

P. Heudel<sup>1</sup>, M. Ahmed<sup>2</sup>, A. Attye<sup>2</sup>, F. Renard<sup>2</sup>

1. Centre Léon Bérard, Lyon, France. 2. GEODAISICS, Grenoble, France.

## Objectif

Démontrer que l'utilisation de l'intelligence artificielle via l'usage de jumeaux numériques permet d'apporter des informations complémentaires facilitant la prise en charge des patientes de + de 70 ans opérées d'un cancer du sein localisé.

## Population

L'analyse des 370 881 DPI du centre Léon Bérard a retrouvé 28 924 patientes opérées pour un cancer du sein précoce dont 4 444 patientes de plus de 70 ans au moment du diagnostic (15,4%). 1221 patientes ayant un de suivi de + de 5 ans ont été retenues avec un âge médian de 77 ans (70 à 100 ans) au moment du diagnostic du cancer du sein. Respectivement, les tableaux 1, 2 et 3 ci-après résument les caractéristiques des patientes, les caractéristiques des tumeurs et le traitement reçu pour un cancer du sein.

Age au diagnostic	70-74	75-79	80-84	85-89	over 90
n (%)	580 (47%)	331 (27%)	204 (17%)	93 (8%)	20 (2%)
Performance status au diagnostic	0	1	2	3 et 4	missing data
n (%)	339 (28%)	322 (26%)	48 (4%)	23 (2%)	497 (40%)
IMC au diagnostic	< 18,5	18,5-25	25-30	> 30	missing data
n (%)	32 (3%)	446 (36%)	409 (33%)	266 (22%)	76 (6%)
Comorbidités au diagnostic	Clrc creat <40	Ins. cardiaque	Coronaropat.	BPCO	diabete
n (%)	57 (5%)	105 (9%)	123 (10%)	36 (3%)	174 (14%)

Table 1: Caractéristiques des patients

Taille tumoral	T1	T2	T3	T4	missing data
n (%)	567 (46%)	286 (23%)	36 (3%)	250 (20%)	90 (7%)
Statut ganglionnaire	N0	N1	N2	N3	missing data
n (%)	614 (50%)	243 (20%)	55 (4%)	55 (4%)	262 (21%)
Grade SBR	I	II	III	missing data	
n (%)	188 (15%)	648 (53%)	281 (23%)	112 (9%)	
Récepteur œstrogène	Positif	Négatif	missing data		
n (%)	978 (80%)	145 (12%)	106 (9%)		
Récepteur progesterone	Positif	Négatif	missing data		
n (%)	838 (68%)	285 (23%)	106 (9%)		

Table 2: Caractéristiques des cancers

Chirurgie mammaire	Mastectomie	Tumorectomie	missing data
n (%)	610 (50%)	611 (50%)	0
Chirurgie axillaire	Ganglion sentinelle	curage	missing data
n (%)	815 (63%)	492 (40%)	0
Traitement adjuvant	Chimio thérapie	Radio thérapie	Hormono thérapie
n (%)	149 (12%)	307 (25%)	962 (78%)

Table 3: Traitements réalisés

## Méthodes

Les ensembles de données de grande dimension peuvent être très difficiles à visualiser/modéliser. Le MANIFOLD LEARNING (ML) est une solution pour rendre les données plus intuitives en réduisant la dimensionnalité non linéaire grâce à un apprentissage automatique multiple. La technique apprend la structure des données à partir des données elles-mêmes, sans utiliser de classifications prédéterminées. Les dimensions de l'espace réduit obtenu correspondent à des variations jointes des données de départ. Ces dimensions peuvent être considérées comme les principales degrés de liberté. L'étude a reçu l'approbation de la Commission Nationale de l'Informatique et des Libertés (autorisation n°11. 9191415 ; 10 octobre 2019) et a été construit dans le respect des réglementations françaises et européennes.

## Résultats

Avec un AUC à 0,89, le Manifold Learning arrive à préciser très correctement le risque de décès de cette population en différenciant 6 clusters ayant chacun des profils de patientes et des caractéristiques tumorales particulières, associés à un risque spécifique de décès. Les tumeurs RH- sont concentrées sur les clusters 2 et 3. Distincts en terme de pronostic, les clusters 4 et 5 ne contiennent que des tumeurs RH+ ayant un stade identique. Le taux de survie moindre du cluster 5 peut s'expliquer par la présence de comorbidités.

