BMJ Open

Population-based analysis of patients with COPD in Catalonia: a cohort study with implications for clinical management

Emili Vela,[1] Ákos Tényi,[2,3] Isaac Cano,[2,3] David Monterde,[4] Montserrat Cleries,[1] Anna Garcia-Altes,[5] Carme Hernandez,[2,3] Joan Escarrabill,[2,6] Josep Roca[2,3]

Check for updates

[1]Area d'Atenció Sanitària, Servei Català de la Salut, Barcelona, Catalonia, Spain
[2]Hospital Clinic de Barcelona, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Universitat de Barcelona, Barcelona, Spain
[3]Center for Biomedical Network Research in Respiratory Diseases (CIBERES), Madrid, Spain
[4]Institut Català de la Salut, Serveis Centrals, Catalunya, Spain
[5]Agencia de Qualitat i Avaluació Sanitaries de Catalunya (AQuAS), Catalunya, Spain
[6]Master Plan for Respiratory Diseases (PDMAR), Ministry of Health (Catalonia) REDISSEC, Health Services Research on Chronic Patients Network, Instituto de Salud Carlos III, Barcelona, Spain

Correspondence to
Professor Josep Roca;
jroca@clinic.cat

ABSTRACT

Background Clinical management of patients with chronic obstructive pulmonary disease (COPD) shows potential for improvement provided that patients' heterogeneities are better understood. The study addresses the impact of comorbidities and its role in health risk assessment.

Objective To explore the potential of health registry information to enhance clinical risk assessment and stratification.

Design Fixed cohort study including all registered patients with COPD in Catalonia (Spain) (7.5 million citizens) at 31 December 2014 with 1-year (2015) follow-up.

Methods A total of 264 830 patients with COPD diagnosis, based on the International Classification of Diseases (Ninth Revision) coding, were assessed. Performance of multiple logistic regression models for the six main dependent variables of the study: mortality, hospitalisations (patients with one or more admissions; all cases and COPD-related), multiple hospitalisations (patients with at least two admissions; all causes and COPD-related) and users with high healthcare costs. Neither clinical nor forced spirometry data were available.

Results Multimorbidity, assessed with the adjusted morbidity grouper, was the covariate with the highest impact in the predictive models, which in turn showed high performance measured by the C-statistics: (1) mortality (0.83), (2 and 3) hospitalisations (all causes: 0.77; COPD-related: 0.81), (4 and 5) multiple hospitalisations (all causes: 0.80; COPD-related: 0.87) and (6) users with high healthcare costs (0.76). Fifteen per cent of individuals with highest healthcare costs to year ratio represented 59% of the overall costs of patients with COPD.

Conclusions The results stress the impact of assessing multimorbidity with the adjusted morbidity grouper on considered health indicators, which has implications for enhanced COPD staging and clinical management.

Trial registration number NCT02956395.

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is one of the major disorders included in the WHO programme addressing non-communicable diseases.[1] It is estimated that COPD will become the third leading cause of death by 2020.[2] Moreover, projections on COPD prevalence and costs over

Strengths and limitations of this study

► The main strength of the study is that it contributes to risk prediction of relevant clinical events in patients with chronic obstructive pulmonary disease (COPD).

► The study shows high potential to assess health risk factors at population level indicating the high impact of comorbidities. It can contribute to define innovative strategies aiming at reducing the healthcare impact of patients with COPD.

► Full potential of the approach should be proven by integrating registry information and electronic medical records.

► Lack of clinical information, spirometric data and history of tobacco smoking reduces the potential for standardised risk characterisation of patients with COPD.

the next 15 years indicate a rapidly escalating burden, mainly due to population ageing, on both health and social support systems.[3 4]

While acknowledging the progress made in terms of standard of care recommendations,[5] it is accepted that a better understanding of patients' heterogeneities constitutes a key challenge to further enhance both prevention and management of patients with COPD aiming at healthcare value generation.[6 7] Recent studies indicate a high impact of comorbidities on use of healthcare resources in patients with COPD prompting the need for assessing novel integrated care strategies with a patient-oriented approach.[8 9] It is well accepted that several prevalent chronic conditions often occur as clusters of comorbidities in patients with COPD[10–12] and potential explanatory mechanisms for the phenomenon have been proposed.[13 14]

The current study addresses comorbidities in patients with COPD based on the hypothesis that assessment of all patients with COPD living in a given geographical area, a population-based analysis of the patients with the disease, can provide valuable insights

to better understand COPD heterogeneity. Specifically, the research objective was to explore the potential of the health registry information contained in the Health Surveillance System of the region of Catalonia (Spain) (7.5 million inhabitants) to enhance health risk assessment and stratification in the clinical arena.[15 16] To this end, we analysed a total of 264 830 patients with COPD from all healthcare layers, registered in 31 December 2014 and followed up in 2015, to elaborate predictive models for six key health indicators: (1) mortality, (2 and 3) hospitalisations (all causes and COPD-related), (4 and 5) multiple hospitalisations (all causes and COPD-related) and (6) users with high healthcare costs. Ultimately, this study aimed to consider the weight of the different covariates included in the predictive models to provide useful information to enhance clinical management.

