

Identification et analyse des phénotypes des patients BPCO : Cohorte PALOMB.

E. H. Ouaalaya,* A. E. Benna, M. Zysman, É. Berteaud, L. Falque, L. Nguyen, J. M. Dupis, M. Sabatini, A. Bernady, A. Ozier, C. Nocent-Eijnani, C. Roy, F. Le Guillou, E. Monge, M. Aliati, A. Prudhomme, L. Petrov, M. L. Quinquenel, M. Staali, F. Pilard, M. Sapène, J. Casteigt, J. Moinard, Y. Daoudi, E. Iglesias, É. Blanchard, J. Macey, R. Veillon, X. Demant, C. Bon, L. Grassion, M. Molimard, C. Raherison-Semjen

*Auteur correspondant : el-hassane.ouaalaya@u-bordeaux.fr (E. H. Ouaalaya)

Palomb
la cohorte BPCO



e-CPLF 2021

.....
vendredi 29 janvier 2021
.....

Déclaration des liens d'intérêts

J'ai actuellement, ou j'ai eu au cours des trois dernières années, une affiliation ou des intérêts financiers ou intérêts de tout ordre avec les sociétés commerciales suivantes **en lien avec la santé**.

- Liens d'intérêt :
 - ✓ Intérêts financiers: néant
 - ✓ Liens durables ou permanents: néant
 - ✓ Intérêts indirects (co-auteurs): Université de Bordeaux, la Fondation de Bordeaux Université, Novartis Pharma, MundiPharma, Astra Zeneca, Chiesi, ALK, Boehringer Ingelheim, Glaxo Smithkline, MundiPharma.

- Liens d'intérêt en relation avec la présentation : néant



Contexte

- L'identification de **sous-groupes** de patients atteints de BPCO et **partageant des caractéristiques** cliniques et/ou paracliniques communes est à l'origine du concept de phénotypes dans cette pathologie.
- De nombreuses études de cohortes observationnelles sont en cours, pour tenter d'identifier de **nouveaux phénotypes** de patients BPCO.



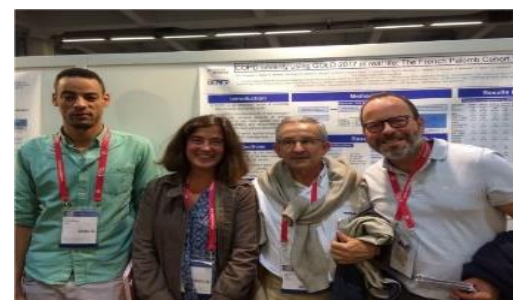
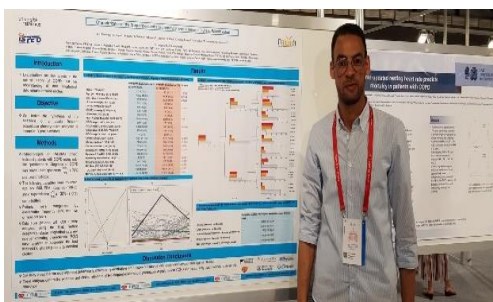
Objectif

- Déterminer l'analyse en cluster la plus optimale pour identifier de façon robuste les phénotypes cliniques dans le cadre de la cohorte PALOMB



Cohorte PALOMB

- Plateforme sur la prise en charge de la BPCO en **Aquitaine** et en **Charentes**.
- **Ces Objectifs:**
 - Mieux connaître la BPCO
 - Améliorer la prise en charge et le suivi des patients.
 - Recenser les données épidémiologiques
 - Créer un portail dans le cadre de la formation médicale continue.
- **Pilotage:**
 - Un comité de pilotage assure la direction opérationnelle du projet.
 - Un conseil scientifique.
 - Une attachée de recherche clinique.
- **L'évolution de la cohorte (depuis janvier 2014):**
 - **3152** : Nombre de patients inclus.
 - **2833** : Nombre de visite des patients inclus et suivis.



Patients, méthodes et analyse statistique (1)

- Analyse des données de la cohorte PALOMB de patients BPCO pris en charge par des pneumologues (hospitaliers et libéraux)
- **2968** patients ayant une BPCO (VEMS /CVF <70% post BD) ont été inclus depuis **Janvier 2014** par leur pneumologue par web questionnaire(<https://www.palomb.fr/>)
- Le temps écoulé jusqu'au décès a été censuré à la fin du suivi (**février 2020**).
- Les données sur leur statut vital ont été obtenues en utilisant le **RNIPP**.



