



**HAL**  
open science

## **Apprentissage d'une mesure pour la stratification de patients atteints de la sclérose latérale amyotrophique**

Guillaume Tejedor, Veronika Peralta, Nicolas Labroche, Patrick Marcel, Hélène Blasco, Hugo Alarcán

### ► **To cite this version:**

Guillaume Tejedor, Veronika Peralta, Nicolas Labroche, Patrick Marcel, Hélène Blasco, et al.. Apprentissage d'une mesure pour la stratification de patients atteints de la sclérose latérale amyotrophique. Masses de Données, Informations et Connaissances en Sciences (MaDICS) 2025, May 2025, Toulouse, France. <hal-05342193>

**HAL Id: hal-05342193**

**<https://hal.science/hal-05342193v1>**

Submitted on 1 Nov 2025

**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.



HAL Authorization

# Apprentissage d'une mesure pour la stratification de patients atteints de la sclérose latérale amyotrophique



Guillaume Tejedor, Veronika Peralta, Nicolas Labroche, Patrick Marcel, H el ene Blasco, Hugo Alarcan

contact : guillaume.tejedor@univ-tours.fr



## Contexte

### Probl ematique

- Maladie impactant le contr ole des mouvements musculaires volontaires.
- Stratification des patients pour identifier des signatures cliniques/biologiques.

### Donn ees

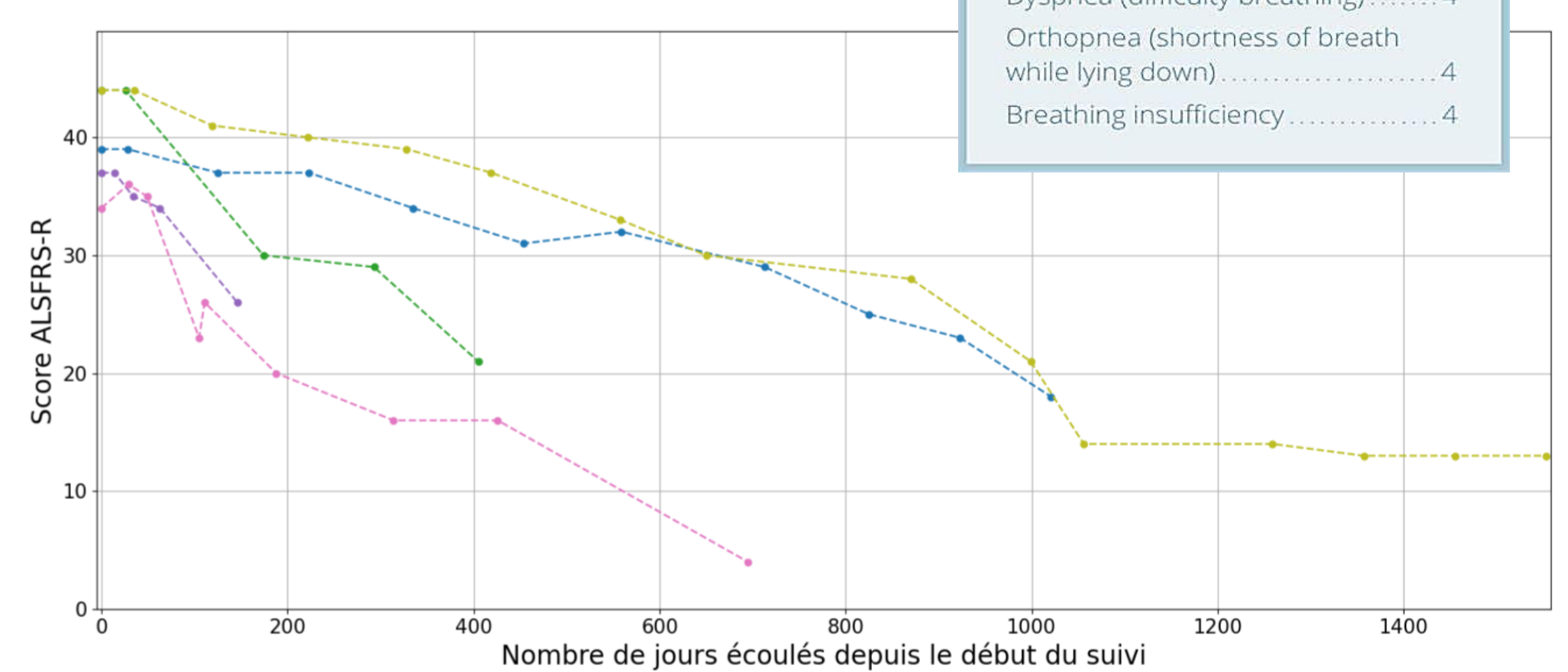
- D emographiques :  ge, localisation du d ebut de la maladie, etc.
- Longitudinales : S equences temporelles (cliniques, biologiques).
- Repr esentation progression maladie : Score ALSFRS-R
  - 12 questions (Chaque question not ee entre 0 et 4)
  - Score total : entre 0 (Perte totale) et 48 (Aucune Perte).

