Diagnostics of the AML with immunophenotypical data

A. Plesa\textsuperscript{a}, G. Ciuperca\textsuperscript{b}
V. Louvet\textsuperscript{b}, L. Pujo-Menjouet\textsuperscript{b} \textsuperscript{1}
S. Génieys\textsuperscript{b}, C. Dumontet\textsuperscript{a}
X. Thomas\textsuperscript{a} and V. Volpert\textsuperscript{b}

\textsuperscript{a} Service d’Hématologie clinique, Hospices Civils de Lyon,
69003 Lyon, France
\textsuperscript{b} Université de Lyon, Université Lyon1, CNRS UMR 5208 Institut Camille Jordan
F - 69200 Villeurbanne Cedex, France

Abstract. Patients with acute myeloblastic leukemia (AML) are divided according to the French American British (FAB) classification into eight subgroups (M0 to M7) on the basis of their degree of maturation/differentiation. However, even if immunophenotypical characterization by flow cytometry is routinely used to distinguish between AML and acute lymphoblastic leukemia (ALL), it is not yet well established for the identification within the AML subgroups. Here we show that certain subgroups of AML can be identified with a good accuracy (80\% of the patients we analyzed) by means of the expression of specific markers.

Key words: acute myeloblastic leukemia, diagnostic, FAB classification, cell markers, mathematical modelling.

AMS subject classification: 62-07, 62P10, 92B15, 92B05

1. Introduction

We analyze here the immunophenotypical data obtained with flow cytometry in a group of 213 patients diagnosed with acute myeloblastic leukemia (AML). The diagnosis of this disease is currently established on the basis of cytomorphological analysis accordingly to the French-American-British (FAB) classification, and supported by enzymatic measurements. Immunophenotyping

\textsuperscript{1}Corresponding author. E-mail: pujo@math.univ-lyon1.fr
is routinely used to distinguish acute lymphoblastic leukemia (ALL) from acute myeloblastic leukemia (AML)
([9], [10]). This is done by using markers common to AML (CD13, CD33, MPO, CD117, CD65) or to ALL. Some correlations between immunophenotyping and diagnosis and prognosis are discussed in [4] and [6]. Flow cytometry is used in [5] and [2] to characterize leukemic blast cells. Our objective is to determine a possible correlation between the FAB classification and the immunophenotypic profiles. Indeed, from the best of our knowledge, we realized that despite numerous reports devoted to this question, there has been no consensus yet regarding a standardized immunophenotypic profile of AML subgroups. Consequently, it appears natural to us to suggest here as a first step a simple approach which allows a rather specific identification of the FAB subgroups by means of expression of certain markers which corresponds to the hematopoietic lineages. The markers used in the immunophenotypical analysis are the following: CD2, CD4, CD5, CD7, CD9, CD10, CD19 (lymphoid lineage), CD13, CD15, CD33, CD34, CD36, CD41, CD65, CD11c (myeloid lineage), cytoplasmicCD3, CD11b, CD13c, CD14, cytoplasmicCD22, MPO, TdTc, CD4, CD5, CD38, cytoCD79a, CD117, HLA-DR (not lineage specific). Expression of each marker is characterized by the percentage of cells expressing this marker with the fluorescence level exceeding a background threshold, determined with an isotype control. Therefore each patient is characterized by 30 numbers from 0 to 100 which correspond to the expression of these 30 markers. Hence we need to analyze the set of 213 points corresponding to the number of patients in the 30-dimensional space and to find some subdomains of this space that could correspond to the subgroups M0 to M7 of the FAB classification.

Statistical studies using a cluster analysis as well as a specific software to analyze the data in the 30-dimensional space have been carried out (data not shown here). It came out that the data analysis was quite efficient for some aspects (to predict survival or relapse of a patient for instance), but not accurate enough to give a clear idea of the importance of the role of each marker for each type of leukemia. We develop here another approach which suggests a simple algorithm of identification of some groups of patients in accordance with the FAB classification.