Patients, méthodes et analyse statistique (2)

- Données relatives à la **symptomatologie respiratoire, l'EFR, comorbidités et aux traitements**
- Pré-traitement des données (**Outliers**)
- Analyse bivariée (Analyse de la **variance**, test de **Chi-2, Wilcoxon, Kruskal-Wallis, Fisher**)
- Réduction de dimension (**ACP**)
- Trois méthodes de **classification non supervisée** ont été utilisées :
 - Classification Hiérarchique sur Composantes Principales (**HCPC**)
 - Partitionnement sur les k-Moyennes (**k-means**)
 - Partitionnement autour des médoïdes (**PAM**)
- Les critères de **robustesse**:
 - coefficients de détermination **R2**
 - **homogénéité** des clusters
 - l'indice de **connectivité**,
 - l'indice de **Dunn**
 - l'indice de **Silhouette**



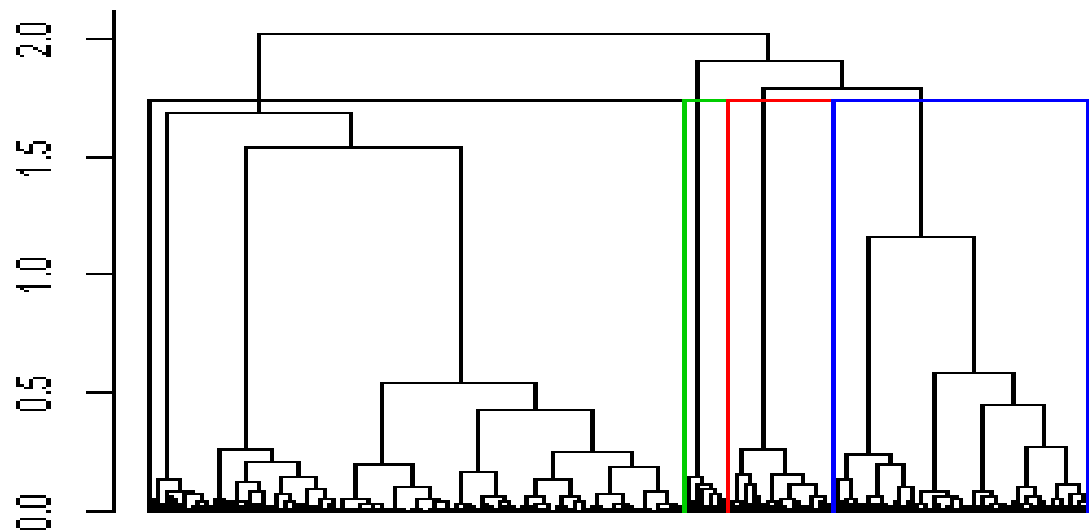
	Classification GOLD 2007				p-value
	L N=540 18,20%	M N=1535 51,70%	S N=715 24,10%	TS N=178 6%	
Hommes, 1892 (63.74%)	307 (56.9)	981 (63.9)	480 (67.1)	124 (69.7)	0.001
Âge, 70.07 (10.89)	66.5 (10.7)	70.5 (11.0)	72.0 (10.5)	69.3 (9.66)	<0.001
Décès, 335 (11.30%)	21 (3.89)	160 (10.4)	116 (16.2)	38 (21.3)	<0.001
IMC (kg/m²), 26.30 (5.84)	25.6 (4.88)	27.1 (5.99)	25.9 (5.95)	23.2 (5.26)	<0.001
Tabagisme, N (%) :					
Jamais, 148 (5%)	19 (3.52)	76 (4.95)	43 (6.01)	10 (5.62)	
Ancien Fumeur, 1671 (56.30%)	267 (49.4)	855 (55.7)	427 (59.7)	122 (68.5)	<0.001
Fumeur Actif, 1149 (38.70%)	254 (47.0)	604 (39.3)	245 (34.3)	46 (25.8)	
Toux, 1722 (58.02%)	289 (53.5)	889 (57.9)	432 (60.4)	112 (62.9)	0.047
Expectoration, 1251 (42.15%)	206 (38.1)	619 (40.3)	337 (47.1)	89 (50.0)	0.001
mMRC, N (%) :					
0-1 1295 (42.15%)	368 (68.1)	747 (48.7)	162 (22.7)	18 (10.1)	<0.001
≥ 2 1673 (56.36%)	172 (31.9)	788 (51.3)	553 (77.3)	160 (89.9)	
Exacerbations, N (%) :					
0-1 2198 (74.05%)	439 (81.3)	1186 (77.3)	473 (66.2)	100 (56.2)	
≥ 2 770 (25.95%)	101 (18.7)	349 (22.7)	242 (33.8)	78 (43.8)	<0.001
VR/CPT, 137.17 (36.73)	113 (54.4)	131 (22.4)	158 (25.2)	182 (28.8)	<0.001
VEMS, 61.31 (20.08)	90.7 (9.33)	64.7 (8.27)	41.1 (5.69)	24.4 (3.82)	0.0001
GOLD 2017, N (%)					
A 1059 (35.68%)	308 (57.0)	612 (39.9)	124 (17.3)	15 (8.43)	
B 1139 (38.37%)	131 (24.3)	574 (37.4)	349 (48.8)	85 (47.8)	
C 236 (7.95%)	60 (11.1)	135 (8.79)	38 (5.31)	3 (1.69)	<0.001
D 534 (18%)	41 (7.59)	214 (13.9)	204 (28.5)	75 (42.1)	



