



**HAL**  
open science

## Leveraging Random Survival Forest (RSF) and PET images for prognosis of Multiple Myeloma at diagnosis

Ludivine Morvan, Thomas Carlier, Clément Bailly, Bastien Jamet, Caroline M Bodet-Milin, Philippe Moreau, Cyrille Touzeau, Françoise M Kraeber-Bodere, Diana Mateus

### ► To cite this version:

Ludivine Morvan, Thomas Carlier, Clément Bailly, Bastien Jamet, Caroline M Bodet-Milin, et al.. Leveraging Random Survival Forest (RSF) and PET images for prognosis of Multiple Myeloma at diagnosis. International Conference on Information Processing in Computer-Assisted Interventions (IPCAI), CARS, Jun 2019, Rennes, France. hal-02174921

**HAL Id: hal-02174921**

**<https://hal.science/hal-02174921>**

Submitted on 5 Jul 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# LEVERAGING RANDOM SURVIVAL FOREST (RSF) AND PET IMAGES FOR PROGNOSIS OF MULTIPLE MYELOMA AT DIAGNOSIS

Ludivine Morvan<sup>1,2</sup>, Thomas Carlier<sup>2,3</sup>, Clément Bailly<sup>2,3</sup>, Bastien Jamet<sup>3</sup>, Caroline Bodet-Milin<sup>2,3</sup>, Philippe Moreau<sup>2,4</sup>, Cyrille Touzeau<sup>2,4</sup>, Françoise Kraeber-Bodéré<sup>2,3</sup>, Diana Mateus<sup>1</sup>

1. Centrale Nantes, LS2N, CNRS UMR 6004, Nantes, France,

2. CRCINA, INSERM, CNRS, University of Angers, University of Nantes, Nantes, France

3. University Hospital of Nantes, Nuclear Medicine Department, Nantes, France

4. University Hospital of Nantes, Haematology Department, Nantes, France

## Context

**Multiple myeloma** (MM) is a bone marrow cancer that accounts for 10% of all haematological malignancies. It was reported that **full-body FDG PET imaging** provides prognostic information for both baseline and therapeutic follow-up of MM patients (MM).

## Aims

- ✓ Predict Progression-Free Survival (PFS).
- ✓ Provide predictive features (Clinics and Radiomics).

## Contribution

There is yet much to discover in the survival analysis of MM. However, the **Random Survival Forest (RSF)**[2] has demonstrated robustness but is not studied in the PET imaging and MM context. We developed a two-stage computer-assisted method based on PET imaging features towards **assisting current diagnosis and treatment decisions** for MM patients, with RSF and **"Variable importance" (VIMP)** [2].

## Definitions

**Right censoring**: When no event (death/relapse) has taken place at the end of the evaluation period.

**C-index**: The concordance probability is the frequency of concordant pairs among all pairs of subjects.

$Error\ prediction = 1 - C-index$

**Survival curve**: Survival rates of a specific population, over a period of time.

## A. Material and method

### 1) The data

- ✓ Prospective multi-centric french IMAJEM study [1], 66 patients
- ✓ 132 Clinical and imaging features (textural and conventional).

### 2) The textural features : Radiomics

Example of the GLCM (Gray-Level Co-Occurrence) Matrix [3]:

Translate the joint probability  $P(i,j|\sigma,\theta)$  of the ROI. The position (i,j) represents the number of times that the **combination of the pixels with i and j levels** appears in the I matrix, separated by  $\sigma$  pixels distance and a  $\theta$  angle.

For a distance of 1 and an angle of  $0^\circ$  (plan horizontal):