2. CD36-34 immunophenotypical classification

As indicated above, each patient has been characterized by the expression of 30 markers. The conventional approach to analyze these data consists in the introduction of positive or negative expressions of the markers, which are greater or less than some (arbitrarily) given value. This threshold could be for example 20% in accordance with the EGIL (European Group for Immunophenotyping of Leukemias) criteria (see [3]). After this process, patients are divided into subgroups accordingly to positive or negative expressions of some markers. In this work we do not impose a priori critical values of expression of the markers. And so we consider all their actual values between 0 and 100%. Therefore we consider a set of points in the 30-dimensional space where the coordinates of these points are expressions of the corresponding markers. As specified above, this set of points can be characterized with cluster analysis. This means that it can be divided into a relatively small number of subsets separated from each other by maximal distance. Some statistical software allows the determination of the optimal number of clusters. In the ideal case we could expect that
the clusters are clearly identified and that each of them corresponds to some of the FAB subgroups. However, this is not what happens in a concrete way. The clusters are not clearly identified and they do not correspond to the FAB classification. And in fact, the very idea of cluster analysis can be rather restrictive. We can easily imagine the situation where two sets of points are separated not by space distance but by some other criteria. More particularly, the expression of a marker that can be greater or less than some given value will divide a set of points into two subsets. This can be completely independent of the existence and the structure of clusters. At the same time, in some cases clusters can be clearly identified.

At this point, it is also important to specify that we apply our analysis not to all markers at the same time or to some groups of markers but consecutively to one marker after another. The advantage of this approach is that we can apply different markers and different criteria for different subgroups of patients which is in some sense an adapting method. Thanks to the methods described above, we are able to characterize the FAB subclasses by immunophenotypical analysis. We base our investigations on the antigenic analysis of hematopoietic lineage in normal hematopoiesis. This analysis has been developed in the past decades and is well known in normal blood formation [7], [8] (see figure 1 based on the work done by Loken et al. in [8]).

---

Figure 1: Schematic diagrams of lineage specific development. Fig.1a: Schematic diagram of monocytoid development. Fig 1b: Neutrophil maturation. Fig 1c: Erythroid maturation. Each stages of maturation can be defined based on the expression of multiple myeloid markers on these cells in the bone marrow. These figures are based on the work of Loken et al.
Our approach is based on the difference in the expressions of antigens for different cell lineages (see figure 1).

Let us start with CD36. In normal hematopoiesis this marker is not expressed by the myeloid lineage (also called L0 lineage) (see figure 2). This lineage consists of the undifferentiated cells, the myeloid progenitor cell (MPC) and the myelo-monocytic progenitor cell (MMPC). It is not expressed in the lineage-committed cells (also called L2 lineage) that consists of myeloblasts, promyelocyte, myelocyte and granulocyte cells (see figure 2). However, it is expressed in the other three lineages, L1 (monoblast, promonocyte, monocyte, macrophage), L3 (proerythroblast, erythroblast, reticulocyte, erythrocyte), L4 (immature megakaryoblast, megakaryoblast, megakaryocyte, platelets). Thus, in the "normal" bone marrow, expression of CD36 distinguishes L0 (MPC, MMPC) and L2 from L1, L3, L4. We first consider the patients M0-M5 of the FAB classification.
8 cases M2-M3 and 3 cases M4-M5. It can be identified with the M2-M3 diagnostics of the FAB classification. We will see below that there are in fact only M2 cases. In the second domain there are all three groups of patients without a clear preference to one of them. In the next section we will use additional markers to distinguish them.

The second domain (II) contains basically M0-M1 patients (79%) while the third domain (III) basically M4-M5 patients (73%). It can be also useful to split the third domain in two subdomains: with CD36 greater than 35% and less than 35%. The first subdomain contains 37 M4-M5 cases from 44 (84%) and can be identified with this diagnostics. There are more M0-M1 and M2-M3 cases in the second subdomain. We will discuss below how to distinguish them with additional markers.

Thus, M0-M1 patients are basically in the first and second domain. The second domain with high expression of CD34 and small expression of CD36 corresponds to the L0 lineage of the normal hematopoiesis. In the first subdomain, where the expression of CD34 is also small corresponds to the beginning of cell maturation in the direction of L1 and L2 lineages where its expression decreases. M2-M3 cases are basically located in the first domain with low expressions of CD34 and CD36 as it is the case at the L2 lineage for the normal hematopoiesis. There are also some M2-M3 cases with a greater expression of CD34 (lower subdomain of the third domain) which corresponds to a partial cell differentiation in the direction of this lineage. Finally, M4-M5 cases are basically located in the third domain with a partial overlapping with the first and the second domains.