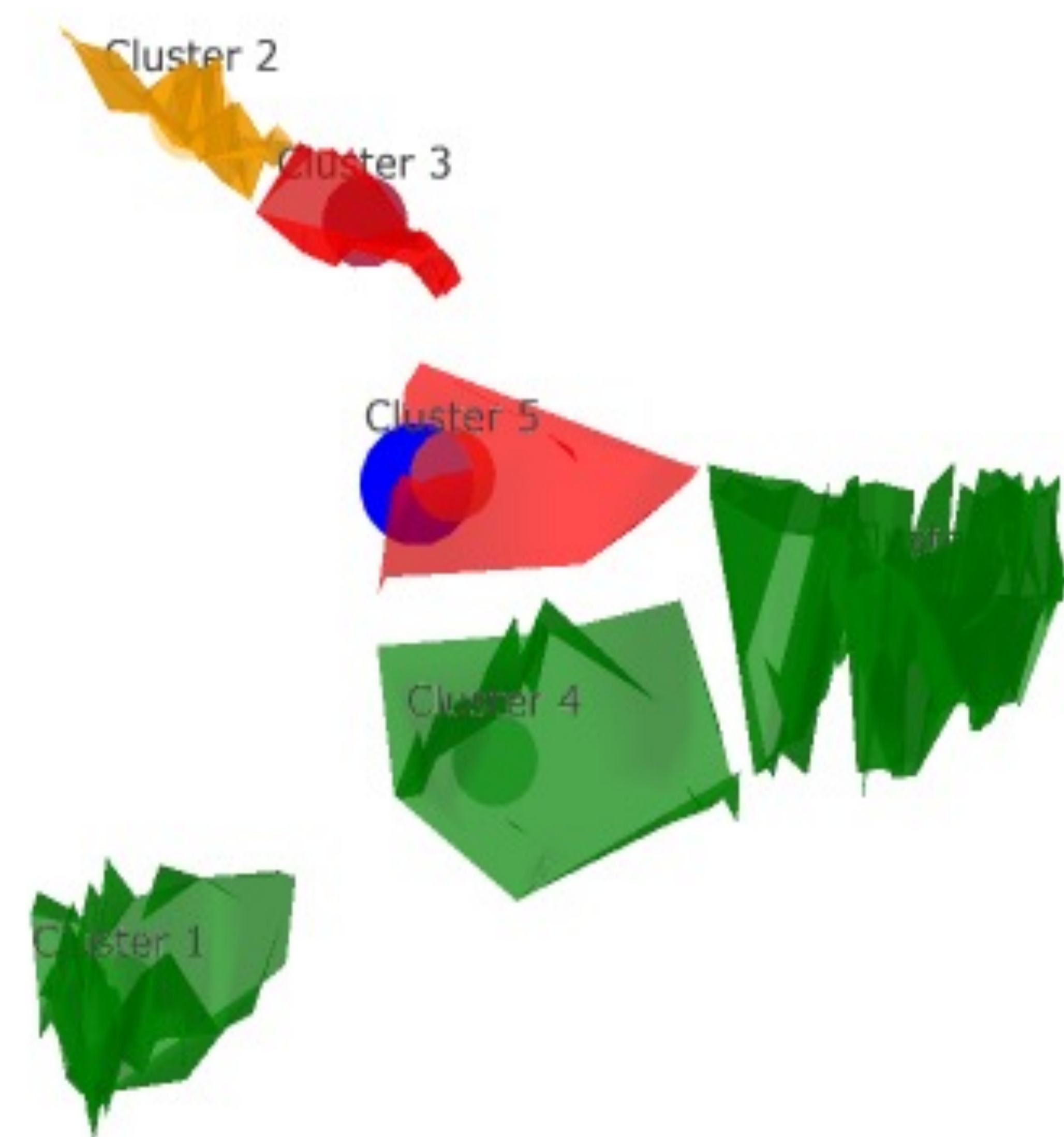


Figure 1. Représentation graphique des 6 clusters définis par le ML

