Hospitalizations due to exacerbations of COPD: A big data perspective

Xavier Serra-Picamal[a], Ramon Roman[a], Joan Escarrabill[b,c,d], Anna García-Altés[a], Josep Maria Argimón[e], Nestor Soler[b,f,g], Rosa Faner[f,g], Elisenda Martinez Carbonell[a], Antoni Trilla[a,b,f], Alvar Agusti[b,f,g,*]

[a] Agencia de Qualitat i Avaluació Sanitaria de Catalunya (AQuAS), Spain
[b] Hospital Clinic, Univ. Barcelona, Spain
[c] Master Plan for Respiratory Diseases (PDMAR), Departament de Salut, Govern de la Generalitat de Catalunya, Spain
[d] REDISSEC, Health Services Research on Chronic Patients Network, Instituto de Salud Carlos III, Spain
[e] Departament de Salut, Govern de la Generalitat de Catalunya, Spain
[f] Institut d'Investigacions Biomediques August Pi I Sunyer (IDIBAPS), Barcelona, Spain
[g] CIBER Enfermedades Respiratorias (CIBERES), Instituto de Salud Carlos III, Spain

ARTICLE INFO

ABSTRACT

Background: Patients with Chronic Obstructive Pulmonary Disease (COPD) may suffer episodes of exacerbation (ECOPD) that require hospitalization and worsen their health status, and prognosis. We hypothesized that a detailed interrogation of health-care "big data" databases can provide valuable information to better understand the risk factors and outcomes of these episodes.
Material and methods: We interrogated four databases of the Catalan health-care system (> 8,000,000 registries) to identify patients hospitalized because of ECOPD for the first time (index event) between 2010 and 2012. Analysis was carried forward since the index event until the end of 2014 or the death of the patient. The two years that preceded the index event were also investigated.
Results: We identified 17,555 patients, (≥ 50 years of age) hospitalized because of ECOPD (ICD9 v.9 codes at discharge) for the first time between 2010 and 2012. In this population we observed that: *(1)* 23% of patients die within a year after being discharged from their first ECOPD hospitalization; *(2)* in the remaining patients, all-cause mortality was related to the number of re-hospitalizations, particularly with early (< 30 days) readmissions; *(3)* despite this being a 'respiratory' cohort, prescription and dispensation of drugs for cardiovascular diseases was higher than for obstructive airway diseases; and, finally, *(4)* lower winter ambient temperatures are associated with hospital admissions for ECOPD particularly in early re-admitters.
Conclusions: Overall these results indicate under appreciation of the burden of COPD in patients hospitalized for the first time because ECOPD.

1. Introduction

Patients with Chronic Obstructive Pulmonary Disease (COPD) often suffer episodes of exacerbation of their disease (ECOPD) that impact negatively their health-status and prognosis [1]. Some of these episodes require hospitalization, and a proportion of them early readmission after hospital discharge [2,3]. The cost associated with ECOPD hospitalization events constitutes the major share of the total cost of COPD care [4]. A better understanding of the determinants and outcomes of ECOPD hospitalizations (and re-hospitalizations) has the potential to facilitate the identification of patients at risk and, eventually, prevent or reduce re-hospitalizations, hence lowering the economic toll that these episodes impose on the health-care system [1].

A vast amount of health-care related information (often referred to as "big-data") is currently digitized, stored in administrative databases and available for analysis that converts information into knowledge [5,6]. We hypothesized that a detailed interrogation of health-care ("big data") databases can provide valuable information to better understand the risk factors (age, gender, co-morbidities, drug treatments, weather conditions) and outcomes (mortality, readmissions) of patients requiring hospitalization because of ECOPD for the first time.

2. Material and methods

2.1. Sources of information

We interrogated four administrative databases in Catalunya, Spain: *(1)* the Hospital discharge database (*CMBD-HA*), which has information on ECOPD hospitalizations events since 2005; *(2)* the *Pharmaceutics Activity Registry*, with information on the number and type of drug dispensations until level 5 of the Anatomical Therapeutic Chemical (ATC) Classification System. For this analysis, we focused on drugs dispensed for obstructive airway diseases (R03) and the cardiovascular system (C) since these diseases often co-occur in the same patients and may be difficult to dissect clinically [4]; *(3)* the *Population Registry* (*RCA*), with information on the vital/mortality status of the population; and, *(4)* the *Servei Meteorològic de Catalunya* database, which collects weather information in Catalonia. In this analysis we focused on the metropolitan area of the city of Barcelona.

2.2. Patients

After analyzing more than 8,000,000 registries in these four databases we identified 17,555 COPD patients in CMBD-HA who: *(1)* were diagnosed of COPD according the ICD9 v9 diagnostic codes (Table S1); *(2)* were hospitalized for the first time because of ECOPD (index event)

between 2010 and 2012 (albeit they might have been hospitalized before (since 2005) for other causes); and, *(3)* were 50 years of age (or older) at the index event (Fig. S1).

2.3. Data analysis

Analysis was carried forward since the index event (first hospitalization because ECOPD) until the end of 2014 or the death of the patient, whatever occurred first (patients with unregistered death date were excluded from analysis). We also investigated registries from the two years that preceded the index event.