The splitting of all M0-M5 patients into three subgroups corresponds to the hematopoietic lin-
eages and simplifies the representation. M6 and M7 subgroups will be discussed below. Here we briefly discuss how to distinguish M0 and M1, M2 and M3, M4 and M5 inside each subgroup. Figure 4 shows all M0 and M1 patients on the CD34 - HLA-DR plane. M0 cases are basically located at the right upper corner where the expression of both markers is high. There is an intermediate region where there are less M0 cases, and finally the lower left corner where there are only M1 cases. This corresponds to the decrease in the expression of both markers in the process of cell maturation. It is interesting to note that the center region contains only few points. This means that the maturation occurs consecutively, first CD34 decreases, then HLA-DR or vice versa but not both at the same time.

Consider next M2 and M3 cases (figure 5). Most of M3 cases are in the lower left corner while M2 cases are distributed all over the plane. There are two M3 cases for larger values of CD36.
HLA-DR can be additionally used to separate them from M2 cases: all M3 cases correspond to small expression of this marker (figure 6).

The M4 and M5 cases are shown in figure 7. The number of M4 cases decreases with the increase in the expression of CD36. This corresponds to cell maturation along the L1 lineage. On the other hand, even for low expression of CD36, M4 and M5 cases are mixed. We have not found any markers yet that would allow us to separate them.

Figure 5: Subgroups M2 (red) and M3 (green) on the CD34 - CD36 plane.

Figure 6: Subgroups M2 (red) and M3 (green) in the CD34 - CD36 - HLA-DR space.
3. **CD34 < 20, CD36 < 20**

In this section we consider patients with the expression of CD34 and of CD36 less than 20%. There are 63 M0-M5 cases: 25 cases M0-M1, 26 cases M1-M2 and 12 cases M5-M6. Figure 8 shows all three subgroups on the CD36 - HLA-DR plane. For HLA-DR < 6 there are all M3 cases (9/9), some M2 cases (3/9), about a half of M0-M1 cases (13/25) and only one M4-5 case (1/10). Consider first the small expression of HLA-DR. Since M2 and M3 cases can be effectively separated, it is sufficient to separate M3 and M0-M1 cases. There is a number of markers that can be used here, in particular CD9 and CD65 (figure 9). Other markers that separate M0-1(HLA-DR
< 6) and M3 are:

\[
\begin{align*}
\text{CD13} &< 76 : M0 - 1 = 11/13, \; M3 = 2/9; \\
\text{cytoMPO} &< 47 : M0 - 1 = 9/13, \; M3 = 3/9; \\
\text{cytoCD13} &< 80 : M0 - 1 = 12/13, \; M3 = 2/9.
\end{align*}
\]

Figure 9: Subgroups M0-M1 (red) and M3 (green) on the CD65 - CD9 plane.

There are only M2 cases from the M2-M3 subgroup for HLA-DR > 6%. Figure 10 shows M2 and M4-M5 cases on the CD65 - CD13 plane. Figures 11 and 12 show M0-M1 (HLA-DR > 6%) cases versus M2 and M4-M5, respectively, on the CD19 - CD10 and CD13 - CD9 planes.

Up to now we have discussed only the M0-M5 cases. For CD34 < 20% and CD36 < 20% there are also 2 M6 cases and 13 M7 (unidentified) cases. Since two cases are not sufficient to study M6 patients we restrict ourselves to the analysis of M7 cases with respect to other patients:

1. For M7 - M3 : CD36 < 5 (M3=7/9, M7=2/13), HLA-DR < 10 (M3=2/9, M7=3/13), CD117 < 80 (M3=7/9, M7=2/13)
2. For M7 - M2: cytoCD22 < 4 (M2=2/9, M7=11/13)
3. For M7 - M4 : CD15 < 20 (M4-5=5/10, M7=12/13)
4. For M7 - M0-1: no marker found

