

Discovering clinically relevant COPD patient subtypes in CALIBER

Maria Pikoula m.pikoula@ucl.ac.uk

UCL Institute of Health Informatics 12th November 2018

Big Data Science BAHIA 2018

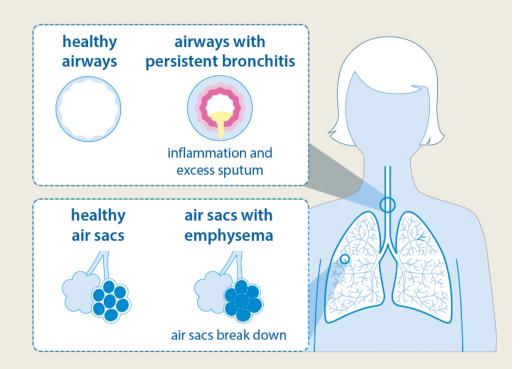
7th_



Chronic obstructive pulmonary disease (COPD)

COPD is a lung disease characterized by chronic obstruction of lung airflow that interferes with normal breathing and is not fully reversible.

- Bronchitis: **airways** are inflamed and narrowed.
- Emphysema affects the air sacs at the end of the airways in the lungs.





COPD burden

- Not simply a "smoker's cough" but an underdiagnosed, life-threatening lung disease.
- Prevalence: 251 million cases globally in 2016.*
- More than 90% of COPD deaths occur in low and middle-income countries.

*Global Burden of Disease Study 2016



Severity

GOLD grade (spirometry)

Severity	FEV1 % predicted
Mild (GOLD 1)	>= 80
Moderate (GOLD 2)	50 - 79
Severe (GOLD 3)	30 - 49
Very severe (GOLD 4)	< 30

MRC shortness of breath scale

Grade	Activity affected				
1	Only strenuous activity				
2	Vigorous walking				
3	With normal walking				
4	After a few minutes of walking				
5	With changing clothing				

FEV₁ (Liters):

Volume that has been exhaled at the end of the first second of forced expiration

FEV₁ % predicted:

 FEV_1 of the patient divided by the average FEV_1 in the population for any person of similar age, sex and body composition.



Acute Exacerbations of COPD

• Acute exacerbations are a major driver of the disease

• The factors that govern AECOPD and disease progression are not well-understood



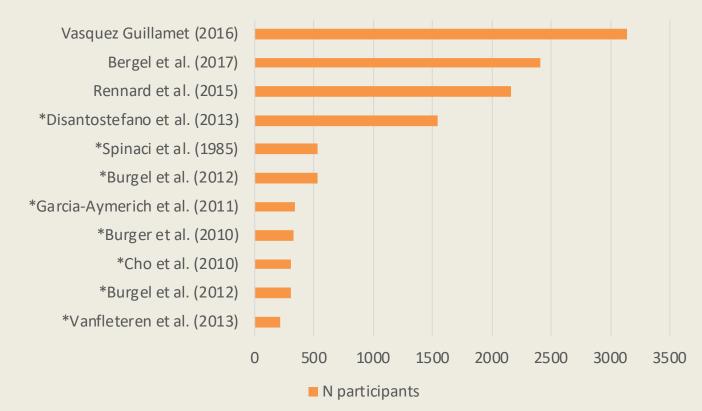
Study aims

- Use electronic health records data to discover new subtypes of COPD with in a hypothesis – free analysis.
- II. Evaluate subtypes with regards to clinically relevant outcomes such as AECOPD