1	2	5	2	3
3	2	1	3	1
1	3	5	5	2
1	1	1	1	2
1	2	4	5	5

Fig. 2: Intensity matrix I

6	4	3	0	0
4	0	2	1	3
3	2	0	1	2
0	1	1	0	0
0	3	2	0	2

Fig. 3: GLCM Matrix p

$$Energy = \sum_{i=1}^{N_p} \sum_{j=1}^{N_p} (p(i,j))^2 \quad (1)$$

with  $N_p * N_p$  the size of the GLCM matrix



Fig. 1: FDG-PET image of a multiple myeloma patient

### 3) The method

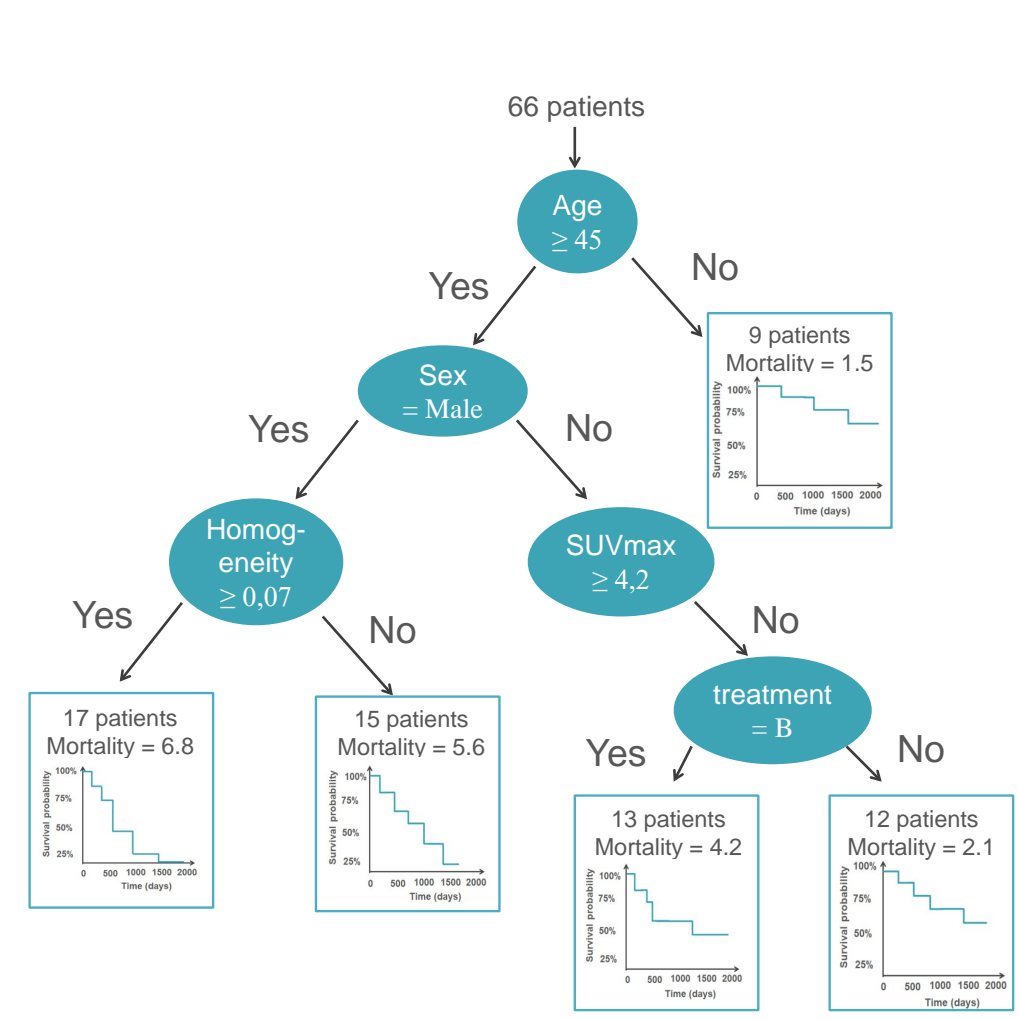


Fig. 4: The RSF differs from the random forest in the target value (the ensemble mortality) but also in the way to separate branches (log-rank test on the mortality)

