more aggressive or treatment-refractory form of the disease with an associated worsening of prognosis (7).

The cell in which the leukaemia transformation occurs could be a lymphoid precursor, a myeloid precursor or a pluripotent haemopoietic stem cell capable of differentiating into both myeloid and lymphoid cells. Regarding the precursor's type, acute leukaemia are divided in two main groups:

- Acute myeloid leukaemia (AML) which can emerge from a lineage-restricted cell, a multipotent stem cell capable of differentiating into cells of granulocytic, monocytic, erythroid and megakaryocytic lineages, or a pluripotent lymphoid-myeloid stem cell.

- Acute lymphoid leukaemia (ALL) which can emerge from a B or T lineage stem cell, or occasionally from a lymphoid-myeloid stem cell.

In addition, there are also the mixed-phenotype acute leukaemia (MPAL). It is a heterogeneous category that compromises acute leukaemia with discrete admixed populations of myeloid and lymphoid blasts ("*bilineal*") or with extensive coexpression of lymphoid and myeloid markers in a single blast population ("*biphenotypic*") (8). However, 1-2% of acute leukaemia could not be classified, known as undifferenced acute leukaemia.

Independent of the precursor's type, acute leukaemia could be *de novo*, if there is no such previous event to justify their appearance, or *secondary*, as an evolutive stage of other previous diseases or due to antineoplastic therapies, such as chemotherapy or radiotherapy.

Acute leukaemia represent the 11th and 10th most frequent cause of cancer occurrence and death worldwide, respectively, with more than 300,000 deaths estimated in 2018. Acute leukaemia are aggressive cancers with 1 to 3 new cases per 100,000 habitants per year with a predominance in masculine sex. The incidence of AML increases with age, from 1.3 cases per 100,000 habitants in patients less than 65 years old, to 12.2 in those over 65 years old. Even with the current treatments, as much as 70% of patients 65 years or older will die of their disease within 1 year of diagnosis (9,10). With respect to ALL, they represent the 15% of the total cases in adults.

Regarding the clinical and biological manifestations leading to suspicion of acute leukaemia include pallor, fever or other signs of infection, petechiae and other haemorrhagic manifestations, bone pain, hepatomegaly, splenomegaly, lymphadenopathy, gum hypertrophy and skin infiltration. The proliferation of blasts and their spread all over bone marrow have in consequence a medullar failure, along with cytopenia in PB (anaemia, neutropenia and thrombocytopenia). Haemorrhagic episodes are frequent especially in acute promyelocytic leukaemia patients who can suffer severe bleeding and disseminated intravascular coagulation. Leukocyte value is variable, although in general an increased value indicates an unfavourable prognosis. The most frequent biochemical parameters are hyperuricemia and the increment of lactate dehydrogenase. The gingival hypertropia is associated with acute monocytic leukaemia because of the tendency of these blasts to infiltrate the gingival tissue (6,7). A suspicion of acute leukaemia generally leads to perform a blood count and film and, if this shows relevant abnormalities, the bone marrow aspiration is proceeded.

Making an accurate diagnosis of haematological neoplasms is crucial for the selection of their most appropriate treatment. Since there are many different types of leukaemia differing in the cell lineage affected, natural history, optimal choice of treatment and prognosis; their classification is essential as it permits to develop selective evidence-based therapeutic approaches.

The diagnosis and classification of acute leukaemia was initially based on morphology. Between 1976 and 1999, a collaborative group of French, American and British haematologists (the FAB group) proposed a classification of acute leukaemia based on morphology supplemented by cytochemistry and to some extend by immunophenotyping (11–13). Over the years, the FAB classification has been replaced by the World Health Organization (WHO) *Classification of Tumour of Haematopoietic and Lymphoid Tissues* (14,15). In 2001, the WHO expert group, in conjunction with the Society for Hematopathology and the European Association of

Hematopathology, proposed an updated system for the classification of leukaemia and lymphoma not only based on morphology, but also it incorporated clinical features, immunophenotyping and the results of cytogenetic and molecular genetic analysis. In 2008 and 2016 further updating of the WHO classification provided new knowledge and great importance to molecular genetic features (16,17).

Summarizing, acute leukaemia are a group of heterogeneous blood-related cancers, differing in their aetiology, pathogenesis, prognosis and response to treatment. The aim of their classification is to reduce this heterogeneity and to identify biologically different groups in order to improve the therapeutic options and patient's prognosis (18).

2.2.1. Acute Myeloid Leukaemia

Acute myeloid leukaemia (AML) are clonal expansions of myeloid blasts in PB, bone marrow or other tissues, which lead to proliferation alongside differentiation block and maturation arrest. The most frequent clinical symptoms of AML are the following: bone marrow failure and in consequence anaemia, neutropenia and thrombocytopenia; extramedullary involvement in other tissues/organs, such as skin, gum, central nervous system or other; proliferative symptoms; coagulopathy, leukostasis and metabolic disorders (lysis tumour syndrome) (17,19).

Myeloblasts are immature cells of variable cell size (15-25 μm) with high nucleus-cytoplasm relation. Their nucleus has a round outline, which sometimes adopts a quadrangular position depending on its cytoplasm, and shows dispersed and immature chromatin with 1-2 visible nucleoli. They have a basophilic cytoplasm, often containing a hint of primary azurophil granulation. Occasionally, blasts show a single long and sharp rod with sharpened ends (Auer rods) (Figure ***2.2***B), or large granules (pseudo-Chédiak-Higashi) suggesting abnormal granules' fusion (20).

In addition to the usual morphological features of myeloid blasts, cytochemistry show myeloperoxidase (MPO) reactivity and the non-specific esterase reaction may be negative or weakly positive, although monoblasts and promonocytes are usually strongly positives. Immunophenotype in AML shows granulocytic (CD13, CD33, CD15, CD65, MPO positive) or monocytic (CD14, CD4, Cd11b, CD11c, CD64, CD36, lysozyme positive) lineages or immature blasts with high CD34 and CD117 expression (17).

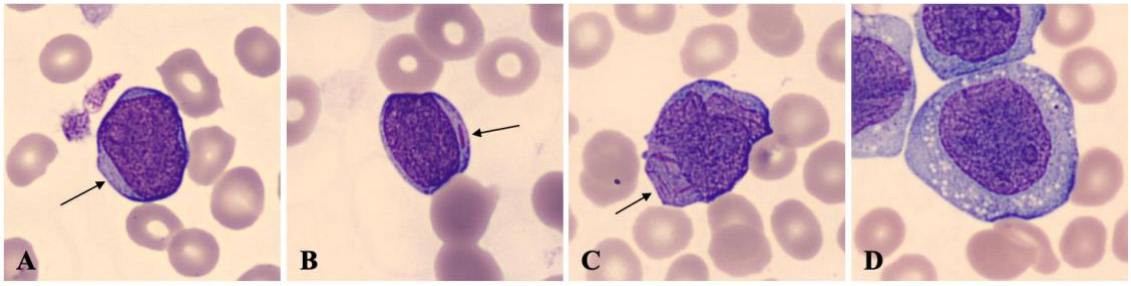


Figure 2.2. Myeloid blast cells. (A) Myeloblast with its nucleus adopting a quadrangular position (arrow); (B) Myeloblast with an Auer rod (arrow); (C) Abnormal promyelocyte with prominent splinters (arrow); (D) Monoblasts with cytoplasmic vacuoles.

The PB smear of AML patients shows leucocytosis, anaemia and thrombocytopenia. Leucocytosis reflects the presence of circulating blast cells, while the number of neutrophils is usually reduced and few cells of intermediate stages of maturation are seen (promyelocytes, myelocytes and metamyelocytes). A minority of patients could present an increased number of eosinophils and, considerably less often, of basophils. Moreover, there may be evidence of dysplastic maturation such as poikilocytosis and macrocytosis, hypolobulated or agranular neutrophils, or hypogranular or giant platelets.

Morphology is the starting point for the diagnosis and classification of acute leukaemia, although immunophenotyping, cytogenetic and molecular genetic analysis also provide crucial diagnostic information. This is why an ideal classification must incorporate all these elements (21,22). The FAB classification published in 1976 was only based on the morphological characteristics of blasts, the cell line affected and their maturation stage, which was further supplemented by cytochemistry and immunophenotyping to distinguish the following morphological types of AML:

- M0 (undifferentiated AML)
- M1 (AML without differentiation)
- M2 (AML with maturation)
- M3 (acute promyelocytic leukaemia)
- M4 (acute myelomonocytic leukaemia)
- M5 (acute monocytic leukaemia)
- M6 (acute erythroid leukaemia)
- M7 (acute megakaryocytic leukaemia)

In 2001, the WHO introduced a new classification system as an effort to integrate advances made in the diagnosis and management of AML. The 2008 WHO classification and its 2016 revision incorporated genetic information with morphology, immunophenotype and clinical presentation to define the seven major disease entities summarised in Table ***2.1*** (14). The WHO criteria for regarding a patient as presenting an AML is to have at least 20% of blast cells in bone marrow or PB, whereas the threshold of the FAB classification was 30% and only applied to the bone marrow. The WHO classification is hierarchical, as shown in Figure 2.3, with therapy-related

cases being assigned first, then cases with specific recurrent genetic abnormalities followed by cases with myelodysplasia-related changes. Finally, when they cannot be assigned to any recognizable cytogenetic/genetic categories, the classification is based on cytological features. Moreover, myeloid neoplasms associated with Down syndrome and blastic plasmacytoid dendritic cell neoplasm are assigned to two separate specific entities (7,17).

Table 2.1. The 2016 revision of the WHO classification of acute myeloid leukaemia (AML).

AML with recurrent genetic abnormalities	AML with t(8;21)(q22;q22.1); *RUNX1-RUNX1T1* APL with t(15;17)(q24.1;q21.2); *PML-RARA* AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22); *CBFB-MYH11* AML with t(9;11)(p21.3;q23.3); *KMT2A-MLLT3* AML with t(6;9)(p23;q34.1); *DEK-NUP214* AML with inv(3)(q21.3q26.2) or t(3;3)(q21.3;q26.2); *GATA2, MECOM* Acute megakaryoblastic leukaemia with t(1;22)(p13.3;q13.1); *RBM15-MKL1* AML with t(9;22)(q34.1;q11.2); *BCR-ABL1* AML with *NPM1* mutation AML with biallelic *CEBPA* mutation AML with *RUNX1* mutation
AML with myelodysplasia-related changes	
Therapy-related myeloid neoplasms	
AML-NOS (not otherwise specified)	AML with minimal differentiation AML without maturation AML with maturation Acute myelomonocytic leukaemia Acute monoblastic/monocytic leukaemia Acute erythroid leukaemia Acute megakaryoblastic leukaemia Acute basophilic leukaemia Acute panmyelosis with myelofibrosis
Myeloid sarcoma	
Myeloid proliferation related to Down syndrome	Transient abnormal myelopoiesis Myeloid leukaemia associated with Down syndrome
Blastic plasmacytoid dendritic cell neoplasm	

APL, acute promyelocytic leukaemia. Adapted from Arber *et al.* 2017 (23).

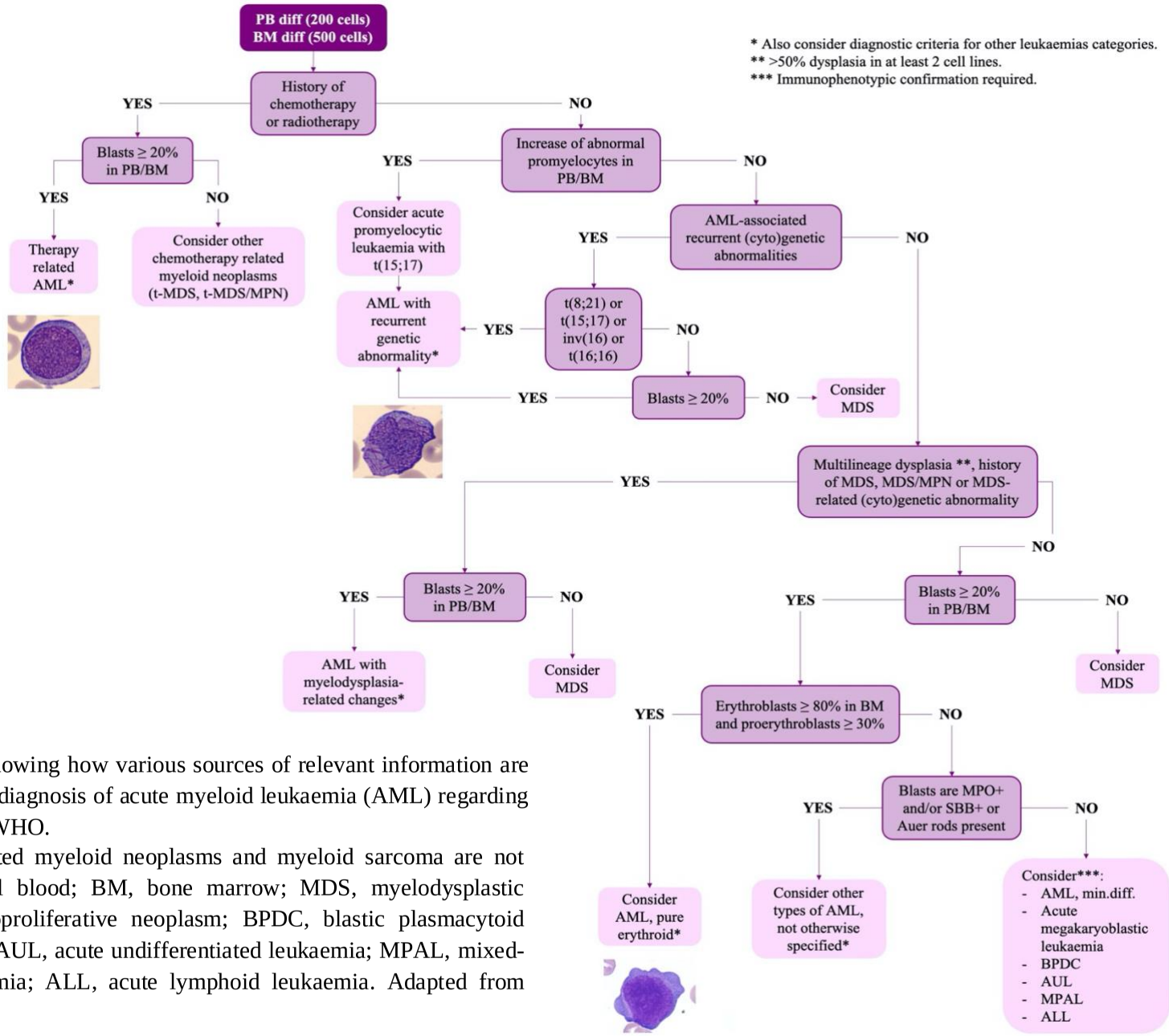


Figure 2.3. Algorithm showing how various sources of relevant information are combined to accomplish diagnosis of acute myeloid leukaemia (AML) regarding the 2016 revision of the WHO.
Down syndrome-associated myeloid neoplasms and myeloid sarcoma are not included. PB, peripheral blood; BM, bone marrow; MDS, myelodysplastic syndrome; MPN, myeloproliferative neoplasm; BPDC, blastic plasmacytoid dendritic cell neoplasm; AUL, acute undifferentiated leukaemia; MPAL, mixed-phenotype acute leukaemia; ALL, acute lymphoid leukaemia. Adapted from Merino A. 2018 (17).

AML with recurrent genetic abnormalities

This entity includes AML with recurrent genetic abnormalities, most of which correspond to balanced translocation/inversions. A bone marrow or PB blast count of 20% is required, except for AML with t(15;17), t(8;21), inv(16) or t(16;16) (23). Moreover, it includes the category of AML with gene mutations. AML with *NMP1* and biallelic *CEBPA* have been incorporated, and AML with *RUNX1* mutation has been added as a new provisional entity.

AML with t(8;21)(q22;q22.1); *RUNX1-RUNX1T1*

This category is characterised by the first balanced translocation discovered in AML by Rowley in 1972 (24). It predominates in younger patients and approximately represents the 8% of the AML cases. The presence of this translocation is diagnostic of AML itself regardless of the blast count. In normal conditions, RUNX1 heterodimerizes to bind DNA and recruit lineage-specific transcription factors to regulate the differentiation of haematopoiesis. The fusion product *RUNX1-RUNX1T1* result of the translocation in this core-binding factor turns genes off, and therefore, blocks the differentiation of blood cells and leads to the production of myeloid blasts (25,26).

Typically, it is characterised by the presence of large blasts with abundant basophilic cytoplasm containing azurophilic granules and occasional Auer rods. Their nuclei are common indented or cleft with large nucleoli. Promyelocytes, myelocytes and mature neutrophils are commonly dysplastic. Eosinophil precursors are frequently increased in bone marrow, but they do now present morphological and cytochemical alterations associated with AML with abnormalities in chromosome 16. It could also be seen an increased number of basophils.

Immunophenotypically blasts express the classical markers of immaturity, CD34 and CD117, together with CD13, CD33. Cytoplasmic MPO is positive. Maturation of a proportion of blasts towards the neutrophil lineage translates into some expression of CD15 and/or CD65 but not CD11b. Many cases show aberrant weak CD19 expression. Other B-lineage antigens that could be found are PAX5, CD79a and TdT (27).

AML with *RUNX1-RUNX1T1* is categorised into a favourable cytogenetic risk group in both children and adults because of the usually good response to chemotherapy. However, several somatic mutations affecting diverse cellular pathways in AML have been identified as poor prognostic factors. These include genes encoding for chromatin modifiers (*ASXL1*, *ASXL2*, *EZH2*, *KDM6A*), cohesin complex (*RAD21*, *SMC3*, *SMC1A*) and signalling pathways (*KIT*, *FLT3*, *NRAS*) (28,29).

Acute promyelocytic leukaemia with t(15;17)(q24.1;q21.2); *PML-RARA*

The predominant cell in this AML type is the abnormal promyelocyte because of the characteristic translocation between chromosomes 15 and 17, in which part of the *PML* gene in chromosome 15 fuses with part of the *RARA* gene in chromosome 17. The resulting fusion protein *PML-RARα* not only is unable to block cell proliferation or induce apoptosis, but it also represses gene transcription. This ends with the differentiation pathway of leukocytes blocked at promyelocyte stage and an abnormal proliferation of these cells (17). The latest 2016 WHO classification emphasised the significance of this cytogenetic rearrangement by renaming this category as acute promyelocytic leukaemia (APL) with *PML-RARA*. The detection of this translocation is diagnostic of APL despite the blast count and it is of utmost importance because patients can suffer severe bleeding and disseminated intravascular coagulation with patients' death (30,31). APL predominates in adults in mild-life, although it can occur at any age, representing the 5-8% of AML cases.

It is of great importance to distinguish between the two morphological variants of APL as they have different immunohistochemical and genetic features, as well as, clinical and prognostic (32–34). The hypergranular APL is the most frequent variant representing the 60-70% of cases. These abnormal promyelocytes present a bilobed nucleus outline and their cytoplasm contains densely packed large granules with sometimes the presence of splinters (see Figure ***2.2***C). The hypogranular APL is also characterised by bilobed nuclei, and there is often a negative image between lobes. There is an absence of granulation and they show a more basophilic cytoplasm than hypergranular abnormal promyelocytes. From a morphological point of view, the differential diagnosis of these hypogranular blasts should be made with respect to monocytic blasts and myeloblasts with *nucleophosmin* (*NMP1*) gene mutation. The white blood cell count is often normal or low in the hypergranular APL, while tends to be elevated in the hypogranular variant. The hypogranular APL has a worse prognosis and represents the 15-20% of the total APL (6).

Regarding the immunophenotype of abnormal promyelocytes, they express characteristic myeloid markers such as CD13, CD33, and often CD117. However, they lack of expression of HLA-DR and CD34, which is not APL specific. They also present strong expression of cytoplasmic MPO, while CD15 and CD16 present on neutrophils are absent. There is also lack of expression of β2 integrins (CD18 β-chain and CD11a, CD11b or CD11 α-chains) (17,27). Abnormal promyelocytes are highly sensitive to anthracycline-based chemotherapy and they differentiate in response to ATRA (all-*trans*-retinoic acid). In addition, those cases of relapse or refractory APL show good response with arsenic trioxide therapy.

AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22); *CBFB-MYH11*

AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22) represents one of the most common subtypes of *de novo* AML cases, compromising 10% of AML in adults and 6% in children. This chromosomal abnormality causes a fusion of the core-binding factor subunit β (*CBFB*) gene at 16q22 and the gene encoding for the smooth muscle myosin heavy chain 11 (*MYH11*) at 16p13, creating the chimeric gene *CBFB-MYH11*. The resulting fusion protein interferes with the formation of the core-binding factor complex and blocks the differentiation of hematopoietic cells (35,36). The presence of this genetic abnormality is diagnostic of AML regardless the blast count.

This AML category is generally associated with acute myelomonocytic leukaemia with an increase of abnormal eosinophils and their precursors showing abundant and large basophilic granules in bone marrow. In PB myeloblasts, monoblasts and promonocytes can be observed. Monocytic blasts have a big size, moderate nucleus-cytoplasmic relation and changeable basophilia. Their nucleus could be round, kidney-shaped or irregular with prominent nucleoli.

The immunophenotype reveals multiple populations, including an immature blast population expressing CD34 and/or CD117 and groups of cells exhibiting granulocytic (CD13, CD33, CD15, MPO) or monocytic (CD4, CD11b, CD11c, CD14, CD64) markers. The incidence of extra-medullary disease is higher than for most types of AML, with a predominance in the central nervous system. Similar to AML with t(8;21), this core-binding factor leukaemia has a favourable prognosis, although *KIT* mutations are present in approximately 30% of cases and negatively impact in adults' prognosis (17).

AML with t(9;11)(p21.3;q23.3); *KMT2A-MLLT3*

This AML is characterised by the rearrangement of the histone-lysine N-methyltransferase 2A (*KMT2A*) gene, previously known as *MLL*, which encodes a transcriptional coactivator that plays an essential role in regulating gene expression during early development and haematopoiesis. This genetic alteration is seen in 5% of *de novo* adults and up to 22% of paediatric patients. Patients suffering from AML with t(9;11) are at increased risk of disseminated intravascular coagulation and have a predominance for extramedullary disease, such as myeloid sarcoma or tissue infiltration, commonly in the gingiva or skin (37).

Although there are not specific morphological features distinctive of this category, AML with t(9;11) is usually associated with those of acute monocytic leukaemia. Monoblasts are large cells with intensely basophilic cytoplasm and round nuclei with immature chromatin and large nucleoli (see Figure ***2.2***D). Promonocytes are often present, with a more mature-shaped nuclei but retaining immature nuclear chromatin. The immunophenotype shows the presence of monocytic

markers (CD14, CD64, CD11b, CD11c and CD4). Moreover, blasts may vary from expressing myeloid-associated antigens (CD13 and CD33) and, are usually CD34 and MPO negative with variable CD117 and CD56 reactivity (17,27).

AML with *KMT2A-MLLT3* is related to intermediate risk disease, with a complete remission rate similar to that of normal AML karyotype. Occasionally, point mutations of *FLT3* may occur in this AML category and are currently of unknown significance. However, an overexpression of *MECOM* (*EVI1*) is associated with a very poor prognosis (37).

AML with t(6;9)(p23;q34.1); *DEK-NUP214*

In AML with t(6;9) occurs a head-to-tail rearrangement with part of the *DEK* gene at chromosome 6 and part of the *NUP214* gene (previously known as *CAN*) in chromosome 9, forming the fusion protein *DEK-NUP214*. It compromises less than 1% of cases of AML. This category sometimes presents as AML with maturation and sometimes as acute myelomonocytic leukaemia. It is associated with basophilia and multilineage dysplasia. It is common an increase of basophils in PB and bone marrow, and in some patients, of eosinophils in bone marrow. The characteristic immunophenotype is the expression of MPO, CD117, CD123, HLA-DR, CD33 and CD13 together with CD9 and CD38, with variable expression of CD15 and CD34. In some cases, TdT is expressed, and those cases with monocytic differentiation may show expression of CD4 and CD64 (7,17).

In AML with t(6;9), this is the unique cytogenetic abnormality in about 80% of patients, being trisomy 8 and 13 the most frequent additional abnormalities. *FLT3-ITD* is found as a second genetic event in up to 70% of adult patients and about 40% of children, whose prognosis is not significantly worsened. The prognosis of this category is poor and the complete remission rate is around 40% being the 5-year survival very low (7,37).

AML with inv(3)(q21.3q26.2) or t(3;3)(q21.3;q26.2); *GATA2, MECOM*

AML with inv(3)(q21.3q26.2) or t(3;3)(q21.3;q26.2) represents about 1% of AML cases. The molecular mechanism that takes place in this category is a juxtaposition of the region surrounding the *RPN1* gene in 3q21 with the *MECOM* (also known as *EVI1*) gene at 3q26.2, causing a reposition of a distal *GATA2* enhancer, which induces the transcription of *MECOM* instead of *GATA2* and thus promotes leukemogenesis. *MECOM* is implicated in the maintenance and expression of normal haematopoietic stem cells and *GATA2* gene encodes an essential protein for the regulation of transcription of genes involved in the development and proliferation of haematopoietic and endocrine cell lineages (17,23).

In this category of AML, multilineage dysplasia and atypical megakaryocytes are common findings. Differentiation can be granulocytic, myelomonocytic or monocytic, although megakaryocytic is over-represented. Regarding blasts' immunophenotype, they express CD34, CD13, CD33 and HLA-DR, with aberrant expression of CD7 in some cases. Megakaryocytic differentiation can be confirmed by CD41, CD42 and/or CD61 positivity (6).

Acute megakaryoblastic leukaemia with t(1;22)(p13.3;q13.1); *RBM15-MKL1*

AML with t(1;22)(p13.3;q13.1) represents less than 1% of cases of AML. This translocation is associated with acute megakaryoblastic leukaemia, and results in the fusion of RNA-binding motif protein-15 (*RBM15*) and megakaryoblastic leukaemia-1 (*MKL-1*) gene. The chimeric protein *RBM15-MKL1* deregulates RNA processing and Ras/MAP kinase signalling and ends altering the normal proliferation and differentiation of megakaryoblasts (38). This type of AML predominantly occurs in infants rather than in adults.

Blast cells show characteristic morphological features of megakaryoblasts. They are very immature and highly polymorphic. Their nucleus is eccentric with dispersed chromatin and 1-3 prominent nucleoli, and their cytoplasm is basophil and reminds of circulant platelets (Figure ***2.4***A). Micro-megakaryocytes and megakaryoblastic elements can also be found in PB with a huge platelet dysmorphia. Immunophenotypically, there is expression of platelet glycoproteins such as CD41 and CD61, and some cases may express CD56. Myeloid antigens (CD13 and CD33) may be expressed while CD34, CD45 and HLA-DR are usually negative (7,17).

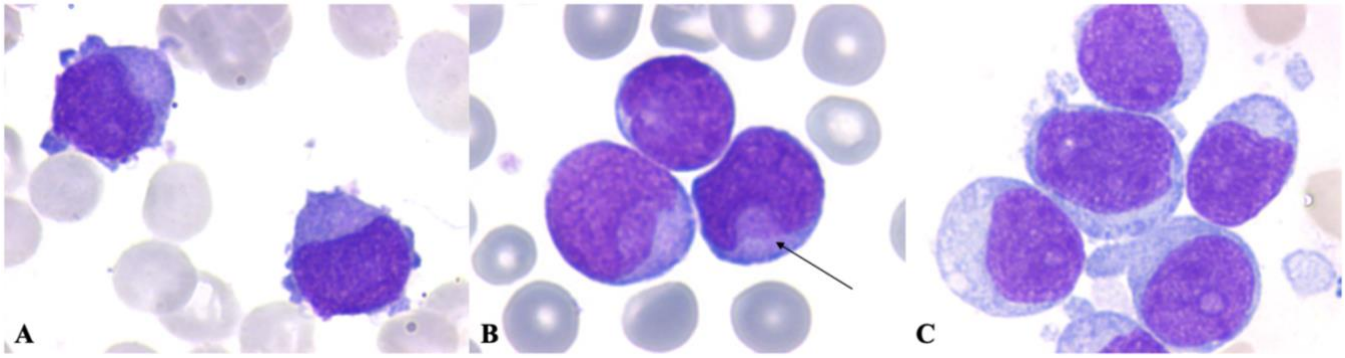


Figure 2.4. (A) Blasts characteristic of acute megakaryoblastic leukaemia. (B) Myeloblasts with the "cup-like" nuclear inclusion (arrow) characteristic of *NMP1* mutation. (C) Blasts from a blastic plasmacytoid dendritic cell neoplasm.

AML with t(9;22)(q34.1;q11.2); *BCR-ABL1*

This is a new provisional category of AML in the WHO classification 2016, because of the significance of distinguishing from a diagnostic point of view *de novo* AML with *BCR-ABL1* from blast transformation of chronic myeloid leukaemia. It is added to recognise those rare *de novo* AML with *BCR-ABL1* cases that may benefit from tyrosine kinase inhibitors therapy (16). This category represents less than 1% of cases of AML. The t(9;22) results in fusion of part of the breakpoint cluster region (*BCR*) gene at 22q11.2 with part of the *ABL1* oncogene translocated from 9q34.1. The chimeric *BCR-ABL1* resulting gene encodes a constitutively activated tyrosine kinase, which is important in leukemogenesis and is a target for treatment (39). The chromosomal abnormality occurs in a pluripotent stem cell, and thus causing abnormality in granulocytic, monocytic, erythroid and megakaryocytic lineages, and also in some precursors of B and T lymphoblasts.

Leukocytosis with blast predominance is observed in most of patients, and basophils count is usually less than 2%. Blast cell morphology show from minimal differentiation to granulocytic maturation. Regarding their immunophenotype, there are not specific features. There is usually an expression of myeloid antigens (CD13, CD33) and CD34, and being also common the aberrant expression of lymphoid markers such as CD19 and CD7 (6).

AML with *NPM1* mutation

In the absence of the recurrent genetic abnormalities described above, the 2016 revision of the WHO classification included as a specific entity those cases of *de novo* AML with *NPM1* mutated (16). The *NPM1* gene encodes nucleophosmin, a phosphoprotein localised in cells' nucleolus that shuttles between nucleus and cytoplasm, and whose mutation leads to a cytoplasmic localization rather than nuclear. Nucleophosmin plays an important role in many cellular functions, such as those involved in protein formation, DNA replication and the progression of the cell cycle (40). Its function in the nucleolus is to attach to a tumour suppressor (*ARF*) to prevent cells from growing in an uncontrolled way, however when nucleophosmin is in the cytoplasm this attachment is not possible and entails leukemogenesis (17).

There is a strong association with acute myelomonocytic and acute monocytic leukaemia. In some cases, blasts show a very characteristic morphology with "cup-like" nuclear inclusions (Figure ***2.4***B). This is particularly seen in those cases with coexisting *FLT3*-internal tandem duplication. Blasts' immunophenotype show expression of CD33, CD117 and MPO in cases with myeloblastic differentiation, lacking expression of CD34 and HLA-DR. Cases with monocytic differentiation could express CD64, CD14 and/or CD11b (7,41,42).

AML with *NPM1* mutation is usually associated with normal karyotype, and it may not be the initial leukemic event as it is generally preceded by *DNMT3A, IDH1, IDH2, KRAS* or *NRAS* mutations. Although AML with *NPM1* mutated show a good response to induction therapy, the high frequency of *FLT3-ITD* as a latter event is associated with worse overall survival. The coexistence of *NPM1, DNMT3* and *FLT3-ITD* mutations has been associated with a very adverse outcome (39).

AML with biallelic *CEBPA* mutation

The 2016 WHO classification included AML with biallelic mutation of *CEBPA*, the gene encoding the myeloid transcription factor CEBPα (CCAAT/enhancer binding protein α). *CEBPA* is a key myeloid transcription factor important for both haematopoietic stem cells self-renewal and for driving transcription during myeloid differentiation. Therefore, *CEBPA* mutations entail a block in myeloid differentiation, transcriptional deregulation and cell cycle alterations (43,44).

Most of these AML have characteristics of AML with or without maturation and multilineage dysplasia is present in 26% of cases, with no adverse prognosis significance. Immunophenotypically blast cells express myeloid antigens such as CD13, CD33, CD65 and CD15, as well as, CD34, HLA-DR and CD7 are usually expressed. With respect to the monoallelic mutation, cases with biallelic mutation are more likely to express HLA-DR, CD7 and CD15 rather than CD56 (7).

AML with biallelic *CEBPA* mutation is associated with a favourable prognosis and a normal karyotype. Nevertheless, some cases could present chromosomal abnormalities being del(9q) common within this group without influence in the prognosis. *FLT3-ITD* is more common in patients with monoallelic *CEBPA*, only between 5-9% has been reported in biallelic *CEBPA* mutation. Moreover, *TET2* mutation is prognostically adverse in biallelic mutated cases, whereas *GATA2* mutation may be prognostically favourable (45).

AML with *RUNX1* mutation

AML with mutated *RUNX1* is other provisional category added to the latest WHO 2016 classification for cases of *de novo* AML with this mutation which are not related to myelodysplastic-related cytogenetic abnormalities (25,26). This type of AML can occur in patients with Fanconi anaemia or severe congenital neutropenia. It represents the 4-16% of AML cases being more frequent in older adults (>60 years) (6).