Cluster analysis in COPD

Literature review 1985 - 2017

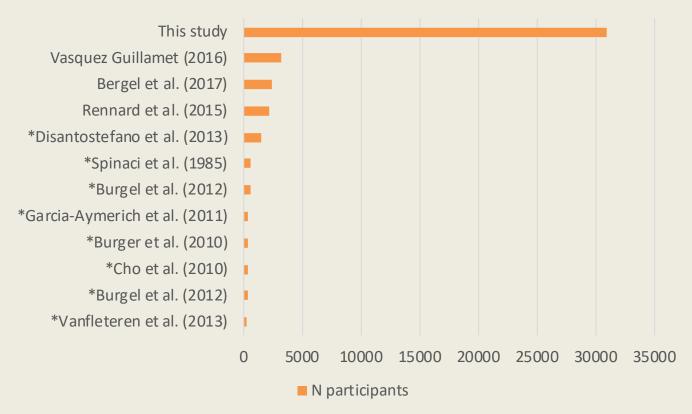


* Pinto et al. Respiratory Research 2015



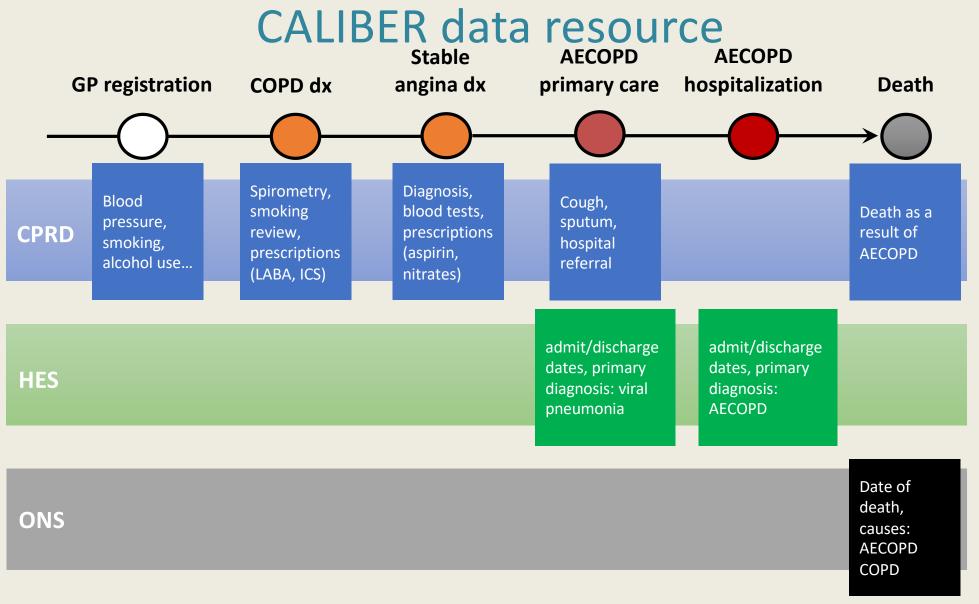
Cluster analysis in COPD

Literature review 1985 - 2017



* Pinto et al. Respiratory Research 2015



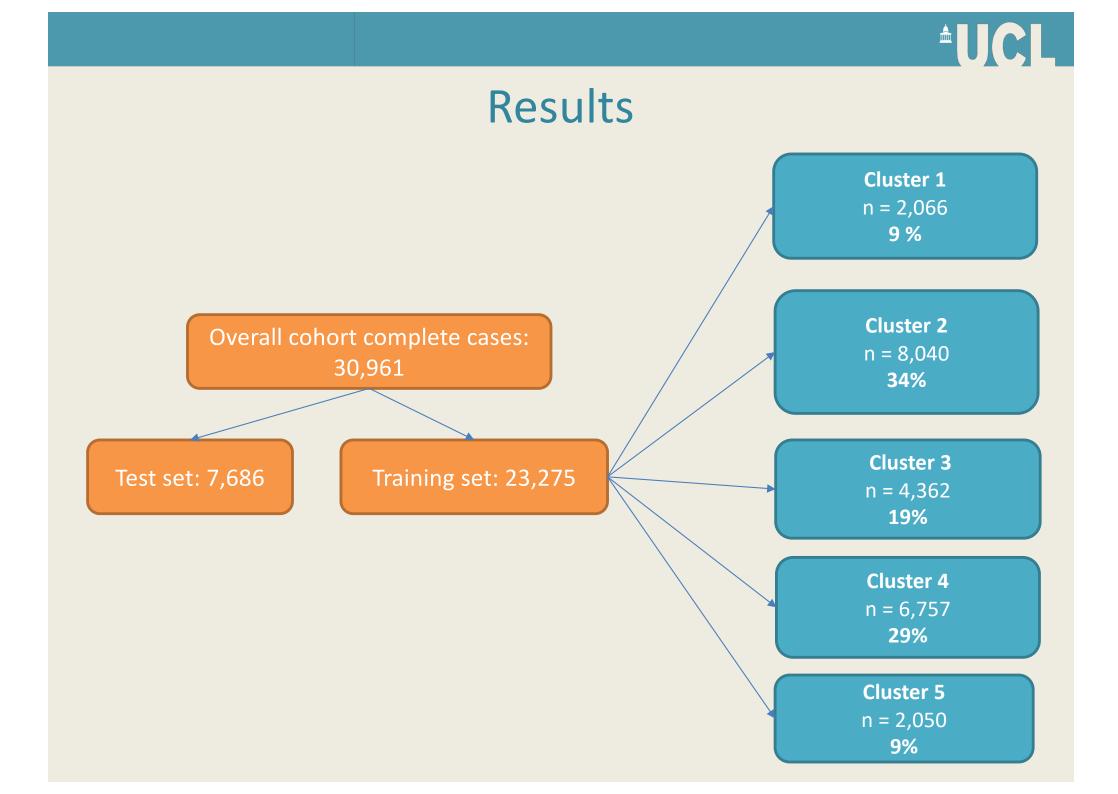


Denaxas et al. IJE 2012



Methods

- COPD phenotype using validated Read codes and smoking status
- 15 risk factors: Sex, BMI, GOLD grade, smoking status, anxiety, depression, atopy, chronic rhinosinusitis, hypertension, heart failure, ischemic heart disease, diabetes, gastroesophageal reflux disease, therapy regimen
- K-means clustering algorithm on complete cases, find optimal solution
- Label clusters based on defining characteristics



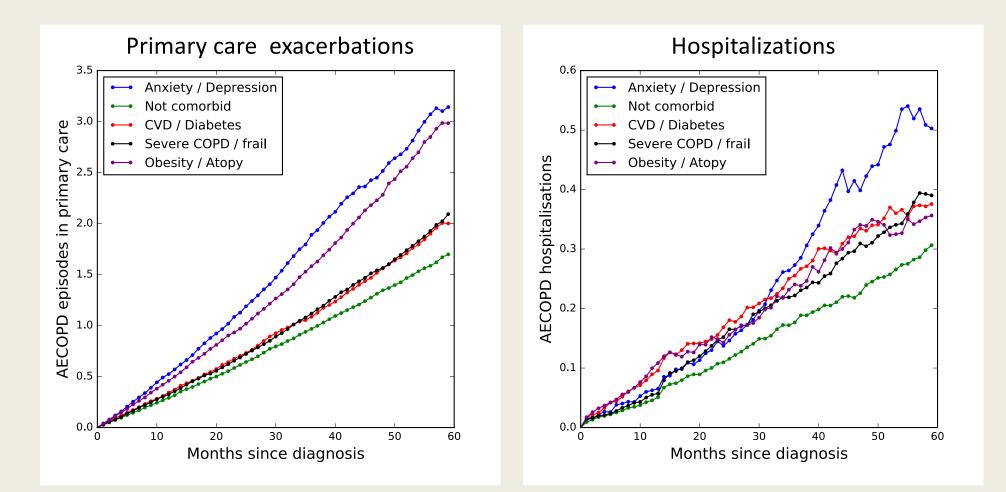


Results – Cluster characteristics

	Overall cohort	1: Anxiety/ Depression	2: Not comorbid	3: CVD / Diabetes	4: Severe COPD / frail	5: Obesity / Atopy
N	30,961	2,066	8,040	4,362	6,757	2,050
Male patients %	55	18	68	81	37	43
BMI % Underweight	4	9	0	0	10	0
BMI % Obese	30	18	32	54	11	53
Depression %	11	66	0	2	3	22
Atopy %	12	15	11	14	9	22
Heart failure %	15	5	10	46	2	24
GOLD % 1 (least severe)	26	35	24	22	27	29
% 4 (most severe)	3	3	2	2	6	1
High Eosinophils %	66	50	73	76	54	66