###  tat de l'art

- Crit eres de recherche :
  - Suivi de la maladie via le score ALSFRS-R.
  - Utilisation des d eclics fonctionnels pour stratifier.
- M ethodes employ ees :
  - Seuils appliqu es sur une variable extraite des s equences ALSFRS-R.
  - Mesure de similarit e (DTW - Dynamic Time Warping).

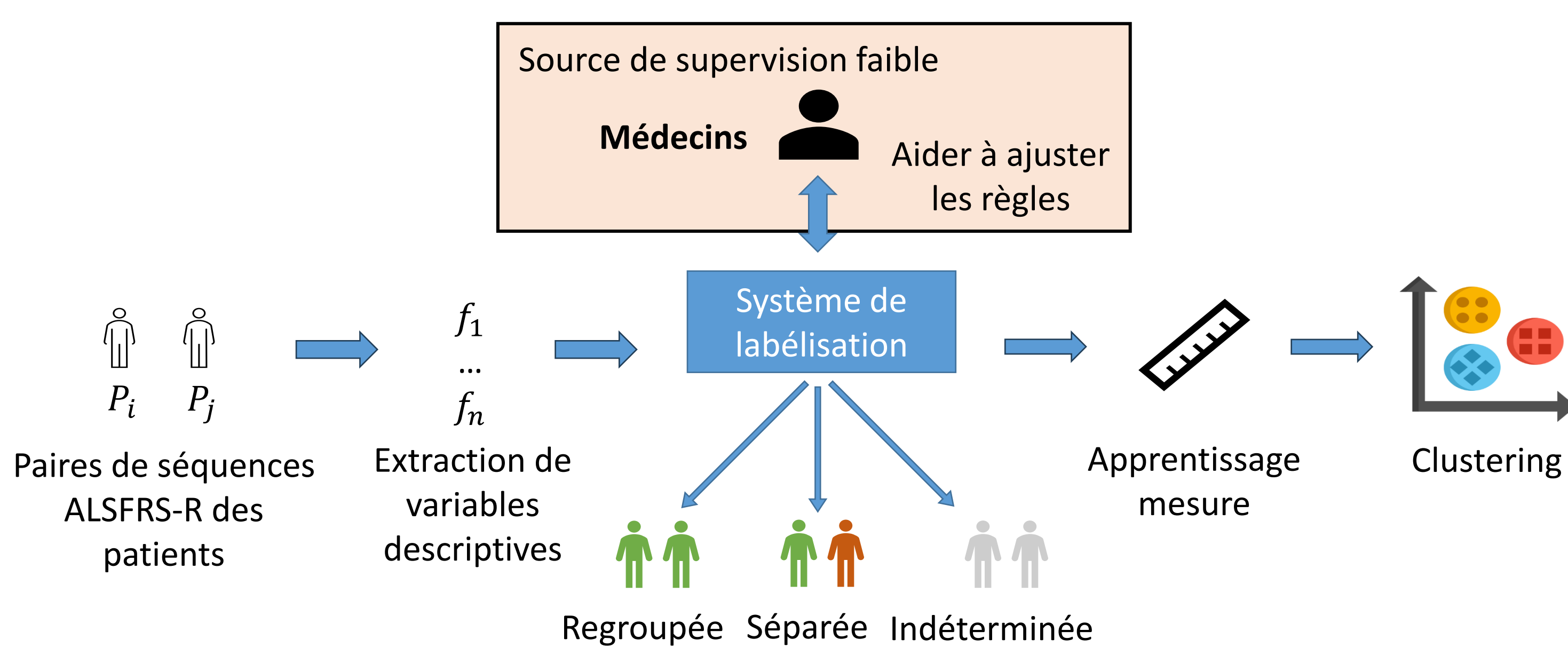
### Jeu de donn ees

- CHRU de Tours.
- 365 patients.
- 66 430 paires.