This category may show morphological features of most AML subtypes, but is most frequent among cases with minimal differentiation cytological features. Immunophenotypically, blasts

express CD13, CD34 and HLA-DR with variable expression of CD33, suggesting immature cells. Most *RUNX1* mutations are monoallelic and could occur with karyotypic abnormalities such as trisomy 8 or trisomy 13. Other genes that may be mutated include *ASXL1*, *KMT2A*, *FLT3*, *IDH1* and *IDH2*. This new provisional entity seems to have a worse prognosis than other AML (7,17).

AML with myelodysplasia-related changes

AML with myelodysplasia-related changes (AML-MRC) is an AML with morphological features of myelodysplasia, which could occur in patients with a prior history of a myelodysplastic syndrome (MDS) or MDS/myeloproliferative neoplasm (MDS/MPN) with MDS-related cytogenetic abnormalities. AML classified in this category must fulfil one of the following criteria: 1) AML arising from previous MDS or MDS/MPN, 2) AML with a MDS-related cytogenetic abnormality and 3) AML with multilineage dysplasia. Furthermore, patients should not have prior history of cytotoxic or radiation therapy (17). Two main amendments were added to the 2016 revision of the WHO classification: 1) multilineage dysplasia in a patient with *NPM1* mutated or with biallelic mutation of *CEBPA* is not considered an AML-MRC, and 2) the presence of del(9q) is no longer regarded as an AML-MRC because of its association with *NPM1* or biallelic *CEBPA* mutations (7,23). In cases lacking these mutations and showing evidence of multilineage dysplasia, which must be present in more than 50% of the cells in at least two myeloid cell lines, remains a poor prognostic indicator and is sufficient to make a diagnosis of AML-MRC. This entity generally has a poor prognosis with a lower rate of complete remission than other AML categories, and mainly occurs in elderly patients and often presents with severe pancytopenia (46).

The immunophenotype reveals features related to MDS with variable expression of myeloid markers such as CD13 and CD33, as well as, CD34, TdT, CD7 and CD56 could also be expressed. The expression of CD14 has been reported to be prognostically adverse (39). With respect to chromosome abnormalities, they are similar to those found in MDS. The most common molecular genetic abnormalities are mutations in *ASXL1*, *RUNX1*, *IDH2* and *TET2* genes (7).

Therapy-related myeloid neoplasms

Therapy-related myeloid neoplasms (t-MNs) is a distinct category in the WHO 2016 classification for patients who develop myeloid neoplasms because of a complication of cytotoxic and/or radiation therapy. This category includes t-AML, t-MDS and t-MDS/MPN. Cytotoxic agents implicated in t-MNs are alkylating agents, topoisomerase II inhibitors, some antimetabolites, anti-tubulin agents, and radiation therapy (23,47).

Morphological features could be those of any FAB category, and being the multilineage dysplasia common particularly after alkylating agents. Auer rods are less common, whereas an increased basophil count, hypocellularity and bone marrow fibrosis are all more common than in *de novo* AML or MDS. The immunophenotype of these blasts is heterogeneous and non-specific. Some cases might show aberrant expression of CD7 or CD56 (7). The prognosis of t-MNs is generally poor, influenced by cytogenetic and genetic mutations, as well as, the underlying malignancy for which prior therapy was received.

AML, not otherwise specified

The WHO category of AML-NOS (not otherwise specified) is a default category for those cases that do not meet criteria for any of the above-mentioned AML categories. It is worth mention that AML-NOS subcategories resemble to specific FAB categories, with the exception that cases meeting the criteria for the specific entities described above are excluded and a blast count of 20% in either blood or bone marrow is sufficient for the diagnosis (6). This category classifies AML based on morphology, cytochemistry and immunophenotype in one of the following subcategories:

- **AML with minimal differentiation** is similar to FAB M0 category. Immunophenotypically blasts show expression of CD34, CD38, HLA-DR, and sometimes CD13, CD117 and CD33. Markers of maturation and MPO are negative and CD7 and TdT could be expressed.

- **AML without maturation** resembles to FAB M1 category. The immunophenotype is similar to that of AML with minimal differentiation with the exception that MPO is positive.

- **AML with maturation** is similar to FAB M2 category. Immunophenotyping usually shows expression of MPO, CD13, CD33 and often CD34, HLA-DR and CD117. Markers of granulocytic maturation are positive (CD11b, CD15 and CD65), and CD7 may be expressed.

- **Acute myelomonocytic leukaemia** resembles to FAB M4 category. The immunophenotype of these blasts is similar to that of AML with maturation with the addition of expression of monocytic markers such as CD4, CD14 and CD64.

- **Acute monoblastic/monocytic leukaemia** is similar to FAB M5 category. Immunophenotyping shows expression of HLA-DR, CD13, CD33, CD15 and CD65, being CD34 and MPO usually negative in monoblasts. Those mature cells show

expression of monocytic markers such as CD11c, CD14, CD64 and CD68, and the expression of CD7 and CD56 is often frequent.

- **Acute erythroid leukaemia** has been the only AML-NOS subcategory with a significant change in the WHO 2016 classification. The erythroid or erythroid/myeloid leukaemia has been removed. In the revised classification, the only current type of acute erythroid leukaemia is the pure erythroid leukaemia. It is characterized by a neoplastic proliferation of undifferentiated or proerythroblastic immature cells committed exclusively to the erythroid lineage, being 80% of the bone marrow cells erythroid with more than 30% proerythroblasts. Those cases previously classified as erythroid/myeloid leukaemia are now classified as MDS with excess of blasts, as now myeloblasts are counted as a percentage of total bone marrow cells and the majority of such cases has less than 20% of total blast cells. This change was based on the close biologic relationship between erythroid/myeloid type and MDS in terms of their clinical presentation, morphological features and genetic abnormalities, as well as, the effort to achieve uniformity in expressing blast percentage across all myeloid neoplasms (6,17). Antigens expressed by the more mature cells are glycophorin A (CD235a) and haemoglobin A, while less mature cells express carbonic anhydrase and CD36. The transferrin receptor and CD71 may be also expressed, but as it happens with CD36, they are not lineage specific. CD34 and HLA-DR are often negative, although CD117 could be positive. Cadherin-E is lineage specific and could be useful for recognition of those cases not expressing CD235a (7). Pure erythroid leukaemia is usually associated with a rapid clinical course with a median survival of only three months.

- **Acute megakaryoblastic leukaemia** resembles to FAB M7 category. Blast cells could present cytoplasmic blebs, suggesting their lineage, or may appear undifferentiated. Differentiation can occur to dysplastic megakaryocytes, and there may be giant or agranular platelets. Immunophenotypically these blasts show expression of CD41 and CD61, as well as, CD13, CD33 and CD36; CD34, CD45 and HLA-DR are often negative (7).

- **Acute basophilic leukaemia** presents differentiation to basophils but there may be blasts with granules characteristic of mast cells. Blasts of basophilic lineage may be packed with basophilic granules, resembling to abnormal promyelocytes, or may have more sparse granules and cytoplasmic vacuolation. Immunophenotyping could show the expression of CD11b, CD123, CD203c and often CD9 and CD25, in addition to CD11b, CD13, CD33 and CD133. Leukemic basophils could be HLA-DR positive and express

CD22. The presence of t(9;22) and *BCR-ABL1* must be excluded. The most common cytogenetic abnormalities that have been reported include t(3;6)(q21;p21) and t(X;6)(p11.2;q23.3) with *MYB-GATA1* (7,48).

- **Acute panmyelosis with myelofibrosis** typically goes with constitutional symptoms such as fever and bone pain. Patients usually present pancytopenia with erythroblasts in PB film, mild poikilocytosis and absence of teardrop cells. Bone marrow is hypercellular with a disproportionate increment of cells of all myeloid lineages and a predominance of immature cells. Blast count is usually relatively low (20-25%), presenting a variable mixture of myeloblasts or monoblasts, proerythroblasts and megakaryoblasts. This category has a poor prognosis (6,7).

Myeloid sarcoma

Myeloid sarcomas are solid extramedullary tumours which can be composed of cells with granulocytic differentiation (granulocytic sarcoma) or with monocytic differentiation (monocytic sarcoma). Common sites of involvement include skin and soft tissues, lymph nodes, the gastrointestinal tract, bones and testis (7,47). Myeloid sarcoma may present *de novo*, accompany PB and bone marrow involvement, present as relapse of AML, or may present as progression of a prior MDS, MPN, or MDS/MPN. Immunohistochemistry is crucial to distinct this category from other tumours such as non-Hodgkin lymphoma.

Myeloid proliferation related to Down syndrome

The WHO 2016 classification include in myeloid proliferation related to Down syndrome transient abnormal myelopoiesis (TAM) and AML associated with Down syndrome. Both are usually megakaryoblastic proliferations, TAM occurring at birth or within days of birth and resolving in one or two months, and AML associated with Down syndrome occurring usually the first three years of life with or without prior TAM and persisting if not treated (23,49). Both conditions are associated with an acquired *GATA1* mutation and mutation of the JAK-STAT pathway (50).

TAM typically presents a high leukocyte count with neutrophilia and the presence of myelocytes and blast cells. Basophilia is common and sometimes there is eosinophilia. Its immunophenotype is characteristic. Blasts are positive for CD7, CD13, CD33, CD34, CD36, CD38, CD71 and CD117. CD41, CD42b and CD61 are usually expressed and CD56 is expressed in about 80% of cases. Regarding blasts from AML associated with Down syndrome, they show platelet antigens (CD41, CD42 and CD61), and also expression of CD13 and CD33, and coexpression of CD7 and CD117 (7).

Blastic plasmacytoid dendritic cell neoplasm

This entity is a rare form of haematologic neoplasm corresponding to a clinically aggressive tumour derived from precursors of plasmacytoid monocytes with a predilection for skin, followed by bone marrow and PB (17). There is an association with AML, MDS and chronic myelomonocytic leukaemia. Its presentation is often with cutaneous lesions and some patients could also present splenomegaly and lymphadenopathy.

Blasts are medium size with irregular nucleus, fine chromatin and showing one or several nucleoli. The cytoplasm is scant, with a grey-blue appearance, without granules and showing microvacuoles or pseudopodia (Figure ***2.4***C). The lack of alpha-naphthyl butyrate esterase activity can help in differentiating it from acute monoblastic leukaemia. The immunophenotype is distinctive and of diagnostic importance. There is almost always expression of CD56 in the absence of B-lineage and most myeloid antigens. The only T-lineage antigens expressed are CD4, often CD7 and CD2. Expression of CD4, CD56, CD123 and CD45RA with lack of expression of CD45RO and CD116 has been found to be specific for this entity. CD43 and plasmacytoid dendritic markers (CD68, CD303, CD304, TCL1A, CLA) and CD123 are expressed. CD34 is usually negative and BCL2, BCL6, MUM1/IRF4 and S100 may be expressed, which they are not in normal plasmacytoid dendritic cells (7,51).

Summarizing, AML is a biologically and clinically heterogeneous disease, whose classification is becoming increasingly complex and it is likely to continue over the years. This will make possible the recognition of new entities with important biological difference. Given the molecular diversity of AML, the new knowledge of these disorders is essential to develop novel therapies, which alongside to the improved genetic profiling and risk stratification could result in incremental gains in remission and survival (52–54).

2.2.2. Acute Lymphoid Leukaemia

Acute lymphoid leukaemia (ALL) and lymphoid lymphoma are both precursor lymphoid neoplasms and thus, they are considered the same biological entity. Their differentiation is made in regards that ALL are malignant transformations and proliferations of lymphoid blasts in the bone marrow and PB, whereas in lymphomas there is only nodal or extranodal infiltration. The most frequent symptoms of ALL are the following: bone marrow failure (anaemia, neutropenia and thrombocytopenia); fever, weight loss, easy bleeding or bruising, fatigue, dyspnoea, infection and extramedullary involvement (lymph nodes, spleen, liver and central nervous system). ALL are the most common childhood cancer, and despite cure rates exceed 90% in children, they remain an important cause of morbidity and mortality in children and adults. The prognosis of

this latest group is very poor and only 30-40% of adult patients with ALL achieve long-term remission (55,56).

Lymphoid blasts are smaller cells than myeloblasts, with high nucleus-cytoplasm relation and regular nucleus' outline (B-lymphoblasts) or with indentations (T-lymphoblasts). Their chromatin is dispersed with few apparent nucleoli or without them. The cytoplasm is scant, slightly basophil and without granulation, and sometimes could contain vacuoles (Figure ***2.5***). Their differential diagnosis should be made with myeloblasts, especially with those of AML with minimal differentiation and AML without maturation. Lymphoblasts are negative for MPO and present positive reactivity for periodic acid-Schiff, while acid phosphatase shows a diffuse positivity with the exception of T-lymphoblasts whose positivity is of centrosomic localization. Immunophenotype in ALL has important diagnostic implications, being CD79a, CD22 and CD19 the most common markers of B-lymphoblasts, and CD7 and cytoplasmic CD3 of T-lymphoblasts (6). Table ***2.2*** shows the most commonly expressed antigens in lymphoid precursors' neoplasia.

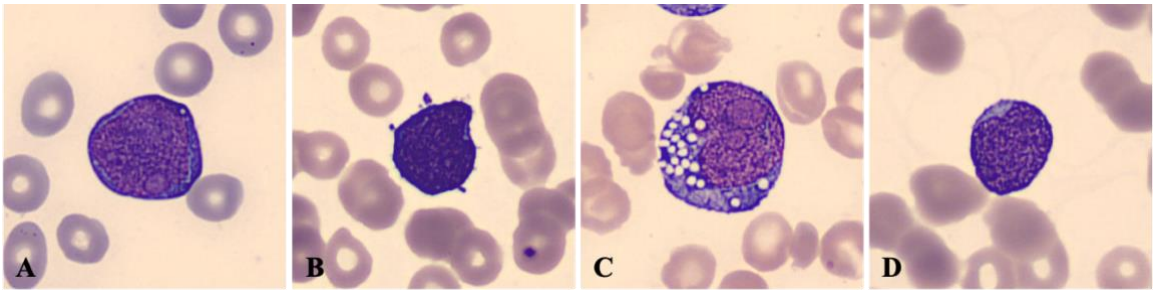


Figure 2.5. Lymphoid blast cells. (A, B) B-lymphoblasts; (C) Burkitt's B-lymphoblast; (D) T-lymphoblast.

Table 2.2. Antigens expressed by blast cells in acute lymphoid leukaemia (ALL).

B-ALL	**Pro-B**	CD79a and/or CD22 and/or CD19
	common	CD79a and/or CD22 and/or CD19, CD10 (CALLA)
	Pre-B	CD79a and/or CD22 and/or CD19, cytoplasmic μ chain
	Mature	CD79a and/or CD22 and/or CD19, surface membrane immunoglobulin
T-ALL	**Pro-T**	CD7 and cytoplasmic CD3
	Pre-T	CD7 and cytoplasmic CD3 and CD2 and/or CD5 and/or CD8
	Thymus-cortical	CD1a
	Mature	Surface CD3 in absence of CD1a

Adapted from Merino A. 2019 (6).

The PB film of ALL patients shows anaemia, neutropenia and thrombocytopenia, although sometimes neutrophil count, platelet count or even both are normal. Leukocyte count could be high, normal or low. In contrast to AML, dysplastic morphological features are not observed, as well as, the presence of micromegakaryocytes is not common (with the exception of cases of lymphoid neoplasia with *Philadelphia* chromosome).

The first attempt to classify ALL was made by the FAB collaborative group who divided ALL into three subtypes: L1, L2 and L3. However, it was replaced by the WHO classification which incorporated morphology, immunophenotyping, cytogenetics and molecular information to group ALL in B or T-lineage specific. Their diagnosis and classification require: 1) the presence of 20% or more lymphoblasts in bone marrow or PB, 2) their assignment to B or T lineage, and 3) their further categorization to recognize real entities which differ in their molecular mechanisms, clinical and haematological features, prognosis and optimal management (7,55). Table ***2.3*** summarizes the 2016 WHO classification.

Table 2.3. The 2016 revision of the WHO classification of acute lymphoid leukaemia (ALL).

B-lymphoblastic leukaemia/lymphoma with recurrent genetic abnormalities	B-lymphoblastic leukaemia/lymphoma with t(9;22)(q34.1;q11.2); *BCR-ABL1* B-lymphoblastic leukaemia/lymphoma with t(v;11q23.3); *KMT2A* rearranged B-lymphoblastic leukaemia/lymphoma with t(12;21)(p13.2;q22.1); *ETV6-RUNX1* B-lymphoblastic leukaemia/lymphoma with hyperdiploidy B-lymphoblastic leukaemia/lymphoma with hypodiploidy B-lymphoblastic leukaemia/lymphoma with t(5;14)(q31.1;q32.1); *IGH/IL3* B-lymphoblastic leukaemia/lymphoma with t(1;19)(q23;p13.3); *TCF3-PBX1* B-lymphoblastic leukaemia/lymphoma, *BCR-ABL1*-like B-lymphoblastic leukaemia/lymphoma with iAMP21
B-lymphoblastic leukaemia/lymphoma, NOS (not otherwise specified)	
T-lymphoblastic leukaemia/lymphoma	Early T-cell precursor lymphoblastic leukaemia
Natural killer cell lymphoblastic leukaemia/lymphoma	

Adapted from Arber *et al.* 2017 (23).

B-ALL with recurrent genetic abnormalities

This entity includes B-ALL with recurrent genetic abnormalities, most of which correspond to balanced translocations and abnormalities in the number of chromosomes.

B-ALL with t(9;22)(q34.1;q11.2); *BCR-ABL1*

B-ALL with t(9;22)(q34.1;q11.2) is characterized by the fusion of part of the *ABL1* oncogene on chromosome 9 with part of the *BCR* gene on chromosome 22 to form the resulting *BCR-ABL1* hybrid gene on chromosome 22. *BCR-ABL1* encodes a chimeric protein with aberrant tyrosine kinase activity. This B-ALL category is also referred to as Ph-positive ALL, because of the implication of chromosome 22 which is known as the Philadelphia (Ph) chromosome. Although this cytogenetic abnormality is also characteristic of chronic myeloid leukaemia, the breakpoint on chromosome 22 differ in both translocations. This ALL is more frequent in adults than in children, representing the 25% of ALL cases in adults and only the 2-4% of cases in children. Immunophenotypically blasts express CD10, CD19 and TdT, and it is not infrequent the association with some myeloid markers such as CD13 and CD33 (6,7).

Regarding secondary cytogenetic abnormalities in Ph-positive ALL, the most common are duplication of the Ph chromosome, del(9p), trisomy 21 and high hyperdiploidy. The presence of them and of a complex karyotype is associated with a worse prognosis. Moreover, in more than 80% of patients *IKZF1* gene is deleted or mutated, which encodes the lymphoid transcription factor *IKAROS*. This may be responsible for the poor prognosis of *BCR-ABL1*-positive ALL patients (57). The treatment with tyrosine kinase inhibitors has led to a marked improvement in the prognosis of this entity.

B-ALL with t(v;11q23.3); *KMT2A*-rearranged

B-ALL with t(v;11q23.3) occurs at all ages but is particularly frequent among babies with congenital ALL and in young patients. The molecular mechanism that takes place in this category is a translocation between *KMT2A* (previously known as *MLL*) in 11q23.3 resulting in the formation of proteins with oncogenic potential (56,58).

This entity is associated with marked leucocytosis (>100 x 10^9/L) and central nervous system infiltration in the moment of diagnosis, which are prognostically adverse. With respect to the immunophenotype, it is strongly associated with early-B-precursor ALL (pro-B), which is the positivity for TdT and pan-B markers such as CD19, and the negativity for CD10 and CD24. Aberrant expression of CD15 and CD65 is also common, and CD33 is sometimes positive. *KMT2A*-rearranged cases show low frequency of secondary somatic mutations, which are often subclonal, indicating that *KMT2A* rearrangement is sufficient to induce leukaemia (7,56).

B-ALL with t(12;21)(p13.2;q22.1); *ETV6-RUNX1*

This category is characterized by the fusion of two transcription factor genes, *ETV6* on chromosome 12 and *RUNX1* on chromosome 21 to form the fusion gene *ETV6-RUNX1* (previously known as *TEL-AML1*). This is one of the commonest subtypes of childhood ALL with a predominance in ages between 2 and 9 years old. It has a very favourable prognosis with a complete remission rate of more than 90% of children (6).

Immunophenotypically blasts show positive expression of CD19, CD10 and usually CD34. There is also high expression of CD40 and HLA-DR and lower expression of CD9, CD20 and CD86. Myeloid antigens such as CD13 and CD33 could also be expressed. Cytogenetic abnormalities associated with B-ALL t(12:21) include del(6q), +8, abnormal 9p, del(11q), i(21q), +21, and translocations between chromosome 12 and its partner chromosomes. However, these secondary abnormalities are not of any prognostic significance (7,59).

B-ALL with hyperdiploidy

The term hyperdiploidy indicates that blasts have more than 50 (but usually fewer than 66) chromosomes, without any translocation or other chromosomal abnormalities. The molecular mechanism of leukemogenesis in hyperdiploidy is unknown. Chromosomes most often gained are 4, 5, 10, 14, 17, 18, 21 and X, in that order. This B-ALL is more frequent in children older than 2 years old than in adults, where it represents only the 6-7% of B-ALL cases. It has a very good prognosis with an overall remission above 90% in children. The immunophenotype is that of common B-ALL, expressing positivity for CD19 and CD10, while cytoplasmic μ chain and surface membrane immunoglobulin are negative. CD34 is often expressed (6,7).

B-ALL with hypodiploidy

Hypodiploidy is defined by blast cells having less than 46 chromosomes. Near haploid (23-29 chromosomes) metaphases usually retain both sex chromosomes and preferentially gain of chromosomes 10, 14, 18 and 21 into the haploid chromosome set. This entity represents the 5% of B-ALL cases, having the same incidence in children than in adults and it is associated with adverse prognosis. The immunophenotype is usually that of common ALL with expression of CD19 and CD10 and without other specific markers. Regarding genetic abnormalities, there is association with deletions of *CDKN2A*, *CDKN2B* and *RB1* and inactivation of *IKZF2* (6,7).

B-ALL with t(5;14)(q31.1;q32.1); *IGH/IL3*

This category of B-ALL is characterized by a translocation between the interleukin 3 (*IL3*) gene on chromosome 5 and an immunoglobulin (*IGH*) gene on chromosome 14. It is associated with

eosinophilia as a result of dysregulation of the *IL3* gene. Although eosinophils do not come from the leukemic clone, they may show morphological abnormalities such as hypolobulation and degranulation. The immunophenotype is that of common ALL, being CD19 and CD10 expressed. This is a rare form of B-ALL accounting of less than 1% of cases and making difficult to predict its prognosis (6,7).

B-ALL with t(1;19)(q23;p13.3); *TCF3-PBX1*

In B-ALL with t(1;19) there is a fusion of the *PBX1* gene from 1q23 with part of the transcription activator gene *TCF3* (previously known as *E2A*) at 19p13.3 to form the hybrid gene *TCF3-PBX1*, which encodes an abnormal transcription factor. This category is more frequent in children (2-5%) than in adults (1-3%). Although it was originally considered to be a high-risk ALL category, its outcome has improved with current more intensive treatments. However, it has been reported a greater risk for central nervous system relapse (60). The t(17;19)(q21-22;p13) is a molecular variant of this category, which leads to the fusion gene *TCF3-HLF*, which is associated with disseminated intravascular coagulation, hypercalcemia and a very poor prognosis (7,56). Blasts show a pre-B phenotype, being cytoplasmic μ chain present but not surface membrane immunoglobulin. The typical immunophenotype is CD19, CD22, CD10 and CD9 positive, and CD21 and CD34 negative.

B-ALL*, BCR-ABL1*-like

B-ALL *BCR-ABL1*-like is not characterized by the t(9;22), but shows a gene-expression profile similar to Ph-positive B-ALL. It is a provisional category in the 2016 revision of the WHO classification due to its clinical relevance and association with an adverse prognosis and responsiveness to tyrosine kinase inhibitors. This category represents about 10-13% of childhood B-ALL, about 21% of adolescent cases and about 27% of young adult cases. Leucocyte count is higher that in other types of B-ALL. The immunophenotype is usually that of common ALL (6,7).

BCR-ABL1-like is associated with high-risk clinical features, a poor response to induction chemotherapy and poor prognosis. Common genetic abnormalities of this entity are related to alterations of B-lymphoid transcription factor genes (*IKZF1* deletions) and genetic alterations deregulating cytokine receptors and tyrosine kinase signalling, such as rearrangements and mutation of *CRLF2*, rearrangements of *ABL*-class tyrosine kinase genes, rearrangements of *JAK2* and the erythropoietin receptor gene *EPOR* (16,56).

B-ALL with iAMP21

This B-ALL is characterised by presenting lymphoblasts with an intrachromosomal amplification of part of chromosome 21 (iAMP21). It particularly occurs in older children (9 years) and is associated with poor prognosis, which is improved with intensive treatment. This provisional entity added in the 2016 WHO classification is detected by FISH with a probe for the *RUNX1* gene that reveals five or more copies of this gene (or three or more extra copies on a single abnormal chromosome 21 in metaphase FISH). It is associated with leukopenia and no specific immunophenotype has been identified (6,7,16). Common associated molecular genetic abnormalities include deletion of *RB1* or *ETV6* and mutation of *RAS* genes, *FLT3* or *CRLF2* (61,62).

B-ALL, not otherwise specified

This category includes all cases of B-ALL which have not been assigned to any of the previously mentioned categories, and thus being a heterogeneous group. Blasts present a variable immunophenotype, generally expressing CD19, CD79a and CD22, and there may be expression of CD10, CD20, CD24, PAX5, CD34 and TdT. CD45 may be weak or negative (7).

The wide range of molecular abnormalities seen in this category may include deletions of genes involved in B-cell differentiation (*RAG1*, *RAG2*, *CD200* and *VPREB1*), transcriptional co-regulators (*BTG1*, *ETV6*, *ERG*, *TBL1XR1*) and others (*TCF3*, *EBF1*, *LEF1*, *RB1*, *CDKN2A*, *CDKN2B* and *IKZF3*) (62,63).

Furthermore, in some cases of Burkitt's lymphoma a leukemization of proliferative lymphoid cells could be observed and this entity is presented as a leukaemia with infiltration in PB and bone marrow. In Burkitt's leukaemia is frequent the central nervous system infiltration. Regarding the morphology of these blasts, they have big size (20-25 μm of diameter) and regular nucleus' outline with disperse chromatin. An abundant cytoplasmic basophilia, prominent vacuoles from lipid origin and high mitotic activity are what most characterise these lymphoblasts. Vacuoles could be also localised in the nucleus (see Figure ***2.5***C). These blasts may show a mature B immunophenotype with expression of surface membrane immunoglobulin. Cytogenetic alterations characteristic of this entity are translocations on chromosome 8, such as t(8;14), t(8;22) and t(8;2) (6).

T-ALL

T-ALL represent only 15% of cases of childhood ALL, typically occurring in adolescents rather than in younger children. Constitutes a quarter of adult ALL cases with a male predominance.

There is involvement of bone marrow and PB or the thymus, as well as, other extranodal sites such as liver, spleen, skin, tonsils and testes. Leucocyte count is typically increased and some T-ALL cases present eosinophilia, which indicates a worse prognosis (6,7).

Regarding the morphology of T-lymphoblasts, they are similar to those of B-lineage, with high nucleus-cytoplasm relation and small-medium size. Nucleus' outline is irregular or convoluted with moderate condensate chromatin and unapparent nucleolus (see Figure ***2.5***D). It is more frequent to observe mitotic figures in T-lymphoblasts rather than in B-lymphoblasts. T-lymphoblasts show an intense centrosomic positivity for acid phosphatases and are TdT positive with variable expression of CD1a, CD2, CD3, CD4, CD5, CD7 and CD8. In about 25% of cases there is rearrangement of the TCR (T-cell receptor) loci, and in consequence in 20% of cases the immunoglobulin gene is affected. A complex karyotype is detected in 50-70% of cases and the most recurrent cytogenetic abnormalities involve the TRα and TRδ in 14q11.2, the TRβ at 7q34 and the TRγ at 7p14.1 (6). Moreover, about 30% of patients have del(9)(p21.3) leading to homozygous loss of the tumour suppressor genes *CDKN2A* and *CDKN2B*. Activating mutations in *NOTCH1* are found in about 50% of patients, and another 30% of patients have a mutation in *FBXW7*. *NOTCH1* encodes a membrane receptor that regulates normal T-cell development and *FBXW7* reduces *NOTCH1* activity. These mutations seem to correlate with a better prognosis (7).

Early T-cell precursor ALL

This provisional category represents about 12-16% of childhood cases and up to 10% of adult cases of T-ALL. Immunophenotypically blasts express CD7 but lack CD1a and CD8, and are positive for one or more myeloid/stem cell markers (CD34, CD117, HLA-DR, CD13, CD33, CD11b or CD65). They also express CD2 and cytoplasmic CD3 and may express CD4. Myeloid-associated gene mutations, such as *FLT3*, *NRAS*/*KRAS*, *DNMT3A*, *IDH1* and *IDH2* are reported at high frequency in this category, whereas more typical T-ALL-associated mutations such as activating mutations in *NOTCH1* or mutations in *CDKN1*/2 are infrequent. Although originally this category was associated with poor prognosis, recent studies with larger number of patients with more effective therapy suggest no prognostic significance (6,16).

Natural killer lymphoblastic leukaemia/lymphoma

This is a rare neoplasm included as a provisional entity in the 2016 revision of the WHO classification within the category of precursor lymphoid neoplasms (64). It has been suggested that those cases previously assigned as T-ALL with expression of cytoplasmic CD3 and sometimes a mediastinal mall, might represent neoplasms of immature natural killer cells. There is expression of CD7 and CD2 and cytoplasmic CD3ε, but without rearrangement of the TCR

loci. Cases are provisionally identified by expression of CD94 1A, an isoform of CD94 which is induced in natural killer cell precursors by IL15 (7).

Summarizing, both AML and ALL are biologically and clinically heterogeneous diseases, whose classification is becoming increasingly complex and it is likely to continue over the years. This will make possible the recognition of new entities with important biological difference. The new knowledge of these disorders will be essential to develop novel therapies, which alongside to the improved genetic profiling and risk stratification will result in incremental gains in remission and survival.

Having said that, the diagnosis of acute leukaemia relies on cytomorphology, cytochemistry, immunophenotyping, cytogenetics and molecular biology. Only the integration of all these diagnostic methods allows for a comprehensive and complementary characterisation of each entity, which is a prerequisite for the optimal diagnosis and management of acute leukaemia.

CHAPTER 3

Digital Morphological Recognition of Blood Cells in the Diagnosis of Acute Leukaemia

This chapter includes two main sections. Section 3.1 focuses on the morphological analysis of peripheral blood cells, which is the first step in acute leukaemia diagnosis. This section also shows the trend of automated methods for the digital morphological analysis of blood smears in clinical laboratories. It is in this stage where we have dedicated effort with the use of computer vision techniques to assist in the diagnostic procedure of acute leukaemia. Section 3.2 introduces the concept of machine learning and describes the main steps of image-based recognition systems.

3.1. Morphological Analysis of Peripheral Blood Cells

Even in the age of molecular analysis, morphological analysis of PB remains the crucial starting point for the diagnostic approach and follow-up of patients with malignant haematological diseases. This is because PB smear examination is considered to be a non-invasive procedure to suggest a differential diagnosis and the identification of further necessary tests in a time and cost-effective way (65,66). This is why cytomorphology was, is and will still be at the forefront of haematological diagnosis.

The morphological analysis of the PB smear provides the percentage of each of the different leukocytes' subpopulations. The blood smear must be requested by the clinician when there is a specific diagnostic suspicion, or unexplained leucocytosis, lymphocytosis or monocytosis. It is also a task of the clinical pathologist to indicate the realization of the PB examination when the results of the complete blood count have quantitative anomalies, or when the flagging system of an automated instrument suggests the presence of blast cells (Figure ***3.1***). Then, the clinical pathologist analyses the morphology of all the cells to detect abnormalities related to different types of diseases and gives a diagnostic orientation.

Figure 3.1. Routine workflow of the morphological assessment of patients' blood samples in clinical laboratories.

Although this purely visual and manual morphological inspection is time consuming, needs expert professionals to review the blood smears in an objective and reliable was and is prone to intra-observer variability; its careful observation provides a lot of information, being a valuable tool in both diagnostic orientation (which guides the complementary tests) and performing some definite diagnoses. Most of the morphological descriptions are given in qualitative terms and there is a lack of quantitative measures. This is particularly true for blasts and abnormal lymphocytes. Indeed, subtle inter-class morphological differences exist among abnormal cell types. For example, the common morphological characteristics shown by myeloblasts include variable cell size with a high nucleus-cytoplasm relation and mostly showing a quadrangular nucleus profile (except abnormal promyelocytes in APL) and a basophilic cytoplasm, often containing azurophilic granules (17,67,68). Some of the described characteristics are shared by lymphoblasts, what makes difficult in most cases recognizing under the microscope the type of acute leukaemia the patient suffers. This gives low specificity scores in routine screening. In addition, the blood smear is of particular importance in Burkitt's leukaemia/lymphoma and APL because it facilitates rapid therapeutic intervention influencing the prognosis. Thus, there is an increasing demand for the development of digital microscopy systems, capable of performing morphological analysis of blood smears in an automated manner (69).