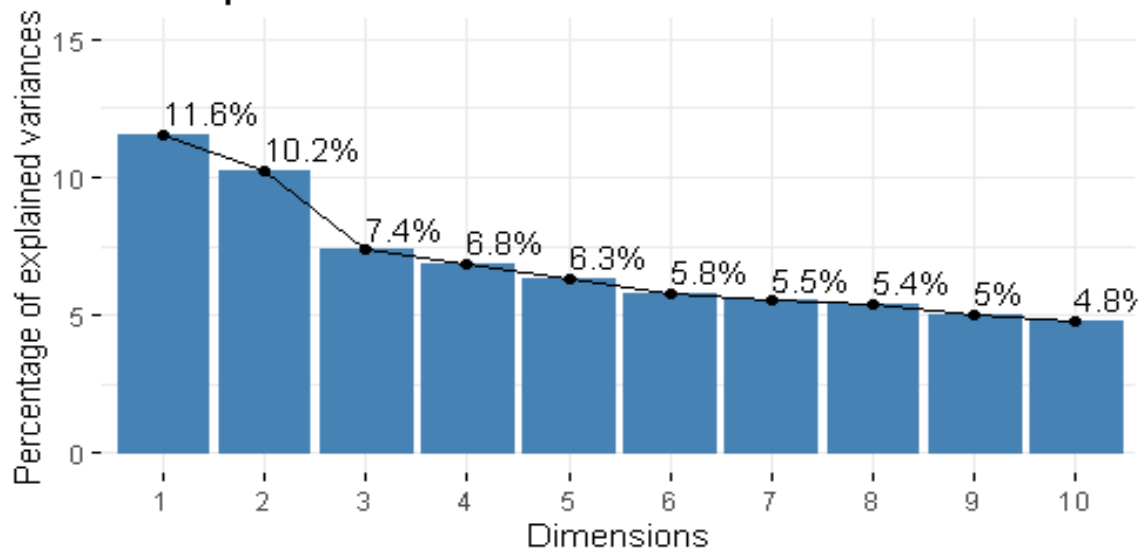
Comorbidités	Classification GOLD 2007				p-value
	L N=540 18,2%	M N=1535 51,7%	S N=715 24,10%	TS N=178 6%	
SAS, 490 (16.51%)	114 (21.1)	278(18.1)	85 (11.9)	13(7.3)	<0.001
DDB, 145 (4.88%)	25 (4.6)	58(3.7)	53 (7.41)	9(5.06)	0.003
Cancer Poumon, 204 (6.87%)	29 (5.3)	120(7.8)	50 (6.99)	5(2.81)	0.034
Autre Cancer, 351 (11.82%)	94 (17.4)	305(19.9)	134 (18.7)	22(12.4)	0.083
Hypertension, 1090 (36.72%)	148 (27.4)	593(38.6)	275 (38.5)	74(41.6)	<0.001
Cardiopathie Ischémique, 502 (17%)	71 (13.1)	284(18.5)	111 (15.5)	36(20.2)	0.014
Trouble du rythme, 346 (11.65%)	39 (7.22)	191(12.4)	89 (12.4)	27(15.2)	0.003
Insuffisance Cardiaque Gauche, 121 (4.07%)	9 (1.6)	55(3.5)	41 (5.73)	16(8.9)	<0.001
Athérome, 170 (5.72%)	20 (3.7)	88(5.7)	50 (6.99)	12(6.7)	0.087
HTAP, 64 (2.15%)	6 (1.1)	17(1.1)	22 (3.08)	19(10.7)	<0.001
Artériopathie, 190 (6.40%)	24 (4.4)	100(6.5)	49 (6.85)	17(9.5)	0.083
Diabète, 367 (12.36%)	42 (7.7)	219(14.3)	87 (12.2)	19(10.7)	0.001
Syndrome Métabolique, 105 (3.53%)	12 (2.2)	58(3.7)	29 (4.06)	6(3.37)	0.311
Dénutrition, 59 (17%)	3 (0.5)	13(0.8)	25 (2)	18(10.1)	<0.001
Manifestation Sinusiennes, 160 (5.40%)	40 (7.4)	70(4.5)	40 (5.5)	10(5.6)	0.091
Problèmes Digestifs, 75 (2.52%)	12 (2.2)	28(1.8)	31 (4.3)	4(2.25)	0.008
Maladie Thromboembolique, 36 (1.21%)	6 (1.1)	20(1.3)	6 (0.8)	4(2.25)	0.442
Ostéoporose, 97 (3.26%)	14 (2.5)	33(2.1)	37 (5.1)	13(7.3)	<0.001
Ostéodensitométrie, 97 (3.26%)	14 (2.5)	33(2.1)	37 (5.1)	13(7.3)	<0.001
Anxiété, 430 (14.48%)	47 (8.7)	211(13.7)	120 (16.8)	52(29.2)	<0.001
Dépression, 382 (12.87%)	70 (13)	191(12.4)	87 (12.2)	34(19.1)	0.081



