Risk and temporal order of disease diagnosis of comorbidities in patients with COPD: a population health perspective

Ákos Tényi,1,2 Emili Vela,3 Isaac Cano,1,2 Montserrat Cleries,3 David Monterde,4 David Gomez-Cabrero,5,6,7 Josep Roca1,2

ABSTRACT

Introduction Comorbidities in patients with chronic obstructive pulmonary disease (COPD) generate a major burden on healthcare. Identification of cost-effective strategies aiming at preventing and enhancing management of comorbid conditions in patients with COPD requires deeper knowledge on epidemiological patterns and on shared biological pathways explaining co-occurrence of diseases.

Methods The study assesses the co-occurrence of several chronic conditions in patients with COPD using two different datasets: Catalan Healthcare Surveillance System (CHSS) (ES, 1.4 million registries) and Medicare (USA, 13 million registries). Temporal order of disease diagnosis was analysed in the CHSS dataset.

Results The results demonstrate higher prevalence of most of the diseases, as comorbid conditions, in elderly (>65) patients with COPD compared with non-COPD subjects, an effect observed in both CHSS and Medicare datasets. Analysis of temporal order of disease diagnosis showed that comorbid conditions in elderly patients with COPD tend to appear after the diagnosis of the obstructive disease, rather than before it.

Conclusion The results provide a population health perspective of the comorbidity challenge in patients with COPD, indicating the increased risk of developing comorbid conditions in these patients. The research reinforces the need for novel approaches in the prevention and management of comorbidities in patients with COPD to effectively reduce the overall burden of the disease on these patients.

INTRODUCTION

Projections on healthcare impact of chronic obstructive pulmonary disease (COPD) over the next 15 years indicate a rapidly escalating health and societal burden mainly due to population ageing and comorbidities.1 2 It is well-known that highly prevalent chronic conditions such as cardiovascular disorders, type 2 diabetes mellitus—metabolic syndrome and/or anxiety—depression often occur as comorbid conditions in patients with COPD.3

Whereas the current standards on COPD management4 acknowledge the adverse effects of comorbidities on COPD prognosis, they suggest that ‘presence of comorbidities should not alter COPD treatment, and comorbidities should be treated per usual standards regardless of the presence of COPD’. However, recent evidence prompts the need for novel approaches in the prevention and management of comorbidities in patients with COPD to effectively reduce the overall burden of the disease.5 6 Identification of such cost-effective strategies aiming at preventing and enhancing management of comorbid conditions in patients with COPD requires deeper knowledge on epidemiological patterns and shared biological pathways explaining co-occurrence of diseases.7 Recently, Gomez-Cabrero et al8 reported the higher risk of developing certain comorbidities in patients with COPD, as compared with patients without COPD. The study used a data-driven analysis of Medicare registries from 13 million hospitalised patients over 65 years. The authors also proposed underlying biological mechanisms that may explain the identified comorbidities. Another direction of comorbidity research aims to uncover temporal disease co-occurrence patterns, showing great potential to explain

Key messages

► Patients with COPD show higher risk for developing certain comorbidities than non-COPD patients, independent of the population (ES, USA) and the specificities of the healthcare system.

► In elderly patients, COPD diagnosis tend to precede the appearance of other comorbid conditions, rather than be diagnosed after them.

► A better knowledge on the underlying shared molecular mechanisms of comorbidities in patients with COPD emerges as major priority for the development of cost-effective strategies to prevent comorbidities.
the dynamics of disease co-occurrence and to highlight characteristic disease sequences potentially caused by underlying mechanisms and common risk factors. As an example, a recent study identified COPD as a central disease with rapid progression to many other conditions, stressing the importance of its early diagnosis.9

In order to gain deeper knowledge on epidemiological patterns explaining co-occurrence of diseases,7 the primary aim of the current study is to reinforce previous evidence on the higher risk of comorbidities in patients with COPD.8 To this end, we conducted a similar analysis to the recent work by Gomez-Cabrero et al9 on an independent dataset retrieved from the Catalan Healthcare Surveillance System (CHSS) in Spain,10 which accounts for 1.4 million patients over 65 years with chronic conditions recruited across all healthcare tiers. The research also explored the temporal order of disease diagnosis of COPD and comorbidities at a population level which might help to further understand the dynamics of comorbidity clustering often seen in patients with COPD.3,11

METHODS
Dataset and study population
The study is based on registry data of over 7.6 million inhabitants retrieved from the CHSS, including: primary care consultations, hospital-related events (hospitalisations, emergency room consultations and specialised outpatient visits), pharmacy, mental health events, socio-sanitary services and other items, such as home-based respiratory therapies, dialysis, outpatient rehabilitation and non-urgent healthcare transportation.12 13

The study used a cross-sectional analysis, incorporating all patients over 65 years and registered in the CHSS who were active and alive during 2016 (n=1 433 376).

