Identification and analysis of clinical phenotypes in COPD patients: PALOMB Cohort

E. H. Ouaalaya[1], É. Berteaud[1], L. Falque[2], E. Monge[2], L. Nguyen[3], A. Ozier[3], J. M. Dupis[4], M. Sabatini[5], C. Nocent-Ejnaini[5], L. Petrov[5], A. Bernady[6], C. Roy[7], F. Le Guillou[8], M. Aliati[9], A. Prudhomme[10], M. L. Quinquenel[11], M. Staali[12], F. Pilard[12], E. Iglesias[12], M. Sapène[13], J. Casteigt[14], J. Moinard[15], Y. Daoudi[16], É. Blanchard[17], J. Macey[17], R. Veillon[17], X. Demant[17], C. Bon[17], L. Grassion[17], M. Molimard[18], C. Raherison-Semjen[1]

[1]Bordeaux University, INSERM, Bordeaux Population Health Research Center, team: EPICENE, UMR1219 - Bordeaux (France), [2]Cabinet Medical, Avenue Thiers - Bordeaux (France), [3]Polyclinique Saint-Augustin - Bordeaux (France), [4]Cabinet Medical Pessac - Pessac (France), [5]CHG Côte Basque - Bayonne (France), [6]Medical Center TOKI EDER - Cambo les bains (France), [7]CH Libourne - Libourne (France), [8]Cabinet Medical la Rochelle - La Rochelle (France), [9]Cabinet Medical, Rue Louis Beydts - Lormont (France), [10]Cabinet Medical, Route de Lourdes - Aressy (France), [11]Cabinet Medical, Avenue de la Brède - Léognan (France), [12]Rehabilitation Center, Avicenne Clinic - Libourne (France), [13]Bel Air Clinic, Avenue de la République - Bordeaux (France), [14]Cabinet Medical, Avenue Montaigne - Saint-Médard-en-Jalles (France), [15]Cabinet Medical, Rue de Rivière - Bordeaux (France), [16]Cabinet Medical, Rue Francis Poulenc - Saint-Médard-en-Jalles (France), [17]CHU de Bordeaux, Pole cardiothoracique, Respiratory Diseases Departement - Bordeaux (France), [18]Univ. Bordeaux, INSERM, Bordeaux Population Health Research Center, team Pharmaco-epidemiology, UMR 1219 - Bordeaux (France)

Background & Aim

- In recent years, several researchers have attempted to identify COPD phenotypes using different cluster analysis.
- This study aimed to determine the most optimal cluster analysis (supervised vs unsupervised) to robustly identify clinical phenotypes.

Methods

- 2,968 COPD patients have been included from January 2014 until February 2020.
- General information (age, BMI, smoking, comorbidities), lung function, exacerbations and symptoms were collected. After 5 years of follow-up, vital status was recorded.
- A hierarchical classification on the principal components (HCPC) was performed, followed by two unsupervised classification algorithms: k-means and PAM (Partition Around Medoids).
- Robustness was defined according to three different indices of validation (connectivity, Dunn and silhouette).

Results (1/3)

- The mean age was 70 years, 63.7% of males, current smokers: 38.7%, mean FEV_1: 61.3% predicted, ≥2 exacerbations: 43.6%, mMRC dyspnea grade≥2: 56.3%, chronic cough: 58%. The 5-year mortality rate was 11.3%.
- The computed indices of validation showed that PAM was robust as compared to HCPC and k-means.
- Based on our hypothesis, four phenotypes were described, using the PAM method.

Results (2/3)

1) Outcomes of the four groups of patients as per the 2007 GOLD recommendations.