The study is a relevant component of the programme for collaborative management of complex chronic patients being deployed in the region during the period 2016–2020.[17] An ancillary aim of the research was to assess factors determining the economic impact of the patients with the disease, as well as to identify areas of action to increase healthcare efficiencies in the management of these patients.

METHODS
Population-based risk assessment: adjusted morbidity grouper
The Catalan Health Surveillance System (CHSS) includes updated registries of the region of Catalonia (Spain) (7.5 million inhabitants) from Primary Care, Hospital-related events (hospitalisations, emergency room consultations and specialised outpatient visits), Pharmacy, Mental Health, Socio-sanitary services and other items (home-based respiratory therapies, dialysis, outpatient rehabilitation and non-urgent healthcare transportation) since 2011.[18 19] It allows analyses on use of healthcare resources,

pharmacy consumption, prevalence of key disorders and population-based health risk assessment.[15 16] It is of note that although integration of CHSS registry data with electronic medical records is not yet in place, it constitutes the main goal of the PADRIS programme,[20] officially launched on January 2017.

The regional population-based health risk assessment tool, named GMA (Adjusted Morbidity Groups), is used to elaborate the health risk strata pyramid of the general population of Catalonia (figure 1, left triangle).[15 16] The GMA tool predicts individual patient risk, periodically updated on a 6-month basis, based on multimorbidity information gathered from CHSS registry data. The rationale behind the use of GMA, against alternative health risk assessment tools, is that it complies with four main recommended criteria,[15] that is, (1) a population health approach (uses the entire population of 7.5 million inhabitants of the region), (2) publicly owned without licensing constraints, (3) open source computational algorithms and (4) the adjusted morbidity grouper relies mostly on statistical criteria, as opposed to other tools that include expert-based coefficients, thus facilitating quick transferability to other territories. Detailed descriptions of the GMA, as well as its evaluation, have been reported elsewhere.[15 16] Methodological details of the GMA algorithm are described in the online supplementary figure 1S.

Study dataset and design
The current study design is a fixed cohort analysis of the entire population of 264 830 patients, alive and above 39 years of age, included in the CHSS on 31 December 2014 with the clinical diagnosis of COPD. New patients registered during 2015 were not included in the study. For the study purposes, the diagnosis of COPD was based on the International Classification of Diseases (Ninth Edition) coding[21] (online supplementary table 1S) declared by the patient's responsible physician, either a
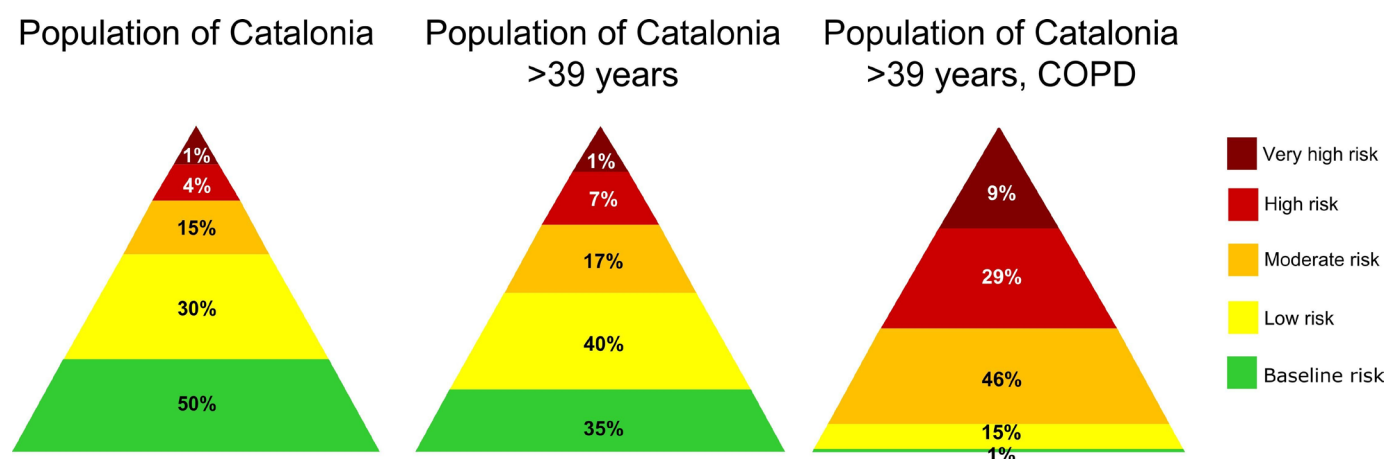