The research considered only chronic conditions of the patients, expressed with the Chronic Condition Indicator for International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) coding (Agency for Healthcare Research and Quality, USA).11 Diagnosis of COPD was based on the ICD-9 coding (online supplementary table S1) declared by the patient’s responsible physician, either a primary care professional or a specialist. The study did not take into account background clinical information nor forced spirometry data. The study identified patient with COPD (n=2 11 418) and their concomitant diseases, which were aggregated in 27 disease groups (DGs) (see online supplementary table S2) representing clinically significant traits.8 DGs with <1% prevalence in the study population were excluded from the study.

Heterogeneities between the Medicare dataset8 and the current study were identified (table 1) and were taken into account in the analysis of the results. Briefly, the Medicare dataset included registries from hospitalised patients over 65 years, considering both acute and chronic conditions, from 1990 to 1993. In contrast, the current study considered chronic diagnosis of patients over 65 years during 2016, obtained from a broader healthcare scenario, including all hospital visits since 2005 as well as diagnoses made in the primary care centres since the first visit of the patient.

Statistical analysis
For each DG, period prevalence was computed as the proportion of existing DG cases between 1 January 2016 and 31 December 2016 compared with the total population in the dataset. Age-associated prevalence was computed for each DG in patients with COPD and without COPD for 5-year age windows between ages 65 and 90 (eg, 75 denotes the prevalence between 73 and 77 years, both included).

Comorbidity association between COPD and DGs was measured using relative risk (RR) and phi correlation coefficient (Φ) (for detailed definition, see online supplementary methods).8 15 Significance of these measures were assessed at the stringent threshold p<0.0001, associated with a Bonferroni corrected p-value<0.01. Healthcare system-related differences in comorbidity associations were compared using two-sided t-tests of RR measures, where p<0.0001 was considered significant.

Temporal order of disease diagnosis
Diagnosis history of the patients included in the CHSS registries was used to study the temporal order of disease diagnosis (COPD ↔ disease). Date of diagnosis for a given DG was defined by the first diagnosis of any corresponding disease in the DG. Patient DG diagnoses, made before and after the COPD diagnosis, were counted to characterise their temporal order. Cases in which COPD was diagnosed simultaneously with other DGs for the first time in the same visit were disregarded from the temporal analysis (see online supplementary figure S1).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Description of the datasets and methodological considerations of the current study and the previous study of Gomez-Cabrero and colleagues</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study population</strong></td>
<td><strong>Study period</strong></td>
</tr>
<tr>
<td>Current study</td>
<td>1.4 million (CHSS)</td>
</tr>
<tr>
<td>Gomez-Cabrero et al</td>
<td>13 million (Medicare)</td>
</tr>
</tbody>
</table>

CHSS, Catalan Healthcare Surveillance System.
For significant comorbid conditions, the directionality of temporal order of disease diagnosis was tested. Preferred direction (the one that appears more often) was assigned to those COPD–DG pairs where significantly more patients were diagnosed with the DG before COPD or the other way around, using a binomial test for each direction with a probability of success equals to 0.5 and sample size $N_{DG} + N_{simultaneous} + N_{COPD}$, where $N_{simultaneous}$ indicates the number of patients with simultaneous first diagnosis of COPD and the DG.\(^9\) To estimate the strength of the directional associations, the causal information fraction (CIF) was used, which, in contrast to RR and $\Phi$, considers the order of occurrence of the disease pairs and emphasises possible causative effects.\(^16\) CIF is defined between a pair of diseases $i$ and $j$, as

$$f_{i \rightarrow j} = \frac{n_{i \rightarrow j}}{n_i} - \frac{n_j}{2N}$$

where $n_i$ and $n_j$ denote the prevalence of the diseases, $n_{i \rightarrow j}$ is the number of individuals diagnosed with disease $i$ followed by disease $j$ and $N$ is the population size. In the current analysis, CIF was used to compute COPD–DG associations.

### RESULTS

#### Comorbidity risk in patients with COPD in different healthcare systems

Nine DGs were discarded from the comparative analysis due to containing solely acute diseases (two DGs) and because of showing <1% prevalence (seven DGs) (online supplementary table S3).

The prevalence of 18 DGs included in the analysis is indicated in figure 1 by the size of the bars. Comparison of prevalence results within the CHSS dataset (red bars) indicates that patients with COPD (dark colour) have higher risk of developing most of the DGs compared with a general patient of the healthcare system (light colour). Identical prevalence patterns are observed in the Medicare dataset (cyan bars). These results are also consistent with the analysis of COPD comorbidity risk based on comorbidity measures (ie, RR, $\Phi$-correlation), indicating significant ($p<0.0001$) disease association between COPD and all the DGs in both healthcare systems (online supplementary table S3).

