Diagnosis of Chronic Lymphoid Leukemia on unstained blood smears using Raman microspectroscopy and supervised classification

HAPPILLON Teddy[1], UNTEREINER Valérie[1], GOBINET Cyril[1], BELJEBBAR Abdelilah[1], DALIPHARD Sylvie[2], CORNILLET-LEFEBVRE Pascale[2], KLOSSA Jacques[3], MANFAIT Michel[1]

[1]MéDIAN Unit, University of Reims Champagne-Ardenne, FRE/CNRS 3481 MEDyC, UFR Pharmacie, 51 rue Cognacq-Jay, 51096, Reims Cedex, France
[2]Laboratoire d'hématologie, Centre Hospitalier Universitaire de Reims, 51100 Reims, France
[3]TRIBVN, 39 rue Louveau, 92320 Châtillon, France

Chronic Lymphoid Leukemia

Chronic Lymphoid Leukemia (CLL) is a blood cancerous disease characterized by the proliferation of lymphocytes (lymphocytosis).

This is the most common leukemia, preferentially affecting people aged over 50 years old. It is incurable and in most cases shows no clinical signs.

Thus, it is often discovered by chance during a blood test.

If necessary, morphological and immunological studies are led by analyzing blood smears colored with May-Grünwald Giemsa, by making a complete blood count and by computing a Matutes score.

These studies are necessary because it is practically impossible to distinguish a healthy cell from a cancerous one only using a conventional microscope even if the cells are stained.

Raman microspectroscopy

In order to establish the diagnosis of the disease reliably and automatically, a new approach based on the Raman microspectroscopy is introduced in this study.

Based on the interaction between matter and light, this technology acquires a real molecular and biochemical signature of the analyzed sample in a label-free manner (no staining or any marking of the sample).

Recently, Raman microspectroscopy has shown a promising potential to highlight the pathophysiological states in both animal and human since differences between several pathophysiological states occur at the molecular level.

From unstained blood smears, obtained by spreading a drop of blood of CLL patients and healthy persons on standard glass slides, Raman spectra are acquired on the nucleus of white blood cells.

Variable selection

In order to improve the classifications and to reduce computational time, the idea to select Regions Of Interest (ROIs) on spectra was retained (i.e. the most relevant spectral bands able to classify correctly the spectra).

For this purpose, the MATLAB randfeatures function was used, which allowed to attribute a score to each wavenumber. By keeping only those having the highest scores, ROIs could be created. Spectra are then reduced to each ROI and classifications could then be realized.

Classification

Using the supervised classification algorithm "Support Vector Machine" (SVM), prediction models are established to distinguish spectra obtained from healthy or pathological cells.

These models are computed considering a population of spectra called "training set" and are then used to classify another population of spectra, the "validation set".

The best prediction model will then be retained and a third set, the "test set", composed of unlabelled spectra will be used to determine a sensitivity and a specificity.