The field of laboratory haematology diagnostics advanced significantly in the 1980s with the introduction of flow-based cell counters. These analysers were capable of subclassifying leukocytes and detecting abnormalities in leukocytes, red blood cells, and platelets using instrument flags. An example of them is the Advia2120i® (Siemens Healthcare GmbH), which performs the complete blood cell count and provides flags when there are malignant cells circulating in PB, such as reactive lymphocytes, abnormal lymphocytes or blast cells. These flags indicate the examination of the PB smear by the clinical pathologist.

Despite the fact that these cell counters have become an integral part of haematology laboratory nowadays, the need for the morphological review of the blood smear still remains. Thus, at the beginning of XXI century a new inclination started in the development of more advanced automated digital microscopy systems able to perform the morphological analysis of blood smears in an automated and more efficient way. Most of these systems make use of a digital camera coupled to a computer system. Digital images of blood cells are then used as inputs for a computer-aided classification system based on blood cell image analysis parameters such as geometric, colour and texture (69).

The latest automated digital image analysers introduced to the haematology laboratory market are the following:

The Vision Hema® Automated Haematology Imaging Analyser (West Medica, Perchtoldsdorf, Austria) provides a detailed analysis and pre-classification of leukocytes, red blood cells, platelets and reticulocytes. It also identifies cells on bone marrow, body fluid and cervical cytology smears. A separate module can be also used for malaria diagnosis (70).

EasyCell®assistant (Medica Corporation, Bedford, MA, USA) automatically detects white blood cells on a blood smear, and then it classifies normal leukocytes, smudge cells and nucleated red blood cells. This system employs image processing and artificial intelligence for cell identification and pre-classification (71).

Nextslide Digital Review Network (Nextslide Imaging LLC, Cleveland, OH, USA) consists of a high-resolution slide scanner and a hosted software-application suite, which includes image-processing, management and online review applications. The image-processing application accepts completed images from the slide scanner and processes them to identify and classify white blood cells. Afterwards, with the online review the processed and raw images can be analysed, as well as evaluate the morphology of red blood cells and platelets (72).

The Roche Cobas® m511 (Roche Diagnostics, Indianapolis, IN, USA) introduces a digital morphology analyser, a cell counter and a classifier in a unique equipment which includes an

auto-stainer and a microscope analyser of blood smears. This technology analyses the cell morphology, counts and classifies each cell providing a 5-part differential on 600 nucleated cells on the slide. It is also used to classify the types of white blood cells, evaluate red blood cells and platelets, and to review their morphology. This system can simultaneously provide both quantitative and morphological information (73).

The HemaCAM® system (Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany) is able to evaluate PB samples and perform differential blood cell counts. This equipment can automatically pre-classify different types of leukocytes, and provide information about red blood cells and platelets. Further applications under development may include the detection of malaria parasites and the automated recognition of bone marrow cells for the workup of leukaemia (74).

The Sysmex DI-60 system (Sysmex Corp. Kobe, Japan) was the first fully-integrated cell image analyser on the market, which could co-operate with the Sysmex XN haematology analyser and the Sysmex slide maker-stainer SP-10, thus allowing a single sample placed on this automated device to provide a complete blood count, perform a smear preparation and also determine digital cell location. This system can pre-classify leukocytes, pre-characterise red blood cells and estimate the number of platelets on PB smears (75).

CellaVision®DM96 (CellaVision AB, Lund, Sweden) is an automated device which scans stained blood slides, identifies white blood cells and then takes images of them at high magnification (x1,000). It is made for the differential pre-classification of leukocytes, evaluation of red blood cells morphology, platelet estimation on blood smears, and it can also analyse body fluids. This equipment includes a motorized microscope, a camera and a computer containing the acquisition and classification software. Cell type recognition is based on an artificial neural network, which analyses the digital images and pre-classifies the cells. The analyser requires a skilled pathologist to review all cells and to re-classify all unidentified cells before it is possible to release the results. Several studies have evaluated the concordance between the automated leukocyte differential counts on the CellaVision to manual differential counts by the standard reference method:

- Kratz *et al.* (76) reported that the pre-classification made by the CellaVision®DM96 was more accurate for neutrophils, lymphocytes and monocytes (92.5%, 96.4% and 81.4% correctly classified respectively) than for eosinophils (63.2%). They concluded that this automated analyser showed similar performance to the manual microscopy.

- Briggs *et al.* (77) similarly documented poor detection of immature granulocytes, eosinophils and basophils, and that the overall pre-classification accuracy of the

CellaVision was 89.2%. In this study is concluded that DM96 accuracy depends on both pathology and the experience of medical operator.

- Cornet *et al.* (78) and Billard *et al.* (79) highlighted a failure to identify lymphoblasts. They noted that the CellaVision®DM96 had difficulty in reliably identifying immature granulocytes, plasma cells and blasts.

- Eilertsen *et al.* (80) determined the effects of pre- and reclassification, as well as, inter-smear and between-technologist variation in blast counts through generalized linear mixed models. They found a significant difference between blast counts of pre-classification and after being reclassified by the clinician (p=0.009).

- Stouten *et al.* (81) analysed 6,945 blood smears containing approximately 1.4 million cells. They concluded that the CellaVision had excellent performance for the five normal leukocyte classes, as well as for blasts.

- Merino *et al.* (82) reported excellent values of correlation for segmented neutrophils, lymphocytes, monocytes and blasts, and good values for eosinophils, basophils and plasma cells when comparing the pre-classification results of the CellaVision®DM96 and the manual microscopic review. Furthermore, the correlation for the reclassification was very good for promyelocytes and myelocytes, intermediate for reactive lymphocytes and erythroblasts, and low for metamyelocytes. They also highlighted that DM96 is not able to reclassify neoplastic lymphoid cells, which has to be done by the clinical pathologist.

The conclusions from these studies were the good correlation and concordance of the application of the digital image analysis over the traditional method by direct microscopy, but highlighting the need of subsequent validation and review by an expert physician. The reduction of the time spent in the analysis by improving the workflow, the efficiency and the quality were also observed. In the end, several of these papers underlined that the CellaVision®DM96 cannot automatically classify abnormal neoplastic cells, and for example when samples from patients suspected to have leukaemia may require review by manual microscopy. It has been reported that further improvement must be done regarding the detection and classification of specific pathological cell types. This is why, the automatic morphological recognition of blast cells is the problem that has given rise to the study of the machine learning approaches of this thesis.

3.2. Machine Learning Approaches for Blast Cell Classification

As it was mentioned above, automated analysers based on digital image analysis are used in laboratories, which pre-classify PB normal cells based on cell image morphological features such as size, shape, colour and texture. However, the morphological discrimination of abnormal PB cells by these devices is still an unresolved problem. Moreover, these analysers cannot discriminate between myeloid and lymphoid blast cell lineages, which is essential for choosing the suitable initial treatment.

3.2.1. Classical Machine Learning Approaches

Multiple efforts have been made to develop image-based automated systems for leukaemia detection. Indeed, many researchers have highlighted that developing image analysis systems for computer-assisted interpretations of PB smears is of a great importance (83–85). However, based on literature review, it has been found that there are few examples of automated systems which can analyse and classify blast cells, and most of these systems are still in development stages (86,87). In addition, all these previous studies have been mainly focused on comparing various classification techniques, and improving nucleus and cytoplasm segmentation procedures, instead of trying to obtain morphological variables and quantitative characteristics useful for cell type evaluation and standardization.

The main steps in image-based recognition systems include image segmentation, feature extraction/selection and classification. Segmentation is the core step in processing digital images, while steps related to features and classification lie within the area of machine learning. Machine learning is a branch of artificial intelligence since it enables the extraction of meaningful patterns from data (in this case from images). Machine learning algorithms are potentially useful components of computer-aided diagnostic support systems, which can help physicians interpret medical imaging findings and reduce interpretation times (88).

3.2.1.1. Segmentation

The segmentation consists of separating the different parts of an image without overlapping. These parts are known as regions of interest (ROIs). This procedure is crucially important to be successful since the accuracy of the succeeding stages (feature extraction and classification) is totally dependent on it. This is why is one of the most complex tasks in image processing and computer vision. Moreover, when it concerns PB cell segmentation there is extra difficulty because of the complicated morphology of the different types of cells and the problems caused by variations in the staining of blood films.

Regarding PB cells, segmentation aims to separate the whole cell from the background and also separate their elements (nucleus and cytoplasm). To understand the essence of segmentation, digital images have to be seen as grids of rectangular pixels. Typically, colour images are decomposed into three colour components: red, green and blue. Each colour component results in an individual greyscale image with a continuous distribution between black at the weakest intensity (0) to white at the strongest (255). Segmentation techniques are generally based on exploiting two basic properties of pixel intensity values: discontinuity and similarity. Discontinuity looks for abrupt intensity changes to detect the borders of the parts to be segmented. Similarity identifies regions with pixels having similar values according to prescribed thresholds. The final result of any segmentation method is a set of binary images, commonly known as *masks*, one for each ROI. Each *mask* contains the ROI as a bounded white region over a black background (see Figure **3.2**) (89).

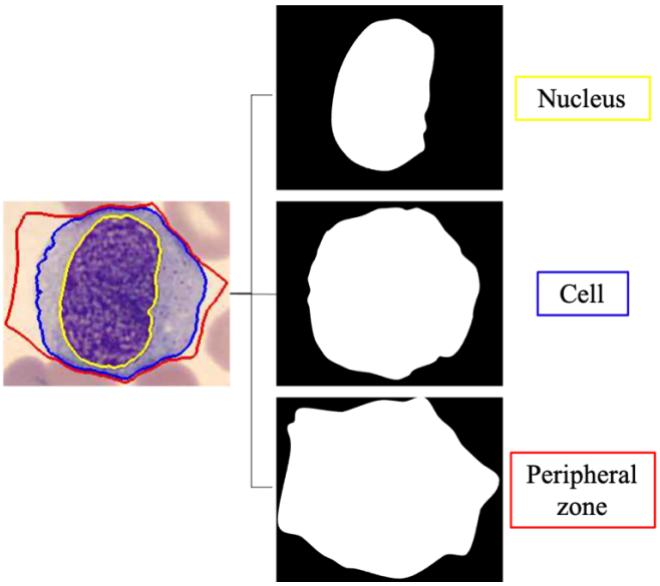


Figure 3.2. Example of a segmented cell monoblast. The original image (left) shows the three regions of interest: nucleus (yellow line), cell (blue line) and peripheral zone around the cell (red line); and their respective masks (right).

3.2.1.2. Feature Extraction/Selection

Feature extraction is the process of obtaining a set of quantitative descriptors for each ROI *mask*. These features allow to obtain quantitative measurements to characterize cells, which are based on morphological variables such as size, nucleus-cell relation, nucleus outline, nucleus eccentricity, chromatin characteristics, basophilic cytoplasm and other. Three main classes of features are used: geometric, colour and texture.

Geometric features are easy to interpret because they have direct intuitive relation with the visual observations and descriptions made by pathologists. They are quantitative geometric interpretations of cell and nucleus shapes, and the most common parameters include *area*, *perimeter*, *circularity*, *diameter*, *nuclear eccentricity*, *elongation*, *roundness*, *convexity*, *nucleus-cell ratio*, and others (1). New geometry-based features may be proposed to quantify more complex characteristics. For instance, our group proposed a novel descriptor in (90) to describe the cytoplasmic profile or *hairiness* in villous lymphocytes. Another new geometric feature named *RBC proximity* was proposed for the first time by our group as an indirect measure of the amount of cell cytoplasm adhered to red blood cells, which resulted useful for the recognition of reactive lymphocytes (91).

Colour is a physical property very helpful in characterising blood cells and closely related to the visual cell appearance. RGB is the colour space most commonly used. However other colour spaces have been used to explore the rich amount of colour information in malignant blood cells, such as CMYK in which the colour components are cyan (C), magenta (M), yellow (Y) and black (B); HSV space is defined by hue (H), saturation (S), and brightness (V); and Lab or Luv spaces use lightness (L) and chromaticity (a, b, u, v). As mentioned before, each colour component results into an individual greyscale image. Colour features are calculated from the histogram obtained from each ROI *mask* of each colour component of its greyscale image. This plot displays the number of pixels corresponding to different intensity values within non-overlapped intervals covering the whole intensity range from 0 (black) to 255 (white). For each histogram, six statistical features are calculated: *mean*, *standard deviation*, *skewness*, *kurtosis*, *energy* (uniformity) and *entropy* (variability) (1,91).

Colour features are involved in the quantification of cytoplasmic basophilia and granulation, both very important morphological characteristics in the classification of myeloid and lymphoid cells (92). However, the interpretation of the nuclear texture by visual observation is not that easy, because variations in fixation and staining have a high influence on chromatin density. Chromatin distribution reflects DNA nucleus organization which contains important cellular diagnostic and prognostic information (91).

Texture describes spatial patterns of material, colour or intensity. In digital image it is quantitatively defined by uniformity, density, pixel tone and their spatial relationships, among others. Such analysis is not an easy task and it is usually performed following two main approaches: the granulometry and the grey level co-occurrence matrix (GLCM). Granulometry measures the size distribution of dark and bright particles in an image through mathematical morphology operations (93). These operations are visualized by means of the so-called granulometric and pseudogranulometric curves. Examples of texture features calculated from

both curves are *mean, standard deviation, skewness* and *kurtosis*. Our group presented a study (94) in which the granulometry was applied for the morphological differentiation of 12 types of abnormal lymphoid cells. The GLCM is defined as the probability that pairs of neighbouring pixels have similar intensities (95). Many measurements can be calculated from the GLCM of a digital image, which provide more information about the texture of the ROI such as *correlation, homogeneity, maximum probability,* and others. This matrix has been widely used for texture quantification in medical imaging, such as ultrasound images for solid neoplasms (96) and in bone marrow images to distinguish erythrocyte precursor cell stages (97).

After feature extraction, it is necessary to reduce the amount of information by selecting a reduced number of features with the purpose of decreasing complexity and computation time. Having too many features can lead to overfitting rather than learning the true basis of a decision (98,99).

3.2.1.3. Classification

The classification step consists of the application of different algorithms to recognize the different cells from their extracted features. There are many techniques, being the most common those involving support vector machines, k-nearest neighbours, decision trees, the naïve Bayes algorithm and neural networks.

In the cell recognition problem, pathologists have to supply the dataset of images truly identified and labelled a priori to train the model in such a way that the model knows the true group to which the cell images belong. This kind of training is called supervised. The training is usually performed through an iterative optimization algorithm which searches for the optimal parameters that minimize the error between the training classification and the true labels. In addition, an important part in any machine learning system is to combine the training process with a validation to evaluate the performance of the classifier using a new set of images. The idea is to assess to what extend the classifier is able to predict the true class for each image of the validation set and modify the structure of the classifier until the validation is considered satisfactory. The method most widely used to do so is the so-called *cross-validation*. It consists of randomly decomposing the training set into a number of equal partitions which do not share images. Then, one subset is left apart and the remaining are used to train the classifier. Once it is trained, the classifier is applied over the first subset to validate its performance. This training/validation process is consecutively repeated until all the subsets are used. The classifier performance is typically measured with the so-called confusion matrix and ROC (Receiver Operating Characteristics) curve. Overall accuracy, sensitivity and specificity are usual quantitative quality indexes to decide whether the classifier is finally ready to be implemented in an operational mode (89).

Accuracy in classification problems is the number of correct predictions made by the model (classifier) over all the predictions made. In our case is the percentage of images correctly classified in their true category. It is a good measure when the target classes are nearly balanced.

Precision measures the proportion of cases which were predicted as having a condition (e.g., acute leukaemia), that actually had acute leukaemia. If we translate this into our case of concern, it provides the percentage of images that were classified, for example, as myeloblasts and that actually were myeloblasts.

Recall, *sensitivity* or *true positive rate* (TPR) measure the proportion of cases which actually had acute leukaemia that was predicted by the algorithm as having acute leukaemia. For example, the percentage of myeloblast images that were correctly classified by the model. Sensitivity provides information about a classifier's performance with respect to false negatives (how many the classifier missed), while precision gives information about its performance with respect to false positives (how many the classifier caught).

Specificity or *true negative rate* (TNR) provide the proportion of cases that did not present the condition of study and were predicted by the model as non-having it (e.g., not having acute leukaemia).

ROC curves give us the ability to assess the performance of a classifier over its entire operating range. The most widely-used measure is the area under the curve (AUC). ROC curves can be used to select a threshold for a classifier which maximises true positives while minimising false positives.

3.2.2. Deep Learning Approaches

Deep learning is a new and popular area of research that is yielding impressive results and growing fast. It is a sub-field of machine learning, whose approach is learning representations from data using the so-called neural networks which emulate the structure and function of human brain neurons. Deep learning does not rely on the extraction and selection of handcrafted features, saving most of the computational steps. However, it requires more sophisticated hardware resources and larger datasets to train a model. The late explosion of deep learning has shifted the focus towards new classification models for leukaemia detection. The application of these new automated systems will increasingly become a part of clinical practise in the coming years in the field of haematological malignancy (100,101).

To date, the most successful models for image analysis are convolutional neural networks (CNNs). CNNs are multi-layer architectures designed to recognize visual patterns. Their structure

mainly contains three different types of layers: 1) convolutional layers; 2) pooling layers and 3) fully connected layers.

The convolutional layer transforms an input to an output through a convolution-based mathematical operation. As inputs are images, both inputs and outputs are volume structures with fixed width, height and depth. Width and height are defined by the number of pixels of images, while depth is defined by the number of planes (colour components). For example, a RGB image has a depth of three. Once the input layer reads the pixels of the image, convolutional layers detect specific patterns using multiple filters that slide across the image. The filter size (*kernel*) determines how many pixels are looked at the same time, whereas the *stride* indicates how many pixels the filter moves to the right when detecting patterns in the image. The higher the *stride* is, the less detailed patterns can be captured. At the end, all these matrices resulting from sliding a filter over the three planes of the image are summed to obtain one feature map. The final output of the convolutional layer is the set of all the feature maps resulting from all the layer's filters. Figure ***3.3*** illustrates the structure of a convolutional layer with a RGB image as example.

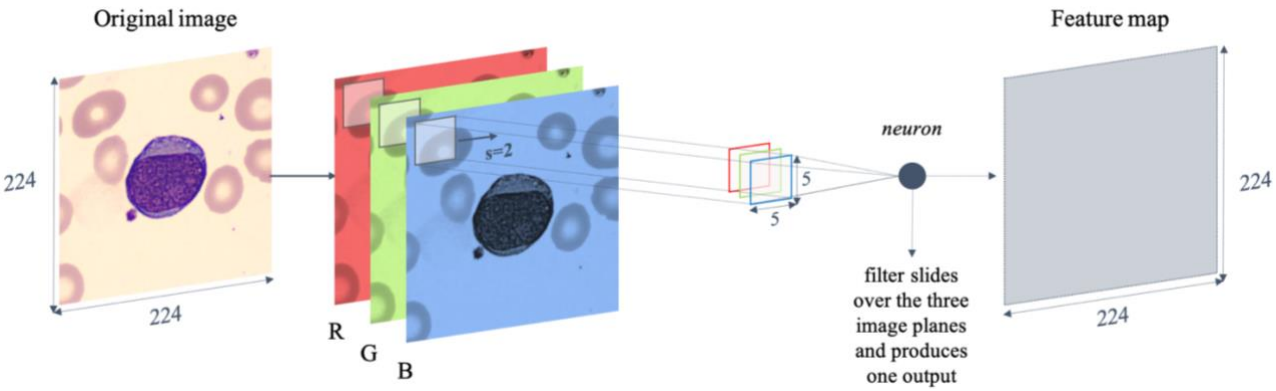


Figure 3.3. Convolutional layer structure. An original RGB image of size 224x224 (*width x height*) is decomposed into its number of planes (*depth*=3). In this example the *filter size* is 5x5 and *stride* (s) is equal to 2, meaning that the filter scans 25 pixels each time while moving 2 pixels to the right until the entire plane is completed. As a result of sliding the filter over the three planes a feature map is produced.

After a convolutional layer it is usual to find specialized layers (*activation layers*) able to amplify the important features. The most common activation function is the rectifier linear unit (ReLU). It converts all negative input values to zero while keeping all positive inputs the same, which helps to increase the learning speed of the network and the classification accuracy.

CNNs could contain one or consecutive convolutional layers creating convolutional blocks. After each convolutional layer or block, it usually comes a *pooling layer* to reduce the size of feature maps. This reduction preserves important information while eliminates irrelevant details, reducing the sensitivity of the model to shifts and distortions. The pooling function most widely used is the so-called *max pooling*, which reduces feature maps size by taking the maximum values of the convolution and, thus, rewarding the convolution function that best extracts the important features of the image. After the successive convolutional and pooling layers, the final result is a set of relevant quantitative features representative of the input image. Afterwards, fully connected layers coupled at the end of the network are trained to learn how to combine these features to perform the final classification. This is done by assigning a probability to each possible class and predict the class with the highest score. Although these are basic elements in all CNNs, the number and how these layers are arranged can vary enormously among the different architectures.

Regarding the training of a neural network, it is an iterative process which typically consists of two phases. Firstly, in a forward phase all input images are passed through the network. Once each image is classified, the difference between their label (ground truth) and the prediction made by the network is measured using a loss function. Secondly, in a backpropagation phase the parameters (*weights*) of the network are updated iteratively so that the class scores are consistent with the true labels. The goal is to minimize the loss function by calculating its gradient. This two-phase process usually needs to be repeated for several cycles (*epochs*) to obtain an optimum trained model. Moreover, the effectiveness of training is influenced by the *learning rate*, an important hyperparameter which indicates how well we are adjusting the *weights* of our network while minimizing the loss function. Small values of *learning rate* make the model converge slowly, whereas large values make it to diverge.

Similar to the above-mentioned classical models, the training process needs to be combined with a validation to assess the performance of the network. This could also be done by using the *cross-validation* approach. A common approach in deep learning is the *hold-out*, which consists of splitting the whole dataset into training and validation, so that the training set is used to fit the model and the validation set is used to provide a punctual evaluation at the end of each *epoch*, which helps to set the hyperparameters of the model. At the end, the performance of models is evaluated with confusion matrices and quality indexes such as accuracy, sensitivity and specificity.

CHAPTER 4

Automatic Recognition of Different Types of Acute Leukaemia in Peripheral Blood by Image Analysis

This chapter presents a paper which develops a methodology based on traditional machine learning algorithms to predict the diagnosis of acute leukaemia using peripheral blood cell images. In this publication, we propose a system which has integrated feature selection after image segmentation for the automatic recognition of myeloid and lymphoid blasts, abnormal promyelocytes, reactive lymphocytes, lymphocytes and monocytes. We also present a detailed analysis of the most relevant features to understand their importance in blast cell detection.

Title: Automatic recognition of different types of acute leukaemia in peripheral blood by image analysis

Authors: Laura Boldú [(1)], Anna Merino [(1)], Santiago Alférez [(2)], Angel Molina [(1)], Andrea Acevedo [(2)] and José Rodellar [(2)]

Affiliations: [(1)] Biochemistry and Molecular Genetics, Biomedical Diagnostic Centre, Hospital Clinic of Barcelona. Barcelona, Catalonia. Spain. [(2)] Mathematics, EEBE, Technical University of Catalonia. Barcelona, Catalonia. Spain.

Citation: Boldú L, Merino A, Alférez S, Molina A, Acevedo A, Rodellar J. Automatic recognition of different types of acute leukaemia in peripheral blood by image analysis. *Journal of Clinical Pathology*, 2019;72(11):755-761. doi: 10.1136/jclinpath-2019-205949. PMID: 31256009.

Quality index: JCR 2019 Pathology, *Impact factor*: 2.460, Q2 (30/78).

Original article

Automatic recognition of different types of acute leukaemia in peripheral blood by image analysis

Laura Boldú,[1] Anna Merino,[1] Santiago Alférez,[2] Angel Molina,[1] Andrea Acevedo,[2] José Rodellar[2]

[1]Biochemistry and Molecular Genetics, Biomedical Diagnostic Center, Hospital Clinic of Barcelona, Barcelona, Catalonia, Spain
[2]Mathematics, EEBE, Technical University of Catalonia, Barcelona, Catalonia, Spain

Correspondence to
Dr Anna Merino, Biochemistry and Molecular Genetics. Biomedical Diagnostic Center, Hospital Clinic of Barcelona, Barcelona, Catalonia, Spain; amerino@clinic.cat

Received 3 May 2019
Revised 7 June 2019
Accepted 8 June 2019

To cite: Boldú L, Merino A, Alférez S, *et al*. *J Clin Pathol* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/jclinpath-2019-205949

ABSTRACT

Aims Morphological differentiation among different blast cell lineages is a difficult task and there is a lack of automated analysers able to recognise these abnormal cells. This study aims to develop a machine learning approach to predict the diagnosis of acute leukaemia using peripheral blood (PB) images.

Methods A set of 442 smears was analysed from 206 patients. It was split into a *training set* with 75% of these smears and a *testing set* with the remaining 25%. Colour clustering and mathematical morphology were used to segment cell images, which allowed the extraction of 2,867 geometric, colour and texture features. Several classification techniques were studied to obtain the most accurate classification method. Afterwards, the classifier was assessed with the images of the *testing set*. The final strategy was to predict the patient's diagnosis using the PB smear, and the final assessment was done with the cell images of the smears of the *testing set*.

Results The highest classification accuracy was achieved with the selection of 700 features with linear discriminant analysis. The overall classification accuracy for the six groups of cell types was 85.8%, while the overall classification accuracy for individual smears was 94% as compared with the true confirmed diagnosis.

Conclusions The proposed method achieves a high diagnostic precision in the recognition of different types of blast cells among other mononuclear cells circulating in blood. It is the first encouraging step towards the idea of being a diagnostic support tool in the future.

INTRODUCTION

Acute leukaemia (AL) are some of the most common neoplastic diseases affecting both adults and children. The starting point for clinical pathologists in the diagnosis of most AL is the detection of blast cells in peripheral blood (PB). For this reason, morphology is still important since a reliable quick diagnosis can be made using basic haematological parameters combined with microscopic investigation.

Automated analysers based on digital image analysis are used in laboratories, which pre-classify PB normal cells. However, the morphological discrimination of abnormal PB cells by these devices is still an unresolved problem, and some authors reported accuracies in blast classification around 76%.[1–3] Moreover, these analysers cannot discriminate between myeloid and lymphoid blast lineages, which is essential for choosing the suitable initial treatment.

Multiple efforts have been made to develop image-based automated methods for classifying blasts, some of them proposing the use of machine learning techniques.[4–16] Most of these previous papers focused on the automatic recognition of myeloblasts vs. lymphoblasts,[9 16] lymphoblasts vs. lymphocytes,[5 10 11] and lymphoblasts or myeloblasts vs. leukocytes.[7 12 13 15] Nevertheless, automatic distinction between blasts and reactive lymphocytes could be also difficult since both share some morphological similarities, as it has been highlighted in some external quality control surveys.[17 18] This is why a previous paper by the authors[19] included for the first time reactive lymphocytes for their automatic recognition alongside myeloblasts and lymphoblasts.

The objective of this paper is to develop a machine learning system to predict the diagnosis of acute leukaemia using PB images. The system input is a set of cell images of an individual smear, and the output is the prediction of one of the following diagnoses: myeloid leukaemia, promyelocytic leukaemia, lymphoid leukaemia or infection. It will be achieved through the automatic recognition of different blasts (myeloid and lymphoid origin) and pathological promyelocytes, reactive lymphocytes and other normal mononuclear cells, such as lymphocytes and monocytes.

Such a high number of groups has not been jointly addressed before. This adds more complexity to the challenging automatic classification of blasts because of the overlapping morphological characteristics they exhibit.[20–22] A particular major challenge is the detection of pathological promyelocytes, whose patients could present serious haemorrhagic accidents and die if no treatment is initiated promptly.[23 24]

MATERIALS AND METHODS

This study was performed following two stages: (1) classifier development, and (2) diagnostic system design and assessment.

Image sets

Blood samples, collected in EDTA, were automatically prepared using the slide maker–stainer SP1000i (Sysmex, Kobe, Japan) and stained with May Grünwald-Giemsa. Digital images of PB cells were acquired by the CellaVision®DM96 (CellaVision, Lund, Sweden) (363×360 pixels) from the routine workload of the Core Laboratory of the Hospital Clinic of Barcelona. PB cell images were identified by pathologists (AMe and AMo)

Table 1 Diagnosis of the patients in the respective sets of study: AL according to the WHO 2016 classification and viral or other infections.

			Training	Testing
Diagnosis	**AL subtype and complementary tests**	**P**	**I**	**I**
AML with recurrent genetic abnormalities	APL with *PML-RARA* (HLA-DR–, CD34–, CD13+, CD33+)	9	597	157
	AML with t(6;9) (p23;q34.1); *DEK-NUP214* (CD45 weak, CD34+, HLA-DR weak, CD117 weak, CD13+, CD33+)	1	76	25
	AML with inv(3) (q21.3q26.2) (CD45 weak, CD34+, HLA-DR+, CD117+, CD13 weak, MPO weak)	1	80	30
	AML with t(9;11) (p21.3;q23.3); *MLLT3-KMT2A*	3	144	65
	AML with mutated *NPM1* (HLA-DR+, CD117+, CD13+, CD33+, CD123+)	4	182	77
AML with myelodysplasia-related changes	CD45 weak, CD34+, HLA-DR+, CD117+, CD13+, CD33+, CD123+	9	376	140
AML, NOS	Acute monoblastic/monocytic leukaemia (HLA-DR+, CD13+, CD33+, CD64+, CD4+, CD36+, CD11b+)	17	657	357
ALL-B/lymphoma with recurrent genetic abnormalities	ALL-B/lymphoma with t(9;22) (q34.1;q11.2); *BCR-ABL1* (CD123+, CD19+, CD22+, CD79a+)	7	135	99
ALL-B/lymphoma, NOS	CD19+, CD22+, CD79a+, CD10+, CD20+, CD34+	8	402	143
	CD45 weak, CD19+, CD79a+, TdT+, CD10–, CD20–	8	400	146
Viral or other infections		53	944	264
Total		120	3,993	1,503

The table also indicates the diagnosis and AL subtype with results of the most relevant complementary tests.
AL, acute leukaemia; ALL-B, acute lymphoid leukaemia;AML, acute myeloid leukaemia; APL, acute promyelocytic leukaemia;HLA-DR, Human leucocyte antigen; I, number of images; MPO, Myeloperoxidase; NOS, not otherwise specified; P, number of patients.

according to their morphological characteristics. Patients' diagnoses were confirmed by the integration of all supplementary test results: flow cytometry, cytogenetics and molecular biology. Patients with AL were diagnosed by the Clinic Haematology Service of the Hospital following the WHO 2016 classification.[25] Table 1 shows the number of images and patients with AL and infections included in *training* and *testing sets*.

A total of 442 PB smears from 206 patients were analysed: 86 corresponding to outpatients (not admitted to the hospital) without any haematological disease, in which cell blood count and smear review were normal, 53 with viral or other infections and 67 with AL: myeloid (18) promyelocytic (9) monocytic (17) and lymphoid (23). From each smear, pathologists selected those cell images corresponding to the confirmed diagnosis, obtaining a total of 7,468 images to work with. The *training set* was arranged with 75% of these smears (332 with 5,493 images), distributed in seven cell groups: lymphocytes (n=844) and monocytes (n=656) as the control group, reactive lymphocytes (n=944), myeloblasts (n=858), monoblasts (n=657), pathological promyelocytes (n=597) and B-lymphoblasts (n=937).

The *testing set* included the remaining 25% of the smears (110 with 1,975 images), distributed as follows: lymphocytes (n=245), monocytes (n=227), reactive lymphocytes (n=264), myeloblasts (n=337), monoblasts (n=357), pathological promyelocytes (n=157) and B-lymphoblasts (n=388).