The comparative analysis between the two datasets (figure 1) shows significant differences in the RR for several DGs that are fully explainable by the heterogeneities in the data sources described in the section Methods. For example, acute diseases, not considered in the current analysis (CHSS), accounted for more than 90% of Medicare cases in endocrine disorders and skin alterations (online supplementary table S4) leading to the visible differences in prevalence.

Figure 2 compares the age-associated prevalence of heart diseases, circulatory disorders and digestive alterations\(^8\) between patients with COPD and without COPD in the two datasets. The figure indicates that in the two datasets the prevalence of the three DGs is consistently higher in patients with COPD (red lines), and that similar age-associated comorbidity patterns are observed. Interestingly, the prevalence of heart diseases for the two groups, COPD and non-COPD, is higher in the Medicare dataset than in the current study, representative of a Mediterranean population with mostly non-hospitalised patients. However, prevalence of heart diseases increases more steeply with age in the CHSS dataset. Similar age-associated prevalence of

---

Figure 1 Prevalence (x axis) of disease groups (DGs) (y axis) in the population of Medicare (light cyan) and Catalan Healthcare Surveillance System (CHSS) (light red), and in patients with chronic obstructive pulmonary disease (COPD) in Medicare (dark cyan) and in CHSS (dark red). The comparative analysis within datasets shows that the prevalence of most of the DGs is higher in patients with COPD (dark colour) than in the entire population (light colour). Differences in the prevalence between datasets are fully explainable by methodological heterogeneities, detailed in the main text. Healthcare system-related differences in comorbidity associations were compared using two-sided t-tests of relative risk measures (*), $p<0.0001$. 

---

©Copyright BMJ Publishing Group Ltd. All rights reserved. This article is freely available online with additional copyright instructions and conditions. Please see http://doi.org/10.1136/bmjresp-2018-000302 for full details.
the remaining DGs in the CHSS dataset is displayed in online supplementary figure S2.

**Temporal order of disease diagnosis**

The analysis of temporal order of disease diagnosis with respect to COPD is shown in figure 3A. Red bars indicate the patients in whom the first diagnosis of a disease from a given DG was done before the diagnosis of COPD. The number of patients in whom the corresponding DG was identified after the diagnosis of COPD are indicated by the cyan bars. It is of note, that for the majority of DGs, COPD was diagnosed first (G1). Interestingly, only heart diseases (G2) were more often diagnosed before COPD than after COPD. A third group of DGs (G3) showed no preferred direction. Figure 3B translates these interactions to a network representation, with directional edges based on the grouping and directional strengths of the association (ie, CIF measure).

**DISCUSSION**

The current research confirms in the CHSS dataset that patients with COPD are in higher risk of developing certain comorbidities than in patients without COPD, along with the results reported in Gomez-Cabrero et al.8 Despite marked methodological heterogeneities, similar age-related prevalence patterns were also observed between the current study and the Medicare dataset in elderly patients, as displayed in figure 2. These results provide a population health perspective of the comorbidity challenge in patients with COPD. It is of note that they are in line with a recent independent report carried out using a similar population-based analysis.17

It is known that clinical prevalence of certain comorbidities is higher in patients with COPD than in those without COPD.18 19 In this study, we show that this effect is also observable using registry data, independent of the population (ES, USA) and the specificities of the healthcare system (figure 1). This relation also persists if studying both acute and chronic (Medicare, USA) or only chronic diseases (CHSS, ES), suggesting the validity of disease interactions on the functional trait level represented by the DGs. Interestingly, age-related patterns of elevated comorbidity risk are similarly observable in the different healthcare systems reinforcing that this effect is persistent in the elderly population (>65 years).

The results also showed that comorbid conditions in elderly patients with COPD tend to appear after the diagnosis of the obstructive disease, which seems to be in line with other studies on disease trajectories.9 20 Different reports17 21 have suggested age-dependent comorbidity patterns in patients with COPD, which indicate an interesting direction for future analysis of temporal order of disease diagnoses. Furthermore, the results showing the distinguished role of cardiovascular health in COPD are also in line with earlier studies on comorbidity clustering21 and with its consistent relation to systemic effects of the
This indicates potential synergies between the management of pulmonary and cardiovascular health and the promotion of physical activity from early stages of the disease with the potential to modulate prognosis in these patients.

The increased risk of developing comorbid conditions in patients with COPD, as well deleterious interactions among concurrent diseases, indicates the need for refining current strategies aiming at reducing the burden of COPD on healthcare systems. In this context, prevention and appropriate treatment of comorbidities arise as central goals in the management of patients with COPD. These considerations are especially relevant in light of reports indicating that the majority of hospital admissions (and increased patients costs) are associated with comorbidities instead of the pulmonary events.