Figure 1 GMA health risk grades. The left triangle depicts the distribution of all Catalan citizens, expressed as percentage, in five arbitrary health risk layers defined using the GMA (Adjusted Morbidity Groups)[15 16] as a population-based health risk stratification tool. The central triangle indicates the distribution in these five health risk layers of the subset of Catalan citizens older than 39 years. The right triangle displays the distribution of the study group of patients with chronic obstructive pulmonary disease (COPD) (264 830 patients) across the GMA health risk grades: baseline, low, moderate, high and very high.

primary care professional or a specialist. The study did not take into account background clinical information, nor forced spirometry data.

The analyses were carried out for the entire population of patients with COPD, but the study considered different subsets of patients based on the GMA health risk grading: (1) baseline, (2) low, (3) moderate, (4) high and (5) very-high-risk patients. The thresholds defining these subsets of patients correspond to the percentiles 50, 80, 95 and 99 of the GMA grading for the general population (figure 1, left triangle). For the purposes of the study, patients with COPD falling into the baseline health risk group (1%) were merged with the low-risk group such that only four GMA grades were considered in the analysis (figure 1, right triangle).

Health risk predictive modelling was elaborated with registry data from year 2015 for the six main dependent variables: (1) mortality, (2 and 3) hospitalisation (all causes and COPD-related), (4 and 5) multiple hospitalisations (all causes and COPD-related) and (6) users with high healthcare costs, as described below. In the study, dependent variables were defined as binary variables. Mortality was defined as true if a patient died during the period from 1 January to 31 December 2015, regardless of whether it occurred in the hospital, at the patient's home or in other settings such as skilled nursing facilities. The category all causes hospitalisations was true if a patient had one or more hospital admissions due to any cause during 2015. COPD-related hospitalisations refer only to events triggered by acute exacerbations of COPD. Multiple hospitalisations were defined as two or more hospital admissions. Users with high healthcare costs refer to subjects above percentile 85 (PCT85) in terms of yearly healthcare costs during 2015.

The study used retrospective deidentified data from administrative databases (CHSS). Therefore, neither informed consent nor ethical committee approval was required according to the current legislation in Catalonia. The analyses were developed under the umbrella of the Nextcare project (http://www.nextcarecat.cat/) (https://clinicaltrials.gov/: NCT02956395).

Assessment of economic burden
Allocation of healthcare expenditure to each patient, including pharmacy, was done through the Personal Health Identification Number since it allows each billing invoice to be attributed to a given patient. The healthcare expenditure includes hospitalisation, primary care, pharmacy, health transport, respiratory home care therapies, outpatient visits and skilled nursing facilities. The key outcome variable of the analysis was 1-year healthcare resource use and expenditure by patient.[19] Calculation of individual healthcare costs was done by addition of costs for each item included in the CHSS registries alluded to above. This methodology allows calculation of the total healthcare expenditure in patients with COPD. Consequently, it allows to perform a holistic analysis of healthcare expenditure.

Statistical analysis
The study outcomes are described for the entire population of patients with COPD. Comparisons among the four subgroups defined by GMA health risk grades were done. The results are summarised in the main manuscript and complementary information is reported in detail in the online supplementary material. In this population-based analysis (table 1), age and number of chronic comorbidities are summarised as mean and SD, the proportion of women and the morbidities are expressed as percentages, while mortality rate, hospitalisation rate and COPD-related hospitalisation rate are expressed per 100 patients with COPD among patient groups. Comparisons among groups were done using analysis of variance for continuous variables, and χ^2 test for binary and nominal variables.

Statistical analyses were performed using SPSS software V.18.0. All statistical tests and confidence intervals were constructed with a type I error (alpha) level of 5%, and P values lower than 0.05 were considered statistically significant.

Predictive modelling
Multiple logistic regression analyses were used to generate health risk predictive models for the six main outcome variables of the study, namely, (1) mortality, (2 and 3) hospitalisation (all causes and COPD-related), (4 and 5) multiple hospitalisations (all causes and COPD-related) and (6) users with high healthcare costs. The following model independent variables were considered: age, sex, GMA health risk grades, previous history of hospital admissions, emergency room consultations and use of social support services. Each independent variable was included in the model as a categorical variable to allow for possible non-linearity in the relationship between variables and the relevant outcomes. For each measure, we collapsed the uppermost categories to ensure there were enough individuals in each cell to allow estimation of the parameters. All covariates were entered in the model one by one and retained when they showed a significant contribution to the predictive accuracy (P<0.10). The predictive role of each covariate into the model was assessed with a log-likelihood ratio test. To evaluate the performance of the resulting predictive models, we calculated the C-statistics (ie, the area under the receiver operating characteristic curve).[22 23]

RESULTS
Population-based analysis
The distribution of the entire population of the region in the health risk strata pyramid is depicted in figure 1, left triangle, wherein citizens are distributed in five health risk grades defined by the GMA percentiles, namely, (1) 50% of the population with baseline health risk (green), (2) 30% of individuals with low health risk (yellow), (3) 15% with moderate health risk (orange), (4) four percent with high risk and (5) 1% with very high health risk (brown).