4. **CD34 > 20, CD36 < 35**

We consider in this section the patients with CD34 greater than 20% and CD36 less than 35%. This part of the CD34-CD36 plane contains the second domain (CD34 > 70%, CD36 < 35%) and the lower part of the third domain. The second domain contains basically M0-M1 cases and only few M2 or M4-M5 cases. On the other hand, there is relatively small number of M0-M1 cases in the lower part of the third domain. Therefore, we restrict ourselves to a short discussion about the separation of M2 and M4-M5 cases for CD34 > 20% and CD36 < 35%. The results for two pairs of markers, CD65 - CD19 and cytoTdT - cytoCD3 are shown in figures 13 and 14. Figures 15, 16 and 17 show M7 cases (unidentified) versus M2, M4 and M5 cases, respectively. Each time we
Figure 10: Subgroups M2 (red) and M4-M5 (green) on the CD65 - CD13 plane.

Figure 11: Subgroups M0-M1 (red) and M2 (green) on the CD19 - CD10 plane.

Figure 12: Subgroups M0-M1 (red) and M4-M5 (green) on the CD13 - CD9 plane.
choose a specific pair of markers best fitted to split different cases. We note finally that there are only 3 M6 cases for CD34 > 20%, CD36 < 35%. This is not sufficient to study them. No markers were found to distinguish M7 cases from M4 and M5 for CD36 > 35%.

Figure 13: Subgroups M2 (red), M4 (green), M5 (blue) on the CD65 - CD19 plane (CD34 > 20%, CD36 < 35%).

5. Immunophenotypical characteristics of death cases

In the previous section we used the immunophenotypical characteristics of the patients to establish their diagnostics in relation with the FAB classification. Here we use immunophenotypical data to evaluate mortality of the patients. We analyze it in each subgroup M0-M7. In each case we choose the markers which allow us to distinguish death cases (if such markers exist). We will see that they can be different for different groups of patients. Therefore they should be studied independently. We note first of all that there are no death cases among M3 patients and that we have not found death specific markers for M2 and M6 cases. Death and survival cases for M7
Figure 15: Subgroups M2 (red) and M7 (green) on the CD15 - CD65 plane (CD34 > 20%, CD36 < 35%).

Figure 16: Subgroups M4 (red) and M7 (green) on the CD34 - CD9 plane (CD34 > 20%, CD36 < 35%).

Figure 17: Subgroups M5 (red) and M7 (green) on the cytoCD3 - cytoCD79a plane (CD34 > 20%, CD36 < 35%).
A. Plesa et al. Diagnostics of the AML

patients are well separated by cytoNII: for 6 death cases its expression is more than 80% while for 5 survival cases less than 30%. Figure 18 shows M4 and M5 patients on the CD36-CD34 plane. Deceased patients are shown with red color. We can observe that the percentage of death cases decreases from the lower right corner to the upper left corner (data not shown here). This is in agreement with the direction of maturation of blood cells in the direction of the L1 lineage: the expression of CD34 specific for myeloid progenitor cells decreases while expression of CD36 specific for the L1 lineage increases. M4-M5 patients are distributed rather uniformly on the CD36-CD34 plane without forming clearly identified clusters neither for any cases the whole group of patients nor for death cases. No more specific markers are found for M4-M5 cases. M0-M1

Figure 18: Death (red) and survival (green) cases for M4-M5 patients on the CD34 - CD36 plane.

patients are presented in Figure 19. The situation here is quite different in comparison with the previous group of patients. First, as we have already discussed before, there are three subgroups of M0-M1 patients. The first subgroup is located in the lower right corner, that is, for small values of the expression of CD34 and for large values of the expression of CD36. This is specific to the myeloid progenitor cells (L0 lineage). The second subgroup is located in the lower left corner where the expression of both markers is small. This is more specific to the L2 lineage where the expression of CD34 decreases while CD36 is not expressed. Therefore we can conclude that there is a partial differentiation of the myeloid progenitor cells in the direction of this lineage. In the first and second domains (left and right lower corners) death and survival cases are mixed, for the left lower corner there are only few death cases. In the third domain death cases are located for larger values of CD36 while survival cases for smaller values. This location of death cases corresponds to a larger discrepancy between morphological characteristics of blasts (M0-M1) and their immunophenotypical characteristics more specific for the L1 lineage. Death and survival cases for M0 patients can also be characterized with cytoCD22 and cytoMPO (Figure 20).