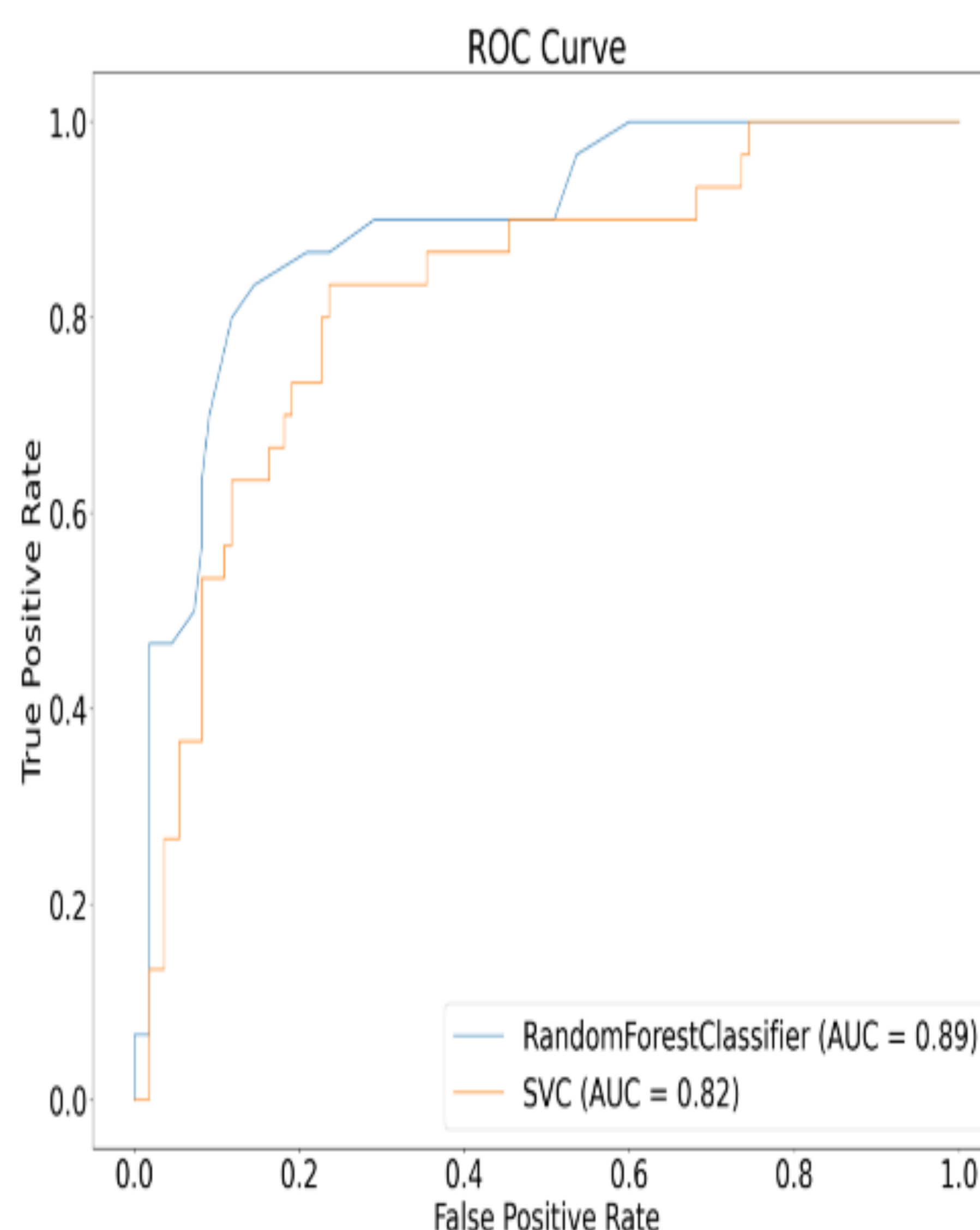


Figure 2. Courbes ROC pour les modèles prédictifs de mortalité à 5 ans.