Clinical Evaluation: AECOPD





Clinical Evaluation: Respiratory /CVD mortality

Age-adjusted Cox regression

Characteristic	Hazard ratio		
Age	1.08 [1.07 – 1.08]		
Cluster			
Not comorbid	1		
Anxiety / Depression	1.28 [1.13 – 1.46]		
CVD / Diabetes	1.49 [1.38 – 1.60]		
Severe COPD / Frailty	1.30 [1.20 – 1.40]		
Atopy / Obesity	1.15 [1.03 – 1.30]		



Conclusions & Impact

- COPD patient subtypes can be identified using routinely generated EHR from primary care.
- Previous findings on CVD and diabetes diabetes were reproduced, and the trend is similar for exacerbations
- Anxiety and depression are distinct comorbidities potentially driving disease progression in younger, female patients
- Atopic and potentially asthmatic patients form a distinct cluster with overall better prognosis



Future directions

- Longitudinal evolution of COPD subtypes
- Models that allow patient membership of more than one subtype
- Genetic associations (UK Biobank)
- The role of comorbid respiratory conditions: Asthma and bronchiectasis





Spiros Denaxas



Natalie Fitzpatrick









Maxine **Mackintosh**

DRUK

Health Data Research UK



Arturo Gonzalez-Izquierdo



Kenan Direk



Ghazaleh Fatemifar





Michalis Katsoulis



Vaclav Papez





Jennifer K Quint

Marcos Barreto

http://denaxaslab.org/