2.3.1. Patient stratification

Following clinical reasoning, we decided *a priori* to stratify the patients included in the analysis in four operational categories: *(1)* **Fragile COPD**: patients who died during the first 12 months after the index event; *(2)* **Non-readmitters**: patients who were not hospitalized again during follow-up and were alive during, at least, the first 12 months; *(3)* **Readmitters**: patients hospitalized because of ECOPD at least once more during follow-up, but never before 30 days after discharge, and were alive during, at least, the first 12 months after the index event; and, finally, *(4)* **Early readmitters**: patients who were hospitalized because of ECOPD at least once more during follow-up after the index event but, in at least one of these occasions, this

Table 1
Main characteristics of the entire cohort (all patients) and four subgroups studied. The main discriminant characteristic of each group is coloured to facilitate identification.

	All patients	Fragile COPD	Non-readmitters	Re-admitters	Early readmitters	p value
Number of patients (%)	17,555 (100%)	4,058 (23%)	8,021 (46%)	3,993 (23%)	1,483 (8%)	
AT INDEX HOSPITALIZATION						
Age, yrs.	75.99 ± 10.05	80.45 ± 9.19*	74.43 ± 10.24	75.15 ± 9.52	74.50 ± 9.08	<0.001
Males , %	70%	74%	66%	73%	79%	<0.001
Charlson Comorbidity Index (CCI)	5.05 ± 2.45	6.41 ± 2.54*	4.51 ± 2.27	4.83 ± 2.27	4.87 ± 2.24	<0.001
Myocardial infarction	7%	9%	6%	7%	7%	<0.001
Congestive Heart failure	21%	31%	16%	19%	19%	<0.001
COPD	50%	61%	41%	54%	60%	<0.001
Length of hospital stay, days	6.60 ± 6.19	7.86 ± 7.95*	6.10 ± 5.32	6.41 ± 5.93	6.32 ± 5.30	<0.001
% hospitalizations October – March	62%	60%	63%	61%	64%	
BEFORE INDEX HOSPITALIZATION						
Number of previous hospitalizations/patient (not due to ECOPD)	2.16 ± 2.60	2.91 ± 3.03*	1.78 ± 2.33	2.10 ± 2.49	2.24 ± 2.62	<0.001
Medications (MPR)						
for obstructive airway diseases (R03)	1.16 ± 1.20	1.13 ± 1.18*	0.95 ± 1.09*	1.41 ± 1.26*	1.69 ± 1.36*	<0.001
for the cardiovascular system (C)	2.24 ± 2.19	2.39 ± 2.14*	2.20 ± 2.18	2.24 ± 2.25	2.10 ± 2.18	<0.001
DURING FOLLOW-UP						
Number hospitalizations/patient						
Any cause (including ECOPD)	2.50 ± 3.00	1.35 ± 1.59*	1.58 ± 2.14*	3.97 ± 2.89*	6.62 ± 4.58*	<0.001
Due to ECOPD	0.81 ±1.64	0.36 ± 0.82*	0.00 ± 0.00*	1.66 ± 1.08*	4.06 ± 3.15*	<0.001
% hospitalizations October – March,						
Due to ECOPD	61%	58%	-	62%	61%	
Due to other causes than ECOPD	54%	54%	54%	55%	54%	
Mean Length of hospital stay, days						
Any cause	7.76 ± 6.88	9.43 ± 9.87*	7.53 ± 6.99	7.01 ± 4.70	7.71 ± 4.43	<0.001
Due to ECOPD	7.05 ± 5.79	8.89 ± 8.29*	-	6.48 ± 5.31*	7.42 ± 4.77*	<0.001
Due to other causes than ECOPD	8.06 ± 7.64	9.54±10.47*	7.53 ± 6.99	7.74 ± 6.22	8.34 ± 6.99	<0.001
Charlson index (at the end of 2014 or death) date)	7.20 ± 2.74	8.12 ± 2.53	6.65 ± 2.70	7.32 ± 2.77	7.35 ± 2.67	<0.001
Myocardial infarction	11%	12%	10%	11%	12%	<0.05
Congestive Heart failure	43%	51%	34%	48%	52%	<0.001

*p < .001 vs. all the other groups in pairwise *post-hoc t*-test.

occurred within the first 30 days that followed the previous discharge and were alive during, at least, the first 12 months after the index event.

2.3.2. Statistics

Variables included in the analysis are specified in the on-line supplement. Results are presented as n (proportion) or mean (± standard deviation) as appropriate. Patient groups were compared using ANOVA (for continuous variables, followed by post-hoc t tests if appropriate) or Chi squared (for categorical variables) or Test of Equal or Given Proportions for proportions. Kaplan-Meier analysis was used to assess survival after the index admission. To identify potential predictors at the index event of the four different groups defined *a priori* (see above), we used a logistic regression model following the methodology proposed by Yu *et al* [7], which is discussed in detail in the on-line supplement. Besides, we also explored if the use of several machine learning techniques [8] improves the predictability of outcomes in this cohort, as detailed also in the on-line supplement. A p value < .05 was considered statistically significant.