Stage 1: classifier development

Image-based recognition systems include image segmentation, feature extraction/selection and classification (see figure 1). In this work, we used an automatic segmentation procedure

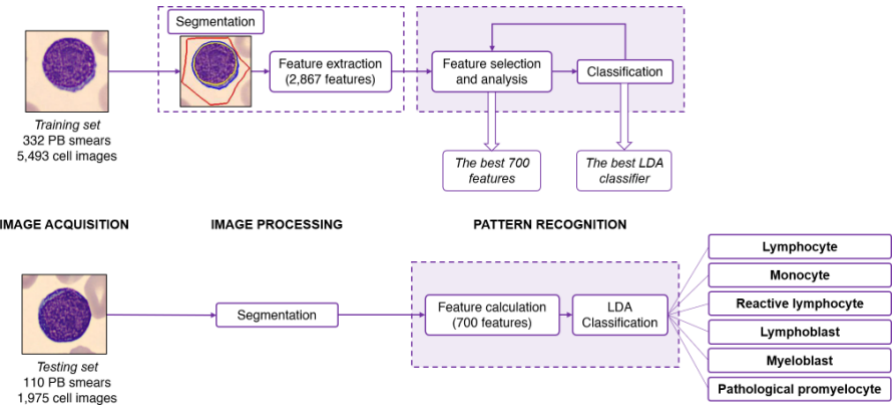


Figure 1 Diagram illustrating the steps followed to develop the classifier. Firstly, the *training set* of 332 smears (5,493 cell images) is processed through the automatic segmentation procedure to obtain a database of features, which is used to tune the best classifier through an iterative process involving 5-fold cross validation, where the most relevant features are determined by the accuracy of the classifier. The final recognition module includes a classifier based on LDA, whose input is the set of the 700 most relevant features. Secondly, the *testing set* of new 110 smears (1,975 cell images) is processed following the preceding steps. This final stage consists of the assessment of the system for cell classification. LDA, linear discriminant analysis; PB, peripheral blood.

recently published by our group. This algorithm uses the image colour information through fuzzy clustering of different colour components and the application of the watershed transformation with markers. A more detailed explanation could be found in Alférez *et al*.[26]. This segmentation allowed us to obtain three regions of interest (ROIs): the nucleus, the whole cell and the peripheral zone around the cell. A fourth ROI was obtained for the cytoplasm by the difference between the cell and the nucleus.

From each ROI, quantitative features were calculated for the purpose of morphological analysis and further classification. In this study, 28 geometric and 2,839 colour and texture features were calculated for each segmented cell. Geometric features give quantitative geometric interpretations of cell and nucleus shapes, whose definitions can be found in Alférez *et al*.[27].

Colour is a physical property very helpful in characterising blood cells. An image is separated into three colour components: red, green and blue. To explore the rich amount of colour information in malignant blood cells, more colour spaces were used. Each colour component results in an individual greyscale image, which contains shades of grey, varying from black at the weakest intensity (0) to white at the strongest (255).[28] .

Colour features are calculated from the histogram of a greyscale image. This plot displays the number of pixels corresponding to different intensity values. For each histogram, six statistical features were calculated: *mean*, *standard desviation*, *skewness*, *kurtosis*, *energy* and *entropy*.[27 29]

Texture in digital images is identified by uniformity, density, pixel tone and their spatial relationships, among others.[30 31] Statistical measures are calculated from the grey level co-occurrence matrix (GLCM), which represents the probability that pairs of neighbouring pixels have similar intensities.[32–34] Examples of these statistical parameters are *correlation*, *cluster shade* and *maximum probability*. In this study, we also considered more mathematical approaches to characterise texture, whose explanation can be found in Alférez *et al*.[27 35 36]

After feature extraction, each cell image is uniquely described by a set of numerical features. Using this set, the classifier output is the prediction of the class to which the cell image belongs. In machine learning, it is important to select a reduced number of features to decrease complexity and computation time.[37] In this study, selection techniques were employed to determine the most relevant features by using conditional mutual information maximisation criteria.[38] This procedure was implemented within the tuning of the classifier parameters, as shown by the feedback arrow in figure 1, to obtain the most relevant features from the best classification.

Several classification techniques were studied to face the challenging number of cell types under study: linear discriminant analysis (LDA), k-nearest neighbour, naive Bayes, support vector machine and random forest. A 5-fold cross validation technique was performed using the images of the *training set*. The overall classification accuracy (ratio of images correctly classified in their true category) was used to choose the best features and classifier. This classifier was further validated using the images of the *testing set*, which were not previously used in the training step.

Furthermore, to interpret the most relevant features involved in the discrimination of the cells under study, statistical analysis was performed using R software[39] over the selected features: Kolmogorov-Smirnov test for residuals normality, Fligner-Killeen test for variance homogeneity, Kruskal-Wallis test to calculate *p*-values and Kruskal-Wallis test after Dunn test by applying Bonferroni adjustment for multiple comparisons to see which pairs of cells could be quantitatively differentiated.

Stage 2: diagnostic system design and assessment

Once the system was ready for the classification of individual cell images, our strategy was to predict patient's diagnosis using the PB smear, trying to emulate the way pathologists interpret results in clinical laboratories. The system input was a set of cell images of an individual smear and the output was the prediction of one of the following diagnoses (see figure 2): myeloid leukaemia, promyelocytic leukaemia, lymphoid leukaemia or infection. The diagnosis was made by identifying the cell class that predominated in the smear. This required a previous step to establish a threshold value such that the diagnosis will be predicted by identifying the cell class with the percentage of images classified above this value. We used all the smears of the *training set* to perform a multiclass Receiver Operating Characteristic (ROC) analysis with the statistical R software. The result was the optimal threshold in terms of the percentage of images correctly classified by the system over the total number of images in the smear, such that it guarantees the largest number of smears correctly predicted in its true diagnosis within the *training set*. Once the threshold was determined, the automatic recognition system was assessed using the smears of the *testing set*, which were not used before.

A confusion matrix was obtained to evaluate the system performance, and sensitivity and specificity were calculated.

RESULTS

Stage 1: classifier development

For the classifier development, we extracted 2,867 quantitative features, which were ranked based on a relevance criterion. As shown in figure 3, we trained and tuned five classifiers using different numbers of the best-ranked features and displayed the corresponding accuracy. The highest accuracy was achieved with the LDA classifier when using the best 700 features.

Table 2 lists the 10 most relevant features. Among them, three features were geometric and seven represented colour and

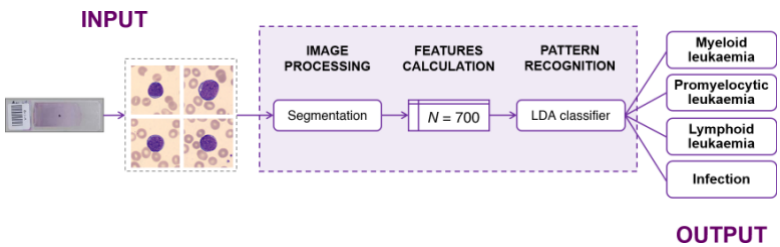


Figure 2 Diagram illustrating the design of the diagnostic system to predict the patient's diagnosis using the peripheral blood smear. The system input is a set of cell images of an individual smear, and the output is the prediction of one of the following diagnoses: myeloid leukaemia, promyelocytic leukaemia, lymphoid leukaemia or infection. LDA, linear discriminant analysis.

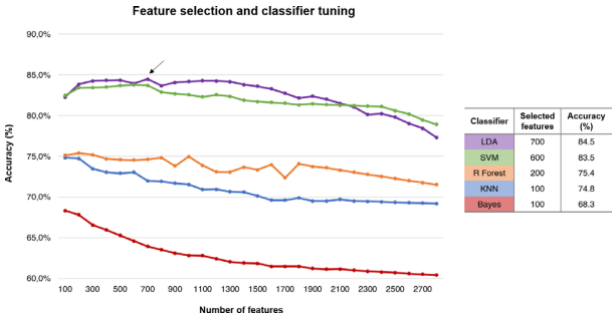


Figure 3 Plot of the five classifiers trained and tuned with all the 2,867 features to obtain the combination of the best number of features and the classifier with the highest overall classification accuracy. The x-axis shows the number of features and y-axis shows the accuracy values (in percentage) for each selected number of features. The highest classification accuracy (marked with an arrow) is achieved with the selection of the most relevant 700 features with LDA classifier. Bayes, naive Bayes; KNN, k-nearest neighbour; LDA, linear discriminant analysis; R forest, random forest; SVM, support vector machine.

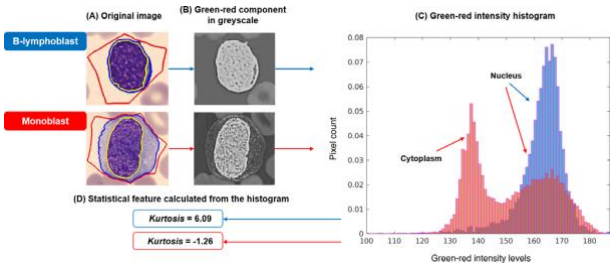


Figure 4 (A) Original images of B-lymphoblast from acute lymphoid leukaemia and monoblast from acute monocytic leukaemia acquired by the CellaVision®DM96. (B) Green–red component in greyscale. (C) Green–red intensity histogram. (D) Statistical feature calculated from the histogram. The whole cell in the B-lymphoblast image is practically formed by bright spots. It correlates with its histogram (blue), which presents a narrow peak with the pixels very localised in a short intensity range between 155 and 175, so that the kurtosis is higher. In contrast, the histogram of monoblast image (red) shows two lower peaks, which cover many pixels with different levels of intensity. This indicates more information variety due to the presence of the cytoplasm that makes the nucleus/cell relation lower; therefore, the kurtosis of the image is lower.

texture. Eight features were calculated from the whole cell and two from the nucleus.

As an example of a colour statistical feature calculated from the histogram, figure 4 shows the *kurtosis of the green–red of the cell*. It was the best feature to distinguish among the seven cell groups. It considers the green–red pixels and measures the histogram's flatness. We observed that *kurtosis* is related to the nucleus/cell relation. In fact, cells with high relation show higher *kurtosis* values. Figure 4 shows two original images of a B-lymphoblast and a monoblast and the histograms from the green–red component. The histogram of a B-lymphoblast (in blue) presents a narrow and flatter peak with the pixels very localised within a short intensity range, so that the *kurtosis* of the cell is higher. In contrast, the histogram of a monoblast (in red) shows two peaks with lower height, indicating more information variety, corresponding to cytoplasm and nucleus, respectively. Table 2 summarises the results of the statistical analysis performed among the cell pairs in which recognition by morphology is difficult. It shows that *kurtosis* was significantly different between reactive lymphocytes and myeloblasts or B-lymphoblasts ($p<0.0001$), between B-lymphoblasts and myeloblasts ($p<0.0001$), and between myeloblasts and monoblasts ($p<0.0001$).

Regarding geometric features, *nuclear area*, *cellular area* and *Nucleus/Cell (N/C) ratio* were the most representative. Those cells with higher nucleus and bigger cellular area, such as pathological promyelocytes and monoblasts, had higher values for *nuclear* and *cellular areas*. In contrast, either myeloblasts or B-lymphoblasts are smaller cells with smaller nucleus, which was correlated with their lower values for these features. The *N/C ratio* was significantly different between myeloblasts and B-lymphoblasts ($p<0.0001$), between myeloblasts and pathological promyelocytes ($p<0.01$) and between myeloblasts and monoblasts ($p<0.0001$) (see table 2).

The *blue correlation of the cell* is the first among the texture features. Figure 5 shows this statistical feature, which is based on the GLCM. *Correlation* measures how correlated a pixel is to its neighbours over the whole cell. As shown in figure 5, the *blue correlation* for a reactive lymphocyte with lower nucleus/cell relation was high (0.94), since bright intensity pixels are close to each other (corresponding to cytoplasm), as well as those of darker intensities (nucleus). Moreover, the majority of values were grouped in the main GLCM's diagonal, which resulted in

Table 2 *P*-values obtained after applying Kruskal-Wallis post-hoc tests after Dunn by applying Bonferroni type-adjustment

Rank	Feature (type)	RL vs. Lymphoblasts	RL vs. Myeloblasts	Lymphoblasts vs. Myeloblasts	Pathol Prom vs. Myeloblasts	Myeloblasts vs. Monoblasts
1	Kurtosis of the *green–red* of the cell (colour)	<0.0001	<0.0001	<0.0001	NS	<0.0001
2	Area of the nucleus (geometric)	<0.0001	NS	<0.0001	<0.0001	<0.0001
3	Area of the cell (geometric)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
4	Correlation of the *blue* of the cell (texture)	<0.0001	<0.0001	<0.0001	NS	<0.0001
5	Mean of the *blue* of the nucleus (colour)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
6	Energy of the *yellow* of the cell (colour)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
7	Cluster shade of the *blue–purple* of the cell (texture)	<0.0001	<0.0001	0.0021	0.0428	<0.0001
8	Skewness of the *green–red* of the cell (colour)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
9	Nucleus/Cell ratio (geometric)	<0.0001	<0.0001	<0.0001	0.026	<0.0001
10	Energy of the *magenta* of the cell (colour)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

Multiple comparisons of the quantitative values from the 10 most relevant features corresponding to five pairs of cell subsets are shown, in which their recognition in peripheral blood smears by morphology usually is difficult. P<0.05 was considered significant.
. NS, non-significant; Pathol Prom, pathological promyelocyte; RL, reactive lymphocyte.

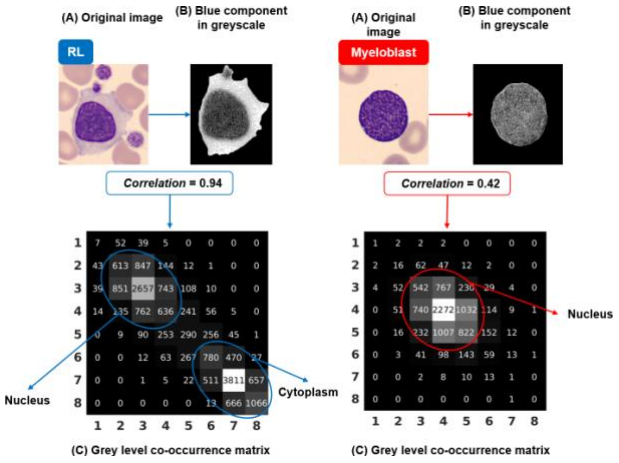


Figure 5 (A) Original images of RL and myeloblast from acute myeloid leukaemia acquired by the CellaVision®DM96. (B) Blue component in greyscale. (C) Grey level co-occurrence matrices obtained in MATLAB®. The RL image shows more bright intensity values due to the presence of its cytoplasm and darker intensities corresponding to the nucleus. The majority of values are concentrated in the main diagonal, which results in a higher correlation value. In contrast, the myeloblast image shows many pixels with different levels of grey intensities. This indicates more information variety, more intensity differences between neighbouring pixels and, therefore, lower correlation. RL, reactive lymphocyte.

a higher *correlation* value. In the opposite way, figure 5 shows a myeloblast with the nucleus occupying almost the whole cell and presenting many pixels with different levels of grey intensities. This indicates more intensity differences between neighbouring pixels and, therefore, a lower *correlation* value (0.42). As seen in table 2, *blue correlation* showed significant differences between reactive lymphocytes and myeloblasts and B-lymphoblasts ($p<0.0001$), between myeloblasts and lymphoblasts ($p<0.0001$), and between myeloid and monocytic blasts ($p<0.0001$).

Classification by individual cells

The last step for classifier development was a validation of the LDA classifier (with the best 700 features). This was done through a blind classification of all the individual images of the *testing set*. Table 3 shows the confusion matrix that summarises the classification results. The true positive rates shown in the main diagonal were: 97.7% for reactive lymphocytes, 97.6% for lymphocytes, 93% for monocytes, 80.8% for myeloid blasts (myeloblasts and monoblasts), 72.6% for pathological promyelocytes and 78.9% for B-lymphoblasts. The overall classification accuracy was 85.8%. Most of the individual cell images were automatically identified in the group of its true class.

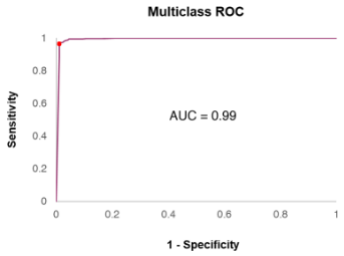


Figure 6 Plot of the multiclass ROC analysis to obtain the best threshold to predict the diagnosis of the subtype of acute leukaemia. The best threshold is defined as the minimum percentage of images required in a cell class to predict a diagnosis. The 50% resulted to be the best threshold (marked with a red dot), obtaining an AUC of 0.99. AUC, area under the curve ; ROC, Receiver Operating Characteristic.

Stage 2: diagnostic system design and assessment

Prior to assessment, figure 6 shows the results of the multiclass ROC analysis performed by analysing all the smears of the *training set*. This curve was obtained by averaging the single ROC curves obtained for each of the six classes under study. It was found that 50% of the cell images correctly classified was the best threshold to predict patient's diagnosis through the smear, obtaining an area under the curve of 0.99.

For example, such threshold may be interpreted in such a way that if more than 50% of the cell images of a smear are classified as myeloblasts, the predicted diagnosis is acute myeloid leukaemia.

Classification by individual smears

Table 4 shows the classification results for the smears of the *testing set*, considering the 50% threshold. As seen in the main diagonal, 100% of smears corresponding to patients with infections were correctly classified, as well as those containing lymphocytes or monocytes (accuracies of 100% and 97%, respectively). In regard to AL, the true positive rates were 88% for myeloid leukaemia, 85% for lymphoid leukaemia and 80% for promyelocytic leukaemia. The overall classification accuracy of individual smears was 94%. Moreover, sensitivity values above 97% were obtained for normal smears and those related to infections and above 80% for all AL subtypes, while specificity values for all categories were above 96%.

Table 3 Confusion matrix of the classification results (in %) for the images of the *testing set*

		Predicted class					
		Lymphoid blasts	Pathological promyelocytes	Myeloid blasts	Reactive lymphocytes	Lymphocytes	Monocytes
True class	**Lymphoid blasts**	**78.9**	1.5	16.5	2.1	1.0	0.0
	Pathological promyelocytes	14.0	**72.6**	13.4	0.0	0.0	0.0
	Myeloid blasts	14.0	1.6	**80.8**	1.6	0.1	1.9
	Reactive lymphocytes	0.4	0.0	1.1	**97.7**	0.0	0.8
	Lymphocytes	0.0	0.0	0.0	2.0	**97.6**	0.4
	Monocytes	1.3	0.4	1.3	4.0	0.0	**93.0**

Rows indicate the true class and columns represent the predicted class supplied by the classifier. Diagonal values are the true positive rates for each cell type. The overall classification accuracy was 85.8%.

Table 4 Confusion matrix of the classification results (in %) for the individual smears of the *testing set*, taking the threshold of 50% into consideration.

		Predicted diagnosis					
		ALL-B	APL	AML	Infections	Control (LY)	Control (MO)
True diagnosis	ALL-B	**85**	0	15	0	0	0
	APL	0	**80**	20	0	0	0
	AML	12	0	**88**	0	0	0
	Infections	0	0	0	**100**	0	0
	Control (LY)	0	0	0	0	**100**	0
	Control (MO)	0	0	0	3	0	**97**

Rows indicate the true diagnosis and columns represent the predicted diagnosis supplied by the classifier. Diagonal values are the true positive rates for each smear. The overall classification accuracy was 94%.
ALL-B, acute lymphoid leukaemia; AML, acute myeloid leukaemia; APL, acute promyelocytic leukaemia; LY, lymphocyte; MO, monocyte.

DISCUSSION

Morphological review of PB smears is still relevant nowadays in that it provides the primary evidence of a specific haematological or non-haematological diagnosis. In this sense, blood cell morphology has become crucial for the initial diagnosis of AL subtype, especially in acute promyelocytic leukaemia, to apply the suitable treatment.

In this paper, for the first time, we employ a LDA classifier combined with our recently published automated segmentation algorithm[26] and with feature extraction for the automatic recognition of blasts, reactive lymphocytes and other mononuclear cells, as well as for the distinction among myeloblasts, B-lymphoblasts and pathological promyelocytes. The automatic recognition of this wide variety of cell types is a new contribution since it has not previously been accomplished. Previous results showed accuracy in the automatic classification of B-lymphoblasts and lymphocytes.[10] However, these two groups of lymphoid cells are usually easier to differentiate than myeloblasts and lymphoblasts. Other studies tried to automatically detect lymphoblasts[7 13] or myeloblasts[6] among other leukocytes, which show a very different morphology since the nucleus of neutrophils, eosinophils and basophils is lobulated and their cytoplasm shows abundant granules. From a diagnostic perspective, our contribution is relevant since it allows us (1) to diagnose malignant and non-malignant diseases and (2) to recognise the AL subtype.

In previous studies, we have described results about the automatic recognition of different types of abnormal lymphocytes, where high precision has been achieved when selecting 140 features.[1] In regard to leukemic cells, the results of this study demonstrate that the best classification is achieved when the number of features is increased compared to our previously published paper.[19] When we considered smaller numbers of features, the accuracy decreased (data not shown). The reason why this study required a greater number of features may be related to the inclusion of a larger image data set, as well as much different types of blasts and other mononuclear cells.

Regarding feature analysis, our results demonstrated that the *N/C ratio* was one of the most important geometric features. This is consistent with our previous publications, which reported that this feature was also the most relevant for the recognition among several abnormal lymphocytes subsets[40] and blasts.[19] Moreover, in[41] were listed the features most commonly used for the automatic recognition of leukaemia, with *N/C ratio* being one of them. Wahhab *et al*.[42] found that the *N/C ratio* was within the best geometric features to distinguish between lymphoblasts and myeloblasts with an accuracy of 93.6%.

The relevance obtained in colour and texture features, such as *Kurtosis* and *Correlation*, can be explained by the large cytoplasm of reactive lymphocytes, monoblasts and some pathological promyelocytes in comparison with the low cytoplasm of myeloblasts and B-lymphoblasts. The *Mean* provides essential information for cell types with dispersed chromatin. The differences observed among different myeloid blasts can be related to the immature chromatin with visible nucleolus in myeloblasts, which is seen as a thinner texture in monoblasts, compared to the abundant azurophil granulation typical in pathological promyelocytes. Moreover, *Cluster Shade* allows the distinction among cells with thinner and condensed texture. Lower values shown in lymphoblasts and myeloblasts can be associated with their nucleus occupying almost the entire cell, resulting in a cell image with a very condensed texture. Higher values are seen in monoblasts and reactive lymphocytes, which exhibit a thinner texture because of their cytoplasm. The results obtained in this work are in accordance with previous studies, such as that of Wahhab *et al*.,[42] who found that information of the nuclear chromatin distribution pattern could be provided by GLCM features, and cluster shade, correlation, mean and skewness were among the most important texture features to distinguish between lymphoblasts and myeloblasts.

The automatic recognition of blasts has always been addressed considering each cell image as a unit,[4 37] without considering individual smears when arranging sets for both training and assessing the classifier. It may happen that images from the same smear are used in both stages, causing an undesired bias. An innovation of this study is that we arranged a set of smears for the system development and a different one for the assessment. For the first time, PB smear was used as a diagnostic unit, which allowed us to achieve an efficient approach for diagnostic prediction in patients with AL or infection using digital images. Future implementation in clinical laboratories of such technology will entail challenges in the standardisation of PB smear staining to ensure high-quality smears and to minimise diversity among images from different sources.

A satisfactory diagnostic ability was achieved as the system differentiated normal smears from those related to infections and with respect to smears with abnormal leukemic cells (as seen in table 4). It is important to note that the lower accuracy for smears with pathological promyelocytes (80%), in relation to the other AL, is because the remaining 20% were classified as smears from patients with myeloid leukaemia, which is true since promyelocytic leukaemia is a subtype acute myeloid leukaemia.

The key contribution of this work is a system which has integrated feature selection to lead to the most relevant quantitative features to achieve the highest percentage of classification accuracy. Moreover, such detailed analysis of the most relevant features to understand their importance has not previously been done. All these improvements related to methodology have allowed us to achieve high overall accuracy not only when classifying cell by cell but also when classifying simultaneously all the images from a single smear for diagnostic purposes.

For the first diagnostic orientation of leukaemia in the clinical laboratory, it is important the detection of quantitative abnormalities in leucocyte count, haemoglobin level or thrombocyte count, which triggers the smear review. The detection of blast cells circulating in blood is the following step. The integration of laboratory parameters, smear review and an automatic classification system able to confirm blast detection and to discriminate among blasts of different origin may improve the diagnostic orientation of patients with AL. In the end, the proposed method would be a practical tool

Take home messages

- Morphological review of the peripheral blood smear is the first step to detect blast cells, being crucial for the initial diagnosis of acute leukaemias.
- The contribution of this paper is to provide an automatic recognition system able to predict the diagnosis of patients with acute myeloid leukaemia, acute promyelocytic leukaemia, acute lymphoid leukaemia and infection using peripheral blood images.
- An innovation of this study is that the smear is used as a diagnostic unit, which enabled the achievement of a satisfactory diagnostic ability when differentiating normal smears from those related to infections and with respect to smears with abnormal leukemic cells.
- In addition, the proposed system would be a practical tool to assist the pathologist in the initial diagnosis of acute leukaemia during the morphological examination in peripheral blood.

to assist the pathologist in the initial diagnosis of acute leukaemia through the morphological examination of PB cells.

Handling editor Prof Mary Frances McMullin.

Contributors All the authors have fullfilled the criteria of authorship and have done substantial contributions to the study and the interpretation of data. They have revised the work, as well as the current revision of the manuscript.

Funding The study was funded by the Directory of Science, Technology and Innovation of the Ministry of Economy and Competitiveness of Spain (grant number: DPI2015-64493-R)

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

REFERENCES

1. Alférez S, Merino A, Mujica LE, *et al*. Automatic classification of atypical lymphoid B cells using digital blood image processing. *Int J Lab Hematol* 2014;36:472–80.
2. Briggs C, Longair I, Slavik M, *et al*. Can automated blood film analysis replace the manual differential? An evaluation of the CellaVision DM96 automated image analysis system. *Int J Lab Hematol* 2009;31:48–60.
3. Huang H-Q, Fang X-Z, Shi J, *et al*. Abnormal localization of immature precursors (ALIP) detection for early prediction of acute myelocytic leukemia (AML) relapse. *Med Biol Eng Comput* 2014;52:121–9.
4. Alsalem MA, Zaidan AA, Zaidan BB, *et al*. A review of the automated detection and classification of acute leukaemia: coherent taxonomy, datasets, validation and performance measurements, motivation, open challenges and recommendations. *Comput Methods Programs Biomed* 2018;158:93–112.
5. Scotti F. Automatic morphological analysis for acute leukemia identification in peripheral blood microscope images. In: CIMSA 2005 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications. IEEE, 2005:96–101.
6. Kazemi F, Najafabadi TA, Araabi BN. Automatic recognition of acute myelogenous leukemia in blood microscopic images using k-means clustering and support vector machine. *J Med Signals Sens* 2016;6.
7. Gumble PM, Analysis RSV. Classification of acute lymphoblastic leukemia using KNN algorithm. *Int J Recent Innov Trends Comput Commun* 2017;5:94–8.
8. Agaian S, Madhukar M, Chronopoulos AT. A new acute leukaemia-automated classification system. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* 2018;6:303–14.
9. Supardi NZ, Mashor MY, *et al*, CSPA. Classification of blasts in acute leukemia blood samples using k-nearest neighbour. In: *Signal processing and its applications*. 2012 IEEE 8th International Colloquium on, 2012: 461–5.
10. Madhukar M, Agaian S, Chronopoulos AT. New decision support tool for acute lymphoblastic leukemia classification. *Int Conf Image Process Algorithms Syst X Parallel Process Imaging Appl II* 2012..
11. Mohapatra S, Patra D, Satpathy S. An ensemble classifier system for early diagnosis of acute lymphoblastic leukemia in blood microscopic images. *Neural Comput & Applic* 2014;24:1887–904.
12. Agaian S, Madhukar M, Chronopoulos AT. Automated screening system for acute myelogenous leukemia detection in blood microscopic images. *IEEE Systems Journal* 2014;8:995–1004.
13. Putzu L, Caocci G, Di Ruberto C. Leucocyte classification for leukaemia detection using image processing techniques. *Artif Intell Med* 2014;62:179–91.
14. MoradiAmin M, Memari A, Samadzadehaghdam N, *et al*. Computer aided detection and classification of acute lymphoblastic leukemia cell subtypes based on microscopic image analysis. *Microsc Res Tech* 2016;79:908–16.
15. Mishra S, Majhi B, Sa PK, *et al*. Gray level co-occurrence matrix and random forest based acute lymphoblastic leukemia detection. *Biomed Signal Process Control* 2017;33:272–80.
16. Jacob A, Mundackal FA. Automated screening system for acute leukemia detection and type classification. *Int J Adv Res Electr Electron Instrum Eng* 2016;5:995–1004.
17. Gutiérrez G, Merino A, Domingo A, *et al*. EQAS for peripheral blood morphology in Spain: a 6-year experience. *Int J Lab Hematol* 2008;30:460–6.
18. Chabot-Richards DS, Foucar K. Does morphology matter in 2017? An approach to morphologic clues in non-neoplastic blood and bone marrow disorders. *Int J Lab Hematol* 2017;39 Suppl 1:23–30.
19. Bigorra L, Merino A, Alférez S, *et al*. Feature analysis and automatic identification of leukemic lineage blast cells and reactive lymphoid cells from peripheral blood cell images. *J Clin Lab Anal* 2017;31:e22024–9.
20. Merino A, Boldú L, Ermens A. Acute myeloid leukaemia: how to combine multiple tools. *Int J Lab Hematol* 2018;00:00–11.
21. Bain BJ. Leukaemia Diagnosis. In: *Chichester*. 4th ed. UK: Wiley-Blackwell, 2010: 68–73.
22. Porwit A, Vardiman JW. Acute myeloid leukemia with expanded erythropoiesis. *Haematologica* 2011;96:1241–3.
23. Breen KA, Grimwade D, Hunt BJ. The pathogenesis and management of the coagulopathy of acute promyelocytic leukaemia. *Br J Haematol* 2012;156:24–36.
24. Sanz MA, Grimwade D, Tallman MS, *et al*. Management of acute promyelocytic leukemia: recommendations from an expert panel on behalf of the European LeukemiaNet. *Blood* 2009;113:1875–91.
25. Arber DA, Orazi A, Hasserjian R, *et al*. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood* 2016;127:2391–405.
26. Alférez S, Merino A, Acevedo A, *et al*. Color clustering segmentation framework for image analysis of malignant lymphoid cells in peripheral blood. *Med Biol Eng Comput* 2019;57:1265–83.
27. Alférez S. Methodology for automatic classification of atypical lymphoid cells from peripheral blood cell images. *Universitat Politècnica de Catalunya* 2015.
28. Johnson S. *Stephen Johnson on digital photography*. 320. O'Reilly Media Inc, 2006.
29. Merino A, Puigví L, Boldú L, *et al*. Optimizing morphology through blood cell image analysis. *Int J Lab Hem* 2018;40(Suppl. 1):54–61.
30. Ruiz M, Mujica LE, Alférez S, *et al*. Wind turbine fault detection and classification by means of image texture analysis. *Mech Syst Signal Process* 2018;107:149–67.
31. Castellano G, Bonilha L, Li LM, *et al*. Texture analysis of medical images. *Clin Radiol* 2004;59:1061–9.
32. Payerle G. *Bound 2*. . Duke University Press, 1974: 3. 45.
33. Yang X, Tridandapani S, Beitler JJ, *et al*. Ultrasound GLCM texture analysis of radiation-induced parotid-gland injury in head-and-neck cancer radiotherapy: an *in vivo* study of late toxicity. *Med Phys* 2012;39:5732–9.
34. Haralick RM, Shanmugam K, Dinstein Its'Hak. Textural features for image classification. *IEEE Trans Syst Man Cybern* 1973;SMC-3:610–21.
35. Gonzalez RC, Woods RE, eds. *Digital Image Processing*. 3rd ed. Prentice Hall, 2007.
36. Mehrotra R, Namuduri KR, Ranganathan N. Gabor filter-based edge detection. *Pattern Recognit* 1992;25:1479–94.
37. Rodellar J, Alférez S, Acevedo A, *et al*. Image processing and machine learning in the morphological analysis of blood cells. *Int J Lab Hem* 2018;40:46–53.
38. Brown G, Pocock A, Zhao M-J, *et al*. Conditional likelihood maximisation: a unifying framework for information theoretic feature selection. *J Mach Learn Res* 2012;13:27–66.
39. Arriaza Gómez AJ, Fernández Palacín F, López Sánchez MA, *et al*. Estadística básica Con R Y R-Commander. *Univ Cádiz*. 2008:1–160.
40. Puigví L, Merino A, Alférez S, *et al*. New quantitative features for the morphological differentiation of abnormal lymphoid cell images from peripheral blood. *J Clin Pathol* 2017;70:1038–48.
41. Agaian S, Madhukar M, Chronopoulos AT. A new acute leukaemia-automated classification system. *Comput Methods Biomech Biomed Eng Imaging Vis* 2016;1163:1163–12.
42. Wahhab HTA. Classification of acute leukemia using image processing and machine learning techniques. 2015:University of Malaya.

CHAPTER 5

A Deep Learning Model (ALNet) for the Diagnosis of Acute Leukaemia Lineage using Peripheral Blood Cell Images

This chapter consist of a paper in which a deep learning-based system is designed and developed for its integration in a real clinical setting. In this publication, we propose a new predictive system designed with two serially connected convolutional networks to assist pathologists in the diagnosis of acute leukaemia during the blood smear review. The concept of fine-tuning is adopted to examine the effects of four well-known CNNs on model performance, which are trained and evaluated using 5-fold cross validation and hold-out. Furthermore, the proposed system is evaluated in a clinical setting with a new dataset of images from two different hospitals.

Title: A deep learning model (ALNet) for the diagnosis of acute leukaemia lineage using peripheral blood cell images

Authors: Laura Boldú [(1)], Anna Merino [(1)], Andrea Acevedo [(1,2)], Angel Molina [(1)] and José Rodellar [(2)]

Affiliations: [(1)] Hospital Clinic of Barcelona-IDIBAPS. Haematology and Cytology Unit. CORE Laboratory. Biomedical Diagnostic Centre, Spain. [(2)] Technical University of Catalonia, Barcelona East Engineering School, Department of Mathematics, Spain.