Considering the high predictive potential of comorbidity groupers on these events, further evaluation of other modalities of disease interactions, such as temporal order of appearance, concomitant clinical characteristics, lifestyle and genetic risk, constitute as interesting next step towards high accuracy health risk prediction, as proposed in Dueñas-Espin et al. Furthermore, a better knowledge of the underlying molecular mechanisms that modulate susceptibility for developing comorbidities in patients with COPD also emerges as major priority. It constitutes an initial step toward elaboration of cost-efficient strategies to prevent comorbid conditions.

However, a large-scale combined approach is indispensable in order to define efficient strategies coping with comorbidity clustering in patients with COPD. It involves meaningful integration of registry data with other information sources reflecting a broader health status, such as electronic health records, environmental and occupational exposures and genetic risks. Furthermore, the evolution of population-based analyses towards personalized approaches, such as identifying personal disease progression or their transcriptional patterns and comparing it with similar patients profiles, is needed to address disease heterogeneity often seen in COPD and to progress towards personalized health risk assessment and service selection.

Finally, it is acknowledged that risk of developing comorbidities can be modulated by confounding risk factors.

---

**Figure 3** (A) Temporal order of pairwise diagnoses in patients with chronic obstructive pulmonary disease (COPD). Red bars show the number of patients whose first diagnosis of a disease from the corresponding disease group (DG) happened before COPD, whereas cyan bars show the cases when such diagnoses were done after COPD. DGs are grouped into preferred directions: (1) G1, DG diagnosis after COPD, (2) G2, disease diagnosis before COPD and (3) G3, no significant directionality. (B) Elderly comorbidity network. Network nodes represent COPD and the different DGs. DGs are coloured by their directionality grouping: cyan for G1, red for G2 and grey for G3. The size of the nodes is proportional to the number of cases affected by both COPD and the DG, colour and thickness of edges are proportional to the strength of the directional association based on the causal information fraction measure. It is of note that simultaneous diagnoses were excluded from the analysis or accounted for when computing binomial directionality. This mainly influenced the data shown on respiratory diseases (>45% of COPD diagnoses were made simultaneously, online supplementary figure S1).
not included in the current analysis, such as degree of airflow limitation and smoking history. It is important to note, however, that risk factors alone cannot explain the observed effects,17,33 which reinforces the need for a large-scale, combined approaches incorporating patient information at population and patient-specific level.

It is acknowledged that registry information alone reflects underdiagnosis of COPD and the lack of forced spirometry data constitutes a significant limitation for accurate diagnosis of COPD. These barriers also indicate the need for speeding-up efforts to facilitate integration between clinical and registry data.

CONCLUSION

The current research confirms that patients with COPD are in higher risk of developing certain comorbidities than patients without COPD. The study results, as well as ongoing research on time-related analyses of disease trajectories, strengthen the need for further investigations on underlying mechanisms of non-pulmonary phenomena observed in patients with COPD with focus on altered regulation of biological pathways likely shared by different comorbid conditions. Furthermore, the study suggests the need for exploring novel modalities for health risk assessment and patient management aiming at consolidating cost-effective strategies to prevent comorbidities.6–10 In this context, current standard of care recommendations3 should necessarily evolve from the current organ-centred orientation to a systems approach.

Author affiliations

1Hospital Clinic de Barcelona, Institut d’Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Universitat de Barcelona, Barcelona, Spain
2Center for Biomedical Network Research in Respiratory Diseases (CIBERES), Madrid, Spain
3Unitat d’Informació i Coneixement, Servei Català de la Salut de la Generalitat de Catalunya, Barcelona, Catalunya, Spain
4Serveis Centrals, Institut Català de la Salut, Barcelona, Spain
5Mucosal and Salivary Biology Division, King’s College London Dental Institute, London, UK
6Unit of Computational Medicine, Center for Molecular Medicine, Department of Medicine Solna, Karolinska Institutet, Karolinska University Hospital and Science for Life Laboratory, Stockholm, Sweden
7Translational Bioinformatics Unit, Navarrabioiomed, Complejo Hospitalario de Navarra (CHN), Universidad Pública de Navarra (UPNA), IdiSNA, Pamplona, Spain

Acknowledgements We want to acknowledge the support of NEXTCARE team (COMRD15-1-0016), AGAUR research groups (2009SGR911 and 2014SGR661) and CERCA Programme/Generalitat de Catalunya.

Contributors Study conception and design: AT, IC, DG-C and JR. Data acquisition: EV, DM and MC. Data analysis: AT, EV, DM and MC. Manuscript preparation: AT, EV, IC, DG-C and JR. Manuscript revision: all authors.

Funding This work was supported by the European Commission grant CONNECARE (H2020-689892).

Competing interests None declared.

Patient consent Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All commercial use is permitted unless otherwise expressly granted.

REFERENCES