Table 1 Main characteristics of the study group by GMA health risk grades

	Low risk	Moderate risk	High risk	Very high risk	Total
Patients (n)	41 809	121 918	76 237	24 866	264 830
Age (years)*	60.8±11.8	69.4±11.8	75.1±10.9	77.4±10.1	70.5±12.5
Women (%)*	36.0	36.2	35.9	35.3	36.0
Morbidity					
No of chronic comorbidities*	2.9±1.1	5.2±1.4	7.0±1.6	8.4±1.7	5.6±2.1
Diabetes (%)*	5.2	22.7	39.4	52.6	27.5
Heart failure (%)*	0.2	5.2	30.7	63.7	17.2
Hypertension (%)*	24.0	60.4	81.1	88.9	63.3
Renal failure (%)*	0.3	5.8	23.4	47.8	13.9
Dementia (%)*	0.3	2.6	7.8	13.6	4.7
Cirrhosis (%)	0.5	1.6	3.0	4.8	2.1
Depression (%)	8.8	20.1	28.6	34.2	22.1
Stroke (%)*	0.5	5.9	19.2	31.8	11.3
Ischaemic coronary disease (%)*	0.9	9.0	28.0	42.7	16.4
Malignancy (%)*	4.2	17.0	32.1	40.9	21.6
Locomotor system (%)*	29.3	56.3	67.9	72.0	56.8
Osteoporosis (%)*	3.9	10.0	14.3	17.9	11.0
Arthrosis (%)*	9.8	32.8	47.6	51.0	35.1
2015 events	n (%)	n (%)	n (%)	n (%)	n (%)
Mortality*	391 (0.9)	3213 (2.6)	6792 (8.9)	5955 (23.9)	16 351 (6.2)
Patients with hospitalisations (all causes)*	1785 (4.3)	12 451 (10.2)	19 417 (25.5)	12 496 (50.3)	46 149 (17.4)
Patients with hospitalisations (COPD related)*	404 (1.0)	2769 (2.3)	4697 (6.2)	3600 (14.5)	11 470 (4.3)
Patients with multiple hospitalisations (all causes)*	366 (0.9)	3026 (2.5)	6687 (8.8)	6035 (24. 0)	16 114 (6.1)
Patients with multiple hospitalisations (COPD related)*	71 (0.2)	524 (0.4)	1169 (1.5)	1210 (4.9)	2974 (1.1)
Users with high healthcare costs*	1119 (2.7)	10 064 (8.3)	17 333 (22.7)	11 609 (46.7)	40 125 (15.2)

Age and number of chronic comorbidities expressed as mean±SD; gender and morbidities are expressed as percentages; comparisons among risk grades were done using analysis of variance for continuous variables and χ^2 test for binary and nominal variables.
*P<0.01.
COPD, chronic obstructive pulmonary disease; GMA, Adjusted Morbidity Groups.

These risk strata show strong associations with mortality, hospital admissions, use of healthcare resources and expenditures, as reported in detail elsewhere.[15 16]

The central triangle (figure 1) indicates the distribution of the general population older than 39 years in the five health risk grades, described above. It constitutes the reference risk strata pyramid to be compared with the population-based COPD analysis. Finally, the distribution of the study group of patients with COPD is depicted in the right triangle that indicates the effects of the chronic pulmonary disease on health risk stratification. As expected, most of the patients with COPD (84%) were distributed between moderate (46%), high (29%) and very high (9%) health risk layers. Only 15% of the patients fell in the low risk level (yellow) and 1% were allocated in the baseline health risk level (green).

The study showed a COPD prevalence of 6.6% of all subjects above 39 years of age. The mean age was 70 (SD 12.5) years. Women represented 36% of the study group. On average, these patients presented 5.6 (SD 2.1) comorbid conditions. The mortality rate was 6.2%. The hospitalisation rate, all causes, was 17.4% (n=46 149), whereas the rate of COPD-related hospitalisations was 4.3% (n=11 470), which represents 24.9% of the total number of hospitalisations registered in these patients. As expected, the rate of comorbidities showed a consistent increase with GMA grading, which was the most apparent for cardiovascular disorders, type 2 diabetes mellitus–metabolic syndrome and/or anxiety–depression, as displayed in table 1.