Citation: Boldú L, Merino A, Acevedo A, Molina A, Rodellar J. A deep learning model (ALNet) for the diagnosis of acute leukaemia lineage using peripheral blood cell images. *Computer Methods and Programs in Biomedicine*, 2021;202:105999. doi: 10.1016/j.cmpb.2021.105999. PMID: 33618145.

Quality index: JCR 2019 Computer Science, Theory & Methods, *Impact factor*: 3.632, Q1 (16/108).

Computer Methods and Programs in Biomedicine 202 (2021) 105999

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

A deep learning model (ALNet) for the diagnosis of acute leukaemia lineage using peripheral blood cell images

Laura Boldú[a], Anna Merino[a,1,*], Andrea Acevedo[a,b], Angel Molina[a], José Rodellar[b]

[a] Hospital Clínic de Barcelona-IDIBAPS, Haematology and Cytology Unit, CORE Laboratory, Biomedical Diagnostic Centre, Spain
[b] Technical University of Catalonia, Barcelona East Engineering School, Department of Mathematics, Spain

ARTICLE INFO

Article history:
Received 31 July 2020
Accepted 8 February 2021

Keywords:
Leukemia
Deep learning
Convolutional neural networks
Morphological analysis
Blood cell automatic recognition

ABSTRACT

Background and objectives: Morphological differentiation among blasts circulating in blood in acute leukaemia is challenging. Artificial intelligence decision support systems hold substantial promise as part of clinical practise in detecting haematological malignancy. This study aims to develop a deep learning-based system to predict the diagnosis of acute leukaemia using blood cell images.

Methods: A set of 731 blood smears containing 16,450 single-cell images was analysed from 100 healthy controls, 191 patients with viral infections and 148 with acute leukaemia. *Training and testing sets* were arranged with 85% and 15% of these smears, respectively. To find the best architecture for acute leukaemia classification VGG16, ResNet101, DenseNet121 and SENet154 were evaluated. Fine-tuning was implemented to these pre-trained CNNs to adapt their layers to our data. Once the best architecture was chosen, a system with two modules working sequentially was configured (ALNet). The first module recognised abnormal promyelocytes among other mononuclear blood cell images, such as lymphocytes, monocytes, reactive lymphocytes and blasts. The second distinguished if blasts were myeloid or lymphoid lineage. The final strategy was to predict patients' initial diagnosis of acute leukaemia lineage using the blood smear review. ALNet was assessed with smears of the *testing set*.

Results: ALNet provided the correct diagnostic prediction of all patients with promyelocytic and myeloid leukaemia. Sensitivity, specificity and precision values of 100%, 92.3% and 93.7%, respectively, were obtained for myeloid leukaemia. Regarding lymphoid leukaemia, a sensitivity of 89% and specificity and precision values of 100% were obtained.

Conclusions: ALNet is a predictive model designed with two serially connected convolutional networks. It is proposed to assist clinical pathologists in the diagnosis of acute leukaemia during the blood smear review. It has been proved to distinguish neoplastic (leukaemia) and non-neoplastic (infections) diseases, as well as recognise the leukaemia lineage.

1. Introduction

Acute leukaemia are a heterogeneous group of blood-related cancers characterized by the abnormal proliferation of blast cells in bone marrow, where causes a replacement of normal cells and a decrease of the three haematopoietic lines in peripheral blood. They represent the 11th and 10th most frequent cause of cancer occurrence and death worldwide, respectively, with more than 300,000 deaths estimated in 2018. Acute myeloid leukaemia is an aggressive cancer with 3.7 new cases per 100,000 habitants per year and with a 5-year survival of only 19% in Europe [1]. A timely and accurate diagnosis is crucial for an effective disease management.

The World Health Organization (WHO) considers morphology, along with other complementary tests such as immunophenotype, cytogenetic and molecular biology, essential for the integrated diagnosis of acute leukaemia [2]. This is why the starting point in their diagnosis is still the detection of blasts in blood. Nevertheless, smear review is time consuming, requires well-trained personnel and is prone to intra-observer variability, which is particularly true when dealing with blasts. Indeed, subtle interclass morphological differences exist for leukaemia types, which turns into low specificity scores in the routine screening [3]. They are well-known the difficulties that clinical pathologists have in the discrimination among different blasts and the subjectivity associated with

* Corresponding author.
E-mail address: amerino@clinic.cat (A. Merino).
[1] Postal address: Villarroel Street, 170. ZIP Code: 08036 City: Barcelona. Country: Spain.

https://doi.org/10.1016/j.cmpb.2021.105999
0169-2607/

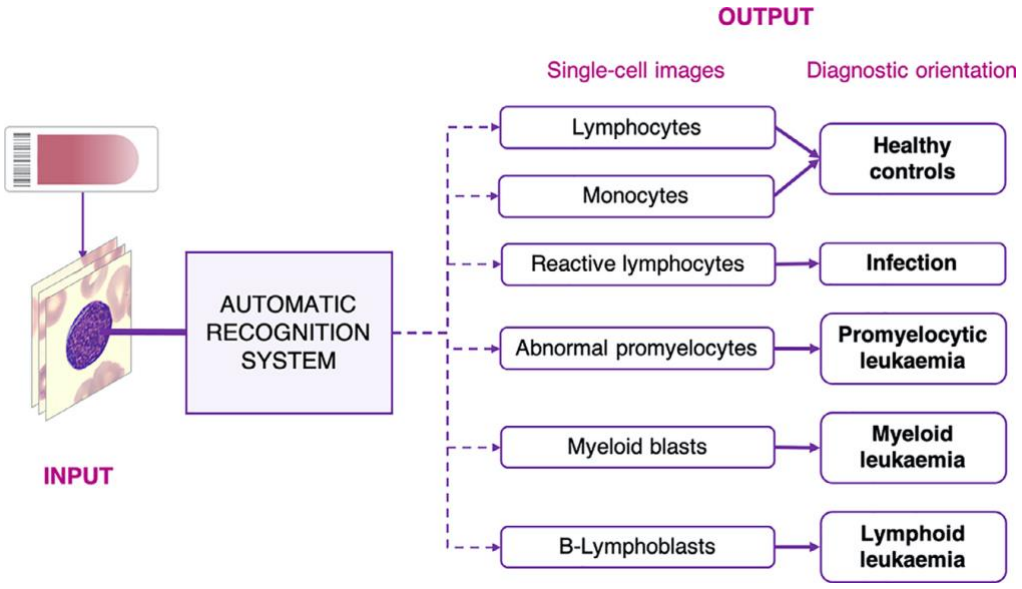


Fig. 1. Scheme of the automatic classification system proposed as a support diagnostic tool for the diagnostic prediction of: 1) infection, 2) acute promyelocytic leukaemia, 3) acute myeloid leukaemia and 4) acute lymphoid leukaemia.

their morphological recognition. Identifying the leukaemia lineage is crucial since the prognostic and immediate therapeutic consequences drastically depend on this differentiation. This challenging problem has been scarcely addressed in the literature in spite that it is well known that automated blood cell image analysers tend to underestimate the number of blast cells [4,5].

Image analysis, quantitative morphological features and machine learning approaches have been the main technological tools adopted in the last decade to overcome such drawbacks [6,7]. The late explosion of deep learning has shifted the focus towards new classification models using convolutional neural networks (CNNs) [7,8]. The application of these automatic classification systems will increasingly become a part of clinical practise in the coming years in the field of haematological malignancy [9,10].

Recently, CNNs have proven to be successful for the classification of normal white blood cells [11–16], the distinction of specific cell subtypes, such as erythroid and myeloid precursors [17], and also for the diagnosis of lymphoma from lymph node digital images [18]. In the context of the morphological classification of cells circulating in peripheral blood in acute leukaemia, most of the previous works used CNNs with transfer learning techniques [19–26], while others used them for feature extraction and afterwards classify images with other well-known traditional machine learning techniques such as support vector machine [27,28].

Literature reveals that lymphoid leukaemia has received more attention, as the automatic recognition of acute leukaemia with CNNs has mainly been addressed in two cases: 1) to differentiate lymphoblasts and leukocytes with very diverse cell morphology, including neutrophils, eosinophils, basophils and lymphocytes [19,24,26,27]; and 2) to separate lymphoblast subtypes [20,21]. Furthermore, CNNs have also been employed in bone marrow images in two problems: 1) the differentiation of the subtypes of acute lymphoid leukaemia regarding the French-American-British (FAB) classification [29], which has recently been replaced by the WHO classification for clinical practise; and 2) the discrimination between acute and chronic leukaemia [30]. It is remarkable that none of these previous works have considered the presence of reactive lymphocytes, which could be difficult since these lymphoid cells share some morphological similarities with blasts [31,32], nor the distinction among blasts from different origin [2,33]. Moreover, the detection of abnormal promyelocytes is of utmost importance because acute promyelocytic leukaemia patients can suffer severe bleeding and die if no treatment is initiated promptly [34].

The objective of this paper is to develop a new CNN-based system (ALNet) to predict the initial diagnostic orientation of acute leukaemia using blood cell images, which must be highly sensitive and specific for clinical application. The system input is a set of cell images of an individual patient's smear, and the output is the prediction of one of the following diagnoses: acute promyelocytic leukaemia (APL), acute myeloid leukaemia (AML non-APL), acute lymphoid leukaemia (ALL) or infection.

The recognition of lymphoblasts from myeloblasts by morphology in the blood smear review is a challenging problem because they share morphological characteristics. To the authors' knowledge, it is the first time in the literature that reactive lymphocytes, abnormal promyelocytes alongside myeloid blasts and B-lymphoblasts have been jointly considered for their automatic recognition using CNN approaches. A previous publication by the authors [35] addressed the recognition of these cell groups using a linear discriminant analysis classifier in which the overall accuracy was 85.8%. Our new contribution is clinically relevant since the proposed sequential system ALNet is able to distinguish neoplastic (leukaemia) and non-neoplastic (infections) diseases, as well as recognise the leukaemia lineage with an overall accuracy of 94.2%. This improvement makes possible its integration as a decision support system to assist pathologists when there is suspicion of acute leukaemia.

2. Material and methods

2.1. Overview

The main objective of this study is to design an automatic classification system to work within the general scheme shown in Fig. 1 and based on CNNs. The rationale for this scheme is twofold: 1) the first step in the diagnosis of acute leukaemia is the morphological review of the peripheral blood smear; and 2) the clinical pathologist could select those abnormal cell images responsible for

Table 1
Diagnosis of the patients in the respective sets of study: acute leukaemia (AL) according to the WHO 2016 classification and viral infections. The table also indicates the diagnosis and AL subtype with results of the most relevant complementary tests. ALL-B acute lymphoid leukaemia; AML, acute myeloid leukaemia; APL, acute promyelocytic leukaemia; HLA-DR; human leucocyte antigen; MPO, myeloperoxidase; NOS, not otherwise specified; P, number of patients.

DIAGNOSIS	AL SUBTYPE AND COMPLEMENTARY TESTS	P	Training N° images	Testing N° images
AML with recurrent genetic abnormalities	APL with *PML-RARA* (HLA-DR-, CD34-, CD13+, CD33+)	15	2,575	1,358
	AML with t(6;9) (p23;q34.1); *DEK-NUP214* (CD45 weak, CD34+, HLA-DR weak, CD117 weak, CD13+, CD33+)	5	144	65
	AML with inv(16) (p13.1q22) or t(16;16) (p13.1;q22); *CBFB-MYH11* (CD45 weak, CD117+, MPO+, CD13+, CD33+, CD4+)	1	21	12
	AML with t(8;21) (q22;q22.1); *RUNX1-RUNX1T1* (CD45 weak, HLA-DR+, CD34+, CD117+, MPO weak, CD13+, CD33+)	2	47	29
	AML with inv(3) (q21.3q26.2) (CD45 weak, CD34+, HLA-DR+, CD117+, CD13 weak, MPO weak)	2	62	25
	AML with t(9;11) (p21.3;q23.3); *MLLT3-KMT2A*	5	138	41
	AML with mutated *NPM1* (HLA-DR+, CD117+, CD13+, CD33+, CD123+)	13	669	68
AML with myelodysplasia-related changes	CD45 weak, CD34+, HLA-DR+, CD117+, CD13+, CD33+, CD123+	51	1,153	390
AML, NOS	Acute monoblastic/monocytic leukaemia (HLA-DR+, CD13+, CD33+, CD64+, CD4+, CD36+, CD11b+)	17	636	144
ALL-B/lymphoma with recurrent genetic abnormalities	ALL-B/lymphoma with t(9;22) (q34.1;q11.2); *BCR-ABL1* (CD123+, CD19+, CD22+, CD79a+)	9	141	337
ALL-B/lymphoma, NOS	CD19+, CD22+, CD79a+, CD10+, CD20+, CD34+	13	199	285
	CD45 weak, CD19+, CD79a+, TdT+, CD10-, CD20-	15	277	401
Viral or other infections		191	1,852	459
TOTAL		339	7,914	3,614

Table 2
Total number of blood cell images and smears used in this work. Images and smears are grouped by class for each dataset and the number of healthy controls and patients for each entity is provided.

Entity	Cell type		Number of images		Number of smears		Number of controls and patients
			Training	Testing	Training	Testing	
Healthy controls	**Lymphocytes**		3,288	312	82	13	100
	Monocytes		1,205*	117	83	4	
Infections	**Reactive lymphocytes**		1,852*	459	168	46	191
Acute leukaemia	**Abnormal promyelocytes**		2,575	1,358	19	14	148
	Myeloid	**Myeloblasts**	2,234	630	200	25	
	blasts	**Monoblasts**	636	144	24	5	
	B-Lymphoblasts		617*	1,023	48	27	
	TOTAL		12,407	4,043	623	108	439

* Cell groups with less than 2,500 images for training, which were up-sampled with data augmenting techniques to balance all the classes

suspecting of acute leukaemia and upload them into the system to obtain the assistance of an automatic classification.

Our design is focused on achieving the following specifications:

1 The prompt and accurate detection of promyelocytic leukaemia, motivated by the emergency of avoiding severe bleeding and disseminated intravascular coagulation with patients' death.
2 The discrimination between myeloid and lymphoid leukaemia lineages, motivated by the different strategical therapies they require.

The system development and assessment involve two main stages:

1 In the first stage, we develop a CNN model to automatically classify images of six cell groups: 1) lymphocytes, 2) monocytes, 3) reactive lymphocytes, 4) abnormal promyelocytes, 5) myeloid blasts (myeloblasts and monoblasts) and 6) B-lymphoblasts.
2 Based on the results obtained when testing this classifier, in the second stage we design and evaluate a recognition system where the input will be a set of cell images of an individual patient's blood smear, and the output will be the prediction of one of the following diagnoses: APL, AML, ALL or infection.

The remaining of this section will be devoted to describe the most relevant issues of the CNN classifier development, including the used image database. To keep a continuous thread in the presentation, all the details concerning the second stage will be presented in Section 3.

2.2. Image datasets

Blood samples, collected in EDTA, were automatically prepared using the slide maker-stainer SP10i (Sysmex, Kobe, Japan) and stained with May Grünwald-Giemsa. Digital images of blood cells were acquired by the CellaVision®DM96 (CellaVision, Lund, Sweden) (363 × 360 pixels) from smears compiled during daily work in the Core Laboratory of the Hospital Clinic of Barcelona. Each digital image contained a single cell. Cell images were identified by pathologists according to their morphological characteristics. Patients' diagnoses were confirmed by the integration of all supplementary data: clinical data, morphology, flow cytometry, cytogenetics and molecular biology. Patients with acute leukaemia were diagnosed by the Clinic Haematology Service of the Hospital following the WHO 2016 classification [36]. We regard these confirmed diagnoses as the ground truth for training and evaluating the models. Table 1 shows the number of images and patients with acute leukaemia and infections included in *training* and *testing sets*.

Table 2 details the number of images and smears corresponding to the different cell classes included in *training* and *testing sets* and the number of healthy controls and patients from which they were obtained. Images and smears are grouped in six classes for each dataset: lymphocytes, monocytes, reactive lymphocytes, abnormal promyelocytes, myeloid blasts [myeloblasts and monoblasts) and

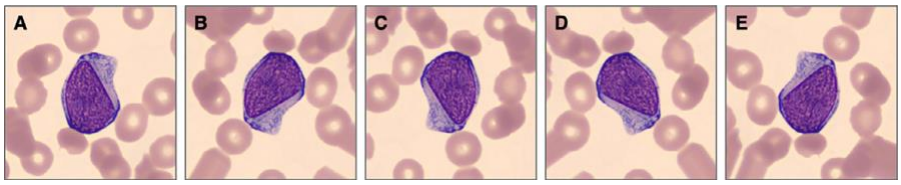


Fig. 2. Examples of some of the transformations applied to increase the number of images of the *train set*. (A) Original image of a B-lymphoblast; (B, C, D, E) Versions of the same image rotated and flipped.

B-lymphoblasts. Smears were obtained from 100 healthy controls, 191 patients with viral infections and 148 patients with acute leukaemia.

The *training set* was arranged with 85% of the smears [623 with 12,407 images), distributed into the above groups of interest. The *testing set,* reserved for the final assessment of the classification system, included the remaining 15% of the smears [108 with 4,043 images).

Data augmenting was performed to balance the *training set* by incrementing the number of images and obtaining the same number per each cell class. Individual images were modified applying randomly transformations as vertical and horizontal flips and rotations from 0 to 60 degrees [37]. Up to four different versions of each original image were obtained from those groups containing less than 2,500 images (monocytes, reactive lymphocytes and B-lymphoblasts) as illustrated in Fig. 2.

2.3. Development and testing of the first CNN classifier

CNNs are multi-layered architectures able to automatically extract complex and high-dimensional features from a large set of images [38]. Through scanning images, they detect and learn patterns by the interaction of their elementary units (neurons) analogising the learning process of human brain. Unlike traditional computer vision models where hand-crafted features are extracted and used to train a classifier, CNNs can learn how to extract relevant features and use them for classification purposes. Moreover, in contrast to previous machine learning methods, the automatic classification of blood cells will not explicitly rely on complex segmentation of the cell regions of interest and the further feature selection.

The first element of a network is always the input layer, which reads image's pixels. Following it are convolutional layers, which are composed by a set of filters responsible for detecting specific patterns in images and extract features. Its output is a set of feature maps. Depending on its architecture, CNNs could contain one or consecutive convolutional layers creating convolutional blocks. After each convolutional layer or block, it usually comes a pooling layer, which reduces the dimension of feature maps while preserving the relevant information and eliminating irrelevant details. Max pooling function is the most widely used, which reduces feature maps size by taking the maximum values. This combination of convolutional and pooling layers is being repeated throughout all the network structure to extract more complex features each time. The depth of a network depends on how many repetitions of these layers are. At the end of them, a set of feature maps representative of the input images is obtained.

In this work, we adopted the concept of fine-tuning, which consists on modifying a CNN architecture previously trained for another task and re-train and adapt some of its layers with our new data to obtain an end-to-end classifier [39]. Its purpose was to take advantage of the knowledge of a CNN pre-trained with a large dataset, which helped to overcome the deficit of training examples and served as an effective weight initialization to, afterwards, use it for the classification of the images of our study. This approach is usually implemented with CNNs with large quantity of parameters because training them from scratch can affect their ability to generalize and may result into low accuracies or overfitting [40].

To examine the effects of different CNN frameworks on model performance, we investigated four well-known CNNs already pre-trained with the ImageNet database [41]: VGG16, ResNet101, DenseNet121 and SENet154 [42–45]. VGG16, ResNet101 and SENet154 were chosen based on their high performance obtained in previous studies by our group. We achieved an overall assessment accuracy of 96.2% with a fine-tuned VGG16 for the recognition of eight groups of normal blood cells [16]. We used a pre-trained ResNet101 to differentiate up to 11 cell groups including normal, reactive, abnormal lymphocytes and blasts, achieving an overall accuracy of 82.8% [46]. We also implemented a SENet154 for the diagnostic prediction among lymphoma, acute leukaemia and infection with an overall assessment accuracy of 94.5% [47]. DenseNet121 was chosen from literature because of its 95.3% accuracy for predicting leukaemia diagnosis [30].

For the transfer learning approach, we removed their last layers, which were fully connected layers trained for classifying 1,000 categories. Afterwards, two fully connected layers were coupled at the end of each network. Fully connected layers learn how to combine the feature maps to perform the final classification of the input images. The design of these layers for each evaluated architecture is detailed in the following paragraph.

A total of 512 feature maps were obtained from VGG16 and converted to an array of 25,088 features, which were the input to the two fully connected layers, the first layer with 1,024 nodes and the second with 512 nodes. From ResNet101 and SENet154, a number of 2,048 feature maps were obtained and converted to an array of 100,352 features. These features were fed to a first fully connected layer of 4,096 nodes and the output to the second of 512 nodes. Regarding DenseNet121, a total of 1,024 feature maps were obtained and converted to an array of 50,176 features. Its first fully connected layer had 2,048 nodes and the second had 512 nodes. To obtain the final classification, a third fully connected layer with six nodes was configured for all the four architectures, one node for each cell class (see Fig. 1). This last layer predicted the class of the input images as probability values by applying a softmax operation.

Training a neural network is an iterative process that typically consists of two phases: a forward and a backpropagation. In the first one, all input images are passed through the network at each iteration (epoch). Once each image is classified, a loss function is calculated by the difference (error) between the prediction made by the network and the label assigned by the clinical pathologist according to the ground truth, which is the diagnosis based on all complementary information following the WHO 2016 (see Table 1) [36]. Loss increases when the predicted probability diverges from the true label. Afterwards, in the second phase, the error is propagated backwards through the network and all the weights are updated to minimize the loss function and obtain a high accuracy (rate of images correctly classified). This two-phase process usually needs to be repeated for several epochs to obtain an optimum trained model.

Table 3
Overall accuracies of the pre-trained CNNs when changing the number of convolutional blocks to be trained with 470 iterations and with the two evaluation approaches: hold-out and 5-fold cross-validation. For the hold-out approach, validation and testing accuracies are shown. For the 5-fold cross validation, we present the mean and the standard deviation of the testing accuracy computed among the five folds. The last two columns show the training time (in minutes) achieved when fine-tuning the entire convolutional blocks, and the classification time (in seconds) when evaluating the selected models with the respective *testing sets*.

		Number of convolutional blocks fine-tuned					Testing	Training time (min)	Class. time (s)
	Model	1	2	3	4	Whole			
Hold-out	**VGG16**	89.2%	88.0%	88.2%	88.6%	94.6%	**88.4%**	6.78	12
	ResNet101	90.1%	89.7%	90.3%	90.2%	93.3%	84.0%	12.52	22.4
	DenseNet121	84.4%	86.8%	87.5%	88.0%	93.6%	84.8%	7.9	14
	SENet154	89.3%	89.8%	89.8%	88.9%	94.6%	**85.0%**	48.2	41
5-fold cross validation	**VGG16**	76.3%±0.93	76.5%±0.67	76.2% ±0.97	76.1% ±1.36	**86.9%**±0.68		123.1	29.5
	ResNet101	80.8%±1.99	81.2%±0.54	80.2% ±0.56	81.8% ±0.41	86.8%±0.62		136.2	32
	DenseNet121	82.6%±1.10	82.9%±1.33	82.4% ±1.20	83.1% ±1.19	86.3%±0.93		262.5	63
	SENet154	79.5%±0.72	80.2%±0.83	80.5% ±0.58	80.3% ±0.60	**87.2%**±1.22		426.7	81.5

Table 4
Confusion matrix of the classification results (in %) of the VGG16 for the images of the *testing set*. Rows indicate the true class and columns represent the predicted class supplied by the network. Diagonal values are the true positive rates for each cell type. The overall classification accuracy was 88.4%.

		Predicted class					
		Lymphocytes	Monocytes	Reactive lymphocytes	Abnormal promyelocytes	Myeloid blasts	B-Lymphoblasts
True class	**Lymphocytes**	**98**	0	1	0	0	1
	Monocytes	0	**91**	3	1	4	1
	Reactive lymphocytes	0	0	**97**	0	2	1
	Abnormal promyelocytes	0	3	0	**94**	3	0
	Myeloid blasts	0	2	0	2	**87**	9
	B-Lymphoblasts	0	0	0	3	22	**75**

In our work, cross-entropy was employed as loss function [37] and ADAM (Adaptive Moment Estimation) as optimizer [48]. Training was performed using a batch size of 64 along 470 iterations, implementing the cycling learning rate policy to obtain optimal classification results with fewer iterations [49]. Whereas different CNNs might require different number of iterations during training, we fixed this value because we wanted to select the best CNN architecture for the automatic classification of the different leukaemia types considered in this work. Moreover, the number of convolutional blocks to fine-tune was considered an hyperparameter to be selected. Several tests were performed varying the number of convolutional blocks to be trained: last block, last two, last three, last four and the entire network.

To select the model, CNN frameworks were trained and evaluated using two different approaches: 5-fold cross validation and hold-out. For the first case, we performed a random split of all the images of the whole dataset [16,450 single-cell images) into five equal subsets, ensuring that all the data from the same patient's smear was kept in the same subset. The same smear did not appear in two different folds. Moreover, folds were approximately balanced in the sense that the number of distinct smears was approximately the same in each fold. The mean and the standard deviation of the testing accuracy were computed among the five folds. For the hold-out approach, the *training set* was split into *train* and *validation sets* with 80% and 20% of the images, respectively. After training, the best CNN models were further evaluated using the *testing set*. The overall accuracy was selected as the main target performance parameter, so that we selected the networks that showed the highest value.

All the experiments were performed using PyTorch software libraries [37] and a Nvidia Titan XP GPU.

3. Results

3.1. First CNN classifier

Table 3 shows the overall classification accuracies of the four evaluated CNNs with 5-fold cross validation and hold-out when changing the number of convolutional blocks trained with 470 iterations. Based on these results, the highest accuracies were obtained when fine-tuning the entire models. This means to train all the convolutional blocks with our own image dataset. Moreover, the highest testing accuracies were achieved with VGG16 and SENet154 architectures (above 85% using hold-out and above 86.8% with 5-fold cross validation) as it is shown in Table 3.

VGG16 showed an accuracy value of 88.4% using hold-out with respect to the best accuracy (87.2%) obtained with SENet154. Finally, we decided to select VGG16 together with the hold-out approach as the best model because of the following: 1) less overfitting was obtained as VGG16 showed better performance when making predictions with the new images of the *testing set*; and 2) it is a simpler architecture compared to the other CNN frameworks and had the best training and classification time, which is an advantage for a potential real-time implementation (see Table 3).

Table 4 shows the confusion matrix of the classification results of the six cell groups with the VGG16. The true positive rates (main diagonal) obtained for myeloid blasts (87%) and B-lymphoblasts (75%) were concluded not to be high enough for our purpose of building a tool to predict the diagnostic orientation of acute leukaemia in a clinical setting.

3.2. Sequential CNN classification system: ALNet

In order to improve the true positive rates obtained for the recognition of myeloid blasts and B-lymphoblasts presented above, we proposed a new strategy with a two-step classification scheme (see Fig. 3), where two separate classifiers work in series: ALNet. The first step (module 1) consists of a VGG16 trained to distinguish abnormal promyelocytes among lymphocytes, monocytes, reactive lymphocytes and blast cells, which encompass both myeloid blasts and B-lymphoblasts. The second step (module 2) required a VGG with an increased number of convolutional layers (VGG19) to discriminate between myeloid blasts and B-lymphoblasts.

Although good results were obtained with VGG16 (testing accuracy of 85.7%), we tried to increase sensitivity for these two last

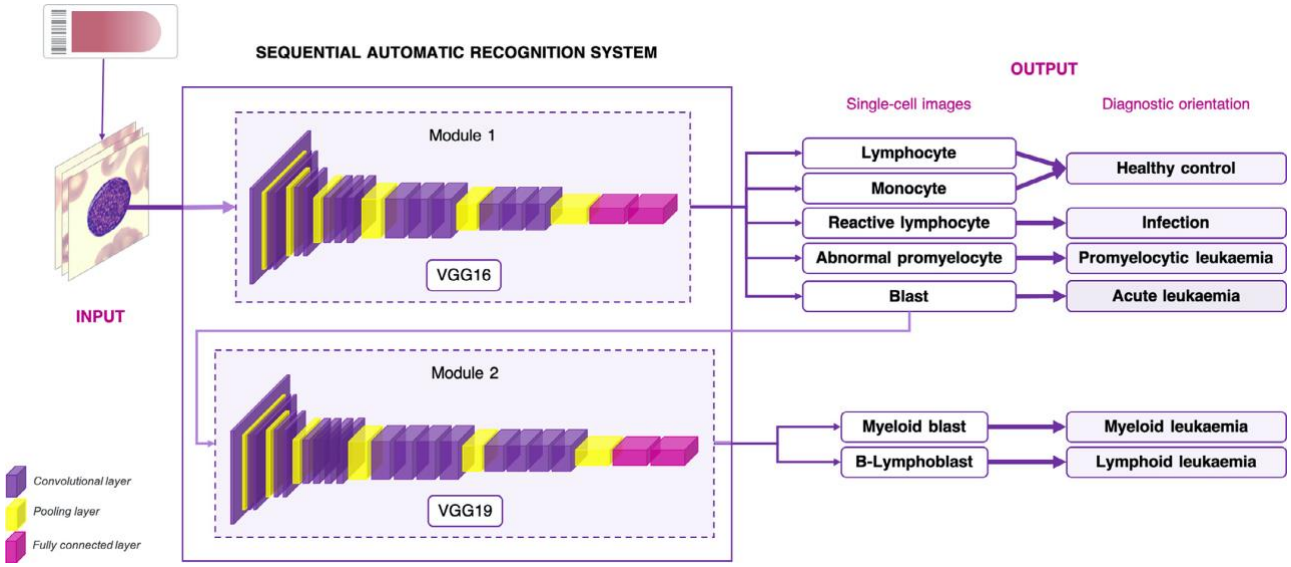


Fig. 3. Diagram illustrating the proposed sequential CNN-based system (ALNet) for the automatic recognition of acute leukaemia lineage. It starts with the selection of the cell images of an individual smear by the clinical pathologist, which are the inputs to the system. ALNet includes two consecutive modules. The first module recognises acute promyelocytic leukaemia (APL) and acute leukaemia (non-APL) from the remaining groups. Only the smears corresponding to patients with non-APL leukaemia (myeloid or lymphoid) are used for lineage classification through the second module.

Table 5
Overall training and testing accuracies obtained for module 2. The last two columns show the training time (in minutes) achieved when fine-tuning the entire models, and the classification speed (in seconds) when evaluating both models with the images of the *testing set*.

	Accuracy		Training time (min)	Class. speed (s)
Model	Training	Testing		
VGG16	95.4%	85.7%	2.42	8
VGG19	99.4%	89.5%	2.64	7

cell groups by using VGG19, which was the same architecture but with 16 convolutional layers instead of the 13 of VGG16. Regarding the design of the fully connected layers for the VGG19, 512 feature maps of 14 × 14 pixels were obtained from its convolutional blocks and converted to an array of 100,325 features, which were the input to the two fully connected layers after an average pooling. The first layer had 1,024 nodes and the second 512 nodes. The third fully connected layer consisted of two nodes, one node for each cell class (see Fig. 3). This increase of layers by using a VGG19 resulted into higher accuracies for both training and testing (99.4% and 89.5%, respectively) without compromising training time and classification speed (see Table 5).

Both VGGs were trained as explained in Section 2.3, using the images arranged as shown in Table 6 and performing fine-tuning to the whole convolutional blocks using hold-out. The first module exhibited a validation accuracy of 99.6% during training, while that of the second module was 99.4%.

The first assessment of ALNet was done through a blind classification of all the single-cell images of the *testing set* (see Table 6). We calculated the sensitivity or true positive rate (TPR), specificity or true negative rate (TNR), precision or positive predictive value (PPV) and overall accuracy as follows:

$$Sensitivity\ (TPR) = \frac{TP}{TP + FN}$$

$$Specificity\ (TNR) = \frac{TN}{TN + FP}$$

$$Precision\ (PPV) = \frac{TP}{TP + FP}$$

$$Overall\ accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Table 7A shows the confusion matrix that summarizes the classification results for module 1. The true positive rates shown in the main diagonal were: 99.9% for lymphocytes, 97.6% for monocytes, 97.2% for reactive lymphocytes, 95.3% for abnormal promyelocytes and 91.7% for blasts (myeloid blasts and B-lymphoblasts). The overall classification accuracy was 94.2%. Table 7B summarizes the classification results for blast cells. As seen in the main diagonal, 99.4% of images corresponding to myeloid blasts and 82% to B-lymphoblasts were correctly classified. The overall classification accuracy was 89.5%. Specificity and precision values for each cell class of both modules are shown in Table 8.