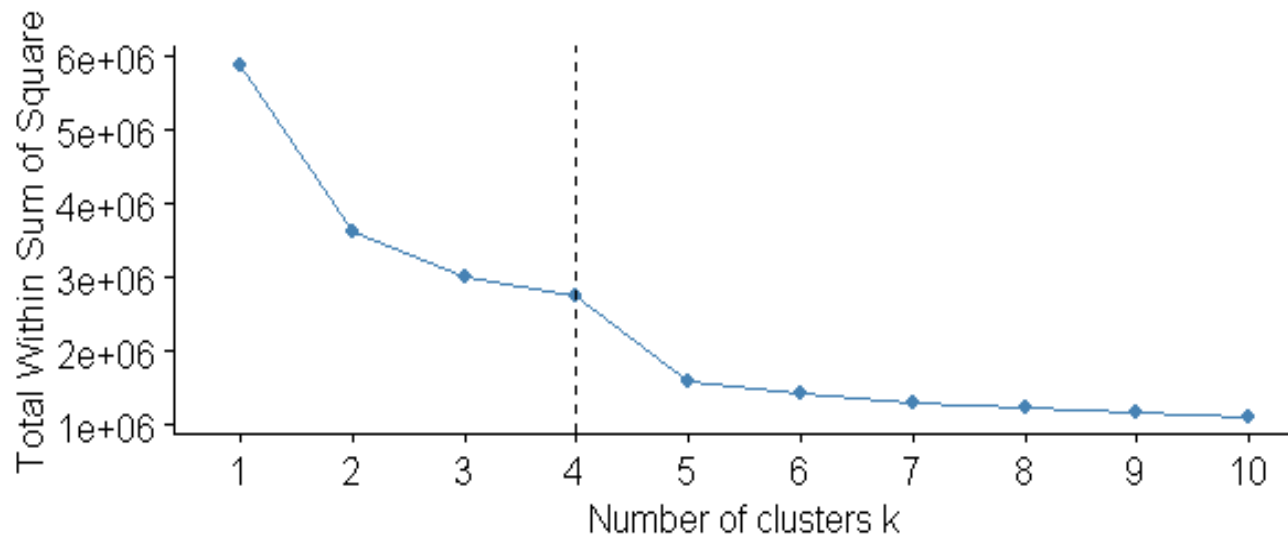
■ ACP + HCPC



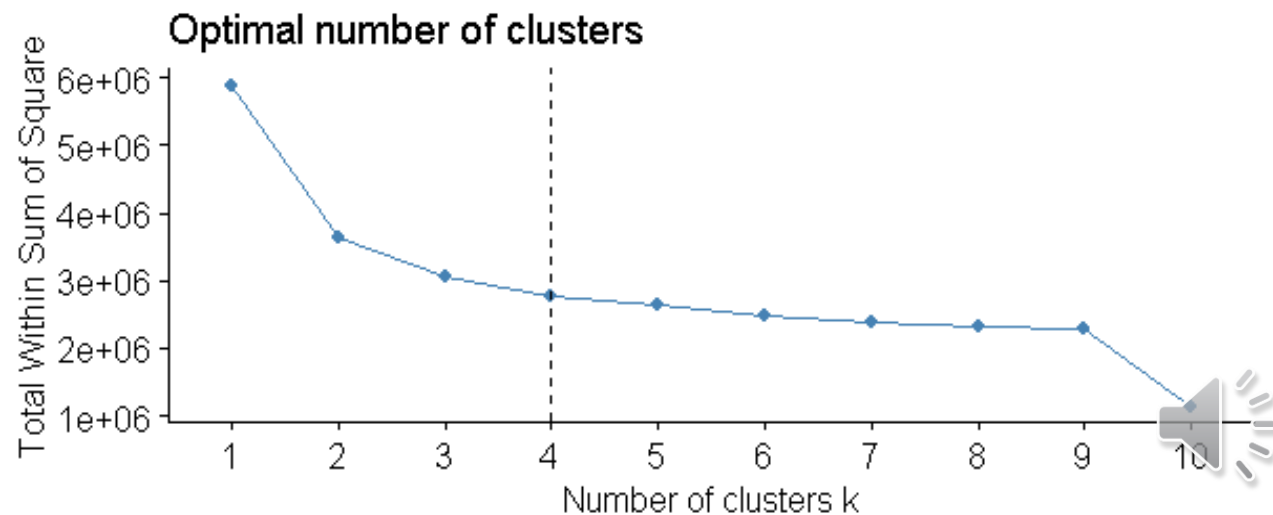
Scree plot



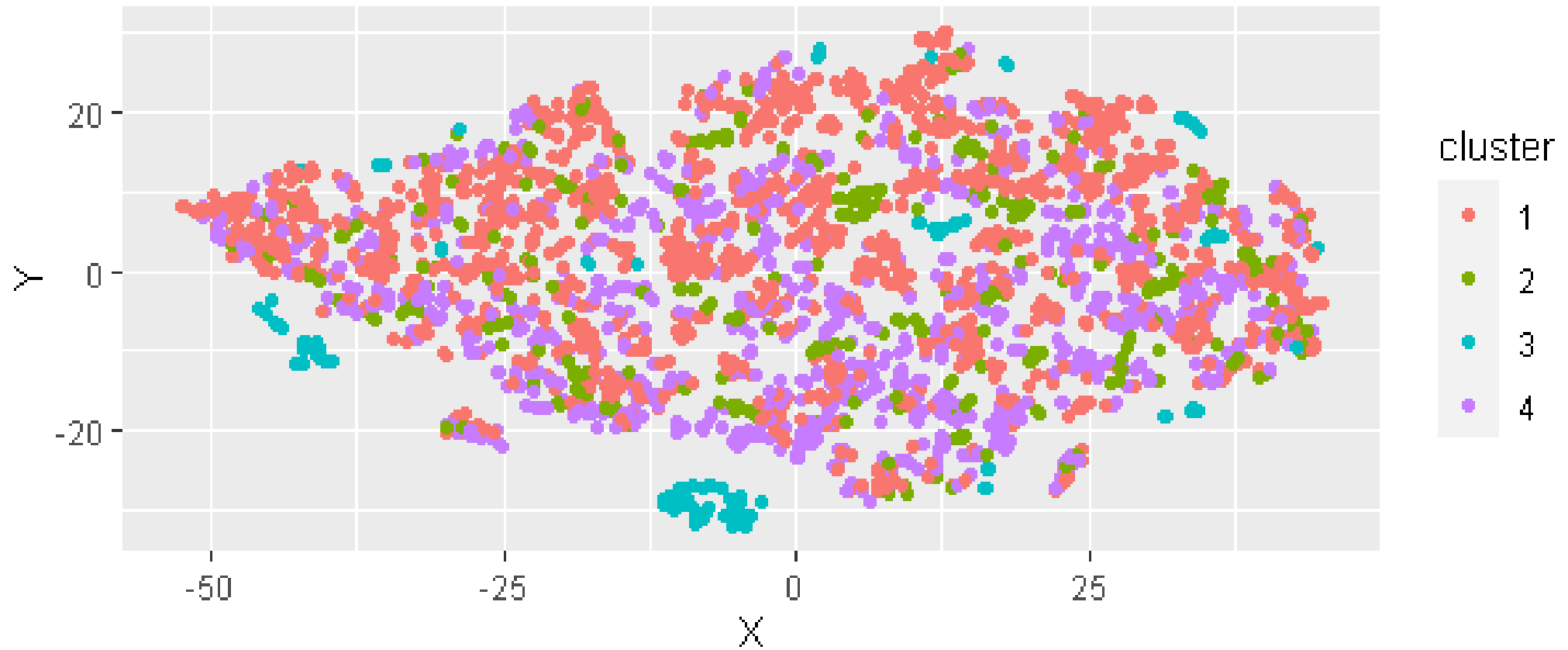
■ K-means



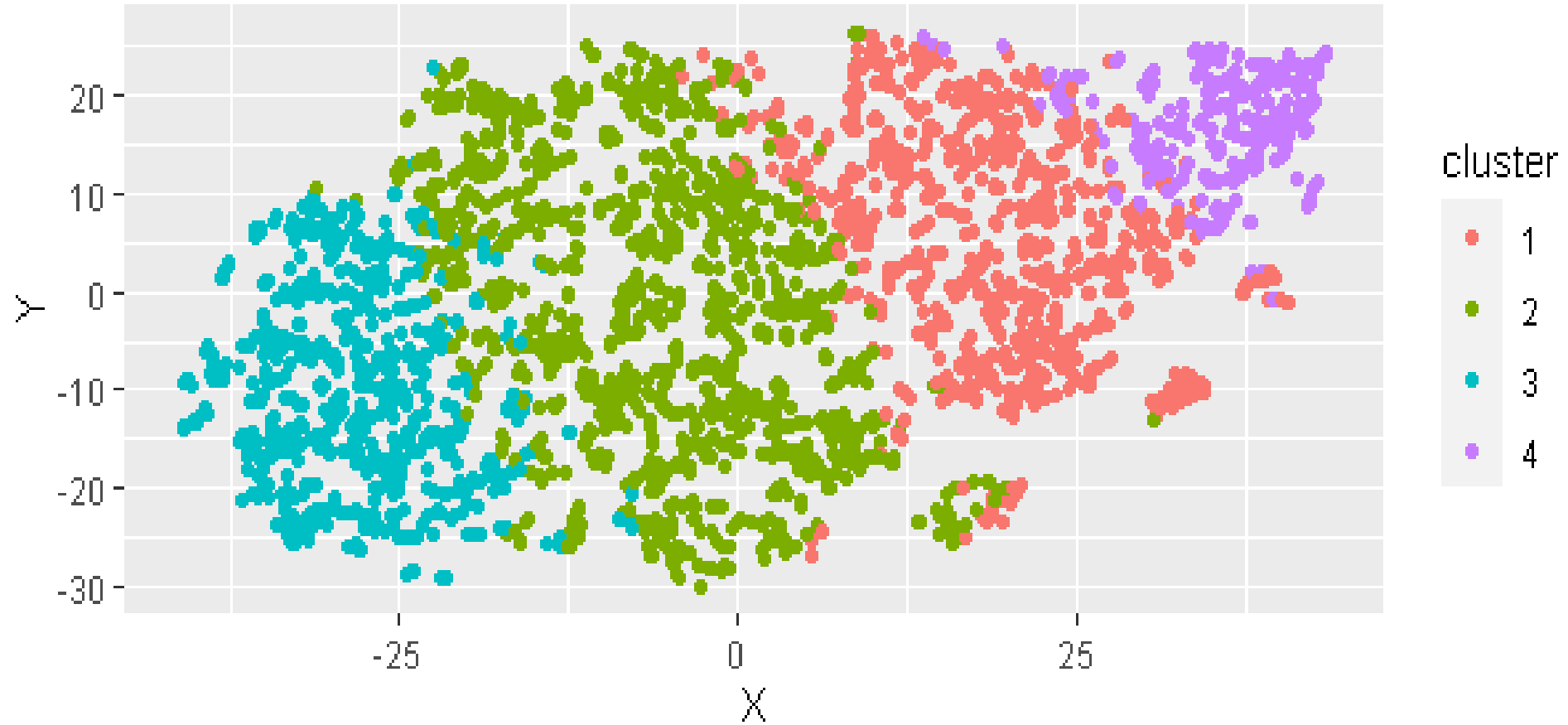
■ PAM



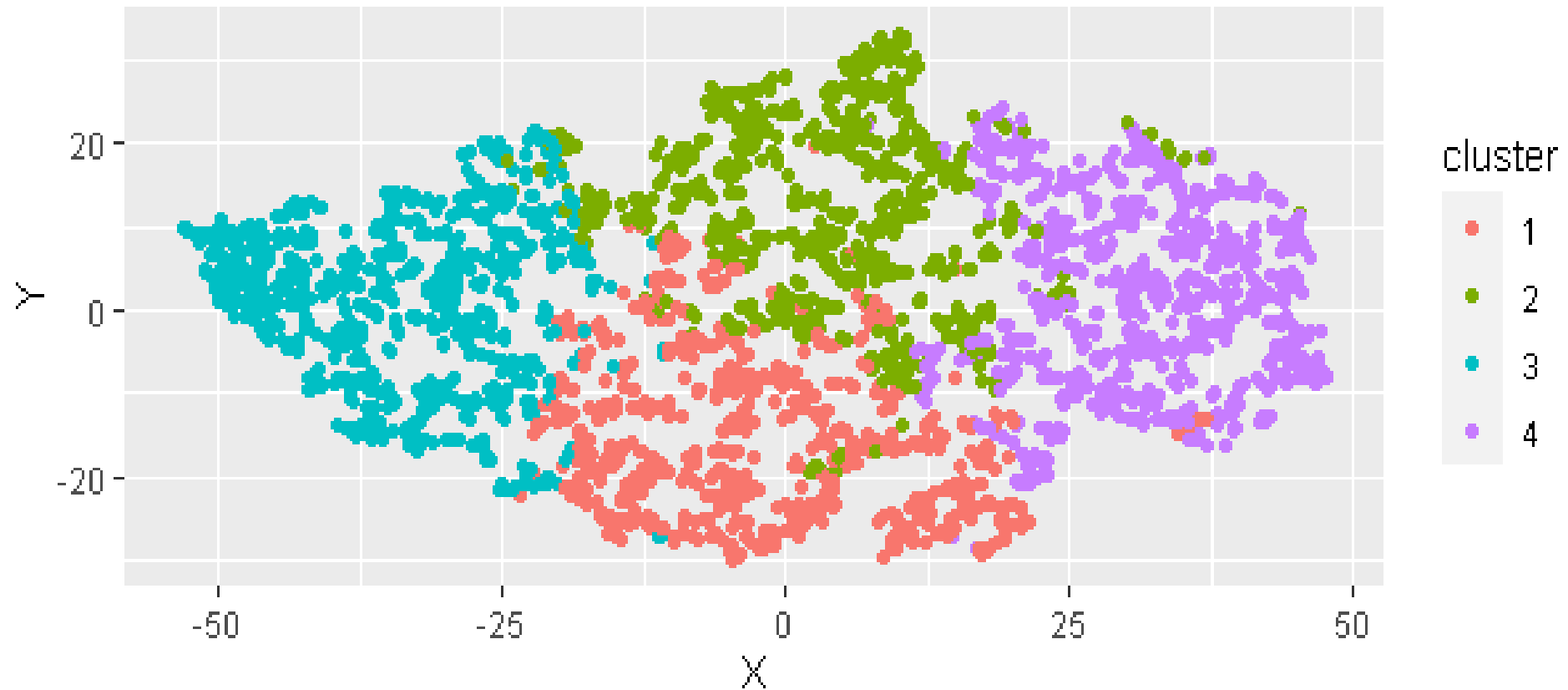
- HCPC (classification hiérarchique sur composantes principales)



- K-means (partitionnement sur les k-Moyennes)



- PAM (partitionnement autour des k-Médoïdes)



Indice	Score	Méthode
Connectivity	617.72	PAM
Dunn	0.0197	PAM
Silhouette	0.143	k-Means

- **L'indice de connectivité** d'une donnée dans l'échantillon est d'autant **plus élevé** que cette données appartient à un nombre important de voisinages des autres données.