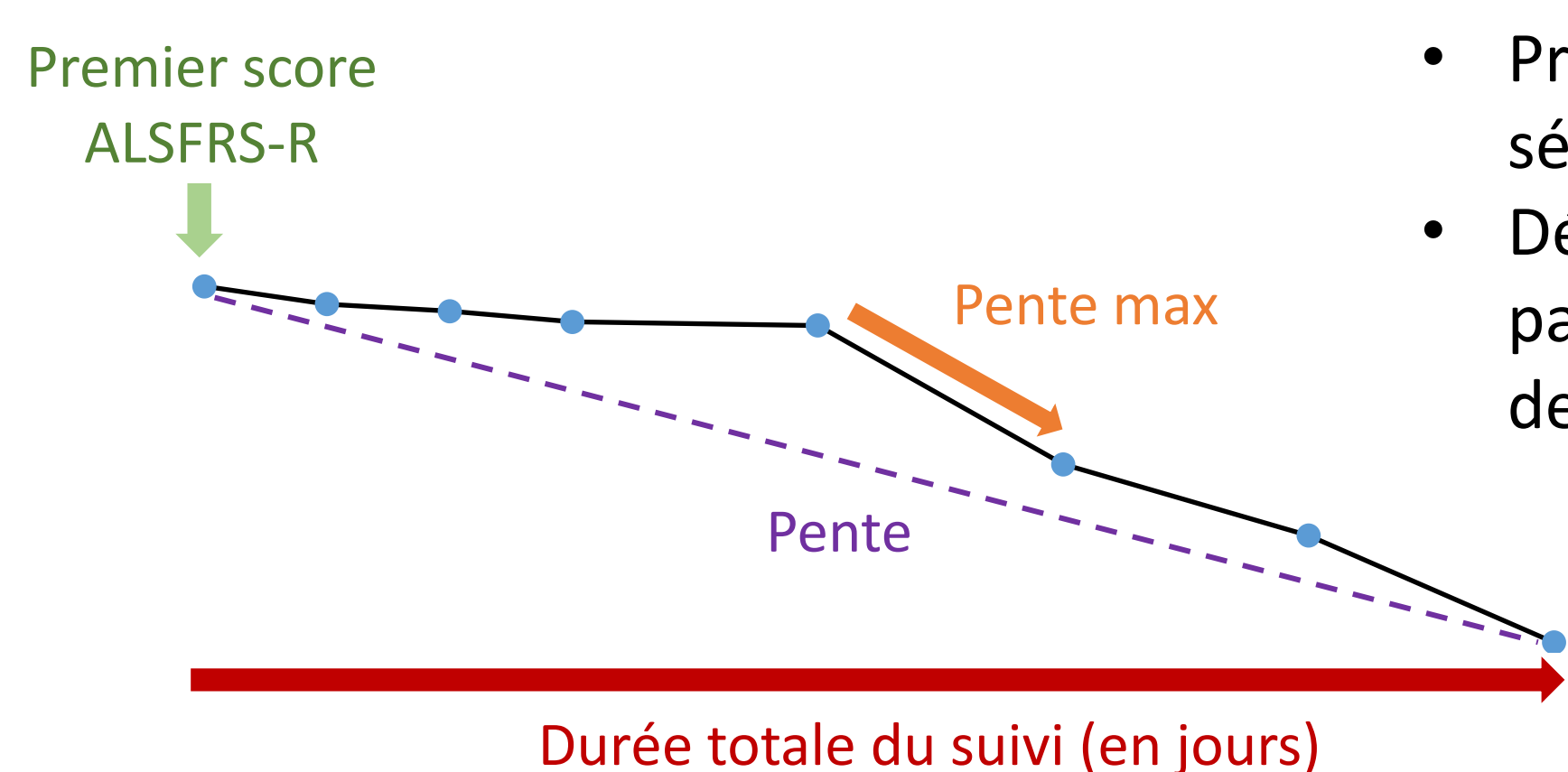


## Approche

### M ethodologie

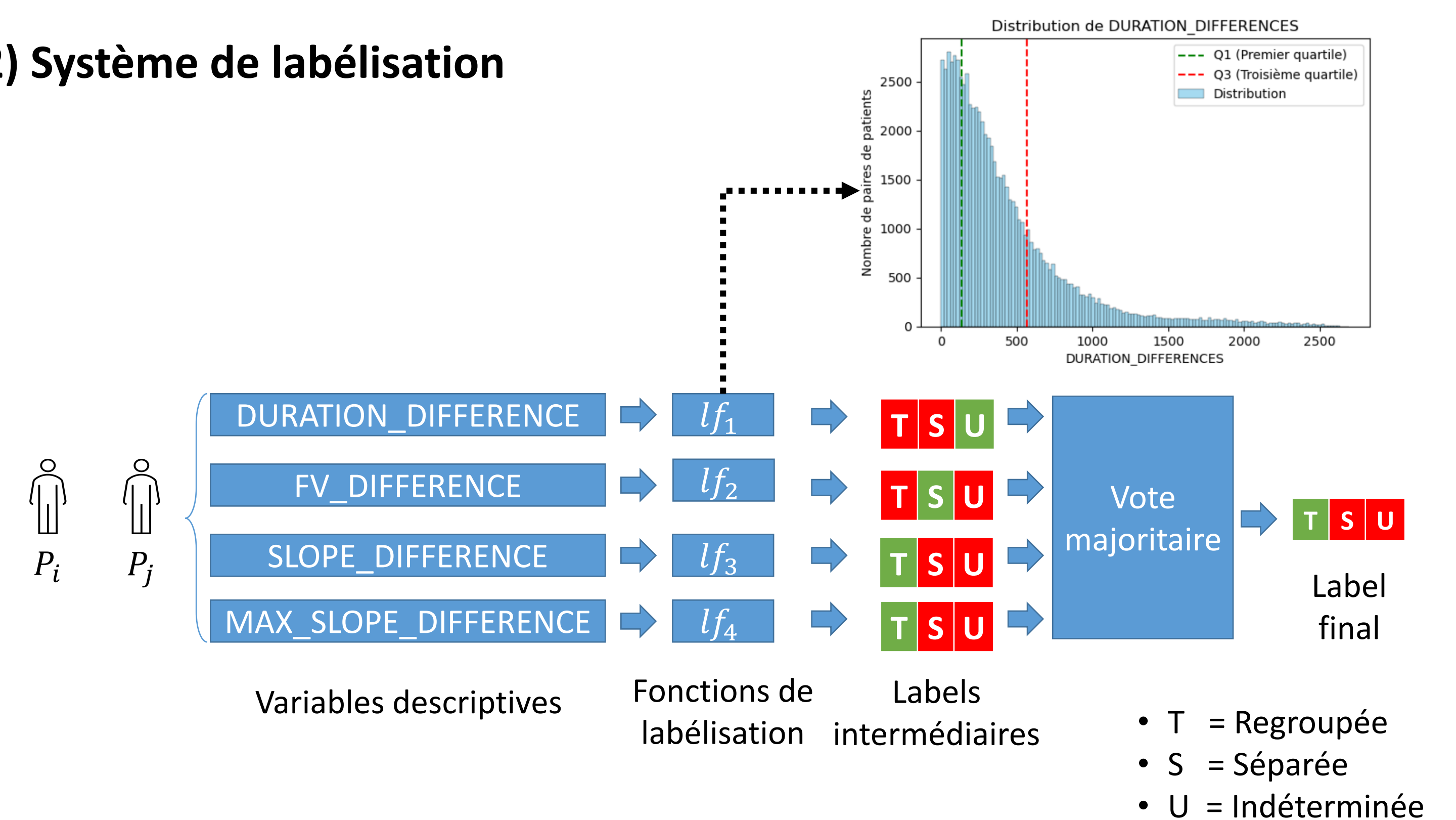


### 1) Extraction des variables descriptives



- Pr e-calcul d'informations extraites des s equences ALSFRS-R.
- D efinition de variables descriptives par paire de s equences (ex : diff erence de dur ee de suivi).

### 2) Syst eme de lab elisation



### 3) Apprentissage de la mesure

Support Vector Machine (SVM) lin eaire pour apprendre les poids de chaque variable.

- $P_i$  et  $P_j$  repr esentent respectivement les patients  $i$  et  $j$ .
- $k$  est le nombre de variables descriptives.
- $f_z$  d esigne la fonction de lab elisation  $z$ .