3.3. Classification system ALNet in a clinical setting

Going towards the final objective of this work, we assessed ALNet to predict patient's diagnosis using the smear as a classification unit, trying to emulate the way pathologists interpret results in clinical laboratories. The system input was a set of cell images of an individual smear selected by the clinical pathologist and the output was the prediction of one of the following diagnoses: APL, AML, ALL or infection (see Fig. 3). The diagnosis was predicted by identifying the cell class that predominated in the smear. A threshold was determined such that the diagnosis will be predicted by the cell class with the percentage of images classified above this value. We used all the smears of the *training set* to perform a multiclass Receiver Operating Characteristics (ROC) analysis with the statistical R software [50]. Results are shown in Fig. 4. It was found that 50% of the cell images correctly classified was the best threshold to predict patient's initial diagnosis through the smear, obtaining values of 1 for the area under the curve. Such satisfactory result could be explained because only 91 out of 14,985 (0.6%) single-cell images of the *training set* were misclassified, and

Table 6
Distribution of the cell images and smears used to train and test the sequential classification system with two modules (ALNet). Images and smears are grouped by class for each dataset and the number of healthy controls and patients is provided.

		Number of images			Number of smears		Number of controls and patients
		Training		Testing	Training	Testing	
	Cell type	Train	Validation				
Module 1	**Lymphocytes**	2,510	778	312	82	13	100
	Monocytes	945*	260	117	83	4	
	Reactive lymphocytes	1,462*	390	459	168	46	191
	Abnormal promyelocytes	2,035*	540	1,358	19	14	15
	Blasts	2,790	697	1,797	272	57	133
Module 2	**Myeloid blasts**	2,296+	574	774	224	30	96
	B-Lymphoblasts	494+	123	1,023	48	27	37

* Cell groups up-sampled to 2,500 images with data augmenting techniques
+ Cell groups up-sampled to 2,900 images with data augmenting techniques

Table 7
Confusion matrix of the classification results (in %) for the images of the *testing set* of modules 1 (A) and 2 (B). Rows indicate the true class and columns represent the predicted class supplied by the network. Diagonal values are the true positive rates for each cell type. The overall classification accuracies of the first (A) and second module (B) were 94.2% and 89.5%, respectively.

A		Predicted class				
		Lymphocytes	Monocytes	Reactive lymphocytes	Abnormal promyelocytes	Blasts
True class	**Lymphocytes**	**99.9**	1	0	0	0
	Monocytes	0.8	**97.6**	0	0.8	0.8
	Reactive lymphocytes	0.2	0.2	**97.2**	0	2.4
	Abnormal promyelocytes	0	0.4	0	**95.3**	4.3
	Blasts	2.5	0.2	1.3	4.3	**91.7**
B		Predicted class				
		Myeloid blasts	**B-Lymphoblasts**			
True class	**Myeloid blasts**	**99.4**	0.6			
	B-Lymphoblasts	18	**82**			

Table 8
Sensitivity, specificity and precision values of module 1 and 2 of ALNet regarding the classification results of individual cell images.

	Module 1					Module 2	
	Lymphocytes	Monocytes	Reactive lymphocytes	Abnormal promyelocytes	Blasts	Myeloid blasts	B-Lymphoblasts
Sensitivity	99.9%	97.6%	97.2%	95.3%	91.7%	99.4%	82.0%
Specificity	98.7%	99.7%	99.3%	97.1%	96.9%	82.0%	99.4%
Precision	86.8%	90.5%	94.9%	94.2%	95.9%	80.7%	99.4%

this number was nearly imperceptible when focusing on the whole smear as classification unit.

Such threshold may be interpreted in such a way that if more than 50% of the cell images of a smear are classified as myeloid blasts, for example, the predicted diagnosis of the patient to whom this smear belongs is acute myeloid leukaemia. Whereas if similar percentages are obtained for more than one class, in this case it is not possible to predict a diagnosis and the system considers the smear as belonging to an "unknown" diagnostic group.

Once the threshold was determined, the automatic recognition system was assessed using one by one all the smears in the *testing set*, which were not previously used (see Table 6). Confusion matrix in Table 9A shows the classification results of the first module. Sensitivity, specificity and precision values of 100% were obtained for all the categories of the first module. In consequence, using the first classification module of ALNet, we correctly detected the following groups: 1) healthy controls, 2) patients with infection, 3) patients with APL and 4) patients with acute leukaemia non-APL.

The smears corresponding to patients with non-APL leukaemia (AML or ALL) were classified through the second module, and the results are shown in the confusion matrix in Table 9B. Regarding AML, sensitivity, specificity and precision values of 100%, 92.3% and 93.7%, respectively, were obtained. As for ALL, a sensitivity of 89% and specificity and precision values of 100% were obtained. The overall accuracy of individual smears was 94.7% (see Table 9B).

At the end of the two-step classification with ALNet, we obtained the correct diagnostic prediction for all patients with APL and AML. With respect to the ALL group, a total of 24 from 27 smears were correctly classified, being the remaining three recognized as AML (two smears) or unknown (one smear).

In addition, to further evaluate the ALNet performance in a clinical setting, we used a new image dataset from two other hospitals (*Josep Trueta* and *Germans Trias i Pujol*). This dataset contained a total of 381 images of blast cells (322 myeloid blasts and 59 B-lymphoblasts) stained with May Grünwald-Giemsa and acquired by the CellaVision analyser used in each laboratory. Fig. 5 shows representative examples of myeloid blast and B-lymphoblast images from the three datasets used to evaluate ALNet.

When classifying by single-cell images, 98% of blast cells were correctly classified by module 1. Module 2 correctly classified 82% of myeloid blasts and 64% of B-lymphoblasts. When using the threshold of 50% to predict the diagnosis from the smear, in the first classification module of ALNet all the smears (100%) were classified as acute leukaemia non-APL. Regarding the leukaemia

Table 9
Confusion matrix of the classification results (in %) for the smears of the *testing set* of modules 1 (A) and 2 (B) taking the threshold of 50% into consideration. Rows indicate the true diagnosis and columns represent the predicted diagnosis supplied by the classifier. Diagonal values are the true positive rates for each smear (in brackets the number of smears). The overall classification accuracies of the first (A) and second module (B) were 100% and 94.7%, respectively. LY, lymphocytes; MO, monocytes; APL, acute promyelocytic leukaemia; AL, acute leukaemia; AML, acute myeloid leukaemia; ALL, acute lymphoid leukaemia; UNK, unknown.

A		Predicted diagnosis				
		Control (LY)	Control (MO)	Infection	APL	AL non-APL
True diagnosis	**Control (LY)**	**100% (13)**	0	0	0	0
	Control (MO)	0	**100% (4)**	0	0	0
	Infection	0	0	**100% (46)**	0	0
	APL	0	0	0	**100% (14)**	0
	AL non-APL	0	0	0	0	**100% (57)**
B		Predicted diagnosis				
		AML	**ALL**	**UNK**		
True diagnosis	**AML**	**100% (30)**	0	0		
	ALL	7% (2)	**89% (24)**	4% (1)		

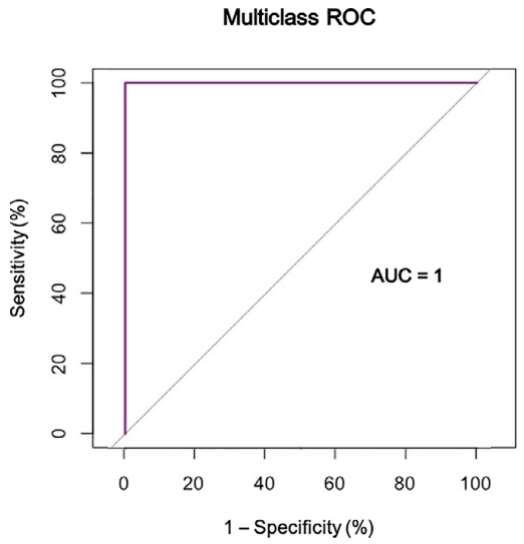


Fig. 4. Plot of the multiclass ROC analysis to obtain the best threshold to predict the diagnosis of the subtype of acute leukaemia. It was obtained by averaging the single ROC curves obtained for each of the classes under study. The best threshold is defined as the minimum percentage of images required in a cell class to predict a diagnosis. The 50% resulted to be the best threshold, obtaining an AUC of 1. AUC, area under the curve; ROC, Receiver Operating Characteristic.

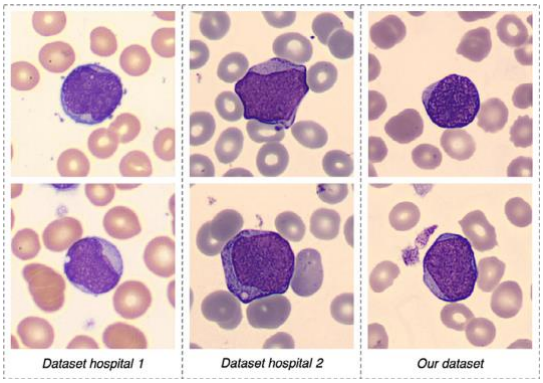


Fig. 5. Examples of B-lymphoblast images (first row) and myeloid blast images (second row) of the three datasets included in this study to evaluate the performance of ALNet.

lineage, 3/5 smears corresponding to AML (60%) and 1/2 ALL smears (50%) were correctly classified.

4. Discussion

For the first diagnostic orientation of acute leukaemia in clinical laboratories, it is important the detection of quantitative abnormalities in leucocyte count, haemoglobin level or thrombocyte count, which triggers the smear review. The detection of blasts circulating in blood is the subsequent step. Thus, cell morphology is crucial for the initial diagnosis of leukaemia lineage to apply the suitable treatment and avoid delays in medical procedures, primarily in acute promyelocytic leukaemia. This is why the automatic recognition of blasts has been a target for automated solutions, tools and methods within the artificial intelligence framework [6,7].

Recently, previous studies described deep learning approaches to recognise lymphoblasts from normal leukocytes [19,22–24,26,27]. Table 10 shows a comparison of the overall accuracy obtained with our proposed system with that obtained in other works in the literature. It also summarizes the number of images and the CNN frameworks used by these previous authors. These studies achieved accuracy values above 88% for lymphoblasts, but it is fair to mention that the significant morphological differences with normal leukocytes make the classification problem more accessible. However, when other authors tried to distinguish B-lymphoblasts from mononuclear cells such as lymphocytes, the sensitivity for B-lymphoblasts decreased to 81.5% [21]. Many of these studies focused on a 'binary' classification (disease vs. normal), which is not a realistic approach and does not reflect the real-life complexity of haematological malignancies diagnosis [7]. Moreover, those studies which classified lymphoblasts subtypes [20,21] were based on the FAB classification, which currently has been replaced by the WHO 2016 classification [36] for clinical practise. The sensitivity for B-lymphoblasts (82%) obtained in this work (Table 7B) is the best result achieved nowadays in the automatic classification among blasts using deep learning techniques. Moreover, the high sensitivity (95.3%) obtained with abnormal promyelocytes (Table 7A) when differentiating them from other blasts overcomes the accuracy of 41% (Table 10) obtained by previous authors [25] when differentiating abnormal promyelocytes from myeloid blasts. Besides this, nearly none false positives were obtained when discriminating normal mononuclear and reactive cells from blasts (as seen in Table 7A).

In our work, the challenge was twofold: 1) differentiate blasts among other mononuclear cells, and 2) discern between myeloid

Table 10
Comparison of the overall accuracy obtained with the proposed system (ALNet) with that of other works in the literature. CNN, convolutional neural network; fc, fully connected layers; NR, not reported; LDP, local directional pattern; SCA, sine cosine algorithm; BL, blast cells; Ab. prom, abnormal promyelocytes; SVM, support vector machine.

Work	Original N° images	Feature extraction	Classification	Accuracy Leukaemia detection	Accuracy Subtype of leukaemia
Thahn *et al.*[19]	108	CNN	fc	96.6%	NR
Shafique *et al.*[20]	260	CNN (AlexNet)	fc	99.5%	96.06%
Pansombut *et al.*[21]	363	CNN	fc	81.74%	81.5% B-lymphoblasts 68.9% T-lymphoblasts
Ahmed *et al.*[22]	354	CNN	fc	88.25%	NR
Jha *et al.*[23]	260	LDP	SCA - fc	98.7%	NR
Prellberg *et al.*[24]	12,528	CNN (ResNeXt50)	fc	88.91%	NR
Matek *et al.*[25]	18,365 (3,312 BL)	CNN (ResNeXt)	fc	90%	94% myeloblasts 41% Ab. prom
Loey *et al.*[26]	564	CNN (AlexNet)	fc	100%	NR
Vogado *et al.*[27]	377	CNN (AlexNet, CaffeNet, VGG)	SVM	99%	NR
Di Ruberto *et al.*[28]	33	CNN (AlexNet)	SVM	94.1%	NR
Rehman *et al.*[29]	330[+]	CNN (AlexNet)	fc	97.78%	NR
Huang *et al.*[30]	1,322[+]	CNN (DenseNet121)	fc	95.3%	95.25%
Proposed system (ALNet)	**16,450 (4,825 BL)**	**CNN (VGG)**	**fc**	**94.2% (cell) 100% (smear)**	**89.5% (cell) 94.7% (smear)**

[+] Datasets of bone marrow digital images

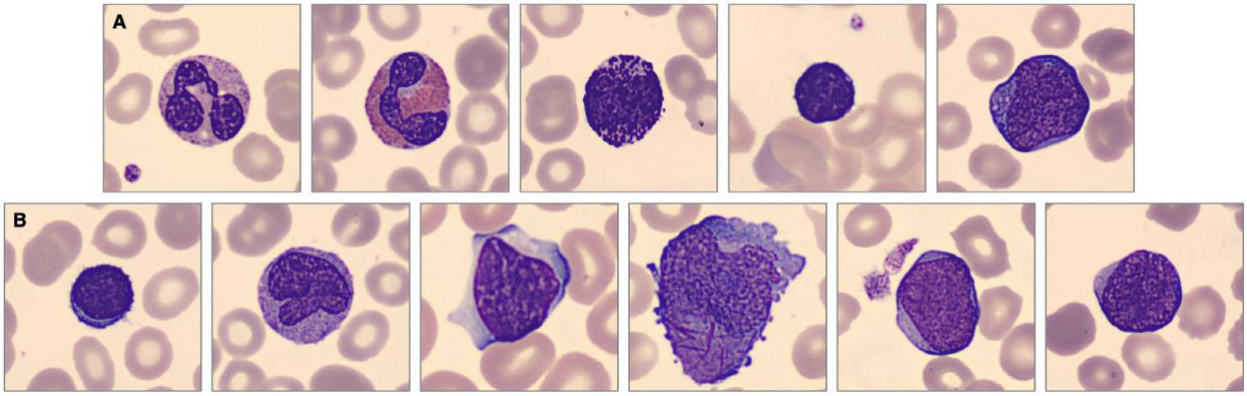


Fig. 6. (A) Examples of cell types included in previous studies from the literature for leukaemia classification (from left to right): neutrophil, eosinophil, basophil, lymphocyte and B-lymphoblast. (B) Cell groups included in this study (from left to right): lymphocyte, monocyte, reactive lymphocyte, abnormal promyelocyte, myeloblast and B-lymphoblast.

blasts and B-lymphoblasts because of the overlapping morphological characteristics that they exhibit. It may be debatable why this work did not include neutrophils or eosinophils or erythroblasts in the classification problem when they may be present in the smear. As shown in Fig. 6A, lymphoblasts exhibit very different morphology with respect to mature leukocytes, such as neutrophils, eosinophils and basophils since their nucleus is lobulated and their cytoplasm shows abundant granules. Fig. 6B shows representative examples of the groups addressed in this study (from left to right: lymphocyte, monocyte, reactive lymphocyte, abnormal promyelocyte, myeloblast and B-lymphoblast). By comparing the cell groups in Fig. 6A (from left to right: neutrophil, eosinophil, basophil, lymphocyte and B-lymphoblast) with the groups in Fig. 6B, it is visible that the automatic differentiation in the second case is more complex.

In a previous work [16] using CNNs, our group achieved high performance with a fine-tuned VGG16 for the recognition of eight groups of normal cells circulating in peripheral blood. We focused on normal leukocytes, obtaining accuracies for neutrophils, eosinophils, immature granulocytes (metamyelocytes, myelocytes and promyelocytes) and erythroblasts of 99.6%, 99.6%, 92.8% and 91.8%, respectively, among the other cell groups. The current paper was focused on the differentiation of the lineage of blast cells with the aim of designing a model to assist the clinical pathologist. Computational assistance is more needed for blast lineage discrimination than for blast versus normal leukocyte differentiation. The integration of the first classifier [16] and the current ALNet is a next step which can be taken in the near future.

In a previous publication by the authors [35], the objectives of this paper were addressed using traditional machine learning algorithms, where a single model based on linear discriminant analysis achieved the highest classification accuracy. In the present work, our purpose was to improve the previous accuracy for the leukaemia lineage differentiation, which was 85.8%. The new strategy herein was to design two separated classifiers working sequentially. The second classifier was specialized to face the most difficult recognition task, which is the automatic differentiation between myeloid blasts and B-lymphoblasts. Using this strategy, sensitivity values were increased to 95.3% for abnormal promyelocytes, 99.4% for myeloid blasts and 82% for B-lymphoblasts (see Table 11). Furthermore, traditional approaches required to segment and extract features manually. Indeed, segmentation presents many challenges because of the complex cell morphology, the variability of the blood smear staining and the general conditions of image acquisition. All these aspects may result in a decrease of the classification accuracy of these traditional models. Table 12 shows a com-

Table 11
True positive rates for each cell type, comparing the proposed deep learning system (ALNet) with the approach previously published by our group. In bold are shown the improved sensitivity values achieved with ALNet.

	True positive rates	
Type of cell	Boldúet *al.* [35]	Proposed system
Lymphocytes	97.6%	**99.9%**
Monocytes	93.0%	**97.6%**
Reactive lymphocytes	97.7%	97.2%
Abnormal promyelocytes	72.6%	**95.3%**
Myeloblasts	80.8%	**99.4%**
B-Lymphoblasts	78.9%	**82.0%**

Table 12
Performance comparison between traditional machine learning approaches previously published by our group and the deep learning models proposed in this study. LDA, linear discriminant analysis; SVM, support vector machine; RF, random forest; KNN, k-nearest neighbours; Bayes, naïve Bayes.

	Method	Accuracy
Machine learning approaches [35]	**LDA**	85.8%
	SVM	83.5%
	RF	75.4%
	KNN	74.8%
	Bayes	68.3%
Deep learning approaches	**VGG16**	94.6%
	ResNet101	93.3%
	DenseNet121	93.6%
	SENet154	94.6%

parison of the accuracies when using those models and the CNN models investigated herein, which are higher in accordance to the results reported by previous authors [21,22,26].

The results of this study demonstrate that the best classification of blast cells is achieved with the proposed sequential system, being VGG the best architecture for both modules. This is consistent with the most frequently used networks in the literature, where simple architectures predominate to address the recognition of acute leukaemia [20,27,29]. Moreover, we concluded from our experience that using hold-out as validation technique with 2,500 images per group made possible obtaining satisfactorily high accuracies as similar accuracies were obtained with the 5-fold cross validation approach. This was consistent with our previous publication [16] and results published from previous authors [26].

An important aspect to remark in this work is the number of images involved and their high quality, being the largest dataset used for leukaemia classification as it is illustrated in Table 10. We used a dataset with over 16,000 peripheral blood cell images, 4,825 being blast cell images, in comparison with the 3,312 blast cell images used by Matek *et al.* [25]. Furthermore, we guarantee the confirmation of their labels through other complementary tests (ground truth). Not only high-quality and a large number of images are required for developing diagnostic systems, but also the availability of images properly annotated by experts, which is scarce in literature [7,8,19]. The quality of a dataset is essential to observe morphological characteristics which can lead towards a diagnosis, not only for daily clinical practise but also to obtain robust models avoiding overfitting [16].

Visualizing the feature maps generated by the intermediate convolutional layers of a CNN framework could give us an idea of which patterns the network extracts, and thus help to interpret classification results. With this purpose, both our ALNet system and the original VGG16 ImageNet model were fed with two representative examples of cell images. Afterwards, we extracted the feature maps from the same convolutional layers to compare them, as it is illustrated in Fig. 7.

Fig. 7A shows original images of a myeloid blast and an abnormal promyelocyte. Fig. 7B displays the feature map 202 of 256 from the convolutional block 3, layer 1. In the feature maps generated by the original VGG16 (left) we can barely notice the cell nucleus, and in the case of the abnormal promyelocyte neither the cell outline. At this depth the original network cannot detect diagonal and horizontal lines to differentiate parts of the cells of the input images. With respect to the feature maps obtained from the VGG16 of ALNet (right), our model is able to precisely detect the cell outline, nucleus shape and texture, along with the recognition of the red blood cells and their central pallor and the platelet next to the myeloid blast. The histogram of the myeloid blast (Fig. 7C) presents a narrow and flatter peak with the pixels very localized within a short intensity range between 100 and 150, and also having a greater number of darker pixels mostly localized in the cell nucleus. In contrast, the histogram of the abnormal promyelocyte shows a peak with lower height, which covers many pixels with different intensity levels. This indicates more information variety due to the presence of the intense azurophil granulation and splinters on its larger cytoplasm.

Fig. 7D shows another case where the improvement of ALNet in detecting relevant characteristics of the cells is visually interpretable. At this stage ALNet is capable of detecting the bilobed nucleus of the abnormal promyelocyte and the nucleolus of the myeloid blast, while the feature map from the original network does not capture almost any information from the images. It is important to mention that subtle morphological differences in the nucleus shape, chromatin texture and cytoplasmic granulation can make the difference among the cell types included in this study.

When evaluating ALNet with images acquired in two other hospitals, sensitivity for the leukaemia lineage discrimination decreased. Previous authors [22] reported that the overall accuracy of their CNN model decreased 6.51% when it was evaluated with a different image dataset. Although all datasets of this work were stained with the 'gold standard' May Grünwald-Giemsa technique used in clinical laboratories, we observed that variations in the optical conditions and resolution of microscopic images affected the classification accuracy of the CNN models (see Fig. 5). It is known that colour information is involved in the quantification of cytoplasmic basophilia and granulation, both being very important characteristics in the classification of myeloid blasts and lymphoblasts [51]. This is why changes related to image clarity, colour scale or resolution could mislead their differentiation.

When developing new diagnostic support tools for laboratory practise, it is important to consider individual patients when organising datasets for training and assessing the system. Using images from the same smear for both training and assessment could cause an accuracy overestimation, what happens in almost all the studies in the literature. To avoid this, we followed the procedure established in [35], by splitting the initial dataset into two sets of different smears. An excellent diagnostic prediction was achieved as the system differentiated all normal smears from those related to infections and with respect to smears containing blast cells (as seen in Table 9A). This satisfactory performance was also obtained with smears corresponding to patients with acute leukaemia from two other hospitals. Moreover, the system presented a very high sensitivity (100%) for the detection of myeloid leukaemia, and high specificity and precision (100%) for promyelocytic and lymphoid leukaemia for the smears of our own dataset.

These results showed that the approach proposed in this work could be suitable for clinical laboratory practise. A current limitation for the practical implementation of ALNet in other laboratories may be that a single-institution data source has been used to train the model, which could result in accuracy variations related to the image staining procedure, as well as optical and resolution of microscopic images. To deal with this, our group has some work in

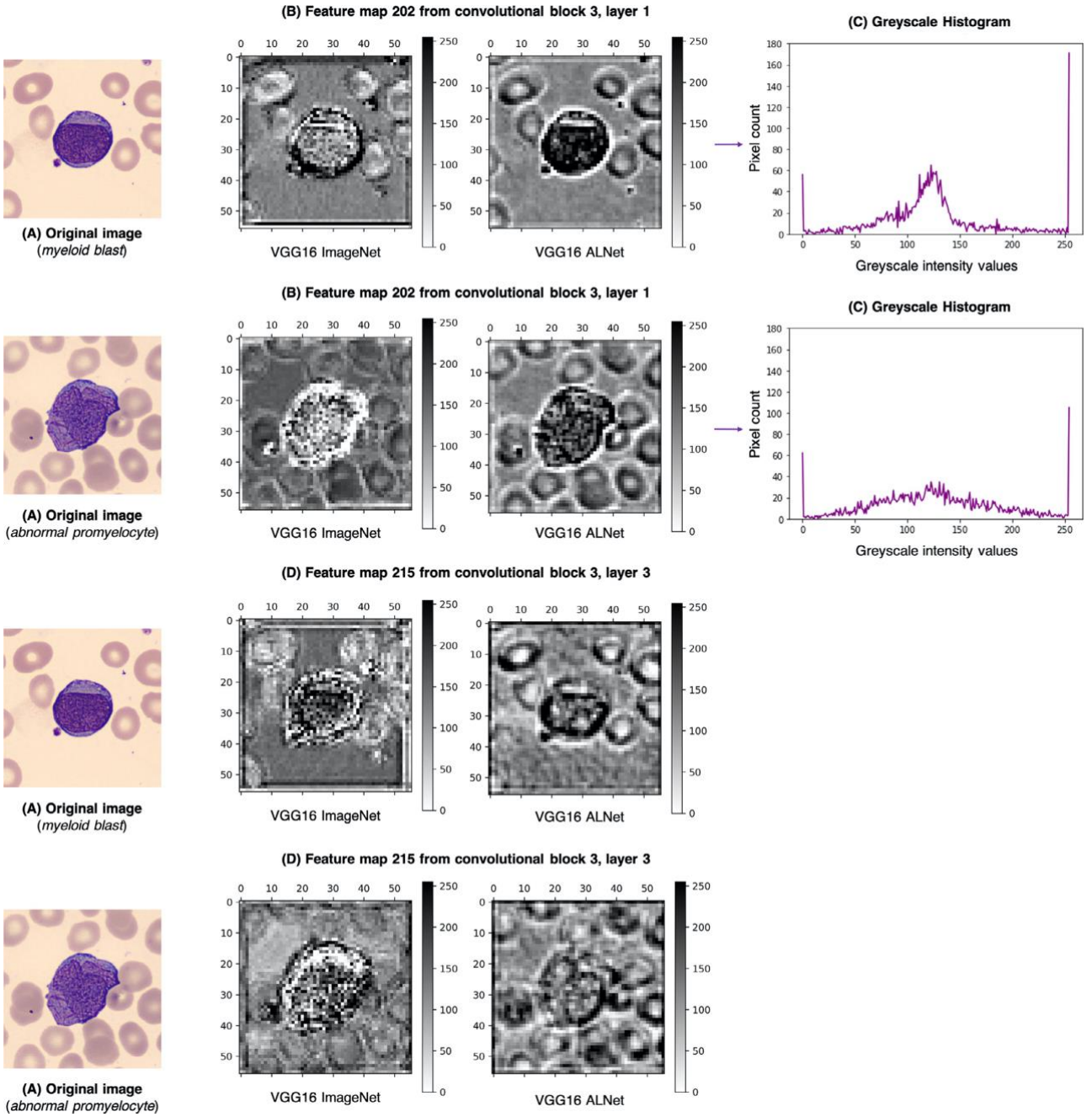


Fig. 7. (A) Original images of a myeloid blast from AML and an abnormal promyelocyte from APL. (B and D) Feature maps from VGG16 ImageNet weights (left) compared with feature maps from VGG16 of ALNet (right) for the classification of five classes of peripheral blood cells. (C) Greyscale histograms calculated from the features maps obtained from the VGG16 of ALNet. AML, acute myeloid leukaemia; APL, acute promyelocytic leukaemia.

progress using new models based on Generative Adversarial Networks (GANs) to standardize the images that feed the CNNs.

5. Conclusions

The final contribution of this paper is a predictive model designed with two serially connected convolutional networks and trained using a dataset with over 16,000 blood cell images obtained from clinical practise. It is proposed to assist clinical pathologists in the diagnosis of acute leukaemia during the blood smear review. It has been proved to distinguish neoplastic (leukaemia) and non-neoplastic (infections) diseases, as well as recognise the leukaemia lineage.

Authors' contributions

Laura Boldú designed the datasets, developed the classifiers, performed and evaluated the experiments, reviewed the literature and contributed to the manuscript writing.

Anna Merino supervised the overall project and contributed to the manuscript writing and editing.

Andrea Acevedo contributed to the algorithmic implementations and analysis of classification results.

Angel Molina contributed to data collection and morphological annotation.

José Rodellar advised on the deep learning classification models and contributed to the manuscript writing and editing.

Declaration of Competing Interest

We wish to confirm that there are not known conflicts of interest associated with this publication and that this research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors that could have influenced its outcome.

Acknowledgments

This work is part of a research project funded by the Ministry of Science and Innovation of Spain, with reference PID2019-104087RB-I00.

We thank Maite Serrando and Cristian López-Indiano for sharing the images acquired in their hospital.

Laura Boldú thanks financial support for the scholarship received from Siemens Healthcare S.L.U as a part of a collaborative agreement with the Hospital Clinic of Barcelona.

References

[1] A Miranda-Filho, M Piñeros, J Ferlay, I Soerjomataram, A Monnereau, F. Bray, Epidemiological patterns of leukaemia in 184 countries: a population-based study, Lancet Haematol. 5 (1) (2018) e14–e24.
[2] A Merino, L Boldú, A. Ermens, Acute myeloid leukaemia: How to combine multiple tools, Int. J. Lab. Hematol. 00 (2018) 1–11 Available from, doi:10.1111/ijlh.12831.
[3] J Rodellar, S Alférez, A Acevedo, A Molina, A. Merino, Image processing and machine learning in the morphological analysis of blood cells, Int. J. Lab. Hematol. (40) (2018) 46–53.
[4] S Alférez, A Merino, L Bigorra, L Mujica, M Ruiz, J. Rodellar, Automatic recognition of atypical lymphoid cells from peripheral blood by digital image analysis, Am. J. Clin. Pathol. 143 (2015) 168–176.
[5] C Briggs, I Longair, M Slavik, K Thwaite, R Mills, V Thavaraja, et al., Can automated blood film analysis replace the manual differential? An evaluation of the CellaVision DM96 automated image analysis system, Int. J. Lab. Hematol. 31 (1) (2009) 48–60.
[6] MA Alsalem, AA Zaidan, BB Zaidan, M Hashim, HT Madhloom, ND Azeez, et al., A review of the automated detection and classification of acute leukaemia: Coherent taxonomy, datasets, validation and performance measurements, motivation, open challenges and recommendations, Comput. Methods Programs Biomed. 158 (2018) 93–112.
[7] HT Salah, IN Muhsen, ME Salama, T Owaidah, SK. Hashmi, Machine learning applications in the diagnosis of leukemia: current trends and future directions, Int. J. Lab. Hematol. 41 (6) (2019) 717–725.
[8] H El Achi, JD. Khoury, Artificial Intelligence and digital microscopy applications in diagnostic hematopathology, Cancers (Basel) 12 (4) (2020) 797.
[9] N Radakovich, M Nagy, A. Nazha, Machine learning in haematological malignancies, Lancet Haematol. 7 (7) (2020) e541–e550.
[10] R Shouval, JA Fein, B Savani, M Mohty, A. Nagler, Machine learning and artificial intelligence in haematology, Br. J. Haematol. (2020).
[11] M Habibzadeh, A Krzyżak, T. Fevens, White blood cell differential counts using convolutional neural networks for low resolution images, Int. Conf. Artif. Intell. Soft Comput. (2013) 263–274.
[12] M-C Su, C-Y Cheng, P-C. Wang, A neural-network-based approach to white blood cell classification, Sci. world J. (2014).
[13] J Rawat, A Singh, HS Bhadauria, J Virmani, JS. Devgun, Application of ensemble artificial neural network for the classification of white blood cells using microscopic blood images, Int. J. Comput. Syst. Eng. 4 (2–3) (2018) 202–216.
[14] F Qin, N Gao, Y Peng, Z Wu, S Shen, A. Grudtsin, Fine-grained leukocyte classification with deep residual learning for microscopic images, Comput. Methods Programs Biomed. 162 (2018) 243–252.
[15] AI Shahin, Y Guo, KM Amin, AA. Sharawi, White blood cells identification system based on convolutional deep neural learning networks, Comput. Methods Programs Biomed. 168 (2019) 69–80.
[16] A Acevedo, S Alférez, A Merino, L Puigví, J. Rodellar, Recognition of peripheral blood cell images using convolutional neural networks, Comput. Methods Programs Biomed. 180 (2019) 105020.
[17] JW Choi, Y Ku, BW Yoo, J-A Kim, DS Lee, YJ Chai, et al., White blood cell differential count of maturation stages in bone marrow smear using dual-stage convolutional neural networks, PLoS One 12 (12) (2017) e0189259.
[18] H El Achi, T Belousova, L Chen, A Wahed, I Wang, Z Hu, et al., Automated diagnosis of lymphoma with digital pathology images using deep learning, Ann. Clin. Lab. Sci. 49 (2) (2019) 153–160.
[19] TTP Thanh, C Vununu, S Atoev, S-H Lee, K-R. Kwon, Leukemia blood cell image classification using convolutional neural network, Int. J. Comput. Theory Eng. 10 (2) (2018) 54–58.
[20] S Shafique, S. Tehsin, Acute lymphoblastic leukemia detection and classification of its subtypes using pretrained deep convolutional neural networks, Technol. Cancer Res. Treat. 17 (2018) 1533033818802789.
[21] T Pansombut, S Wikaisuksakul, K Khongkraphan, Phon-on A. Convolutional neural networks for recognition of lymphoblast cell images, Comput. Intell. Neurosci. (2019).
[22] N Ahmed, A Yigit, Z Isik, A. Alpkocak, Identification of leukemia subtypes from microscopic images using convolutional neural network, Diagnostics 9 (3) (2019) 104.
[23] KK Jha, HS. Dutta, Mutual information based hybrid model and deep learning for Acute Lymphocytic Leukemia detection in single cell blood smear images, Comput. Methods Programs Biomed. 179 (2019) 104987.
[24] J Prellberg, O. Kramer, Acute lymphoblastic leukemia classification from microscopic images using convolutional neural networks, in: ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging, Springer, 2019, pp. 53–61.
[25] C Matek, S Schwarz, K Spiekermann, C. Marr, Human-level recognition of blast cells in acute myeloid leukaemia with convolutional neural networks, Nat. Mach. Intell. 1 (11) (2019) 538–544.
[26] M Loey, M Naman, H. Zayed, Deep transfer learning in diagnosing leukemia in blood cells, Computers 9 (2) (2020) 29.
[27] LHS Vogado, RMS Veras, FHD Araujo, RR V Silva, KRT. Aires, Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification, Eng. Appl. Artif. Intell. 72 (2018) 415–422.
[28] C Di Ruberto, A Loddo, G. Puglisi, Blob detection and deep learning for leukemic blood image analysis, Appl. Sci. 10 (3) (2020) 1176.
[29] A Rehman, N Abbas, T Saba, SI UR Rahman, Z Mehmood, H. Kolivand, Classification of acute lymphoblastic leukemia using deep learning, Microsc. Res. Tech. 81 (11) (2018) 1310–1317.
[30] F Huang, P Guang, F Li, X Liu, W Zhang, W. Huang, AML, ALL, and CML classification and diagnosis based on bone marrow cell morphology combined with convolutional neural network: a STARD compliant diagnosis research, Medicine (Baltimore) 99 (45) (2020) 1–8.
[31] G Gutiérrez, A Merino, A Domingo, JM Jou, JC. Reverter, EQAS for peripheral blood morphology in Spain: a 6-year experience, Int. J. Lab. Hematol. 30 (6) (2008) 460–466.
[32] D Chabot-Richards, K. Foucar, Does morphology matter in 2017? An approach to morphologic clues in non-neoplastic blood and bone marrow disorders, Int. J. Lab. Hematol. 39 (2017) 23–30.
[33] BJ. Bain, in: Leukaemia Diagnosis, 4th ed., Wiley-Blackwell, Chichester, UK, 2010, pp. 68–73.
[34] KA Breen, D Grimwade, BJ. Hunt, The pathogenesis and management of the coagulopathy of acute promyelocytic leukaemia, Br. J. Haematol. 156 (1) (2012) 24–36.
[35] L Boldú, A Merino, S Alférez, A Molina, A Acevedo, J. Rodellar, Automatic recognition of different types of acute leukaemia in peripheral blood by image analysis, J. Clin. Pathol. 72 (11) (2019) 755–761.
[36] DA Arber, A Orazi, R Hasserjian, J Thiele, MJ Borowitz, MM Le Beau, et al., The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia, Blood 127 (20) (2016) 2391–2405.
[37] V. Subramanian, Deep Learning with PyTorch: a Practical Approach to Building Neural Network Models Using PyTorch, Packt Publishing Ltd, 2018.
[38] G Litjens, T Kooi, BE Bejnordi, AAA Setio, F Ciompi, M Ghafoorian, et al., A survey on deep learning in medical image analysis, Med. Image Anal. 42 (2017) 60–88.
[39] J Yosinski, J Clune, Y Bengio, H. Lipson, How transferable are features in deep neural networks? in: Advances in Neural Information Processing Systems, 2014, pp. 3320–3328.
[40] A. Géron, Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, O'Reilly Media, 2019.
[41] O Russakovsky, J Deng, H Su, J Krause, S Satheesh, S Ma, et al., Imagenet large scale visual recognition challenge, Int. J. Comput. Vis. 115 (3) (2015) 211–252.
[42] K Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, 2014 arXiv Prepr arXiv14091556.
[43] K He, X Zhang, S Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[44] G Huang, Z Liu, L Van Der Maaten, KQ. Weinberger, Densely connected convolutional networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4700–4708.
[45] J Hu, L Shen, S Albanie, G Sun, E. Wu, Squeeze-and-Excitation Networks, 2017.