Predictive modelling
The summary information of the predictive modelling for each of the six outcome variables—mortality, hospitalisations (all causes and COPD-related hospitalisations), multiple hospitalisations (all causes and COPD-related

Table 2 Summary description of the six predictive models

	Mortality	Hospitalisations All causes	Hospitalisations COPD related	Multiple hospitalisations All causes	Multiple hospitalisations COPD related	Users with high healthcare costs (PCT85)
C-statistics (AUC)	0.829	0.766	0.807	0.803	0.865	0.763
Covariates	OR					
Sex						
Male	1	1	1	1	1	1
Female	0.701	0.825	0.885	0.818	0.823	0.84
Age (years)						
40–54	1	1	1	1	1	1
55–64	1.693	1.254	1.435	1.308	1.428	0.884
65–74	2.103	1.506	1.675	1.576	1.7	0.91
75–84	3.749	2.054	2.017	1.983	1.634	0.898
>84	9.911	2.827	2.049	2.335	1.272	0.681
Admissions						
Group A	1	1	1	1	1	1
Group B	1.394	1.903	4.699	1.952	6.952	1.59
Group C	1.368	2.039	8.592	2.266	15.214	1.577
GMA grade						
Low	1	1	1	1	1	1
Moderate	1.826	1.896	1.644	2.128	1.649	3.125
High	4.316	3.998	2.646	5.419	2.852	8.858
Very high	8.919	7.851	3.548	11.042	3.891	20.285
Emergency room visits						
0	1	1	1	1	1	1
1–2	1.106	1.405	1.37	1.526	1.505	1.353
3–5	1.276	1.865	1.684	2.179	2.013	1.771
>5	1.643	2.615	2.235	3.354	3.032	2.428
Social support						
No	1	1	1	1	1	1
Yes	2.289	1.008	0.892	0.917	0.922	0.877

The intensities of grey background colour reflect the magnitude of ORs, being white when value is one and stronger grey when they are closer to 0 or have a higher positive value.

Admissions:

Group A corresponds to patients with no registries of hospital admissions within the period 2011–2014; Group B includes patients with history of admissions before 2014, but without admissions in that year; Group C includes patients with hospital admissions during 2014. Graphical representation and details of the six predictive models with the corresponding 95% CIs of the ORs are shown in the online supplementary figures 2S–7S and tables 2S–7S.

AUC, area under the receiver operating characteristic curve; COPD, chronic obstructive pulmonary disease; GMA, Adjusted Morbidity Groups.

repeated admissions) and users with high healthcare costs—are depicted in table 2, wherein significant covariates, the corresponding ORs and the C-statistics are indicated for each predictive model.

It is of note that age, closely followed by GMA grading, showed the two highest independent associations in the mortality model. Likewise, GMA grading depicted the highest predictive role in three out of the six models: hospitalisations (all causes), multiple hospitalisations (all causes) and users with high healthcare costs. Interestingly,

for COPD-related events (hospitalisations and multiple hospitalisations), the covariate with highest predictive role was history of hospitalisations, whereas the ORs of the covariate GMA grades were markedly lower than in the other predictive models.

Economic impact of patients with COPD

Figure 2 compares the costs of the different items, expressed as percentages, for the general population of Catalonia (outer circle) and those generated by the study
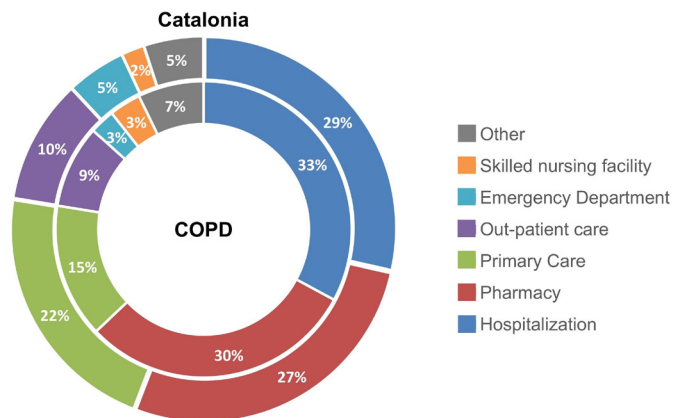
Figure 2 Indicates the distribution of costs of main items, expressed as percentages. The outer circle corresponds to overall cost for the Catalan Health System, whereas the inner circle indicates the corresponding relative costs ascribed to patients with chronic obstructive pulmonary disease (COPD). The absolute values are (1) Hospitalisation (€2291.8 million and €356.6 million, respectively), (2) Pharmacy (€2193.4 million and €325.8 million), (3) Primary care (€1745.0 million and €158.9 million), (4) Outpatient specialised care (€842.9 million and €98.1 million), (5) Emergency department (€401.5 million and €29.7 million), (6) Skilled nursing facility (Catalonia €155.1 million; COPD €37.5 million), and (7) Other (€404.0 million and €78.0 million). The last item, Others, includes home-based respiratory therapies, dialysis, outpatient rehabilitation and non-urgent healthcare transportation.

group of patients with COPD (inner circle). Briefly, COPD hospitalisations, pharmacy, skilled nursing care and other costs (respiratory therapies) are above the mean cost of the corresponding items in the general population of the region.