- $w_{f_z}$  d esigne le poids appris par classifieur.
- $m_{f_z}$  la valeur de la variable descriptive.

$$mesure(P_i, P_j) = \sum_{z=1}^k w_{f_z} * m_{f_z}$$

## Exp erimentations

Mesure	UMAP	Algorithme de clustering	c=3	c=4	c=5
DTW	NO	AHC	0.304 ± 0.263	0.234 ± 0.241	0.22 ± 0.243
DTW	NO	KMEDOID	0.299 ± 0.251	0.211 ± 0.266	0.157 ± 0.254
DTW	YES	AHC	0.449 ± 0.243	0.43 ± 0.248	0.43 ± 0.192
DTW	YES	KMEANS	0.475 ± 0.192	0.456 ± 0.174	0.413 ± 0.204
DTW	YES	KMEDOID	0.473 ± 0.194	0.452 ± 0.184	0.426 ± 0.178
DSHF	NO	AHC	0.627 ± 0.182	0.442 ± 0.193	0.443 ± 0.198
DSHF	NO	KMEDOID	0.343 ± 0.223	0.274 ± 0.199	0.229 ± 0.2
DSHF	YES	AHC	0.494 ± 0.291	0.471 ± 0.263	0.373 ± 0.272
DSHF	YES	KMEANS	0.483 ± 0.183	0.436 ± 0.183	0.451 ± 0.169
DSHF	YES	KMEDOID	0.485 ± 0.181	0.411 ± 0.178	0.443 ± 0.191

- AHC : Agglomerative Hierarchical Clustering.
- Silhouette Moyenne ±  cart-type variant le nombre de clusters  $c$  entre 3 et 5.

## Conclusion et perspectives

### R esum e

- Approche combinant syst eme de lab elisation et SVM pour stratifier les patients via les s equences ALSFRS-R.
- Comparaison DTW : Meilleure performance de notre m ethodologie.

### Perspectives

- Nouvelles caract eristiques pour affiner les clusters.
- Analyse statistique : Survie, diff erences entre clusters.
- Pr ediction du cluster d es le diagnostic.