3. Results

3.1. Cohort characteristics

As shown in Table 1, at the index event the mean age of the population studied (n = 17,555) was 76 years, 70% of patients were males and their mean CCI was 5.05 ± 2.45. Before the index event (first hospitalization because of ECOPD), patients had already required 2.16 ± 2.6 hospitalizations/patient (by design, due to reasons other than ECOPD), although COPD had been previously identified as a co-morbidity in many of them (Table 1). MPR of drugs for obstructive airway diseases (R03) was 1.16 and that of drugs for the cardiovascular system (C) 2.24. Sixty two per cent of index hospitalization events occurred between October and March, both months included.

3.2. Stratification of patients at index event

According to the operational stratification explained in Methods, 23% of patients were classified as Fragile COPD patients, 46% as non-readmitters, 23% as readmitters and 8% as early readmitters (Fig. 1, panel A). To visualize the contrast between these four groups, in Table 1 we highlighted in yellow differences of potential clinical relevance. At index event, Fragile COPD patients were oldest, had the highest CCI and prevalence of cardiovascular co-morbidities (which, interestingly, were similar in the other three groups), required more prolonged hospitalization and had been hospitalized more often before (for reasons other than ECOPD). Also of interest was the observation that, like in other three groups, in Fragile COPD patients cardiovascular drugs were dispensed much more often than drugs for obstructive respiratory diseases. On the other hand, the group of non-readmitters included the highest prevalence of females, while the group of early readmitters showed a higher incidence of hospitalizations during follow-up, particularly because of ECOPD (Table 1). Finally, the proportion of winter hospitalizations was similar in all four groups during follow-up.
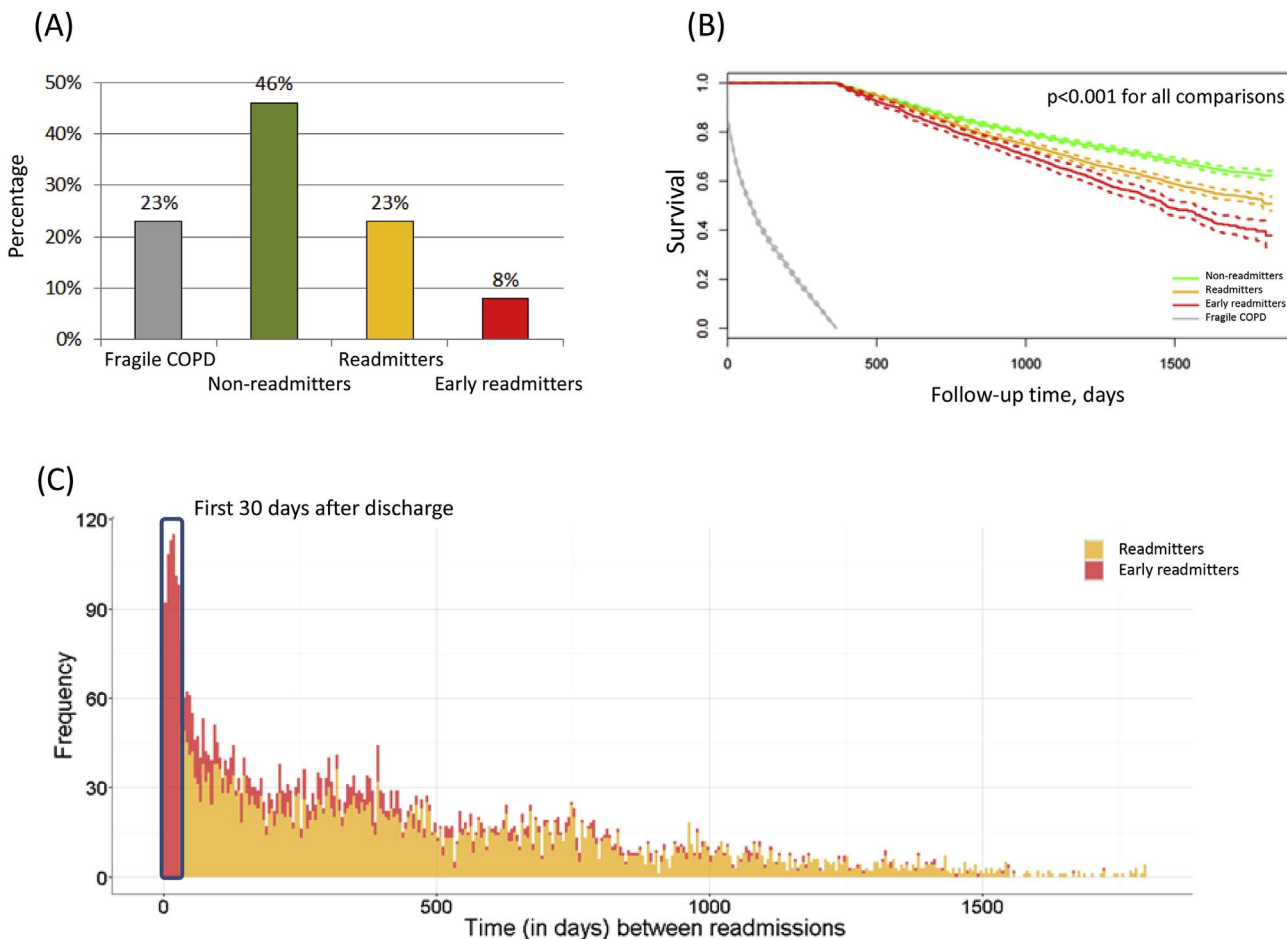