6. Discussion

We study in this work the inter-relation between the FAB classification of acute myeloid leukemia and immunophenotypical characteristics of blast cells obtained with flow cytometry. The main
Figure 19: Death (red) and survival (green) cases for M0-M1 patients on the CD34 - CD36 plane.

Figure 20: Death (red) and survival (green) cases for M0 patients on the cytoCD22 - cytoMPO plane.
conclusion is that the FAB groups M0-M7 can be rather well identified by using properly chosen markers. This choice is based on the correlation in the expression of certain markers with cell maturation and differentiation (Figure 1). The key point in our reasoning is related to differentiation of myelo-monocytic progenitor cell (MMPC) into either monoblasts or myeloblasts. In both cases expression of CD34 specific for MMPC decreases. Differentiation into monoblasts is accompanied by the increase of CD36 expression while it is not the case for differentiation into myeloblasts. Thus, in normal hematopoiesis these two markers allow the identification of three lineages: L0 (MPC-MMPC), L1 (monoblasts, promonocytes, monocytes, macrophages), L2 (myeloblasts, promyelocyte, myelocyte, granulocyte). In the ideal case, they can be characterized by the pairs (1,0), (0,1), (0,0) where 1 shows the expression and 0 the absence of expression of CD34 (first number) or of CD36 (second number). Some additional markers should be used for a more detailed characterization. In the leukemic case there is an imbalance between cell differentiation, self-renewal and apoptosis. If, for cells of some particular type, differentiation and apoptosis are less than what they are supposed to be and moreover they begin to self-renew (or to self-renew more intensively if they do it in non pathological conditions) then this can result in a substantial increase of the number of immature cells (blasts). The FAB classification of AML is based on the cytomorphological characterization of malignant cells. It is M0 for MPC, M1 for MMPC, M2 and M3 for myeloblasts and promonocytes, M4 and M5 for monoblasts and promonocytes. On the other hand, as it is discussed above, these cells can be characterized by the expression of certain markers. Hence we can expect the correlation between the FAB classification and immunophenotyping. This hypothesis is confirmed by the analysis of medical data for the group of 213 patients. This leads to the following characterization of the FAB groups:

M0 - M1 : two clearly identified subgroups, one for large CD34 expression, another one for small. M0 cases are more specific for large expression, M1 for small expression. There are some intermediate cases,
M2 - M3 : basically located in the region with small CD34 and small CD36 expressions. There are some M2 cases outside,
M4 - M5 : distributed rather uniformly on the CD34 - CD36 plane with a weak overlapping with other groups basically located at the corners,
M6 : not enough cases for a reliable analysis,
M7 : contains unidentified FAB cases. They can be characterized with the immunophenotypical data.

We remind the reader that each patient is characterized by expressions of 30 markers. We have used in our analysis 3 main markers (CD34, CD36, HLA-DR) and about ten additional markers. All markers which we have found to specify the FAB groups are presented in this work. This means that the expressions of the majority of markers are not related to the FAB classification or at least, to the best of our knowledge, it is not clear for us how they can be related. This is also due to the lack of clinical data. Another remark, which is essential for this analysis, is that different groups of patients should be characterized by different sets of markers. We begin with CD34 and CD36 for all patients which allows us to identify three main groups. After that each group is considered
independently and the markers which can characterize them are chosen. It is known that leukemia often begins with a single malignant cell. It begins to self-renew creating a rapidly growing population of immature cells. If we assume that all malignant cells for a given patient are similar to each other, then expressions of all markers should be either 0% or 100% depending on whether these cells express given marker. However in most cases expressions of markers have some intermediate values. Moreover, even the expression of the markers specific for certain cells may not correspond to what we can expect according to the FAB classification. For example, M1 patients can have low expression of CD34 while M2 patients can have high expression of this marker. More often it is opposite and this opposite case correlates with the FAB classification. Nevertheless, abnormal cases are not rare and it would be interesting to explain how they appear. In the following we try to give a contribution to this explanation using a multi-agent modelling (see [1] for the details).