- **L'indice de Dunn**: Le rapport entre la distance maximum qui sépare deux éléments classés ensemble et la distance minimum qui sépare deux éléments classés séparément. Plus cette distance est **grande** meilleure est la partition
- **L'indice de Silhouette**: Il travaille à l'échelle microscopique, il s'intéresse aux patients en particulier et non pas aux classes. Quand **il est proche de 1, le patient est bien classé**, Par contre, si il est proche de -1, cela veut dire que le patient est mal classé. Mais si il est proche de 0 alors il pourrait également être classé dans la classe la plus proche. Il est calculé à partir de la moyenne entre les indices de ses éléments.



Résultats (8)

Description des variables selon les groupes obtenus par la méthode PAM.

	PAM				p-value
	1 N=719 24.22%	2 N=711 23.95%	3 N=759 25.57%	4 N=779 26.24%	
Hommes, 1892 (63.74%)	527 (73.3)	372 (52.3)	474 (62.5)	519 (66.6)	<0.0001
Âge, 70.07 (10.89)	77.3 (9)	65.7 (10.2)	68.0 (9.5)	69.4 (10.9)	<0.0001
Décès, 335 (11.30%)	114 (15.9)	61 (8.5)	120 (15.8)	40 (5.13)	<0.0001
IMC (kg/m), 26.30 (5.84)	27.1 (5.8)	26.1 (5.7)	25.0 (6)	26.9 (5.5)	<0.0001
VEMS, 61.31 (20.08)	50.9 (9)	71.0 (9.5)	40.8(13.3)	82.0 (13.4)	0.0001
Tabagisme, N (%) :					
Jamais, 148 (5%)	51 (7.09)	26 (3.6)	38 (5)	33 (4.24)	<0.0001
Ancien Fumeur, 1671 (56.30%)	505 (70.2)	332 (46.7)	402 (53)	432 (55.5)	
Fumeur Actif, 1149 (38.70%)	163 (22.7)	353 (49.6)	319 (42)	314 (40.3)	
Toux, 1722 (58.02%)	406 (56.5)	434 (61)	472 (62.2)	410 (52.6)	0.008
Expectoration, 1251 (42.15%)	313 (43.5)	281 (39.5)	354 (46.6)	303 (38.9)	0.007
mMRC, N (%) :					
0-1 1295 (42.15%)	225 (31.3)	419 (58.9)	171 (22.5)	480 (61.6)	
≥2 1673 (56.36%)	494 (68.7)	292 (41.1)	588 (77.5)	299 (38.4)	<0.0001
Exacerbations, N (%) :					
0-1 2198 (74.05%)	543 (75.5)	548 (77.1)	469 (61.8)	638 (81.9)	
≥2 770 (25.95%)	176 (24.5)	163 (22.9)	290 (38.2)	141 (18.1)	<0.0001
VR/CPT, 137.17 (36.73)	133 (13.9)	135 (11)	177 (43.9)	103 (15)	0.0001



0.0001

Résultats (9)

Description des comorbidités selon les groupes obtenus par la méthode PAM.