Fig. 1. *Panel A*: frequency distribution of the four groups studied. *Panel B*: Kaplan-Meier survival of the four groups studied. *Panel C*: Frequency distribution of the number of days elapsed between different ECOPD hospitalization events in readmitters and early-readmitters during follow-up. By definition, the former did not have any admission during the first 30 days that followed hospital discharge (box). For further explanations, see text.
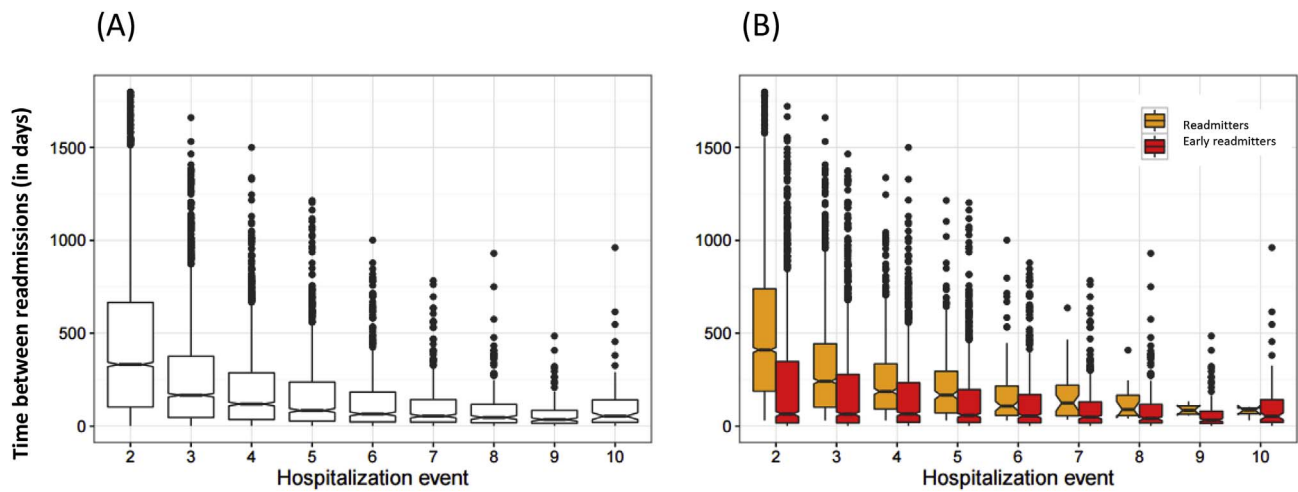
(A)

(B)

Fig. 2. Box plot of time elapsed between hospital readmissions as a function of the number of hospitalization events in the entire cohort (*Panel A*) and comparing readmitters and early readmitters (*panel B*). For further explanations, see text.

3.3. Observations during follow-up

3.3.1. Mortality

Survival of the entire cohort at the end of follow-up was 40% (Fig. S2). When it was analysed separately for each of the four groups considered above we found that (Fig. 1, panel B): *(1)* by design, Fragile COPD patients (23% of the cohort) had 100% mortality within a year after their first ECOPD hospitalization episode; and, that *(2)* in the other three groups, mortality rate was lowest in non readmitters, intermediate in readmitters and highest in early readmitters (p < .001 for all comparisons). In each of the four groups considered, a higher MPR quartile, this is a higher number of drugs dispensed before the index admission, was generally associated with poorer survival both for R03 (Fig. S3) and C (Fig. S4).

3.3.2. Hospitalizations

During follow-up (excluding the index event) patients suffered 2.50 ± 3.00 hospitalizations/patient (Table 1); the majority of them (67%; 1.69 ± 2.23 per patient) were not due to ECOPD. Mean length of stay (LOS) was slightly shorter in hospitalizations due to ECOPD (7.05 ± 5.79 vs. 8.06 ± 7.64 days). The CCI at the end of the study period increased (in survivors) from 5.05 ± 2.45 (at index event) to 7.20 ± 2.74. As shown in Fig. 2 (panel A), time between readmissions decreased progressively in proportion to the number of admissions.

When patients were stratified as above, by design non-readmitters patients did not have any other hospitalization event during follow-up whereas readmitters and early-readmitters patients did (Table 1). The most frequent cause of readmission in the latter group was ECOPD but this was not the case in readmitters (Table 1). Mean LOS of ECOPD related hospitalisations was lowest in readmitters (6.48 ± 5.31 days), followed by early readmitters (7.42 ± 4.77 days) and highest in Fragile COPD (8.89 ± 8.29 days) (p < .001). Fig. 1 (panel C) shows the frequency distribution of the number of days elapsed between ECOPD hospitalizations during follow-up in readmitters and early readmitters. By definition, no re-admitter required early re-hospitalization (< 30 days after discharge) but, as shown by the red peak in Fig. 1 (panel C), 627 of 1506 (41.6%) new hospitalization events in early readmitters occurred during this period. As shown in Fig. 2 (panel B), the time elapsed between hospitalizations was always significantly lower in early readmitters.