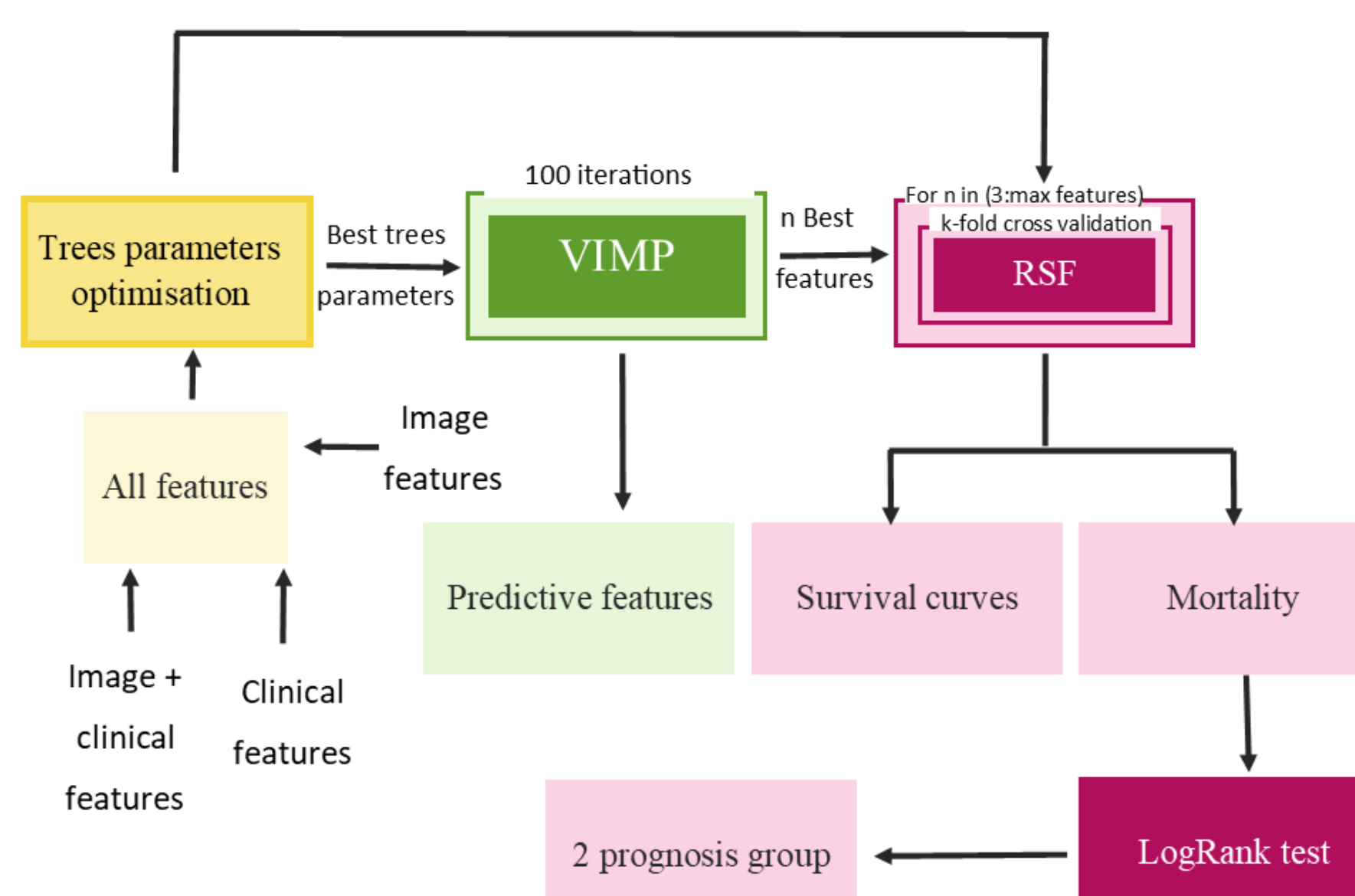


Fig. 5: pipeline of the method

## B. Results

### 1) Evaluation of the method

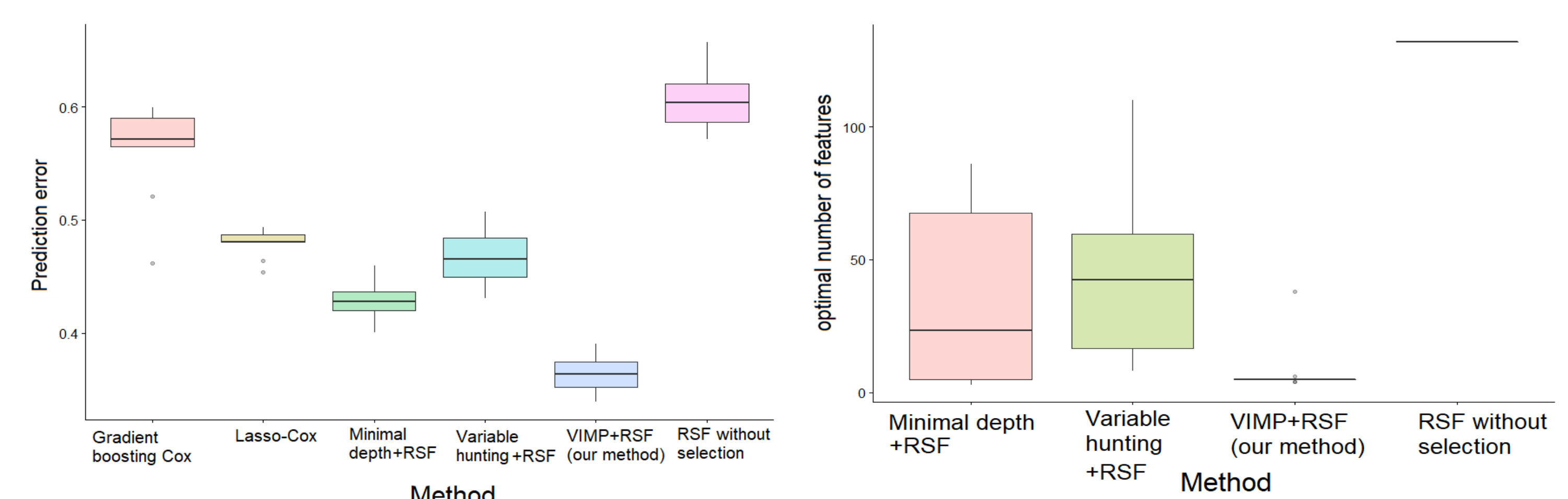


Fig. 6: Prediction error for each method.