The average healthcare costs of patients with COPD per year was €4238 as compared with a mean of €987 a year per citizen in the region. It is of note that the cost generated by patients with COPD represents 13.5% of the overall healthcare costs in the region.
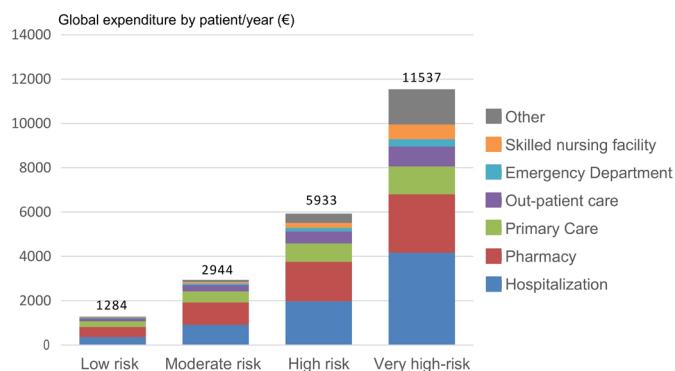

Figure 3 Average patient cost per year and relative contribution of the seven items (see text and figure 2 legend) for the four subgroups of patients with chronic obstructive pulmonary disease classified according to the GMA (Adjusted Morbidity Groups) scoring from low health risk (left column) to very high health risk (right column).

The cost analysis by GMA grading (figure 3) clearly showed a steady increase of costs per patient/year with considered GMA risk grades: (1) low, €1284; (2) moderate, €2944; (3) high, €5933; and (4) very high, €11537. Hospitalisation and pharmacy are the two items with highest impact on costs associated to GMA grading. It is worth to mention that we observed huge differences (7.41 times) between the users with high healthcare costs per year (PCT85: €16 131), which represents 59% of the overall costs of patients with COPD, and patients with COPD below PCT85 (€2177). Online supplementary table 8S indicates use of healthcare resources by GMA category.

DISCUSSION
Interpretation of the main findings
The main finding of the current study was the identification of comorbidities, expressed as GMA grades, as the covariate with highest discriminative impact on target events (mortality, hospitalisations) (table 2) and on yearly healthcare costs per patient (figure 3). These results can be relevant for commissioning of innovative healthcare services aiming at preventing materialisation of recent predictions on increases of healthcare impact of patients with COPD over the next 15 years.[3 4] However, they may also foster enhanced clinical management of individual cases with COPD.

It is of note that the performance of these models (table 2 and online supplementary figure S2–S7) shows quite acceptable goodness of fit (ie, C-statistics), indicating the potential of exploring synergies between population-health risk assessment (ie, GMA grading system) and clinical information to enhance health risk assessment and stratification in the clinical arena, as reported in Dueñas-Espín et al.[15] Ultimately, integration between registry data and electronic medical records of healthcare providers emerge as a high priority goal to properly pave the way towards personalised medicine for patients with chronic disorders.[24]

As expected, the covariate GMA showed a higher contribution in all causes hospital-related events than in only COPD-related admissions (table 2). It is of note that the latter represented only 22% of all admissions (table 1). While current GOLD recommendations address pulmonary events explaining frequent COPD exacerbations, the results of the current study suggest the need for further analyses aiming at identifying specific management needs for complex patients with COPD and comorbid conditions.

Two complementary healthcare strategies
The distribution of patients with COPD in the regional health risk stratification pyramid (figure 1, right triangle) allows identification of two different scenarios with well-defined associated challenges.

Patients with high health risk
The distribution of yearly healthcare costs per patient (figure 3) provides a strong rationale for targeting

patients close to the tip of the risk stratification pyramid as candidates for large-scale deployment of innovative services based on care coordination. However, two main limiting factors should be overcome in order to achieve proper designs of integrated care services.[25] First, poor comparability among interventions assessing effects on integrated care management for patients with COPD indicates an urgent need for standardisation of service workflows. Second is the limited healthcare impact of standard interventions addressed to patients at the tip of the pyramid,[26] that is, the low ratio between magnitude of the interventional effects and the resources devoted to achieve them, which may imply little healthcare value generation.[6 7] These two factors strengthen the need for further evidence on cost-effectiveness of well-defined integrated care interventions for complex chronic patients.[27]

Patients with low and moderate health risk

A better understanding of underlying mechanisms of comorbidity clustering in these patients emerges as a central need to effectively slow down patients' progress towards the tip of the pyramid. A natural consequence should be the development of efficient preventive interventions[28] aiming at delaying patients' worsening in terms of health risk scoring, which is recognised as a central unmet need for enhanced COPD management.