Fig. 3 presents the temporal (2010–2014) relationship between daily ambient temperature and hospital admissions in the metropolitan area of Barcelona. Both baseline and peak values were lowest in non re-admitters, highest in early readmitters and intermediate in readmitters (Fig. 3). As expected, temperature increased during the summer and decreased in winter. These temperature changes did not almost influence hospitalizations in Fragile COPD (who died within the first 12 months after discharge) but clearly influenced them in the other groups where hospital admissions for ECOPD generally (but not always) peaked in winter (see three peaks in the summer of 2011).
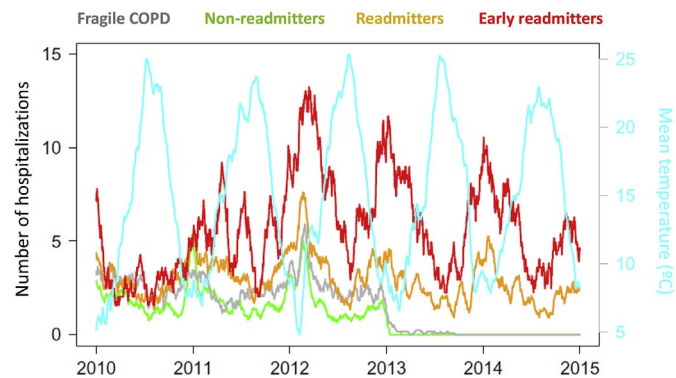
3.4. Predictors of longitudinal outcomes at index event

A logistic regression model that compared Fragile patients (n = 4.058) with the other three groups considered together (n = 13.497) identified the following significant risk factors male gender, older age, CCI, history of previous hospitalizations and prescription of cardiovascular drugs. The model has an AUC of 0.73 (Fig. 4, panel A) and a McFadden R^2 statistic of 0.11, that indicates sub-optimal adjustment of the model. With this caveat in mind, dominance analysis showed that CCI (49.5%), age (36.0%) and number of previous hospitalizations not due to ECOPD (9.8%) were the three strongest predictors of this group. Of note, the proportion of early re-admitters in Fragile patients (~10%) was similar to that of early re-admitters group. On the other hand, a similar analysis that compared Early Readmitters (n = 1.483) *vs.* Readmitters (n = 3.993), identified male gender and use of respiratory drugs two years before the index hospitalization event as significant predictors of early readmission, but the AUC of the model was 0.57 (Fig. 4, panel B) and the R^2 McFadden was 0.02, indicating poor predictive capacity for early readmitters. Finally, the
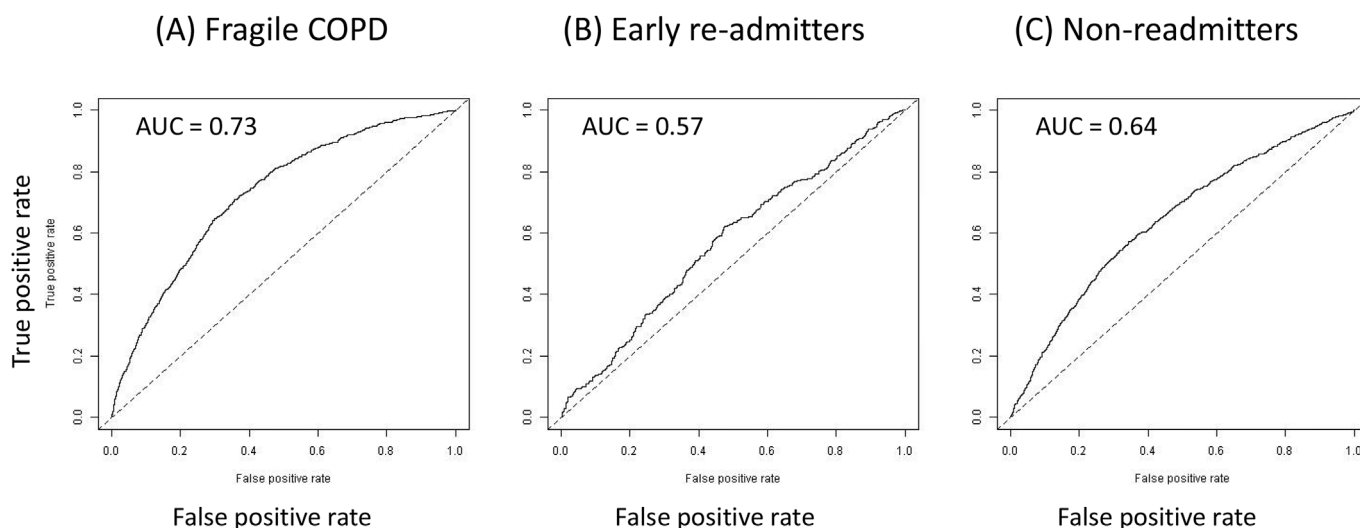

Fig. 3. Relationship between ambient temperature (30 days running average; blue line, right Y axis) and mean number of hospitalizations (30 days running average number of hospitalizations per group/number of patients per group *10.000, left Y axix) through the duration of the study (2010–2014) in the four groups studied in the area of Barcelona. For further explanations see text. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

(A) Fragile COPD   (B) Early re-admitters   (C) Non-readmitters

Fig. 4. Receiver Operating Curves (ROC) and Area under the Curve (AUC) for Fragile COPD patients (*Panel A*), Early Readmitters (*Panel B*) and Non-readmitters (*Panel C*). For further explanations, see text.

comparison of non-readmitters (n = 8.021) *vs.* all readmitters n = 3.993 + 1.483) identified gender, CCI, number of previous hospitalizations and use of respiratory and cardiovascular drugs at index event as significant predictors of readmission with an AUC of 0.64 (Fig. 4, panel C) but an R2 McFadden value of was 0.05, indicating also poor adjustment of the model.