Comorbidités	PAM				p-value
	1 N=719 24.22%	2 N=711 23.95%	3 N=759 25.57%	4 N=779 26.24%	
SAS, 490 (16.51%)	120(16.7)	120(16.9)	68 (8.9)	182 (23.4)	<0.0001
DDB, 145 (4.88%)	50 (6.9)	23 (3.23)	43 (5.6)	29 (3.7)	0.003
Cancer Poumon, 204 (6.87%)	57 (7.9)	53 (7.45)	43 (5.6)	51 (6.5)	0.323
Autre Cancer, 351 (11.82%)	89 (12.4)	80 (11.3)	74 (9.7)	108 (13.9)	0.083
Hypertension, 1090 (36.72%)	330 (45.9)	230 (32.3)	262 (34.5)	268 (34.4)	<0.0001
Cardiopathie Ischémique, 502 (17%)	163 (22.7)	97 (13.6)	110 (14.5)	132 (16.9)	<0.0001
Trouble du rythme, 346 (11.65%)	135 (18.8)	54 (7.6)	65 (8.5)	92 (11.8)	<0.0001
Insuffisance Cardiaque Gauche, 121 (4.07%)	50 (6.9)	16 (2.2)	39 (5.1)	16 (2.05)	<0.0001
Athérome, 170 (5.72%)	55 (7.6)	34 (4.8)	43 (5.6)	38 (4.8)	0.067
HTAP, 64 (2.15%)	23 (3.2)	3 (0.4)	31 (4)	7 (0.9)	<0.0001
Artériopathie, 190 (6.40%)	68 (9.4)	38 (5.3)	54 (7.1)	30 (3.8)	<0.0001
Diabète, 367 (12.36%)	117 (16.3)	73 (10.3)	86 (11.3)	91 (11.7)	0.003
Dyslipidémie, 673 (22.67%)	196 (27.3)	133 (18.7)	133 (17.5)	211 (27.1)	<0.0001
Syndrome Métabolique, 105 (3.53%)	32 (4.45)	19 (2.6)	26 (3.4)	28 (3.6)	0.341
Dénutrition, 59 (2%)	15 (2.09)	8 (1.1)	33 (4.3)	3 (0.4)	<0.0001
Manifestations Sinusiennes, 160 (5.40%)	34 (4.73)	32 (4.5)	35 (4.6)	59 (7.5)	0.019
Problèmes Digestifs, 75 (2.52%)	15 (2.09)	14 (1.9)	30 (3.9)	16 (2)	0.038
Maladie Thromboembolique, 36 (1.21%)	7 (0.97)	12 (1.7)	9 (1.2)	8 (1)	0.591
Ostéoporose, 97 (3.26%)	31 (4.31)	11 (1.5)	38 (5)	17 (2.2)	<0.0001
Ostéodensitométrie, 97 (3.26%)	31 (4.31)	11 (1.5)	38 (5)	17 (2.2)	<0.0001
Anxiété, 430 (14.48%)	104 (14.5)	102 (14.3)	152 (20)	72 (9.2)	<0.0001
Dépression, 382 (12.87%)	60 (8.3)	105 (14.8)	120(15.8)	97 (12.5)	0.027

Discussion - Conclusions

■ Principaux résultats:

- Comparaison de **trois méthodes** d'analyse en clusters non-supervisée (**HCPC, k-means et PAM**)
- **PAM** : Meilleur candidat pour l'identification des phénotypes à l'échelle de PALOMB.
- **PAM (Phénotype 1)** : Sévère (VEMS), moins symptomatique, personnes âgées, comorbidités cardiovasculaires et diabète.
- **PAM (Phénotype 3)** : Très sévère (VEMS), symptomatique, IMC faible, comorbidités métaboliques, ostéo-articulaire et neuropsychiatrique.

■ La confrontation à la littérature:

- **Puissance satisfaisante** des résultats de la classification et **concordance globale** avec les résultats les plus importants de la littérature.
- **P-R. Burgel et al (ERJ, 2010)**:
 - Quatre phénotypes cliniques
 - Différences marquées d'âge, de symptômes, de comorbidités et de mortalité.
- **Validation** : Validation externe de nos résultats.