[46] S Alferez, A Merino, L Boldú, A Acevedo, A Molina, J. Rodellar, A deep learning approach to automatically classify pathological cell images in peripheral blood, in: ISLH 2019 Abstract Proceedings, 2019.
[47] A Merino, S Alférez, L Boldú, A Molina, L Puigví, A Acevedo, et al., Automatic differentiation of acute leukaemia, lymphoma and reactive lymphocytes in peripheral blood using a novel convolutional network, in: ISLH 2020 Abstract Proceedings, 2020.
[48] DP Kingma, Ba J. Adam, A Method for Stochastic Optimization, 2014 arXiv Prepr arXiv14126980.
[49] LN. Smith, Cyclical learning rates for training neural networks, in: 2017 IEEE Winter Conference on Applications of Computer Vision (WACV, 2017, pp. 464–472.
[50] DJ Hand, RJ. Till, A simple generalisation of the area under the ROC curve for multiple class classification problems, Mach. Learn. 45 (2) (2001) 171–186.
[51] A Merino, L Puigví, L Boldú, S Alférez, J. Rodellar, Optimizing morphology through blood cell image analysis, Int. J. Lab. Hematol. 40 (2018) 54–61.

UNIVERSITAT DE BARCELONA

DIRECTORS' REPORT TO DEPOSIT THE THESIS BY COMPENDIUM OF ARTICLES

We declare that the Thesis whose title is "**Automatic recognition of different types of acute leukaemia using peripheral blood cell images**" presented by **Laura Boldú Nebot**, which we have supervised, meet the requirements to be presented as a compendium of articles. In both publications the PhD student is the first author and none of these articles have been previously used for any other doctoral thesis.

The complete reference of the first article is the following:

Boldú L, Merino A, Alférez S, Molina A, Acevedo A, Rodellar J. Automatic recognition of different types of acute leukaemia in peripheral blood by image analysis. *Journal of Clinical Pathology*, 2019;72(11):755-761. doi: 10.1136/jclinpath-2019-205949. PMID: 31256009.

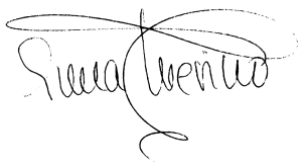
Quality index: JCR 2019 Pathology, *Impact factor*: 2.460, Q2 (30/78).

The complete reference of the second article is the following:

Boldú L, Merino A, Acevedo A, Molina A, Rodellar J. A deep learning model (ALNet) for the diagnosis of acute leukaemia lineage using peripheral blood cell images. *Computer Methods and Programs in Biomedicine*, 2021;202:105999. doi: 10.1016/j.cmpb.2021.105999. PMID: 33618145.

Quality index: JCR 2019 Computer Science, Theory & Methods, *Impact factor*: 3.632, Q1 (16/108).

Date: 30th April of 2021



Anna Merino González

José Rodellar Benedé

CHAPTER 6

Results and Discussion

This chapter consists of a brief summary of the main results and a discussion of the main contributions derived from this thesis, in which an automatic classification system to predict the diagnosis of acute leukaemia using digital images of PB cells was designed, developed and tested. The main goal was to develop a practical support tool to assist the clinical pathologist to reach a more objective and accurate diagnosis when there is a suspicion of acute leukaemia.

For the first diagnosis of acute leukaemia in clinical laboratories, it is important the detection of quantitative abnormalities in leucocyte count, haemoglobin level or thrombocyte count, which triggers the smear review. The detection of blasts circulating in blood is the subsequent step. Thus, cell morphology is crucial for the initial diagnosis of leukaemia lineage to apply the suitable treatment and avoid delays in medical procedures, primarily in APL. This is why the automatic recognition of blasts has been a target for automated solutions, tools and methods within the artificial intelligence framework (102,103).

Most of the previous studies in the literature focused on recognising lymphoblasts from normal leukocytes, which show very different morphology since the nucleus of neutrophils, eosinophils and basophils is lobulated and their cytoplasm shows abundant granules. It is fair to mention that these significant morphological differences with normal leukocytes make the classification problem more accessible. Moreover, those studies which classified lymphoblasts subtypes were based on the FAB classification, which currently has been replaced by the WHO 2016 classification (16) for clinical practise. Many of these previous studies focused on a 'binary' classification (disease vs. normal), which is not a realistic approach and does not reflect the real-life complexity of haematological malignancies diagnosis (103).

The motivation behind this thesis was to develop a system capable of distinguishing those cell groups in which their recognition by morphology is usually difficult in regard to acute leukaemia diagnosis. The challenge was twofold: 1) differentiate blasts among other mononuclear cells, such as reactive lymphocytes and normal lymphocytes and monocytes, and 2) discern between myeloid blasts and B-lymphoblasts because of the overlapping morphological characteristics they exhibit.

6.1. Automatic Recognition of Different Types of Acute Leukaemia in Peripheral Blood by Image Analysis

In this section, the results derived from the first publication (104) of this thesis are summarized. In this first work, five classical machine learning classification techniques were explored for the automatic recognition of six cell groups: normal lymphocytes and monocytes from healthy individuals as the control group, reactive lymphocytes from patients with viral or other infections, myeloid and monocytic blasts from patients with AML, abnormal promyelocytes from patients with APL and B-lymphoblasts from patients with ALL. Figure ***6.1*** illustrates the main steps characteristic of the image-based recognition systems which we followed to develop the first classifier.

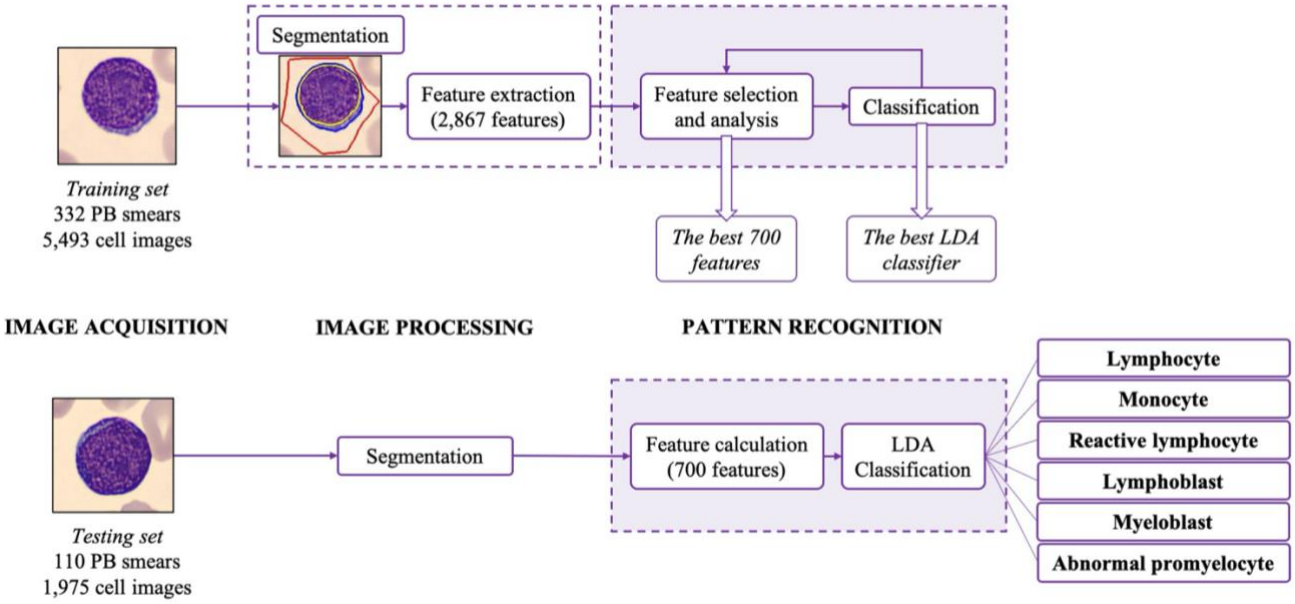


Figure 6.1. Diagram illustrating the steps followed to develop the classifier. Firstly, the *training set* of 332 smears (5,493 cell images) was processed through the automatic segmentation procedure to obtain a database of features, which was used to tune the best classifier through an iterative process involving 5-fold cross validation, where the most relevant features were determined by the accuracy of the classifier. The final recognition module included a classifier based on linear discriminant analysis (LDA), whose input was the set of the 700 most relevant features. Secondly, the *testing set* of new 110 smears (1,975 cell images) was processed following the preceding steps. This final stage consisted of the assessment of the system for cell classification.

In this work, we combined the classification algorithms with the latest automated segmentation algorithm published by our research group (105) and with feature extraction for the automatic recognition of blasts, reactive lymphocytes and other normal mononuclear cells (lymphocytes and

monocytes), as well as for the distinction among myeloid blasts, B-lymphoblasts and abnormal promyelocytes. The automatic recognition of this wide variety of cell types was a new contribution since it had not previously been accomplished.

Five classification techniques were studied to face the challenging number of cell types under study: linear discriminant analysis (LDA), k-nearest neighbour, naïve Bayes, support vector machine and random forest. A total of 2,867 quantitative features (28 geometric and 2,839 of colour and texture) was extracted for each segmented cell, and they were ranked based on relevance criterion by using conditional mutual information maximization criteria (106). A more detailed analysis and interpretation of the most relevant features could be found in (104).

As shown in Figure ***6.2***, we trained and tuned the classifiers using different numbers of the best-ranked features and displayed the corresponding overall accuracy (ratio of images correctly classified in their true category). The highest accuracy was achieved with the LDA classifier when using the best 700 features.

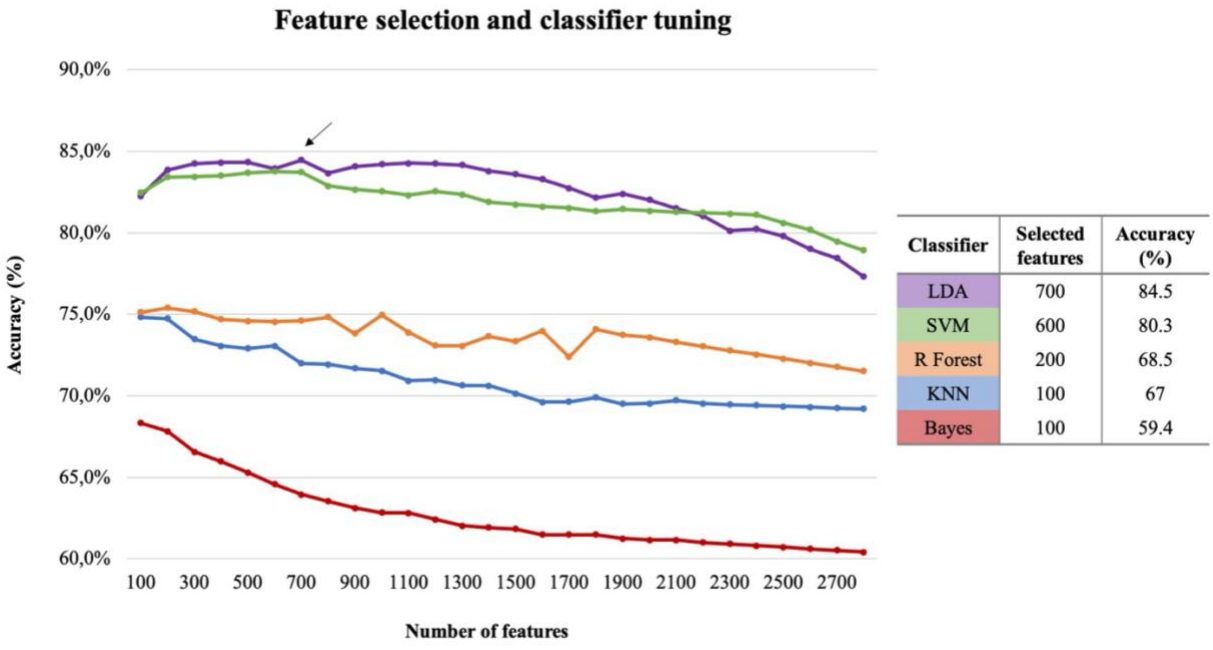


Figure 6.2. Plot of the five classifiers trained and tuned with all the 2,867 features to obtain the combination of the best number of features and the classifier with the highest overall classification accuracy. The x-axis shows the number of features and y-axis shows the accuracy values (in percentage) for each selected number of features. The highest classification accuracy (marked with an arrow) was achieved with the selection of the most relevant 700 features with LDA classifier. Bayes, naïve Bayes; KNN, k-nearest neighbour; LDA, linear discriminant analysis; R forest, random forest; SVM, support vector machine.

This LDA classifier was further validated through a blind classification using all the images of the *testing set*, which were not previously used in any of the training steps.

6.1.1. Classification by individual cells

The confusion matrix that summarises the classification results of this first classifier is shown in Figure **6.3**. The true positive rates shown in the main diagonal were: 97.7% for reactive lymphocytes, 97.6% for lymphocytes, 93% for monocytes, 80.8% for myeloid blasts (myeloblasts and monoblasts), 72.6% for abnormal promyelocytes and 78.9% for B-lymphoblasts. The overall classification accuracy was 85.8%. Most of the individual cell images were automatically identified in the group of its true class.

		Predicted class					
		Lymphoid blasts	**Abnormal Promyelocytes**	**Myeloid blasts**	**Reactive lymphocytes**	**Lymphocytes**	**Monocytes**
True class	**Lymphoid blasts**	**78.9**	1.5	16.5	2.1	1	0
	Abnormal Promyelocytes	14	**72.6**	13.4	0	0	0
	Myeloid blasts	14	1.6	**80.8**	1.6	0.1	1.9
	Reactive lymphocytes	0.4	0	1.1	**97.7**	0	0.8
	Lymphocytes	0	0	0	2	**97.6**	0.4
	Monocytes	1.3	0.4	1.3	4	0	**93**

Figure 6.3. Confusion matrix of the classification results (in %) for the images of the *testing set*. Rows indicate the true class and columns represent the predicted class supplied by the classifier. Diagonal values are the true positive rates for each cell type. The overall classification accuracy was 85.8%.

6.1.2. Classification by individual smears

Once the classifier was ready for the classification of individual cell images, our strategy was to predict patient's diagnosis using the blood smear from an individual patient as a diagnostic unit. The system input was a set of cell images of an individual smear and the output was the prediction of one of the following diagnoses: AML, APL, ALL or infection. The diagnosis was made by identifying the cell class that predominated in the smear. This required a previous step to establish

a threshold value such that the diagnosis was predicted by identifying the cell class with the percentage of images classified above this value. It was found that 50% of the cell images correctly classified was the best threshold to predict patient's diagnosis through the smear, obtaining an area under the ROC curve of 0.99.

Such threshold could be interpreted in such a way that if more than 50% of the cell images of a smear are classified as myeloid blasts, for example, the predicted diagnosis of the patient to whom this smear belongs is AML. Whereas if similar percentages are obtained for more than one class, in this case it is not possible to predict a diagnosis and the system considers the smear as belonging to an "unknown" diagnostic group.

Considering the 50% threshold, 100% of smears corresponding to patients with infections were correctly classified, as well as those containing lymphocytes or monocytes (accuracies of 100% and 97%, respectively). In regard to acute leukaemia, the true positive rates were 88% for AML, 85% for ALL and 80% for APL. The overall classification accuracy of individual smears was 94%. Moreover, sensitivity values above 97% were obtained for normal smears and those related to infections and above 80% for all leukaemia subtypes, while specificity values for all categories were above 96%.

The key contribution of this first work was a system which had integrated feature selection to lead to the most relevant quantitative features to achieve the highest percentage of classification accuracy. Moreover, such detailed analysis and interpretation of the most relevant features to understand their importance in the classification of blast cells had not previously been done.

6.2. A Deep Learning Model for the Diagnosis of Acute Leukaemia Lineage using Peripheral Blood Cell Images

Traditional machine learning approaches required to segment and extract features manually. This meant that all possible variations in the morphology of both normal and abnormal cells had to be considered, which was truly difficult. Specially segmentation was considered the most critical step, mainly because of the complex morphology and subtle differences among cells, the variability of the blood smear staining and the general conditions of image acquisition. All these aspects directly resulted in a decrease of the classification accuracy of these classical models.

This section summarises the results derived from the second publication (107) of this thesis, where we tried to overcome some of the limitations of these traditional frameworks by addressing the use of new emerging deep learning techniques for the automatic recognition of the same six groups of cells. The goal was to improve the previous accuracy for the leukaemia lineage

differentiation by designing and developing a new CNN-based system, which had to be highly sensitive and specific for its integration as a decision support system to assist pathologists when there is a suspicion of acute leukaemia.

The main objective of this second study was to design an automatic classification system based on CNNs and focused on achieving the following specifications: 1) the prompt and accurate detection of APL, motivated by the emergency of avoiding severe bleeding and disseminated intravascular coagulation with patients' death; and 2) the discrimination between myeloid and lymphoid leukaemia lineages, motivated by the different strategical therapies they require.

We adopted the concept of fine-tuning to take advantage of the knowledge of a CNN previously trained for another task and re-train and adapt some of its layers for the classification of the images of our study. We investigated the effects of four well-known CNNs on model performance: VGG16, ResNet101, DenseNet121 and SENet154. To select the best model, CNN frameworks were trained and evaluated using two different approaches: 5-fold cross validation and hold-out.

Table **6.1** shows the overall classification accuracies for the four evaluated CNNs with 5-fold cross validation and hold-out when changing the number of convolutional blocks trained with 470 iterations. The highest accuracies were obtained when fine-tuning the entire models. This means to train all the convolutional blocks with our own image dataset. Moreover, the highest testing accuracies were obtained with VGG16 and SENet154 architectures (above 85% using hold-out and above 86.8% with 5-fold cross validation) as it is shown in Table **6.1**. VGG16 showed an accuracy value of 88.4% using hold-out with respect to the best accuracy (87.2%) obtained with SENet154.

Finally, we decided to select VGG16 together with the hold-out approach as the best model because of the following: 1) less overfitting was obtained as VGG16 showed better performance when making prediction with the new images of the *testing set*; and 2) it is a simpler architecture compared to the other CNN frameworks and had the best training and classification time, which is an advantage for a potential real-time implementation (see Table **6.1**).

	Model	Number of convolutional blocks fine-tuned 1	2	3	4	Whole	Testing	Training time (min)	Class. time (s)
Hold-out	**VGG16**	89.2%	88.0%	88.2%	88.6%	94.6%	**88.4%**	6.78	12
	ResNet101	90.1%	89.7%	90.3%	90.2%	93.3%	84.0%	12.52	22.4
	DenseNet121	84.4%	86.8%	87.5%	88.0%	93.6%	84.8%	7.9	14
	SENet154	89.3%	89.8%	89.8%	88.9%	94.6%	**85.0%**	48.2	41
5-fold cross validation	**VGG16**	76.3% ±0.93	76.5% ±0.67	76.2% ±0.97	76.1% ±1.36	**86.9%** ±0.68		123.1	29.5
	ResNet101	80.8% ±1.99	81.2% ±0.54	80.2% ±0.56	81.8% ±0.41	86.8% ±0.62		136.2	32
	DenseNet121	82.6% ±1.10	82.9% ±1.33	82.4% ±1.20	83.1% ±1.19	86.3% ±0.93		262.5	63
	SENet154	79.5% ±0.72	80.2% ±0.83	80.5% ±0.58	80.3% ±0.60	**87.2%** ±1.22		426.7	81.5

Table 6.1. Overall accuracies of the pre-trained CNNs when changing the number of convolutional blocks to be trained with 470 iterations and with the two evaluation approaches: hold-out and 5-fold cross-validation. For the hold-out approach, validation and testing accuracies are shown. For the 5-fold cross validation, we present the mean and the standard deviation of the testing accuracy computed among the five folds. The last two columns show the training time (in minutes) achieved when fine-tuning the entire convolutional blocks, and the classification time (in seconds) when evaluating the selected models with the respective *testing sets*.

The true positive rates obtained with the VGG16 for myeloid blasts (87%) and B-lymphoblasts (75%) were concluded not to be high enough for our purpose of building a tool to predict the diagnostic orientation of acute leukaemia in a clinical setting.

6.2.1. Sequential CNN classification system: ALNet

In order to improve the recognition between myeloid and lymphoid blasts, we proposed a new strategy with a two-step classification scheme (see Figure **6.4**), where two separate classifiers worked in series: ALNet. The first step (module 1) consisted of a VGG16 trained to distinguish abnormal promyelocytes among lymphocytes, monocytes, reactive lymphocytes and blast cells, which encompassed both myeloid blasts and B-lymphoblasts. The second step (module 2) required a VGG with an increased number of convolutional layers (VGG19) to discriminate between myeloid blasts and B-lymphoblasts.

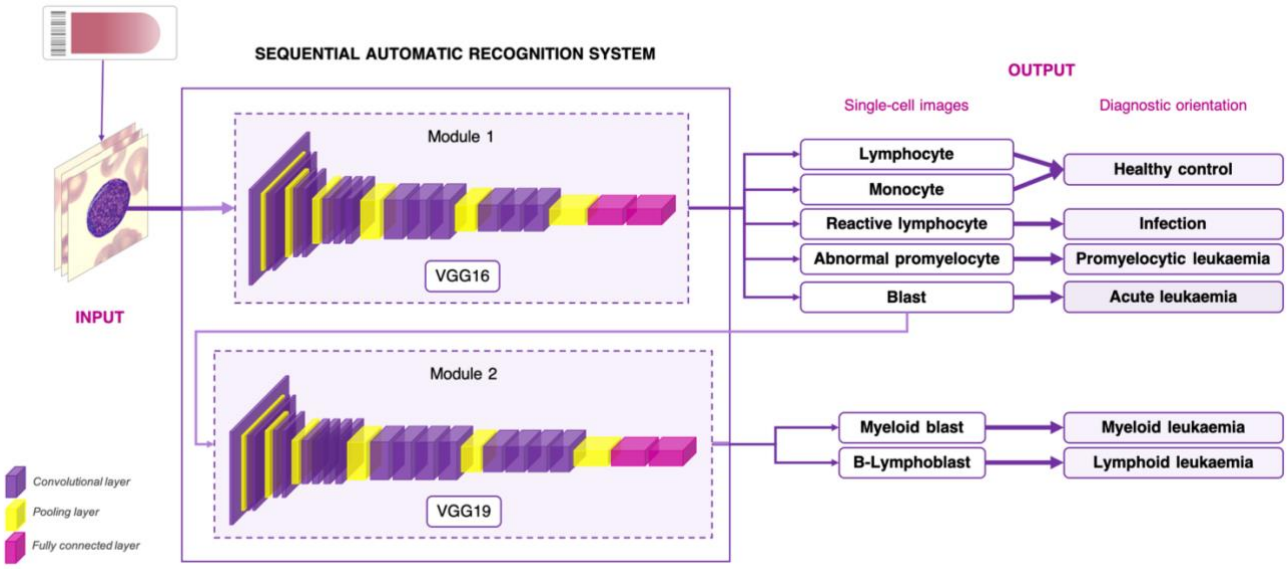


Figure 6.4. Diagram illustrating the proposed sequential CNN-based system (ALNet) for the automatic recognition of acute leukaemia lineage. It starts with the selection of the cell images of an individual smear by the clinical pathologist, which are the inputs to the system. ALNet includes two consecutive modules. The first module recognises acute promyelocytic leukaemia (APL) and acute leukaemia (non-APL) from the remaining groups. Only the smears corresponding to patients with non-APL leukaemia (myeloid or lymphoid) are used for lineage classification through the second module.

The first assessment of ALNet was done through a blind classification of all the single-cell images of the *testing set* (108 smears with 4,043 images). Figure ***6.5***A shows the confusion matrix of the classification results of module 1. The true positive rates shown in the main diagonal were: 99.9% for lymphocytes, 97.6% for monocytes, 97.2% for reactive lymphocytes, 95.3% for abnormal promyelocytes and 91.7% for blasts (myeloid blasts and B-lymphoblasts). The overall classification accuracy was 94.2%. Figure ***6.5***B summarizes the classification results for blast cells. As seen in the main diagonal, 99.4% of images corresponding to myeloid blasts and 82% to B-lymphoblasts were correctly classified. The overall classification accuracy was 89.5%.

A

True class \ Predicted class	Lymphocytes	Monocytes	Reactive lymphocytes	Abnormal Promyelocytes	Blasts
Lymphocytes	**99.9**	1	0	0	0
Monocytes	0.8	**97.6**	0	0.8	0.8
Reactive lymphocytes	0.2	0.2	**97.2**	0	2.4
Abnormal Promyelocytes	0	0.4	0	**95.3**	4.3
Blasts	2.5	0.2	1.3	4.3	**91.7**

B

True class \ Predicted class	Myeloid blasts	B-Lymphoblasts
Myeloid blasts	**99.9**	1
B-Lymphoblasts	0.8	**97.6**

Figure 6.5. Confusion matrix of the classification results (in %) for the images of the testing set of modules 1 (A) and 2 (B). Rows indicate the true class and columns represent the predicted class supplied by the network. Diagonal values are the true positive rates for each cell type. The overall classification accuracies of the first (A) and second module (B) were 94.2% and 89.5%, respectively.

An important aspect to remark of this second work was the number of images involved and their high quality, being the largest dataset used for leukaemia classification. We used a dataset with 16,450 PB cell images, 4,825 being blasts. Furthermore, we guaranteed the confirmation of their labels through other complementary tests (clinical data, morphology, flow cytometry, cytogenetics and molecular biology). Not only high-quality and a large number of images are required for developing diagnostic systems, but also the availability of images properly annotated by experts, which was scarce in literature (103,108,109). The quality of a dataset is essential to observe morphological characteristics which can lead towards a diagnosis, not only for daily clinical practise but also to obtain robust models avoiding overfitting. This is why, visualizing the feature maps generated by the intermediate convolutional layers of ALNet could give us an idea of which patterns the network extracted, and thus helped to interpret the classification results.

6.2.2. ALNet in a clinical setting

ALNet was also assessed to predict patient's diagnosis using the smear as a classification unit. It was also found that 50% of the cell images correctly classified was the best threshold to predict patient's initial diagnosis through the smear, obtaining values of 1 for the area under the ROC curve.

The confusion matrix in Figure **6.6**A shows the classification results of the first module using one by one all the smears in the *testing set*. Sensitivity, specificity and precision values of 100% were obtained for all the categories. In consequence, using the first classification module of ALNet, we correctly detected the following groups: 1) healthy controls, 2) patients with infection, 3) patients with APL and 4) patients with acute leukaemia non-APL.

The smears corresponding to patients with non-APL leukaemia (AML or ALL) were classified through the second module, and the results are shown in the confusion matrix in Figure **6.6**B. Regarding AML, sensitivity, specificity and precision values of 100%, 92.3% and 93.7%, respectively, were obtained. As for ALL, a sensitivity of 89% and specificity and precision values of 100% were obtained. The overall accuracy of individual smears was 94.7% (see Figure **6.6**B).

At the end of the two-step classification with ALNet, we obtained the correct diagnostic prediction for all patients with APL and AML. With respect to the ALL group, a total of 24 from 27 smears were correctly classified, being the remaining three recognized as AML (two smears) or unknown (one smear).

When developing new diagnostic support tools for laboratory practise, it is important to consider individual patients when organising datasets for training and assessing the system. The automatic classification of blasts had always been addressed considering each cell image as a unit, without considering individual smears when arranging sets for both training and assessing the classifier. Using images from the same smear for both training and assessment could cause an accuracy overestimation, what happened in almost all the studies in the literature. An innovation of this thesis was that we arranged a set of smears for the system development and a different one for the assessment. For the first time, the PB smear was used as a diagnostic unit, which enabled the achievement of a satisfactory diagnostic ability when differentiating normal smears from those related to infections and with respect to smears with abnormal leukemic cells.

A

True diagnosis \ Predicted diagnosis	Control (LY)	Control (MO)	Infection	APL	AL non-APL
Control (LY)	**100% (13)**	0	0	0	0
Control (MO)	0	**100% (4)**	0	0	0
Infection	0	0	**100% (46)**	0	0
APL	0	0	0	**100% (14)**	0
AL non-APL	0	0	0	0	**100% (57)**

B

True diagnosis \ Predicted diagnosis	AML	ALL	UNK
AML	**100% (30)**	0	0
ALL	7% (2)	**89% (24)**	4% (1)

Figure 6.6. Confusion matrix of the classification results (in %) for the smears of the *testing set* of modules 1 (A) and 2 (B) taking the threshold of 50% into consideration. Rows indicate the true diagnosis and columns represent the predicted diagnosis supplied by ALNet. Diagonal values are the true positive rates for each smear (in brackets the number of smears). The overall classification accuracies of the first (A) and second module (B) were 100% and 94.7%, respectively. LY, lymphocytes; MO, monocytes; APL, acute promyelocytic leukaemia; AL, acute leukaemia; AML, acute myeloid leukaemia; ALL, acute lymphoid leukaemia; UNK, unknown.