3.5. Machine learning analysis

To investigate if different machine learning algorithms [8] can improve the predictability of the logistic regression model used above, as detailed in the on-line supplement, we compared the performance of nine different algorithms (Logistic regression, Zero R, Support Vector Machine (SVM), IBK, Bagging, J48, Random Forest, Naïve Bayes and Attribute Selection) in five different dataset configurations (Original, Balanced, Missing, Normalized and Standardized) [8]. Results (Table S2) showed that the original logistic regression model provides the highest AUC value (0.73) of all nine algorithms for all five different datasets configurations explored. These results, therefore, support the validity of the standard logistic regression model presented above.

4. Discussion

This study investigates a relevant health problem (hospitalizations for ECOPD) from a novel perspective ('big data analysis') and provides several observations of interest. The most salient ones are that: *(1)* a quarter of COPD patients hospitalized for the first time because of ECOPD die within a year of hospital discharge (Fragile COPD). These patients are mostly old males, with frequent co-morbidities (often of cardiovascular origin) and history of previous hospitalizations (because reasons other than ECOPD); *(2)* in the remaining patients, all-cause mortality during follow-up is tightly related to the number of re-hospitalizations, particularly in early (< 30 days) readmitters; *(3)* despite this being a 'respiratory' cohort, prescription and dispensation of drugs for cardiovascular diseases was higher than for obstructive airway diseases, suggesting under appreciation of COPD; and, finally, *(4)* the effect of low ambient temperatures on hospital admissions for ECOPD is greatest in early re-admitters.

4.1. Previous studies

Many previous studies have investigated risk factors for re-hospitalization and death in ECOPD patients [3,7,9–14]. Our study differs

from them in two important aspects. First, we studied patients at their first hospitalization for ECOPD. Garcia-Aymerich *et al* used a similar approach but in a much smaller cohort (n = 342) [15]. And, second, to our knowledge our study is the first to define *a priori* four different 'outcome trajectories' after hospital discharge. This stratification facilitated the emergence of a significant population of patients (23%) that died during the year that followed their *first* hospital admission because of ECOPD (Fragile COPD). If mortality would have been analysed globally in the entire cohort this group would have been missed and overall survival confounded and misinterpreted.

4.2. Interpretation of findings

Several observations of our study deserve comment. First, despite that this cohort only includes COPD patients hospitalized for the first time because of ECOPD, the number of previous hospitalizations (not due to ECOPD) was remarkable (2.16 ± 2.6 per patient). This is different from the observation that some COPD patients present repeated episodes of ECOPD (i.e., frequent exacerbators) [16] and probably indicates a high level of frailty of the cohort studied [17].

Second, that a significant proportion (23%) of patients hospitalized for the first time because of ECOPD die within a year (Fragile COPD) suggests that the presence of COPD may have been under-estimated (hence under-treated) by the attending physicians [18] and/or under-perceived (hence under-reported) by patients [19]. Supporting this interpretation is that, firstly, COPD was frequently recorded as 'co-morbidity' (but never as a primary diagnosis) in previous hospitalizations and, secondly, that drugs for obstructive airway diseases were less often prescribed in these patients than drugs for the cardiovascular system (Table 1). The proper identification of these Fragile COPD patients would be therefore important to optimize treatment and, hopefully, improve their outcome. Our results indicate that, in real-life, the profile of a "Fragile COPD patient" is that of an old male, with frequent co-morbidities (often of cardiovascular origin) and numerous previous hospitalizations for reasons other than COPD.

Third, older age, airflow limitation severity (not determined here), presence of co-morbidities and previous hospital admissions (for whatever reason, including ECOPD) are well established factors for hospital readmission in COPD patients [3,7,12–14]. In our study we went one step further and compared the characteristics and long-term outcomes of three groups of patients defined *a priori*. We observed that: *(1)* never readmitters included the highest prevalence of females (34%) whereas most early-readmitters were males (79%) (Table 1); *(2)* co-

morbidities and previous hospitalizations were similar across these three groups (Table 1); *(3)* the majority of hospitalizations during follow up were not due to ECOPD, except in early-readmitters (Table 1), in whom almost half of the total number of readmissions were "early" (< 30 days after discharge); *(4)* as shown in Fig. 2 the time elapsed between ECOPD readmissions during follow-up decreased in proportion to the number of hospitalization events and was always lower in early readmitters than in readmitters; *(5)* there was a clear relationship between re-hospitalizations and all-cause mortality, which is lowest in non-readmitters, intermediate in readmitters and largest in early readmitters.

