# Development of a metabolic score predictive of survival in patients with Multiple Myeloma

Philippe Laurent[1], Alizée Steer[1], Elina Alaterre[1], Angélique Bruyer[1], Jérôme Moreaux [2,3,4,5]
1 Diag2tec, Montpellier, France; 2 Institute of Human Genetics, Unité Mixte de Recherche, Centre National de la Recherche Scientifique, Université Montpellier, Montpellier, France; 3 Department of Biological Hematology, CHU Montpellier, Montpellier France; 4 University of Montpellier, UFR Medicine, Montpellier, France; 5 Institut Universitaire de France, Paris, France

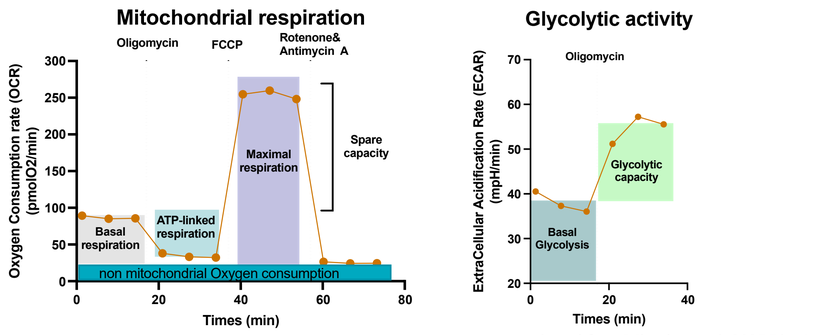
JOBIM 2026 30 JUIN>3 JUIL Strasbourg
ID submission 330

## Background & Aim

Metabolic activity plays an important role in treatment response across various cancers. However, in Multiple Myeloma (MM), no gene-expression-based score integrating functional metabolic data has yet been reported to predict patient survival. Because metabolic measurements are rarely available in patient cohorts, we used functional metabolic profiling from a collection of Human Myeloma Cell Lines (HMCL), representative of the molecular heterogeneity of the MM disease, together with RNA sequencing data to develop a gene-expression–based metabolic score.

## Method

Seven glycolytic or oxidative phosphorylation parameters, measured (Basal ECAR, Glycolytic Capacity, Basal OCR, Maximum Respiration, ATP-Linked Respiration and Spare Respiratory Capacity) or calculated (Glycolytic ATP Production Rate, which is equivalent to the Proton Efflux Rate due to glycoclysis) using the Seahorse XF Mito Stress test Kit, were combined with the raw counts data of 105 metabolic genes from 26 HMCLs. After raw count normalization (RLE method), log2-transformation and standardization, we developed predictive scores derived from machine learning approaches including Elastic Net regression, partial least squares (PLS) regression, and stepwise optimization. Correlations between predictive scores and metabolic measurements were then evaluated, along with their prognostic value, using the maxstat method in both Montpellier (n=102) and CoMMpass (n=669) cohorts of MM patients. These gene expression data were obtained from plasma cells of newly diagnosed multiple myeloma patients.




Adapted from Minming et al, 2027 PlosOne

***Figure 1*** **: Measurement and calculation of seven metabolics data representing the ATP production using the Seahorse XF Analyser.**

## Pipeline

### 1. List of genes involved in metabolic pathways

An a priori selection of 105 genes, including 25 genes involved in glycolytic metabolism and 80 genes involved in oxidative phosphorylation.

### 2. ElasticNet Regression

For each metabolic data, Elastic Net regression was applied to the initial set of 105 genes using 100 repetitions of 5-fold cross-validation. Only genes selected more frequently than expected under permutation were retained (*Figure 2)*.
This is a variable selection method based on penalizing the least informative variables.

*β penalty formula :* $\hat{\beta} = argmin_{\beta}(\|y - \beta\|^2 + \lambda_2\|\beta_1\|^2 + \lambda_1\|\beta_1\|_1)$

### 3. Partial Least Square (PLS) regression

Elastic Net regression may retain correlated predictors despite reducing the number of variables. PLS regression was subsequently applied to address for the remaining multicollinearity, and variables with a Variable Importance in Projection (VIP) score greater than 1 indicating an above-average contribution to the model, were retained for further analysis.

### 4. Stepwise optmisation

A bidirectional stepwise selection procedure based on the Akaike Information Criterion (AIC) was subsequently performed to identify the most parsimonious linear model.