In addition, to further evaluate the ALNet performance in a clinical setting, we used a new image dataset from two other hospitals (*Josep Trueta* and *Germans Trias i Pujol*). This dataset contained a total of 381 images of blast cells (322 myeloid blasts and 59 B-lymphoblasts). When classifying by single-cell images, 98% of blast cells were correctly classified by module 1. Module 2 correctly classified 82% of myeloid blasts and 64% of B-lymphoblasts. When using the threshold of 50% to predict the diagnosis from the smear, in the first classification module of ALNet all the smears (100%) were classified as acute leukaemia non-APL. Regarding the leukaemia lineage, 3/5 smears corresponding to AML (60%) and 1/2 ALL smears (50%) were correctly classified.

When using images from two other hospital, sensitivity for the leukaemia lineage discrimination decreased. Although all datasets were stained with the 'gold standard' May Grünwald-Giemsa technique used in clinical laboratories, we observed that variations in the optical conditions and resolution of microscopic images affected the classification accuracy of the CNN models. It is known that colour information is involved in the quantification of cytoplasmic basophilia and granulation, both being very important characteristics in the classification of myeloid blasts and lymphoblasts (91). This is why changes related to image clarity, colour scale or resolution could mislead their differentiation.

The fact of having used only a single-institution image data source to develop ALNet could represent a limitation for its current implementation in other clinical laboratories. The different staining protocols of blood smears, the optical equipment to perform the blood cell analysis and the methods for image digitalization could result in accuracy variations when using ALNet.

Regarding the scientific and technological impact of this thesis, the approaches developed have a huge clinical impact as they could be new support tools for quick diagnosis of acute leukaemia during the blood smear review. They have been proved to distinguish neoplastic (leukaemia) and non-neoplastic (infections) diseases, as well as recognise the leukaemia lineage. They have relatively low cost since depend on a microscope, which is usual in all laboratories, and software. This could ease their integration in the workflow and information systems of clinical laboratories, and allow strict selection of other diagnostics tests, like immunophenotype and molecular studies, which involve complexity and high cost in specialized hospitals.

Concerning the social and economic impact, people aged 60 years or older will rise from 900 million to 2 billion in the next future (WHO data), increasing the incidence of haematological diseases. This demographic change is a challenge since it may become a big social problem in the context of lower economical resources and therefore lower human resources in the health sector. In this scenario, the development of new computational technologies becomes indispensable to assist pathologists to give a prompt and efficient diagnosis, such as the approaches developed in this thesis. Furthermore, there is a global need to develop new diagnostic tools that support the laboratory daily practise, especially for those with low resources.

CHAPTER 7

Conclusions

This thesis has contributed in developing tools required for the initial diagnostic orientation of acute leukaemia to support clinical pathologists during their daily clinical practise. The thesis has grown through the evolution of various works, starting with a discrimination among normal mononuclear cells, reactive lymphocytes and three types of leukemic cells using traditional machine learning techniques, and ending with a new predictive system designed with two serially connected convolutional networks, which has been proved to distinguish neoplastic (leukaemia) and non-neoplastic (infections) diseases, as well as recognise the leukaemia lineage. Furthermore, this new system has been validated in a real-clinical setting. This thesis has also contributed in advancing the state of the art of the automatic recognition of acute leukaemia by providing a more realistic approach which reflects the real-life complexity of acute leukaemia diagnosis.

The most significant conclusions of this thesis are summarised below:

- Several classification techniques have been implemented to automatically recognise different blasts from myeloid and lymphoid origin, abnormal promyelocytes, reactive lymphocytes, lymphocytes and monocytes. The automatic recognition of this wide variety of cell types is a new contribution since it has not previously been accomplished and provides a more realistic approach reflecting the real-life complexity of acute leukaemia diagnosis.

- The blood smear has been used as a diagnostic unit, which has enabled the achievement of a satisfactory diagnostic ability when differentiating normal smears from those related to infections and with respect to smears with abnormal leukemic cells.

The specific conclusions derived from the first approach can be further summarised as follows:

- The full set of extracted features has been able to describe size, shape, colour and texture of the different cell types included in this thesis, involving the calculation of 28 geometric and 2,839 colour and texture features. The further successful classification results show that this feature set provides excellent cell characterisation.

- Feature selection is a necessary step to decrease complexity and computation time of the algorithm associated to a high number of features. In this thesis many classification experiments have been carried out with different combinations of geometric, colour and texture features, applying conditional mutual information maximisation criteria.

- It has been found that 700 features were the most relevant and less redundant for the classification step, which included a combination of geometric, colour and texture features. Moreover, the most representative colour spaces to distinguish the cell groups of this thesis resulted to be the following: L*a*b, CMYK and RGB.

- The 700 most relevant features offer: 1) quantitative descriptions of cytological characteristics which are usually described in qualitatively terms, and 2) efficiency to automatically distinguish among a significant number of different blast cell lineages.

- The three most representative geometric features for the automatic recognition of blast cells are the *nuclear area*, the *cellular area* and the *Nucleus/Cell ratio*.

- The *kurtosis of the green-red of the cell* and the *blue correlation of the cell* are the first among the colour and texture features, respectively. They are related to the nucleus/cell relation, which allowed us to discriminate between reactive lymphocytes and myeloblasts or lymphoblasts, between lymphoblasts and myeloblasts, and between myeloblasts and monoblasts.

- The *mean of the blue of the nucleus* provides essential information for cell types with dispersed chromatin. Differences were observed among different myeloid blasts with immature chromatin and visible nucleolus, compared to the abundant azurophil granulation typical in abnormal promyelocytes.

- The *cluster shade of the blue-purple of the cell* allows the distinction among cells with thinner and condensed texture, demonstrating that information of the nuclear chromatin distribution pattern could be provided by grey level co-occurrence matrix (GLCM) features, which helped us to distinguish among myeloblasts, monoblasts, abnormal promyelocytes and lymphoblasts.

- In this thesis several traditional classification techniques have been implemented to automatically recognize different blasts from myeloid and lymphoid origin, abnormal promyelocytes, reactive lymphocytes, lymphocytes and monocytes. The best classification results have been achieved using a linear discriminant analysis (LDA) classifier.

- The methodology for the automatic blast cell recognition has been assembled with our own image segmentation algorithm followed by the integration of feature selection within the classification step, which has shown high accuracies in the system development stage.

The specific conclusions derived from the second approach are the following:

- Data augmenting techniques has been used to balance the number of images of each cell class during the development stage of the deep learning-based system, which allowed us to avoid overfitting and obtain satisfactorily high accuracies for cell classification.

- For the transfer learning approach implemented in this thesis, the last layers of the evaluated CNN architectures have been removed and three new fully connected layers have been coupled at the end of each network. With this design we took advantage of the knowledge of the pre-trained CNNs with ImageNet, which helped to overcome the deficit of training images and served as an effective weight initialisation to then classify our own image dataset.

- By fixing the number of iterations at 470 with a batch size of 64 and implementing the cycling rate policy has been possible to select the best CNN architecture for the automatic classification of the different leukaemia types considered in this thesis. Moreover, the highest accuracies have been obtained when fine-tuning the entire CNN models.

- VGG16 together with the hold-out approach has been selected as the best model because of the following: 1) less overfitting was obtained as VGG16 showed better performance when making predictions with new images; and 2) it is a simpler architecture compared to the other CNN frameworks and had the best training and classification time, which is an advantage for a potential real-time implementation.

- The new strategy (ALNet) proposed with the design of two separated classifiers working sequentially has improved the accuracy for leukaemia lineage differentiation obtaining excellent results for its two main specifications: 1) the prompt and accurate detection of promyelocytic leukaemia; and 2) the discrimination of myeloid and lymphoid leukaemia.

- Using hold-out as a validation technique with 2,500 images per group has made possible obtaining satisfactorily high accuracies as similar accuracies were obtained with the 5-fold cross validation approach.

- Visualizing the feature maps generated by the intermediate convolutional layers of ALNet has given us an idea of which patterns the network extracts and helped to interpret the classification results.

- The feature map 202 from the convolutional block 3 (layer 1) is able to precisely detect the cell outline, nucleus shape and texture, along with the recognition of the red blood cells and their central pallor and the circulating platelets, while the feature map 215 from the convolutional block 3 (layer 3) is capable of detecting the bilobed nucleus of the abnormal promyelocyte and the nucleolus of the myeloid blast.

- The relevance of ALNet for the diagnosis of patients with acute leukaemia has been evaluated in a real-clinical practise scenario performing a proof of concept with images from three other centres. A satisfactory diagnostic prediction has been achieved, although we concluded that variations in the optical conditions and resolution of microscopic images affected the classification accuracy of the models.

7.1. Future Perspectives

The results of this thesis are promising and provide hope for new computational technologies in the field of hematopathology to give a prompt and efficient diagnosis. Nevertheless, additional research is needed to further improve the accuracy of the system for diagnostic prediction. This section discusses potential future perspectives that can be studied from the bases of this work.

A current limitation for the practical implementation of ALNet in other laboratories may be that a single-institution data source has been used to train the model, which could result in accuracy variations related to the image staining procedure and the optical conditions and resolution of the microscopic images. To deal with this, our group has some work in progress using new models based on Generative Adversarial Networks (GANs) to generate artificial staining effects with the aim to standardize the images that feed the CNN models.

Regarding the most challenging recognition task, which is the automatic differentiation between myeloid and lymphoid blasts, the use of Bayesian tools such as Bayesian inference and Markov Monte Carlo methods could be useful for incorporating information from patients and diseases in the learning process. It would be interesting to investigate if the incorporation of demographic and biological data, such as haematological and biochemistry parameters, of acute leukaemia patients before its diagnosis, as well as epidemiological data of acute leukaemia diseases into the CNN models may be useful to improve the predictive ability of ALNet. Moreover, it also might be of interest to study new visualization tools to analyse the CNN's filters and neurons activations to incorporate information and clinical interpretation from the feature maps automatically generated in the learning process. Features maps would be analysed to search for patterns which can be associated with some specific morphological characteristics of blast cells. Examples of

these could be the “cup like” in the nucleus of myeloblasts from AML with mutation in the nucleophosmin (*NPM1*) gene or splinters in abnormal promyelocytes.

As further implementation, ALNet would be validated using a web application for real-time on-site operation by integrating it within a system hosted in a server. The whole system would be designed to assist the user (clinical pathologist) when he/she is visually analysing the cell images from a patient’s smear and has to give a diagnostic orientation. The first step would be to upload the set of images to the hosting server. Then, the classification of the cell images would be the output presented back to the user in graphical and numerical form. This workflow would be designed as a prototype to be operative from a computer to work in real time, which would offer the possibility to perform a multicentric proof of concept and increase the number of new patients. Alternatively, the system could be easily integrated within a large workflow in a laboratory that includes other types of analysers and information sources about patients.

References

1. Alférez Baquero ES. Methodology for automatic classification of atypical lymphoid cells from peripheral blood cell images. Universitat Politècnica de Catalunya, 2015.

2. Puigví Fernández L. Caracterització morfològica de cèl·lules limfoides normals, reactives, anormals i blàstiques de sang perifèrica mitjançant processament digital d'imatges. Universitat Politècnica de Catalunya, 2019.

3. Milena A, Lipes A. Deep Learning System for the Automatic Classification of Normal and Dysplastic Peripheral Blood Cells as a Support Tool for the Diagnosis. University of Barcelona, 2020.

4. Ciesla B. Hematology in practice. FA Davis; 2018.

5. Bain BJ, Clark DM, Wilkins BS. Bone marrow pathology. John Wiley & Sons; 2019.

6. Merino A. Manual de Citología de Sangre Periférica y Líquidos Biológicos. Editor Panam Madrid, Spain. 2019.

7. Bain BJ. Leukaemia diagnosis. John Wiley & Sons; 2017.

8. Charles NJ, Boyer DF. Mixed-phenotype acute leukemia: diagnostic criteria and pitfalls. *Arch Pathol Lab Med*. 2017;141(11):1462–8.

9. De Kouchkovsky I, Abdul-Hay M. Acute myeloid leukemia: a comprehensive review and 2016 update. *Blood Cancer J*. 2016;6(7):e441--e441.

10. Miranda-Filho A, Piñeros M, Ferlay J, Soerjomataram I, Monnereau A, Bray F. Epidemiological patterns of leukaemia in 184 countries: a population-based study. L*ancet Haematol*. 2018;5(1):e14--e24.

11. Bennett JM, Catovsky D, Daniel M-T, Flandrin G, Galton DAG, Gralnick HR, et al. Proposals for the classification of the acute leukaemias French-American-British (FAB) co-operative group. *Br J Haematol*. 1976;33(4):451–8.

12. Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DAG, Gralnick HR, et al. Proposed revised criteria for the classification of acute myeloid leukemia: a report of the French-American-British Cooperative Group. *Ann Intern Med*. 1985;103(4):620–5.

13. Bain B. Acute leukemia. Bain B (ed), Leukemia, Diagnosis. A guide to the FAB classification. Gower Medical Publishing. London; 1990.

14. Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri S, Stein H, et al. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. revised 4t. Lyon, France: IARC Press; 2017.

15. Cazzola M. Introduction to a review series: the 2016 revision of the WHO classification of tumors of hematopoietic and lymphoid tissues. *Blood*. 2016;127(20):2361–4.

16. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016;127(20):2391–405.

17. Merino A, Boldú L, Ermens A. Acute myeloid leukaemia: How to combine multiple tools. *Int J Lab Hematol*. 2018;00:1–11. Available from: https://doi.org/10.1111/ijlh.12831.

18. Merino A. Clasificación de las leucemias agudas mieloides. *Rev del Lab Clínico*. 2010;3(3):139–47.

19. Döhner H, Weisdorf DJ, Bloomfield CD. Acute myeloid leukemia. *N Engl J Med*. 2015;373(12):1136–52.

20. Daneshbod Y, Medeiros LJ. Pseudo Chediak-Higashi anomaly in acute monoblastic leukemia. *Blood, J Am Soc Hematol*. 2016;128(21):2583.

21. den Berghe H. Morphologic, immunologic and cytogenetic (MIC) working classification of the acute myeloid leukaemias. *Br J Haematol.* 1988;68(4):487–94.

22. Bain BJ. Classification of acute leukaemia: the need to incorporate cytogenetic and molecular genetic information. *J Clin Pathol*. 1998;51(6):420.

23. Arber D, Brunning R, Le Beau M. Acute myeloid leukaemia and related precursor neoplasms. In: Swerdlow SH, Campo E, Harris NL, et al, eds. WHO classification of tumours of haematopoietic and lymphoid tissues. Revised 4t. Lyon, France: IARC Press; 2017. 130–71 p.

24. Rowley JD, JD R. Identification of a translocation with quinacrine fluorescence in a patient with acute leukemia. 1973.

25. Schnittger S, Dicker F, Kern W, Wendland N, Sundermann J, Alpermann T, et al. RUNX1 mutations are frequent in de novo AML with noncomplex karyotype and confer an unfavorable prognosis. *Blood, J Am Soc Hematol*. 2011;117(8):2348–57.

26. Gaidzik VI, Bullinger L, Schlenk RF, Zimmermann AS, Röck J, Paschka P, et al. RUNX1 mutations in acute myeloid leukemia: results from a comprehensive genetic and clinical analysis from the AML study group. *J Clin Oncol*. 2011;29(10):1364–72.

27. Bain BJ, Béné MC. Morphological and immunophenotypic clues to the WHO categories of acute myeloid leukaemia. *Acta Haematol*. 2019;141(4):232–44.

28. Ishikawa Y, Kawashima N, Atsuta Y, Sugiura I, Sawa M, Dobashi N, et al. Prospective evaluation of prognostic impact of KIT mutations on acute myeloid leukemia with RUNX1-RUNX1T1 and CBFB-MYH11. *Blood Adv*. 2020;4(1):66–75.

29. Shaikh AF, Kakirde C, Dhamne C, Bhanshe P, Joshi S, Chaudhary S, et al. Machine learning derived genomics driven prognostication for acute myeloid leukemia with RUNX1-RUNX1T1. *Leuk Lymphoma*. 2020;1–7.

30. Sanz MA, Grimwade D, Tallman MS, Lowenberg B, Fenaux P, Estey EH, et al. Management of acute promyelocytic leukemia: recommendations from an expert panel on behalf of the European LeukemiaNet. *Blood*. 2009;113(9):1875–91.

31. Breen KA, Grimwade D, Hunt BJ. The pathogenesis and management of the coagulopathy of acute promyelocytic leukaemia. *Br J Haematol.* 2012;156(1):24–36.

32. Sainty D, Liso V, Cantù-Rajnoldi A, Head D, Mozziconacci M-J, Arnoulet C, et al. A new morphologic classification system for acute promyelocytic leukemia distinguishes cases with underlying PLZF/RARA gene rearrangements. *Blood, J Am Soc Hematol.* 2000;96(4):1287–96.

33. Lin P, Hao S, Medeiros LJ, Estey EH, Pierce SA, Wang X, et al. Expression of CD2 in acute promyelocytic leukemia correlates with short form of PML-RARA transcripts and poorer prognosis. *Am J Clin Pathol*. 2004;121(3):402–7.

34. Lo-Coco F, Ammatuna E, Montesinos P, Sanz MA. Acute promyelocytic leukemia: recent advances in diagnosis and management. In: Seminars in oncology. 2008. p. 401–9.

35. Schnittger S, Weisser M, Schoch C, Hiddemann W, Haferlach T, Kern W. New score predicting for prognosis in PML-RARA+, AML1-ETO+, or CBFBMYH11+ acute myeloid leukemia based on quantification of fusion transcripts. *Blood*. 2003;102(8):2746–55.

36. Kurata K, Yamamoto K, Okazaki Y, Noguchi Y, Matsui K, Matsumoto H, et al. Detection of a novel CBFB-MYH11 fusion transcript in acute myeloid leukemia M1 with inv (16)(p13q22). *Cancer Genet*. 2020;241:72–6.

37. Arber DA. Acute myeloid leukemia. In: Hematopathology. Elsevier; 2018. p. 429–66.

38. Ma Z, Morris SW, Valentine V, Martin L, Herbrick J-A, Cui X, et al. Fusion of two novel genes, RBM15 and MKL1, in the t (1; 22)(p13; q13) of acute megakaryoblastic leukemia. *Nat Genet*. 2001;28(3):220–1.

39. Wang RC, Chang C-CJ. Chronic Myeloid Leukemia, BCR-ABL1 Positive. In: Precision Molecular Pathology of Myeloid Neoplasms. Springer; 2018. p. 99–113.

40. Chen W, Rassidakis GZ, Li J, Routbort M, Jones D, Kantarjian H, et al. High frequency of NPM1 gene mutations in acute myeloid leukemia with prominent nuclear invaginations ("cuplike" nuclei). *Blood*. 2006;108(5):1783–4.

41. Nomdedeu J, Bussaglia E, Villamor N, Martinez C, Esteve J, Tormo M, et al. Immunophenotype of acute myeloid leukemia with NPM mutations: prognostic impact of the leukemic compartment size. *Leuk Res*. 2011;35(2):163–8.

42. Haferlach C, Mecucci C, Schnittger S, Kohlmann A, Mancini M, Cuneo A, et al. AML with mutated NPM1 carrying a normal or aberrant karyotype show overlapping biologic, pathologic, immunophenotypic, and prognostic features. *Blood, J Am Soc Hematol*. 2009;114(14):3024–32.

43. Dufour A, Schneider F, Metzeler KH, Hoster E, Schneider S, Zellmeier E, et al. Acute myeloid leukemia with biallelic CEBPA gene mutations and normal karyotype represents a distinct genetic entity associated with a favorable clinical outcome. *J Clin Oncol*. 2010;28(4):570–7.

44. Wilhelmson AS, Porse BT. CCAAT enhancer binding protein alpha (CEBPA) biallelic acute myeloid leukaemia: cooperating lesions, molecular mechanisms and clinical relevance. *Br J Haematol*. 2020;

45. Taskesen E, Bullinger L, Corbacioglu A, Sanders MA, Erpelinck CAJ, Wouters BJ, et al. Prognostic impact, concurrent genetic mutations, and gene expression features of AML with CEBPA mutations in a cohort of 1182 cytogenetically normal AML patients: further evidence for CEBPA double mutant AML as a distinctive disease entity. *Blood, J Am Soc Hematol*. 2011;117(8):2469–75.

46. Xu X-Q, Wang J-M, Gao L, Qiu H-Y, Chen L, Jia L, et al. Characteristics of acute myeloid leukemia with myelodysplasia-related changes: A retrospective analysis in a cohort of Chinese patients. *Am J Hematol*. 2014;89(9):874–81.

47. Hwang SM. Classification of acute myeloid leukemia. *Blood Res*. 2020;55(Suppl):S1.

48. Quelen C, Lippert E, Struski S, Demur C, Soler G, Prade N, et al. Identification of a transforming MYB-GATA1 fusion gene in acute basophilic leukemia: a new entity in male infants. *Blood, J Am Soc Hematol*. 2011;117(21):5719–22.

49. Roberts I, Izraeli S. Haematopoietic development and leukaemia in Down syndrome. *Br J Haematol*. 2014;167(5):587–99.

50. Yoshida K, Toki T, Okuno Y, Kanezaki R, Shiraishi Y, Sato-Otsubo A, et al. The landscape of somatic mutations in Down syndrome-related myeloid disorders. *Nat Genet*. 2013;45(11):1293–9.

51. Deotare U, Yee KWL, Le LW, Porwit A, Tierens A, Musani R, et al. Blastic plasmacytoid dendritic cell neoplasm with leukemic presentation: 10-Color flow cytometry diagnosis and HyperCVAD therapy. *Am J Hematol*. 2016;91(3):283–6.

52. Network CGAR. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med*. 2013;368(22):2059–74.

53. Papaemmanuil E, Gerstung M, Bullinger L, Gaidzik VI, Paschka P, Roberts ND, et al. Genomic classification and prognosis in acute myeloid leukemia. *N Engl J Med*. 2016;374(23):2209–21.

54. Döhner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Büchner T, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*. 2017;129(4):424–47.

55. Terwilliger T, Abdul-Hay M. Acute lymphoblastic leukemia: a comprehensive review and 2017 update. *Blood Cancer J*. 2017;7(6):e577--e577.

56. Iacobucci I, Mullighan CG. Genetic basis of acute lymphoblastic leukemia. *J Clin Oncol*. 2017;35(9):975.

57. Mullighan CG, Su X, Zhang J, Radtke I, Phillips LAA, Miller CB, et al. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. *N Engl J Med*. 2009;360(5):470–80.

58. Bernard OA, Berger R. Molecular basis of IIq23 rearrangements in hematopoietic malignant proliferations. *Genes, Chromosom Cancer*. 1995;13(2):75–85.

59. Moorman A V, Richards SM, Martineau M, Cheung KL, Robinson HM, Broadfield ZJ, et al. TEL-AML1 fusion in childhood acute lymphoblastic leukaemia (ALL): Demographics, clinical features, cytogenetic characteristics and prognosis. 2005.

60. Jeha S, Pei D, Raimondi SC, Onciu M, Campana D, Cheng C, et al. Increased risk for CNS relapse in pre-B cell leukemia with the t (1; 19)/TCF3-PBX1. *Leukemia*. 2009;23(8):1406–9.

61. Strefford JC, van Delft FW, Robinson HM, Worley H, Yiannikouris O, Selzer R, et al. Complex genomic alterations and gene expression in acute lymphoblastic leukemia with intrachromosomal amplification of chromosome 21. *Proc Natl Acad Sci*. 2006;103(21):8167–72.

62. Harrison CJ. Targeting signaling pathways in acute lymphoblastic leukemia: new insights. *Hematology*. 2013;2013(1):118–25.

63. Mullighan CG, Goorha S, Radtke I, Miller CB, Coustan-Smith E, Dalton JD, et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature*. 2007;446(7137):758–64.

64. Borowitz MJ, Béné MC, Harris NL, Porwit A, Matutes E, Arber DA. Natural killer (NK) cell lymphoblastic leukaemia/lymphoma. In: Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri S, Stein H, et al., editors. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. 4th edn. Lyon, France: IARC Press; 2017. p. 213.

65. Bain BJ. Diagnosis from the blood smear. *N Engl J Med*. 2005;353(5):498–507.

66. Haferlach T, Schmidts I. The power and potential of integrated diagnostics in acute myeloid leukaemia. *Br J Haematol.* 2020;188(1):36–48.

67. Bain BJ. Leukaemia Diagnosis. 4th ed. Chichester, UK: Wiley-Blackwell; 2010.68–73 p.

68. Porwit A, Vardiman JW. Acute myeloid leukemia with expanded erythropoiesis. *Haematologica*; 2011.

69. Kratz A, Lee S, Zini G, Riedl JA, Hur M, Machin S, et al. Digital morphology analyzers in hematology: ICSH review and recommendations. *Int J Lab Hematol*. 2019;41(4):437–47.

70. West Medica. VisionHema [Internet]. [cited 2020 Nov 4]. Available from: http://www.green-med.net/analizador-de-imaacutegenes---hematoacuteloga.html

71. Medica Corporation. Hematology Imaging System. EasyCell assistant [Internet]. [cited 2020 Nov 4]. Available from: https://www.medicacorp.com/products/hematology-imaging-analyzers/

72. Yu H, Ok CY, Hesse A, Nordell P, Connor D, Sjostedt E, et al. Evaluation of an automated digital imaging system, Nextslide Digital Review Network, for examination of peripheral blood smears. Arch Pathol Lab Med [Internet]. 2012 Jun [cited 2013 Apr 10];136(6):660–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22646275

73. Winkelman JW, Tanasijevic MJ, Zahniser DJ. A novel automated slide-based technology for visualization, counting, and characterization of the formed elements of blood: A proof of concept study. *Arch Pathol Lab Med.* 2017;141(8):1107–12.

74. IIS" FlflC. Hema CAM 2019 [Internet]. [cited 2020 Nov 4]. Available from: www.iis.fraunhofer.de

75. Kim HN, Hur M, Kim H, Kim SW, Moon H-W, Yun Y-M. Performance of automated digital cell imaging analyzer Sysmex DI-60. *Clin Chem Lab Med*. 2017;56(1):94–102.

76. Kratz A, Bengtsson H-I, Casey JE, Keefe JM, Beatrice GH, Grzybek DY, et al. Performance evaluation of the CellaVision DM96 system: WBC differentials by automated digital image analysis supported by an artificial neural network. *Am J Clin Pathol*. 2005;124(5):770–81.

77. Briggs C, Longair I, Slavik M, Thwaite K, Mills R, Thavaraja V, et al. Can automated blood film analysis replace the manual differential? An evaluation of the CellaVision DM96 automated image analysis system. *Int J Lab Hematol*. 2009;31(1):48–60.

78. Cornet E, Perol J-P, Troussard X. Performance evaluation and relevance of the CellaVisionTM DM96 system in routine analysis and in patients with malignant hematological diseases. *Int J Lab Hematol*. 2008;30(6):536–42.

79. Billard M, Lainey E, Armoogum P, Alberti C, Fenneteau O, Da Costa L. Evaluation of the CellaVision™ DM automated microscope in pediatrics. *Int J Lab Hematol*. 2010;32(5):530–8.

80. Eilertsen H, Henriksson CE, Hagve T-A. The use of CellaVision™ DM 96 in the verification of the presence of blasts in samples flagged by the Sysmex XE-5000. *Int J Lab Hematol*. 2017;39(4):423–8.

81. Stouten K, Riedl JA, Levin MD, van Gelder W. Examination of peripheral blood smears: performance evaluation of a digital microscope system using a large-scale leukocyte database. *Int J Lab Hematol*. 2015;37(5):e137--e140.

82. Merino A, Brugués R, García R, Kinder M, Torres F, Escolar G. Estudio comparativo de la morfología de sangre periférica analizada mediante el microscopio y el CellaVision DM96 en enfermedades hematológicas y no hematológicas. *Rev del Lab Clínico*. 2011;4(1):3–14.

83. Agaian S, Madhukar M, Chronopoulos AT. A new acute leukaemia-automated classification system. *Comput Methods Biomech Biomed Eng Imaging Vis*. 2016;1163:1–12.

84. MoradiAmin M, Memari A, Samadzadehaghdam N, Kermani S, Talebi A. Computer aided detection and classification of acute lymphoblastic leukemia cell subtypes based on microscopic image analysis. *Microsc Res Tech*. 2016;79(10):908–16.

85. Mishra S, Majhi B, Sa PK, Sharma L. Gray level co-occurrence matrix and random forest based acute lymphoblastic leukemia detection. *Biomed Signal Process Control.* 2017;33:272–80.

86. Scotti F. Automatic morphological analysis for acute leukemia identification in peripheral blood microscope images. In: CIMSA 2005 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications, 2005. p. 96–101.

87. Mohapatra S, Patra D, Satpathy S. An ensemble classifier system for early diagnosis of acute lymphoblastic leukemia in blood microscopic images. *Neural Comput Appl*. 2014;24(7–8):1887–904.

88. Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine learning for medical imaging. *Radiographics*. 2017;37(2):505–15.

89. Rodellar J, Alférez S, Acevedo A, Molina A, Merino A. Image processing and machine learning in the morphological analysis of blood cells. *Int J Lab Hematol.* 2018;40:46–53.

90. Alférez S, Merino A, Mujica LE, Ruiz M, Bigorra L, Rodellar J. Automatic classification of atypical lymphoid B cells using digital blood image processing. *Int J Lab Hematol*. 2014;36(4):472–80.

91. Merino A, Puigví L, Boldú L, Alférez S, Rodellar J. Optimizing morphology through blood cell image analysis. *Int J Lab Hematol*. 2018;40:54–61.

92. Angulo J. A mathematical morphology approach to cell shape analysis. In: Progress in Industrial Mathematics at ECMI 2006. Springer; 2008. p. 543–7.

93. Gonzalez RC, Woods RE. Digital image processing. 3rd ed. Prentice Hall; 2007.

94. Puigví L, Merino A, Alférez S, Acevedo A, Rodellar J. New quantitative features for the morphological differentiation of abnormal lymphoid cell images from peripheral blood. *J Clin Pathol*. 2017;70(12):1038–48.

95. Haralick R, Shanmugan K, Dinstein I. Textural features for image classification. Vol. 3, IEEE Transactions on Systems, Man and Cybernetics. 1973. p. 610–21. Available from: http://dceanalysis.bigr.nl/Haralick73-Textural features for image classification.pdf

96. Yang X, Tridandapani S, Beitler JJ, Yu DS, Yoshida EJ, Curran WJ, et al. Ultrasound GLCM texture analysis of radiation-induced parotid-gland injury in head-and-neck cancer radiotherapy: An in vivo study of late toxicity. *Med Phys.* 2012;39(9):5732.

97. Kono K, Hayata R, Murakami S, Yamamoto M, Kuroki M, Nanato K, et al. Quantitative distinction of the morphological characteristic of erythrocyte precursor cells with texture analysis using gray level co-occurrence matrix. *J Clin Lab Anal.* 2018;32(Jan):e22175.

98. Saeys Y, Inza I, Larrañaga P. A review of feature selection techniques in bioinformatics. *Bioinformatics*. 2007;23(19):2507–17.

99. Kuhn M, Johnson K. An introduction to feature selection. In: Applied predictive modeling. Springer; 2013. p. 487–519.

100. Shouval R, Fein JA, Savani B, Mohty M, Nagler A. Machine learning and artificial intelligence in haematology. *Br J Haematol*. 2020;

101. Radakovich N, Nagy M, Nazha A. Machine learning in haematological malignancies. *Lancet Haematol*. 2020;7(7):e541--e550.

102. Alsalem MA, Zaidan AA, Zaidan BB, Hashim M, Madhloom HT, Azeez ND, et al. A review of the automated detection and classification of acute leukaemia: Coherent taxonomy, datasets, validation and performance measurements, motivation, open challenges and recommendations. *Comput Methods Programs Biomed*. 2018;158:93–112.

103. Salah HT, Muhsen IN, Salama ME, Owaidah T, Hashmi SK. Machine learning applications in the diagnosis of leukemia: Current trends and future directions. *Int J Lab Hematol*. 2019;41(6):717–25.

104. Boldú L, Merino A, Alférez S, Molina A, Acevedo A, Rodellar J. Automatic recognition of different types of acute leukaemia in peripheral blood by image analysis. *J Clin Pathol*. 2019;72(11):755–61.

105. Alférez S, Merino A, Acevedo A, Puigví L, Rodellar J. Colour Clustering Segmentation Framework for Image Analysis of Malignant Lymphoid Cells in Peripheral Blood. *Med Biol Eng Comput*. 2018;

106. Brown G, Pocock A, Zhao M-J, Luján M. Conditional likelihood maximisation: a unifying framework for information theoretic feature selection. *J Mach Learn Res*. 2012;13(Jan):27–66.

107. Boldú L, Merino A, Acevedo A, Molina A, Rodellar J. A deep learning model (ALNet) for the diagnosis of acute leukaemia lineage using peripheral blood cell images. *Comput Methods Programs Biomed*. 2021;105999.

108. El Achi H, Khoury JD. Artificial Intelligence and Digital Microscopy Applications in Diagnostic Hematopathology. *Cancers (Basel)*. 2020;12(4):797.

109. Shahin AI, Guo Y, Amin KM, Sharawi AA. White blood cells identification system based on convolutional deep neural learning networks. *Comput Methods Programs Biomed.* 2019;168:69–80.