Fig. 7: Optimal number of features kept per method.

Method	Average p-value
Our method	0.05
Gradient-Boosting Cox	0.27
Lasso-Cox	0.4
Without selection	0.40
Minimal depth	0.24
Variable-Hunting	0.11

Tab. 1: Average p-value according to the method.

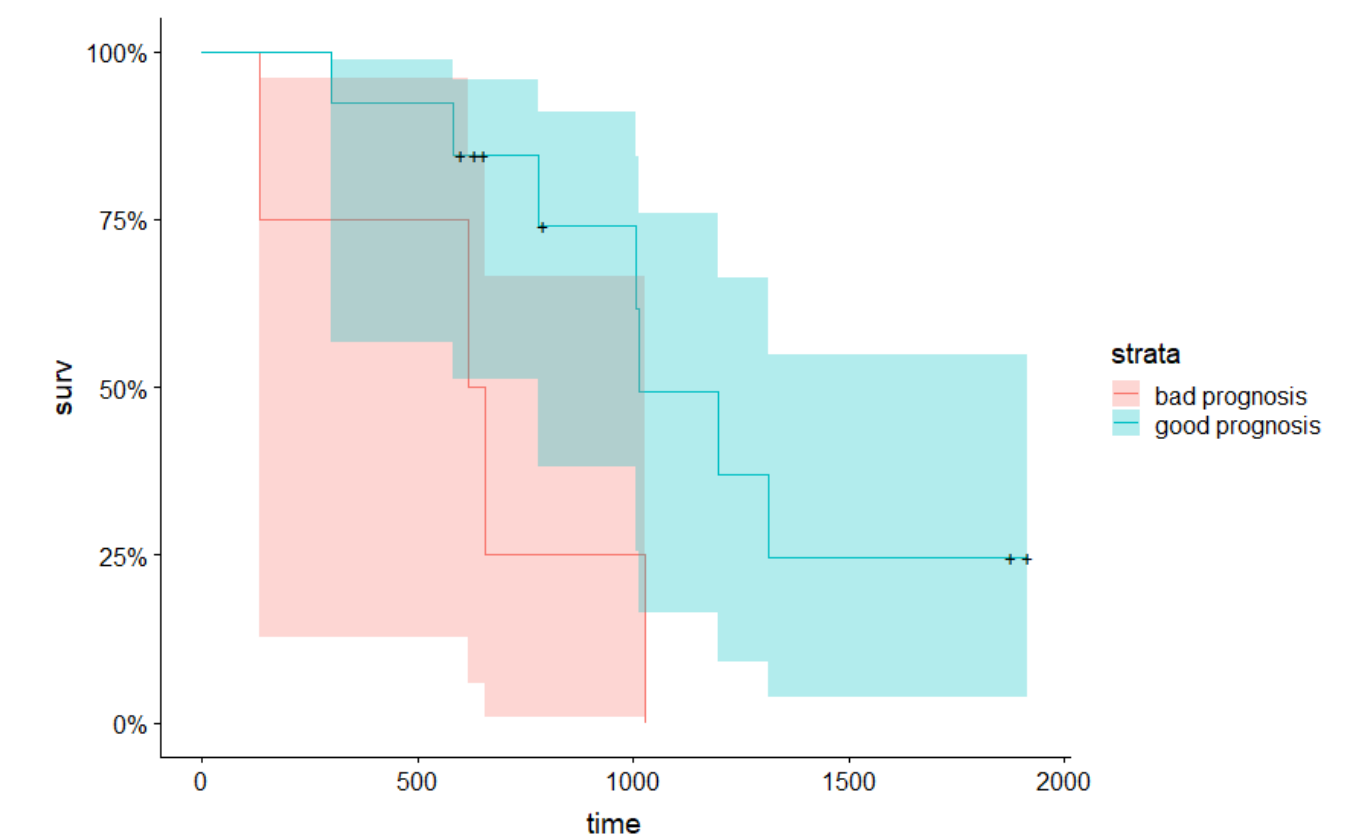


Fig. 8: Kaplan Meier curves of the two groups obtained with the test set (pink : bad prognosis, blue : good prognosis). Error : 0.39, p-value = 0.045