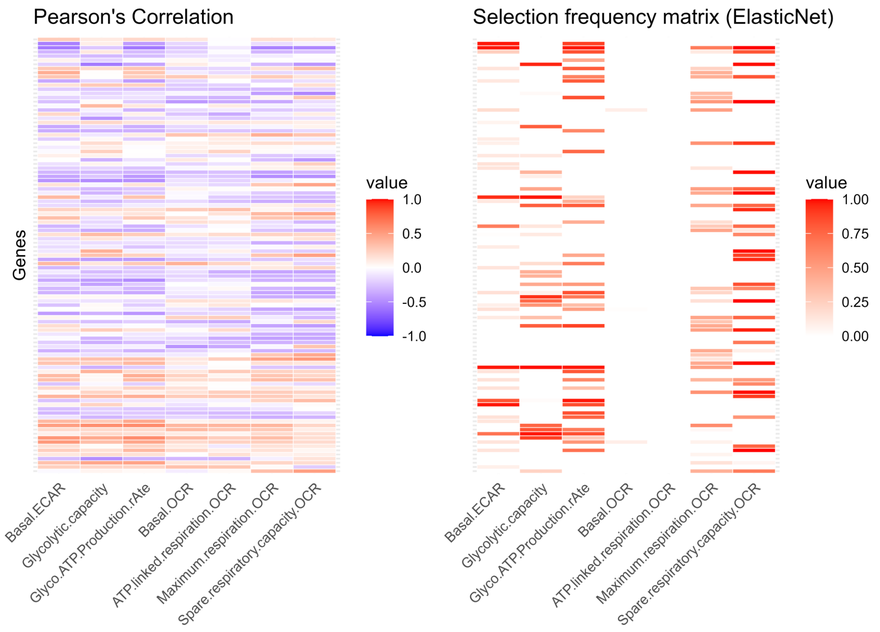
### 5. Performance Criteria

The final model was evaluated based on the adjusted $R^2$ and the Variance Inflation Factor (VIF), which respectively assess model fit and multicollinearity. but also that this score can significantly separate patients into low and high risk groups thanks to the maxstat function

## Results

### ElasticNet Regression

***Figure 2*** **: Comparison of the correlation matrix of genes of interest for each MD with the frequency matrix of use of these genes by elasticNet.**




The selection frequency matrix was generated from 100 repetitions of Elastic Net regression using 5-fold cross-validation for each metabolic data.
The resulting values, divided by 100, provide a selection frequency ranging from 0 to 1.
A strong consistency can be observed between the selection frequency of each gene and its univariate correlation. These results indicate that some genes are not selected for any metabolic data, while some metabolics data are not explained by any of the genes included in this panel.
It can also be observed that certain genes are selected to explain both glycolytic and oxidative phosphorylation-related metabolics data. This dual involvement highlights the complexity of metabolic regulation, where metabolic pathways may be interconnected and interdependent.

### Model selection & cross cohort compatibility

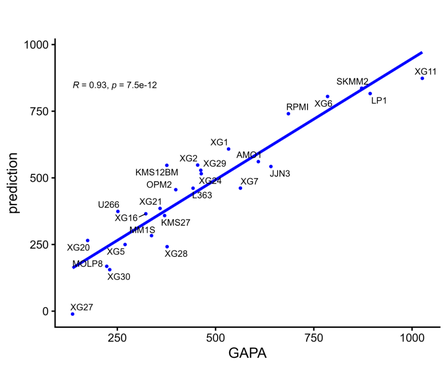



***Figure 3*** **: Correlation between Glycolysis ATP Production rAte (GAPA) and Leave One Oout Cross Validation prediction for the 26 HMCL cell lines.**

Among the evaluated models, the GAPA-based predictive model showed the best fit ($R^2$adj = 0.93, p= <0.001) and the stronghest associations with the other metabolic data. Therefore, this model was selected for the etablishment of the metabolic score.

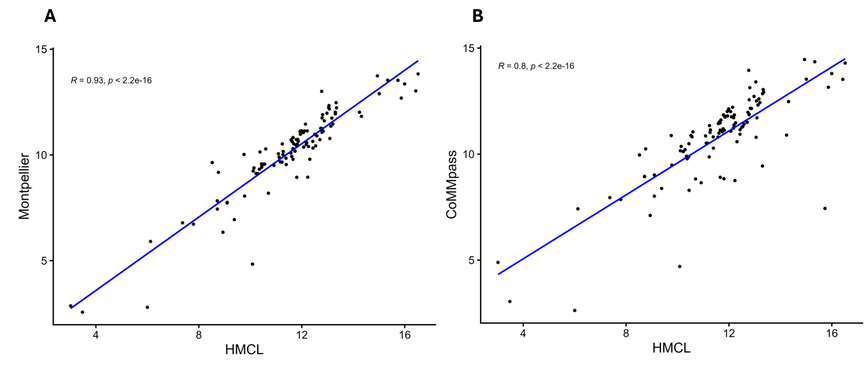



***Figure 4*** **: Spearman's correlation of geometric means of expression of the 105 genes normalized et standardised of the cell lines and cohorts of Montpellier (A) and CoMMpass (B).**