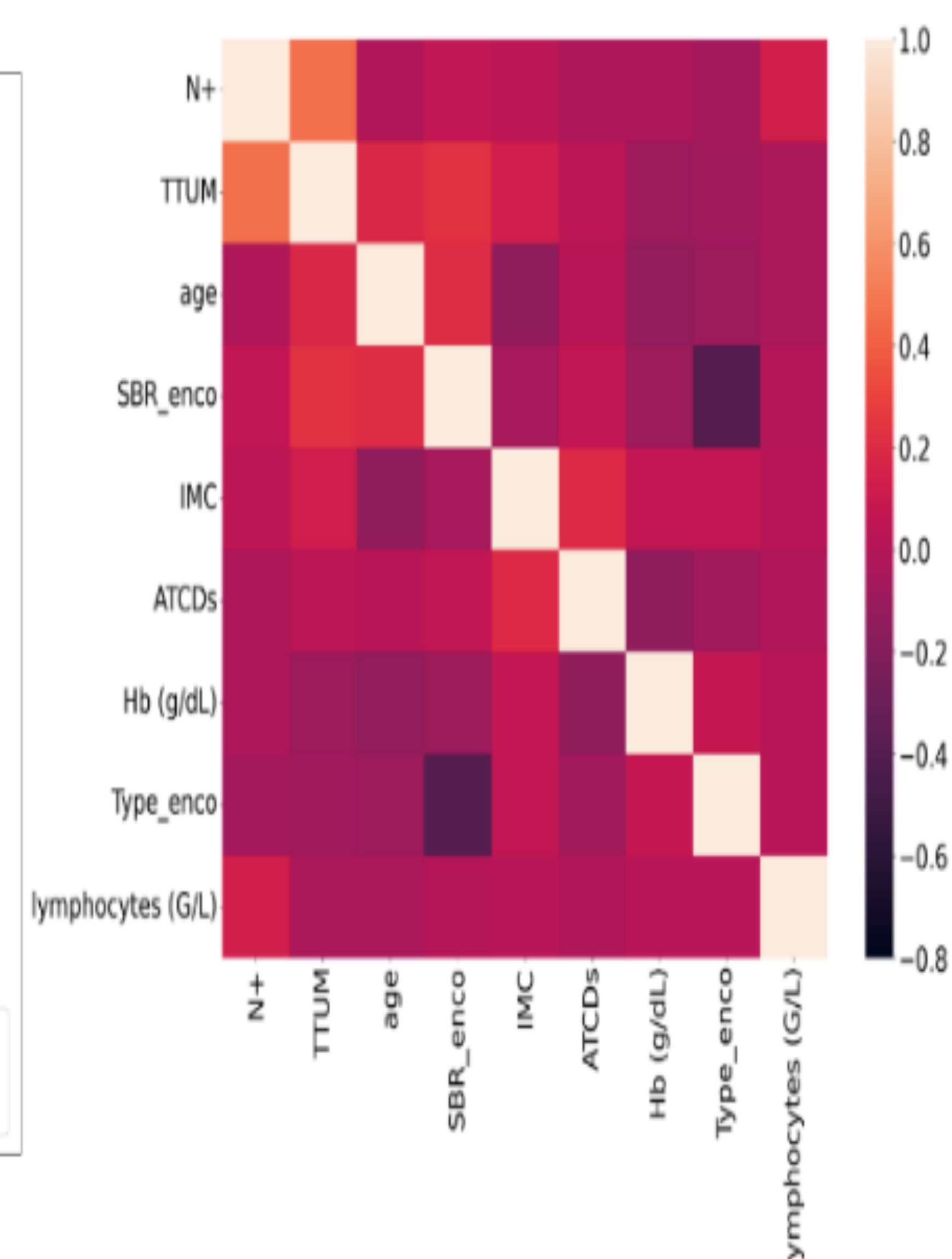


Figure 3. Heatmap montrant les corrélations / paires entre variables

Cluster	0	1	2	3	4	5
Hb (G/dL)	13.4	13.2	13.1	13	13.5	11.7
IMC	25	28.2	25.3	28.3	26	23.1
Nombre de N+	0.6	0.7	1	7.5	1	1.4
Tumeur (mm)	19.6	19.1	24.2	69.7	23.8	36.5
Age (années)	75.5	76.5	77.8	76.8	79.7	80.4
Lymphocy. (G/L)	1.8	2.1	1.9	1.6	3.6	1.5
Comorbidités	0	1	0.4	0.3	0	0.7
RH+/RH-	1	1	0	0.8	1	1
SBR High/low	0	0.1	0.7	0.5	0.7	0.7

Table 4: Caractéristiques des 6 clusters définis par le ML

## Conclusions

L'usage des jumeaux numériques définis par Manifold Learning permet une estimation de la survie spécifique à 5 ans des patientes âgées de plus de 70 ans et ayant un cancer du sein localisé. En intégrant les interactions entre variables clinico-biologiques au sein de la population étudiée, il offre une assistance décisionnelle précieuse dans ce contexte médical complexe. Cette solution vise à guider les médecins vers des stratégies thérapeutiques adaptées au profil parfois complexe des patientes âgées.