Strengths and limitations of the study

The uniqueness of the current study is that it was carried out using registry data that allow population-based analyses of all patients with a given condition(s) in the region. It is of note that complexities involved in implementation and optimisation of large-scale health information technology systems often impede health assessment of the entire population in a real-world setting.[29]

We acknowledge, however, that the use of registry information alone reflects underdiagnosis of COPD and constitutes a significant limitation that may explain a rather low figure for COPD prevalence in the region.[5 30] The lack of clinical information, spirometric data and history of tobacco smoking reduces the potential for a proper characterisation of patients with COPD. As indicated above, the barriers associated to the existence of health information silos further strengthens the need for speeding up the current efforts to achieve real integration between clinical and registry data[5] that will open new avenues to enhance both medical knowledge and clinical practice.

Episodes of inpatient and outpatient care carried out in private hospitals were not available for analysis because private providers do not use the Personal Health Identification Number. Nevertheless, the vast majority (approximately 97.5% in 2015) of COPD hospitalisations are done in public hospitals.

Clinical impact and perspectives generated by the study

The current study confirms that prevalent chronic conditions such as cardiovascular disorders, type 2 diabetes mellitus–metabolic syndrome and/or anxiety-depression often occur in high-risk patients with COPD.[10–12] We believe that the study provides a strong rationale for further research on subject-specific health risk prediction and stratification aiming at early identification of patients with low to moderate health risk who are prone to develop comorbid conditions in order to enhance preventive management in a cost-effective manner.[31 32] The current research may contribute to foster future developments of GOLD recommendations[5] addressing non-pulmonary manifestations of COPD that should have a positive impact on both staging and management of complex chronic patients.

Moreover, the study reinforces the ongoing strategies aiming at speeding up the evolution of the current health surveillance system in Catalonia towards a Digital Health Framework conceptually formulated in Cano et al[33 34] with potential to articulate three categories of data: (1) outcomes from population-based health-risk predictive models; (2) healthcare and biomedical research knowledge resulting from integration of clinical, physiological and biological/molecular information; and (3) informal care information from in-place personal health folders[15 17] encompassing information on lifestyle, adherence profile, socioeconomic status, social support and environmental factors. It is envisaged that inclusion of all these covariates influencing patient health should markedly increase the predictive accuracy and facilitate clinical decision-making based on sound estimates of the prognosis of an individual.

CONCLUSIONS

The current study provides a population-based analysis of 264 830 patients with COPD based on administrative health registries in Catalonia. The results illustrate the high impact of comorbidities on undesirable clinical events. We believe that the results highly encourage further developments fostering interoperability between health registries and electronic medical records to enhance clinical risk prediction.