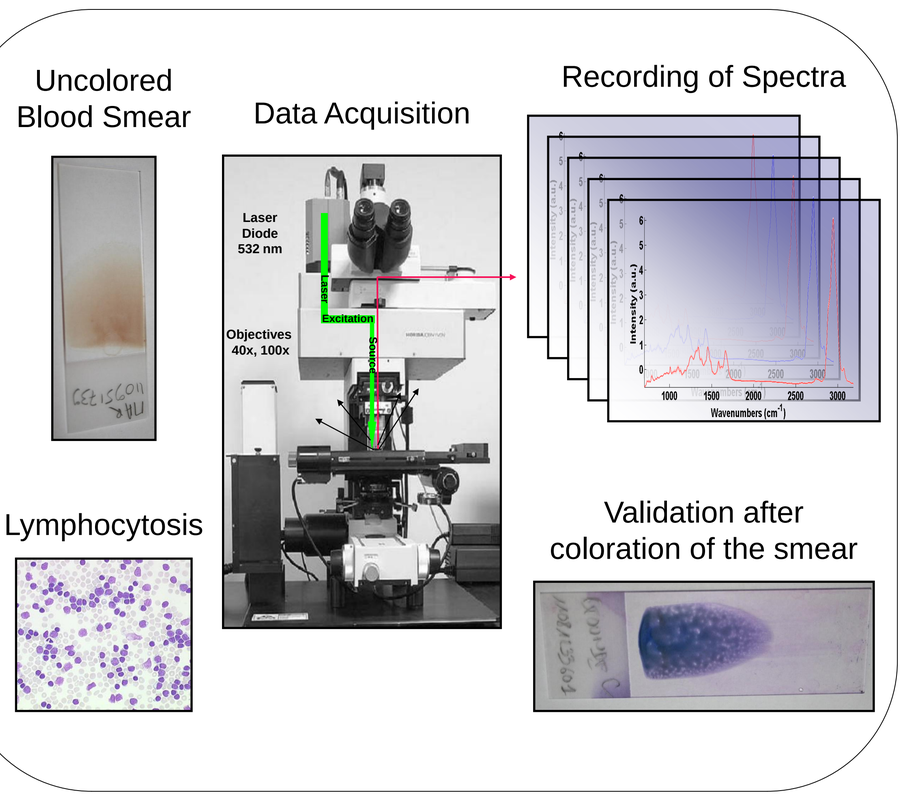
Results

The best prediction model among all gave a sensitivity of 83.3% and a specificity of 100% on the "training set", which implies that the vast majority of the spectra used to create it is well recognized.

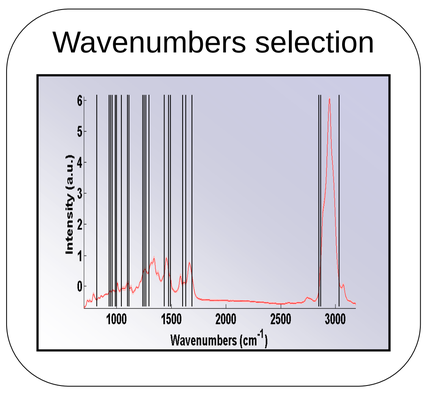
Then a sensitivity of 90.9% and a specificity of 84.6% on the "validation set" are obtained.

And finally, after checking the real membership of every spectrum that belong to the "test set", a sensitivity of 80% and a specificity of 100% are obtained on the patients/persons.

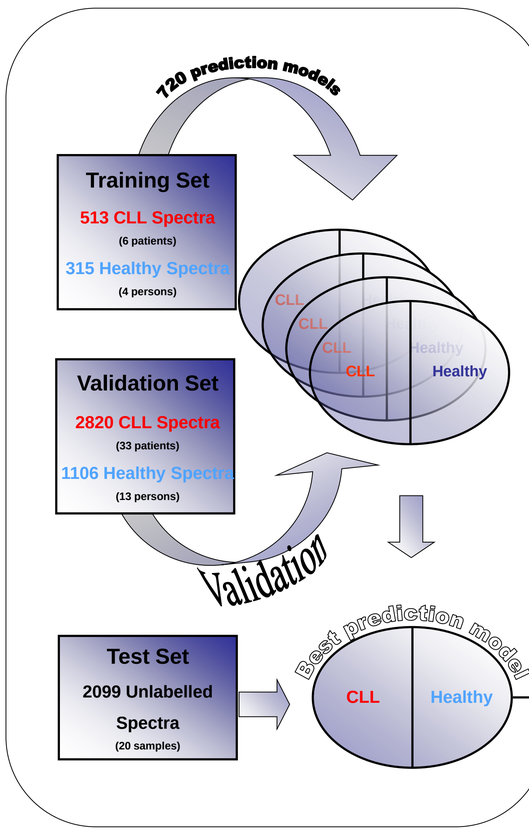
Spectra acquisition



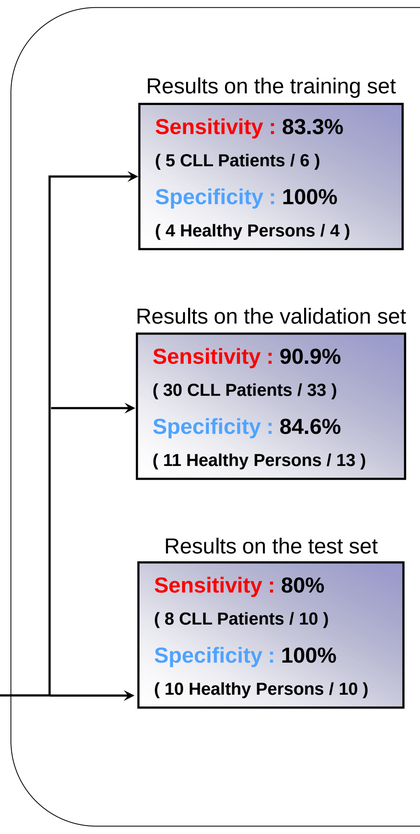
Region Of Interest



Classification Models



Predictions



Conclusion

The results obtained on the discrimination between healthy and pathological cells show the real potential of Raman microspectroscopy coupled with supervised classification for the diagnosis of CLL. Future studies will concentrate on the prognosis, and if possible, the prediction of CLL.