### 3) The interest of using textural features

	All characteristics	textural and conventional characteristics	clinical characteristics
Our method	0.34	0.36	0.45
Minimal depth method	0.45	0.45	0.48
Variable-Hunting method	0.40	0.41	0.45
Without selection	0.51	0.67	0.52

Tab.2: Influence of the features class on the prediction error, with different selection methods. There are the same patients in each sub-databases

### 3) The predictive features

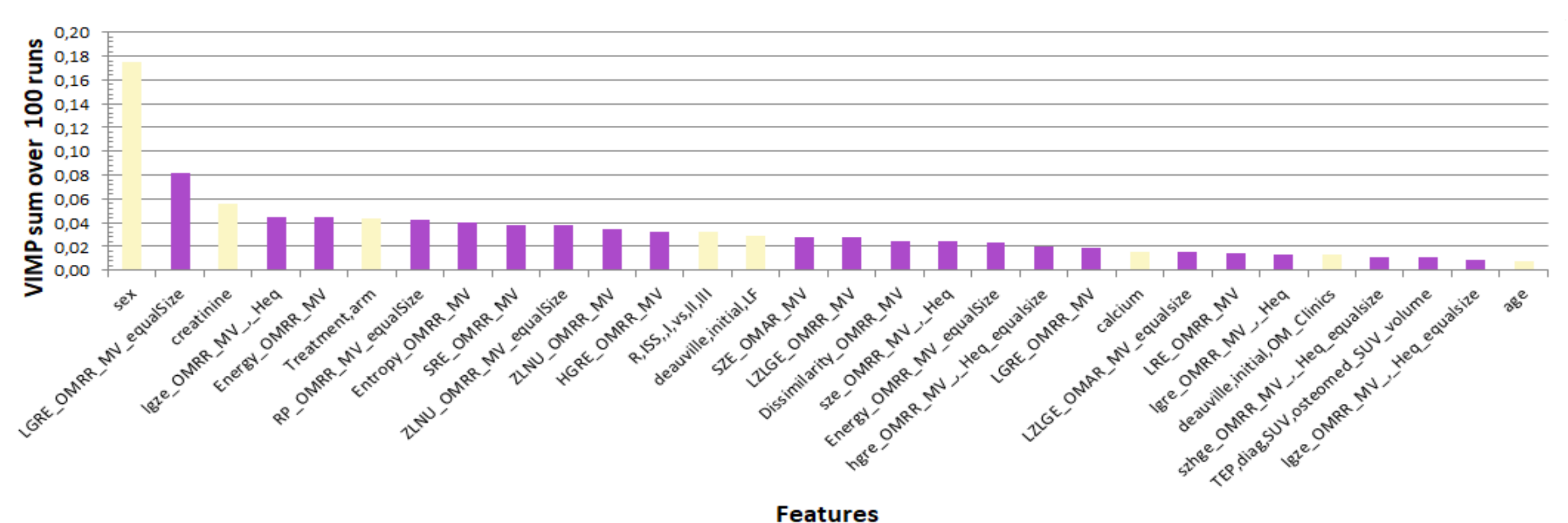


Fig. 8 : Histogram of the features found as predictive with the VIMP method. Yellow: clinical, purple: Imaging. OMRR (One Matrix relative resampling), OMAR (absolute resampling), Heq (histogram equalization), equalsize (equal size of voxels)

## C. Conclusions and perspectives

- ✓ The proposed method is more efficient than conventional approaches.
- ✓ It is possible to correctly separate two classes of patients (good/bad prognosis).
- ✓ It is possible to determine the features that are the most predictive.
- ✓ It shows the interest of using textural features.
- ✓ The relative resampling is more predictive than absolute resampling
- ✓ This approach can be generalised to other diseases.
- ✓ Thereafter, more patients will be included.

## References

- [1] P. Moreau, et al., Prospective evaluation of magnetic resonance imaging and [18F] fluorodeoxyglucose positron emission tomography-computed tomography at diagnosis and before maintenance therapy in symptomatic patients with multiple myeloma included in the ifm/dfei 2009 trial : Results of the imajem study. Journal of Clinical Oncology, 35(25).
- [2] H. Ishwaran, et al. Random survival forest. The annals of applied statistics, 2(3).
- [3] M. Vallières et al. Image biomarker standardisation initiative. Journal of Clinical Oncology, 35(25) :2911-2918.

**Thanks** : This work has been partially funded by the SIRIC ILIAD and the MILCOM Connect Talent.