Fourth, the proportion of early readmitters in our cohort (8%) was lower than that usually reported in other studies (around 20%). Several explanations can be conceived for this observation. On the one hand, we included in our analysis only those patients hospitalized for the first time because of ECOPD, whereas previous studies included all admissions. On the other, in our analysis one patient can be classified only in one group. Hence, Fragile patients (dead within a year from discharge) also included patients with early readmissions (259/4058 = 6.4%), who by design were not included among early readmitters. Finally, previous studies measured the number of early readmissions episodes whereas our analysis quantified the number of patients fulfilling being readmitted during the first 30 days after discharge.

Finally, our study confirms that hospitalizations due to ECOPD occur more frequently (albeit not exclusively) in the winter season (Fig. 3), probably supporting an infectious component [20]. Yet, our results also provide some novel observations in this regard: *(1)* all year long, the baseline and peak levels of hospitalizations were highest in early-readmitters, intermediate in readmitters and lowest in never-readmitters; and, *(2)* most winters are characterized by a single, large hospitalization peak (2010, 2012, 2013) but, sometimes (2011) this can be substituted by several smaller but repeated peaks over the spring, summer and autumn (Fig. 3), suggesting different aetiologies (e.g., allergies).

4.3. Strengths and limitations

The large sample size and long follow-up period, combined with the *a priori* stratification of patients for analysis, are clear strengths of our study. Further, the uses of novel machine learning algorithms validate the observations provided by more conventional statistical analysis. On the other hand, the lack of more granular information on several important clinical features of this population, such as smoking history and lung function, are clear limitations of our study. Likewise, given its large sample size, p values are often statistically significant, so the interpretation of differences between groups needs to be carefully balanced clinically.

4.4. Conclusions

This "big data" approximation to patients hospitalized for the first time because of ECOPD reveals a global under appreciation and under treatment of COPD, as shown by the existence of a significant proportion (23%) of the cohort studied that die within a year of their first hospitalization because of ECOPD, and by the fact that, despite this being a 'respiratory' cohort, prescription and dispensation of drugs for cardiovascular diseases was higher than for obstructive airway diseases. It also shows that repeated hospitalizations, particularly those occurring within the first 30 days after discharge, are associated with higher mortality risk, that some years the winter peak of hospitalizations because of ECOPD is substituted by other smaller peaks occurring from spring to autumn and that the effect of low ambient temperatures on hospital admissions for ECOPD is greatest in early re-admitters.

Contribution of authors

Conception and design: RR, JE, JMA, AA; Analysis and interpretation: all; drafting the manuscript for important intellectual content: all.

Conflicts of interest

No author has any direct conflict of interest with the content of this paper.

Appendix A. Supplementary data

References

[1] B.R. Celli, M. Decramer, J.A. Wedzicha, K.C. Wilson, A. Agusti, G.J. Criner, W. MacNee, B.J. Make, S.I. Rennard, R.A. Stockley, C. Vogelmeier, A. Anzueto, D.H. Au, P.J. Barnes, P.R. Burgel, P.M. Calverley, C. Casanova, E.M. Clini, C.B. Cooper, H.O. Coxson, D.J. Dusser, L.M. Fabbri, B. Fahy, G.T. Ferguson, A. Fisher, M.J. Fletcher, M. Hayot, J.R. Hurst, P.W. Jones, D.A. Mahler, F.o. Maltais, D.M. Mannino, F.J. Martinez, M. Miravitlles, P.M. Meek, A. Papi, K.F. Rabe, N. Roche, F.C. Sciurba, S. Sethi, N. Siafakas, D.D. Sin, J.B. Soriano, J.K. Stoller, D.P. Tashkin, T. Troosters, G.M. Verleden, J. Verschakelen, J. Vestbo, J.W. Walsh, G.R. Washko, R.A. Wise, E.F.M. Wouters, R.L. ZuWallack, An official American thoracic society/European respiratory society statement: research questions in COPD, Eur. Respir. J. 45 (4) (2015) 879–905.

[2] M. Steiner, Hospital admission and readmission for acute exacerbation of COPD. A tough nut to crack, Thorax 70 (12) (2015) 1108–1109.

[3] M. Guerrero, E. Crisafulli, A. Liapikou, A. Huerta, A. Gabarrus, A. Chetta, N. Soler, A. Torres, Readmission for acute exacerbation within 30 Days of discharge is associated with a subsequent progressive increase in mortality risk in COPD patients: a long-term observational study, PLoS One 11 (3) (2016) e0150737.

[4] Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease (GOLD), (2018) http://goldcopd.org , Accessed date: 16 November 2017.

[5] T.B. Murdoch, A.S. Detsky, The inevitable application of big data to health care, JAMA 309 (13) (2013) 1351–1352.

[6] C. Auffray, R. Balling, I. Barroso, L. Bencze, M. Benson, J. Bergeron, E. Bernal-Delgado, N. Blomberg, C. Bock, A. Conesa, S. Del Signore, C. Delogne, P. Devilee, A. Di Meglio, M. Eijkmans, P. Flicek, N. Graf, V. Grimm, H.J. Guchelaar, Y.K. Guo, I.G. Gut, A. Hanbury, S. Hanif, R.D. Hilgers, A. Honrado, D.R. Hose, J. Houwing-Duistermaat, T. Hubbard, S.H. Janacek, H. Karanikas, T. Kievits, M. Kohler, A. Kremer, J. Lanfear, T. Lengauer, E. Maes, T. Meert, W. Muller, D. Nickel, P. Oledzki, B. Pedersen, M. Petkovic, K. Pliakos, M. Rattray, I.M. JR, R. Schneider, T. Sengstag, X. Serra-Picamal, W. Spek, L.A. Vaas, O. van Batenburg, M. Vandelaer, P. Varnai, P. Villoslada, J.A. Vizcaino, J.P. Wubbe, G. Zanetti, Making sense of big data in health research: towards an EU action plan, Genome Med. 8 (1) (2016) 71.