Let us consider the following scheme of the consecutive differentiations

\[ C_0 \rightarrow 2C_1, \quad C_1 \rightarrow 2C_2, \quad C_2 \rightarrow 2C_3. \]

The concentration of \( C_0 \) is assumed to be constant. Let us write the ordinary differential system of equations describing the evolutions of the concentrations \( C_1 \) and \( C_2 \) assuming that the cells differentiate with the rate \( k_d \) and go to apoptosis with the rate \( k_a \):

\[
\frac{dC_1}{dt} = 2k_d C_0 - (k_d + k_a)C_1, \tag{6.1}
\]

\[
\frac{dC_2}{dt} = 2k_d C_1 - (k_d + k_a)C_2. \tag{6.2}
\]

To describe the leukemic case we add self-renewal of \( C_1 \) cells:

\[ C_1 \rightarrow 2C_1 \]

and assume that the rates of differentiation and apoptosis are modified:

\[
\frac{dC_1}{dt} = 2k_d C_0 - (k_d^* + k_a^* - k_r)C_1, \tag{6.3}
\]

\[
\frac{dC_2}{dt} = 2k_d^* C_1 - (k_d^* + k_a^*)C_2. \tag{6.4}
\]

We assume here that the differentiation and apoptosis rates for \( C_2 \) are also modified because they have the same genetic changes as for \( C_1 \). This system has a stationary solution

\[
C_1^0 = \frac{2k_d C_0}{k_d^* + k_a^* - k_r}, \quad C_2^0 = \frac{2k_d^* C_1}{k_d^* + k_a^*}.
\]

Solution of system (6.3), (6.4) has the form

\[
C_1(t) = \frac{b_1}{k_r - k_d^* - k_a^*} e^{(k_r - k_d^* - k_a^*)t} - \frac{2k_d C_0}{k_r - k_d^* - k_a^*},
\]

119
The constants $b_1$ and $b_2$ are determined by the initial condition. If $k_r k_d^* + k_a^*$, then the solution converges to the stationary solution $(C_0^0, C_0^1)$. If $k_r k_d^* + k_a^*$, then the solution exponentially grows. Thus, if the rate of self-renewal is sufficiently small, then the solution remains bounded. This case can be interpreted as chronic leukemia, while the second case where the concentration of blasts exponentially grows corresponds to acute leukemia.

In the case $k_r k_d^* + k_a^*$ we have

$$\frac{C_1(t)}{C_2(t)} \to k_d^* k_r \quad \text{as} \quad t \to \infty.$$ 

Therefore the ratio of the concentrations depends on the rates of self-renewal and differentiation.
both of them correspond to M4 in the FAB classification if the morphology of $C_3$ cells remains
the same as of $C_1$ cells (due to the genetic change evolution in cell immunophenotype is not neces-
ecessarily accompanied by the evolution of its morphology). Another possible situation is shown in
Figure 21c. The hematopoietic scheme is

$$C_0 \rightarrow C_1, \quad C_1 \rightarrow C_2 + C_3, \quad C_2 \rightarrow 2C_1, \quad C_3 \rightarrow C_4 + C_5.$$ 

In this case cells $C_1$ either self-renew or differentiate into two different cells, $C_4$ (blue) or $C_5$ (dark
blue). At the first stage, $C_1$ gives two intermediate cells $C_2$ and $C_3$. The first one leads to self-
renewal, the second one to differentiation. This case describes M1 in the FAB classification where
MMPC can self-renew and differentiate into monoblasts or myeloblasts.

As a conclusion we believe that the development of leukemic hematopoiesis strongly depends
on the ratio of self-renewal and differentiation of malignant cells. This assumption remains to
be confirmed in our future works by the use of experimental and clinical data as well as more
developed models.

References


leukemia by flow cytometric analysis. Use of CD45 and right-angle light scatter to gate on

A. Falkenrodt, G. Lecalvez, M. Maynadie, M.C. Bene, *Immunological classification of acute
myeloblastic leukemias : relevance to patient outcome*, Leukemia, 17 (2003), No.3, 515-527.

No.1, 43-8.

seau, P. Bernard, *Flow cytometry CD45 gating for immunophenotyping of acute myeloid
leukemia*, Leukemia., 11 (1997), No.11, 1878-86.

Casadevall, J.P. Marie, *The immunophenotype of 177 adults with acute myeloid leukemia: prop-

ization of erythroid. lymphoid and monomyeloid lineages in normal human bone marrow*, In
Flow Cytometry in Hematology. Laerum OD, Bjerksnes R. eds. Academic Press, New York,