In order to assess the metabolic score of MM patients, it was essential to ensure that the expression levels of the 105 genes were comparable between the data used to construct the model (in HMCLs) and the expression data from two independent cohorts of MM patients (Montpellier n=102 and CoMMpass n=669 patients respecively). To do this, we ensure that the Spearman correlation is greater than or equal to 0.8 (r = 0.92 (A); r = 0.8 (B)).

### GAPA score predicts survival in MM patients

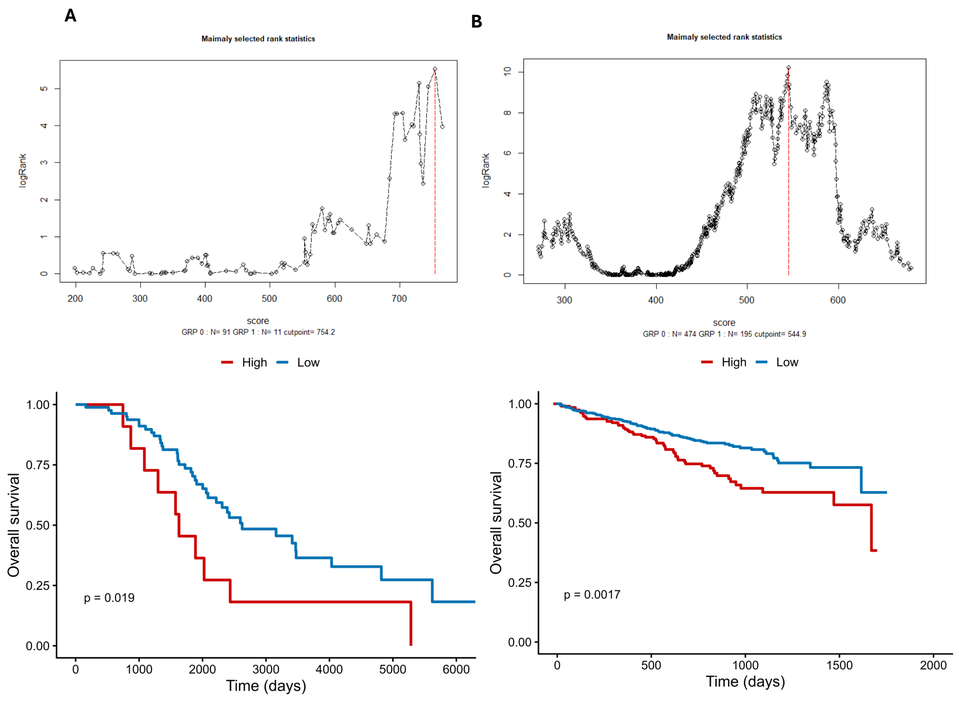



***Figure 5*** **: Prognostic Value of the GAPA Score in Montpellier (A) and CoMMpass (B) cohorts from the threshold calculated by the maxstat function**

The maxstat function generates a threshold that allows observations to be separated into two groups of patients (low and high risk) based on the GAPA score.
Significant differences in overall survival (OS) were observed between low and high-risk groups in both (A) Montpellier n=102 and (B) CoMMpass n=669 cohorts (p = 0.019 ; p = 0.0017, respectively).
These results suggests that the GAPA score can be associated with poor prognosis in MM patients.

## Conclusion

Geometric mean gene expression levels were highly correlated across datasets (Spearman r = 0.80–0.92, p < 0.001), supporting cross-cohort comparability between HMCL and MM datasets. The best predictive score was derived from the "Glycolysis ATP Production Rate" calculated value, which consists of a linear combination of 11 metabolic genes (6 glycolytic and 5 oxidative phosphorylation genes). This predictive score significantly stratified MM patients in both Montpellier (low-risk = 91, high-risk = 11, p = 0.02) and CoMMpass (low-risk = 474, high-risk = 195, p = 0.002) cohorts.

We developed an 11-gene expression-based score associated with high glycolytic metabolism in HMCL and identified high-risk MM patients, suggesting increased glycolytic activity in this subgroup. Beyond its prognostic value, this score may facilitate the identification of patients most likely to benefit from metabolism-directed therapeutic approaches.