[7] T.C. Yu, H. Zhou, K. Suh, S. Arcona, Assessing the importance of predictors in unplanned hospital readmissions for chronic obstructive pulmonary disease, Clin. Outcomes Res. 7 (2015) 37–51.

[8] A. Nachev, T. Reapy, Predictive models for post-operative life expectancy after thoracic surgery, Math. Software Eng. 1 (1) (2015) 1–5.

[9] A.S. Iyer, S.P. Bhatt, J.J. Garner, J.M. Wells, J.L. Trevor, N.M. Patel, d. Kirkpatrick, J.C. Williams, M.T. Dransfield, Depression is associated with readmission for acute exacerbation of chronic obstructive pulmonary disease, Ann. Am. Thorac. Soc. 13 (2) (2016) 197–203.

[10] P. Almagro, C. Hernandez, P. Martinez-Cambor, R. Tresserras, J. Escarrabill, Seasonality, ambient temperatures and hospitalizations for acute exacerbation of COPD: a population-based study in a metropolitan area, Int. J. Chron. Obstruct. Pulmon. Dis. 10 (2015) 899–908.

[11] E. Balcells, E. Gimeno-Santos, J. de Batlle, M.A. Ramon, E. Rodriguez, M. Benet, E. Farrero, A. Ferrer, S. Guerra, J. Ferrer, J. Sauleda, J.A. Barbera, A. Agusti, R. Rodriguez-Roisin, J. Gea, J.M. Anto, J. Garcia-Aymerich, P.-C.S. Group, Characterisation and prognosis of undiagnosed chronic obstructive pulmonary disease patients at their first hospitalisation, BMC Pulm. Med. 15 (2015) 4.

[12] H. Müllerova, D.J. Maselli, N. Locantore, J. Vestbo, J.R. Hurst, J.A. Wedzicha, P. Bakke, A. Agusti, A. Anzuetofor the ECLIPSE Investigators, Hospitalized exacerbations of COPD: risk factors and outcomes in the ECLIPSE cohort, Chest 147 (4) (2015) 999–1007.

[13] S. Hartl, J.L. Lopez-Campos, F. Pozo-Rodriguez, A. Castro-Acosta, M. Studnicka, B. Kaiser, C.M. Roberts, Risk of death and readmission of hospital-admitted COPD exacerbations: European COPD audit, Eur. Respir. J. 47 (1) (2016) 113–121.

[14] L.C. Hunter, R.J. Lee, I. Butcher, C.J. Weir, C.M. Fischbacher, D. McAllister, S.H. Wild, N. Hewitt, R.M. Hardie, Patient characteristics associated with risk of first hospital admission and readmission for acute exacerbation of chronic obstructive pulmonary disease (COPD) following primary care COPD diagnosis: a cohort study using linked electronic patient records, BMJ Open 6 (1) (2016) e009121.

[15] J. Garcia-Aymerich, F.P. Gomez, M. Benet, E. Farrero, X. Basagana, A. Gayete, C. Pare, X. Freixa, J. Ferrer, A. Ferrer, J. Roca, J.B. Galdiz, J. Sauleda, E. Monso, J. Gea, J.A. Barbera, A. Agusti, J.M. Anto, Identification and prospective validation of clinically relevant chronic obstructive pulmonary disease (COPD) subtypes, Thorax 66 (2011) 430–437.

[16] J.R. Hurst, J. Vestbo, A. Anzueto, N. Locantore, H. Mullerova, R. Tal-Singer, B. Miller, D.A. Lomas, A. Agusti, W. MacNee, P. Calverley, S. Rennard, E.F.M. Wouters, J.A. Wedzicha, Susceptibility to exacerbation in chronic obstructive pulmonary disease, N. Engl. J. Med. 363 (12) (2010) 1128–1138.

[17] S.K. Park, C.R. Richardson, R.G. Holleman, J.L. Larson, Frailty in people with COPD, using the national health and nutrition evaluation survey dataset (2003-2006), Heart Lung 42 (3) (2013) 163–170.

[18] M. Bednarek, J. Maciejewski, M. Wozniak, P. Kuca, J. Zielinski, Prevalence, severity and underdiagnosis of COPD in the primary care setting, Thorax 63 (5) (2008) 402–407.

[19] G. Scioscia, I. Blanco, F. Burgos, C. Gistau, E. Arismendi, A. Agusti, Different dyspnea perception in COPD patients with frequent and infrequent exacerbations, Thorax 72 (2017) 117–121.

[20] J.A. Wedzicha, R. Singh, A.J. Mackay, Acute COPD exacerbations, Clin. Chest Med. 35 (1) (2014) 157–163.