REFERENCES

1. WHO. *Innovative care for chronic conditions: building blocks for action*. Geneva: World Health Organization (WHO/MNC/CCH/02.01), 2002. http://www.who.int/chp/knowledge/publications/icccglobalreport.pdf (accessed 9 Mar 2017).
2. Murray CJ, Lopez AD. Measuring the global burden of disease. *N Engl J Med* 2013;369:448–57.
3. Khakban A, Sin DD, FitzGerald JM, *et al*. The projected epidemic of COPD hospitalizations over the next 15 years: a population based perspective. *Am J Respir Crit Care Med* 2016:rccm.201606-1162PP.
4. McLean S, Hoogendoorn M, Hoogenveen RT, *et al*. Projecting the COPD population and costs in England and Scotland: 2011 to 2030. *Sci Rep* 2016;6:31893.
5. Vogelmeier CF, Criner GJ, Martinez FJ, *et al*. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease 2017 Report. GOLD Executive Summary. *Am J Respir Crit Care Med* 2017;195:557–82.
6. Porter ME. What is value in health care? *N Engl J Med* 2010;363:2477–81.
7. Porter ME, Larsson S, Lee TH. Standardizing patient outcomes measurement. *N Engl J Med* 2016;374:504–6.
8. van Boven JFM. Costly comorbidities of COPD: the ignored side of the coin? *Eur Respir J* 2017;50:1700917.
9. Chen W, FitzGerald JM, Sin DD, *et al*. Excess economic burden of comorbidities in COPD: a 15-year population-based study. *Eur Respir J* 2017;50:1700393.
10. Vanfleteren LE, Spruit MA, Groenen M, *et al*. Clusters of comorbidities based on validated objective measurements and systemic inflammation in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2013;187:728–35.
11. Divo MJ, Casanova C, Marin JM, *et al*. COPD comorbidities network. *Eur Respir J* 2015;46:640–50.
12. Vanfleteren LE, Spruit MA, Franssen FM. Tailoring the approach to multimorbidity in adults with respiratory disease: the NICE guideline. *Eur Respir J* 2017;49:1601696.
13. Barnes PJ. Mechanisms of development of multimorbidity in the elderly. *Eur Respir J* 2015;45:790–806.
14. Menche J, Sharma A, Kitsak M, *et al*. Disease networks. Uncovering disease–disease relationships through the incomplete interactome. *Science* 2015;347:1257601.
15. Dueñas-Espín I, Vela E, Pauws S, *et al*. Proposals for enhanced health risk assessment and stratification in an integrated care scenario. *BMJ Open* 2016;6:e010301.
16. Monterde D, Vela E, Clèries M. Los grupos de morbilidad ajustados: nuevo agrupador de morbilidad poblacional de utilidad en el ámbito de la atención primaria. *Atención Primaria* 2016;48:674–82.
17. Department of Health C. Catalonia health plan for 2016–2020. 2016. http://salutweb.gencat.cat/web/.content/home/el_departament/Pla_salut/pla_salut_2016_2020/Documents/Pla_salut_Catalunya_2016_2020.pdf
18. Farré N, Vela E, Clèries M, *et al*. Real world heart failure epidemiology and outcome: a population-based analysis of 88 195 patients. *PLoS One* 2017;12:e0172745.
19. Farré N, Vela E, Clèries M, *et al*. Medical resource use and expenditure in patients with chronic heart failure: a population-based analysis of 88 195 patients. *Eur J Heart Fail* 2016;18:1132–40.
20. Department of Health C. Programa públic d'analítica de dades per a la recerca i la innovació en salut (PADRIS). 2017. http://salutweb.gencat.cat/web/.content/home/ambits_tematics/linies_dactuacio/recerca/enllacos/Programa_analitica_dades_PADRIS_aquas2017_publica.pdf
21. Cherkin DC, Deyo RA, Volinn E, *et al*. Use of the International Classification of Diseases (ICD-9-CM) to identify hospitalizations for mechanical low back problems in administrative databases. *Spine* 1992;17:817–25.
22. Zou KH, O'Malley AJ, Mauri L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation* 2007;115:654–7.
23. Ware JH. The limitations of risk factors as prognostic tools. *N Engl J Med* 2006;355:2615–7.
24. Cano I, Tenyi A, Vela E, *et al*. Perspectives on big data applications of health information. *Curr Opin Syst Biol* 2017;3:36–42.
25. McWilliams JM, Schwartz AL. Focusing on high-cost patients—the key to addressing high costs? *N Engl J Med* 2017;376:807–9.
26. Lewis GH. 'Impactibility models': identifying the subgroup of high-risk patients most amenable to hospital-avoidance programs. *Milbank Q* 2010;88:240–55.
27. Cano I, Dueñas-Espín I, Hernandez C, *et al*. Protocol for regional implementation of community-based collaborative management of complex chronic patients. *NPJ Prim Care Respir Med* 2017;27:44.
28. Escarrabill J, Torrente E, Esquinas C, *et al*. Auditoría clínica de los pacientes que ingresan en el hospital por agudización de EPOC. Estudio MAG-1. *Arch Bronconeumol* 2015;51:483–9.
29. Cresswell KM, Bates DW, Sheikh A. Ten key considerations for the successful optimization of large-scale health information technology. *J Am Med Inform Assoc* 2017;24:182–7.
30. Miravitlles M, Soriano JB, García-Río F, *et al*. Prevalence of COPD in Spain: impact of undiagnosed COPD on quality of life and daily life activities. *Thorax* 2009;64:863–8.
31. Gomez-Cabrero D, Menche J, Vargas C, *et al*. From comorbidities of chronic obstructive pulmonary disease to identification of shared molecular mechanisms by data integration. *BMC Bioinformatics* 2016;17:23–35.
32. Roca J, Vargas C, Cano I, *et al*. Chronic obstructive pulmonary disease heterogeneity: challenges for health risk assessment, stratification and management. *J Transl Med* 2014;12 (Suppl 2):S3.
33. Cano I, Lluch-Ariet M, Gomez-Cabrero D, *et al*. Biomedical research in a digital health framework. *J Transl Med* 2014;12 (Suppl 2):S10.
34. Cano I, Alonso A, Hernandez C, *et al*. An adaptive case management system to support integrated care services: lessons learned from the NEXES project. *J Biomed Inform* 2015;55:11–22.