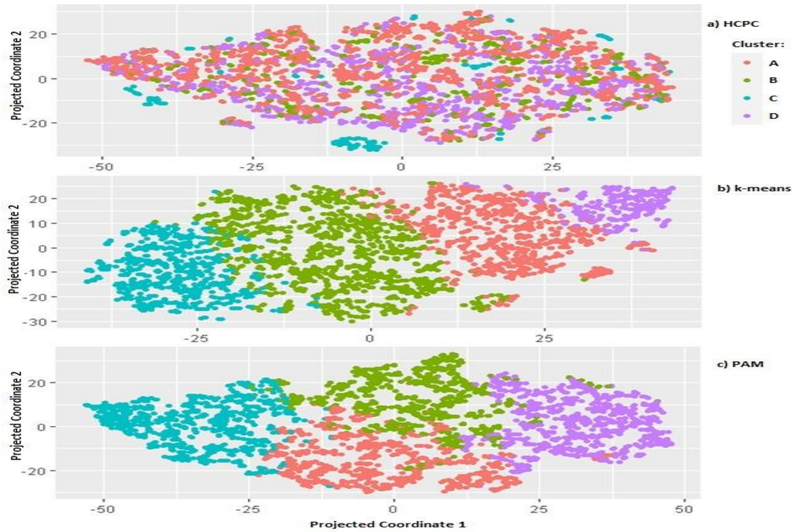
Distribution		GOLD 2007 stages				p-value
		I	II	III	IV	
2,968 COPD patients, n (%)		540 (18.2)	1,535 (51.7)	715 (24.1)	178 (6)	
Demographics,						
Age, yrs. mean (SD)	70 (10.9)	66.5 (10.7)	70.5 (11)	72.0 (10.5)	69.3 (9.6)	<0.001
Males, n (%)	1,892 (63.7)	307 (56.9)	981 (63.9)	480 (67.1)	124 (69.7)	0.001
BMI (kg/m^2), mean (SD)	26.3 (5.8)	25.6 (4.8)	27.1 (5.9)	25.9 (5.9)	23.2 (5.2)	<0.001
Smoking habits,						
Never smokers, n (%)	148 (5)	19 (3.5)	76 (4.9)	43 (6)	10 (5.6)	
Former smokers, n (%)	1,671 (56.3)	267 (49.4)	855 (55.7)	427 (59.7)	122 (68.5)	<0.001
Current smokers, n (%)	1,149 (38.7)	254 (47.0)	604 (39.3)	245 (34.3)	46 (25.8)	
Symptoms,						
Chronic cough, n (%)	1,722 (58)	289 (53.5)	889 (57.9)	432 (60.4)	112 (62.9)	0.047
Chronic sputum, n (%)	1,251 (42.1)	206 (38.1)	619 (40.3)	337 (47.1)	89 (50)	0.001
Functional Performance						
FEV_1, % pred, mean (SD)	61.3 (20)	90.7 (9.3)	64.7 (8.2)	41.1 (5.7)	24.4 (3.8)	0.0001
RV/TLC, % pred, mean (SD)	137.1 (36.7)	113 (54.4)	131 (22.4)	158 (25.2)	182 (28.8)	<0.001
mMRC Dyspnoea Scale,						
0-1, n (%)	1,295 (42.1)	368 (68.1)	747 (48.7)	162 (22.7)	18 (10.1)	
≥ 2, n (%)	1,673 (56.3)	172 (31.9)	788 (51.3)	553 (77.3)	160 (89.9)	<0.001
Exacerbations previous 12 months,						
0-1, n (%)	2,198 (74)	439 (81.3)	1186 (77.3)	473 (66.2)	100 (56.2)	
≥ 2, n (%)	770 (25.9)	101 (18.7)	349 (22.7)	242 (33.8)	78 (43.8)	<0.001
Revised 2017 GOLD ABCD criteria,						
A, n (%)	1,059 (35.6)	308 (57)	612 (39.9)	124 (17.3)	15 (8.4)	
B, n (%)	1,139 (38.3)	131 (24.3)	574 (37.4)	349 (48.8)	85 (47.8)	<0.001
C, n (%)	236 (7.9)	60 (11.1)	135 (8.8)	38 (5.3)	3 (1.6)	
D, n (%)	534 (18)	41 (7.5)	214 (13.9)	204 (28.5)	75 (42.1)	
Mortality,						
Mortality rate, n(%)	**335 (11.3)**	21 (3.9)	160 (10.4)	116 (16.2)	38 (21.3)	<0.001

2) Evaluation results of internal validation: Score of the best classification method compared to the internal validation choice index

index	Score	Method
Connectivity	617.72	PAM
Dunn	0.0197	PAM
Silhouette	0.143	K-means

Results (3/3)

3) Clustering membership of Hierarchical clustering (HCPC), k-means and PAM on 2-D projected coordinates.



- The phenotype A (24.2%) consisted of elderly patients with severe airflow limitation, low symptoms, cardiovascular comorbidities, diabetes and a higher mortality.
- The phenotype B (23.9%) contained more female patients, young patients with moderate airflow limitation and a high rate of current smokers.
- The phenotype C (25.5%) contained patients with very severe airflow limitation, more symptoms and low BMI.
- The Phenotype D (26.2%) was composed of patients with mild airflow limitation and low dyspnoea

Conclusion

These results showed the superiority of PAM classification compared with two other algorithms (k-means and HCPC) in terms of the robustness.

Fundings: Bordeaux University Foundation, Novartis Pharma, Isis Medical, Glaxo SmithKline, Boehringer Ingelheim and Chiesi.