Seasonality, risk-factors and burden of community-acquired pneumonia in COPD patients: A population database study using linked healthcare records

Nicholas P Williams¹,², Ngaire A Coombs³, Matthew Johnson⁴, Lynn Josephs³,⁴, Lucy Rigge⁴,⁵, Karl J Staples², Mike Thomas¹,³,⁴ and Tom MA Wilkinson¹,²,⁴

Affiliations: ¹Southampton NIHR Respiratory Biomedical Research Unit, Southampton General Hospital, Southampton, UK; ²Clinical & Experimental Sciences, University of Southampton Faculty of Medicine, Southampton General Hospital, Southampton, UK; ³Primary Care & Population Sciences, University of Southampton Faculty of Medicine, University of Southampton, Southampton, UK; ⁴NIHR CLAHRC Wessex, Faculty of Health Sciences, University of Southampton, Southampton, UK; ⁵Department of Respiratory Medicine, University Hospital Southampton NHS Foundation Trust, Southampton, UK.

Correspondence: Dr Nicholas P Williams, Clinical & Experimental Sciences, University of Southampton Faculty of Medicine, Sir Henry Wellcome Laboratories, Southampton General Hospital, Mailpoint 810, Tremona Road, Southampton SO16 6YD, UK.

E-mail: NPWilliams@doctors.org.uk
Abstract

Community acquired pneumonia (CAP) is more common in patients with chronic obstructive pulmonary disease (COPD) than in the adult general population, with studies of hospitalised CAP consistently reporting COPD as a frequent comorbidity. However, despite an increasing recognition of its importance, large studies evaluating the incidence patterns over time, risk-factors and burden of CAP in COPD are currently lacking.

A retrospective observational study using a large UK-based database of linked primary and secondary care records was conducted. Patients with a diagnosis of COPD aged ≥40 years were followed for 5 years from the 1st January 2010. CAP and exacerbation episodes were identified from hospital discharge data and primary care coding records, and rates were calculated per month, adjusting for mortality, and displayed over time. In addition, baseline factors predicting future risk of CAP and hospital admission at CAP, were identified.

14513 COPD patients were identified: 13.4% (n=1938) had ≥1 CAP episode, of whom 18.8% suffered from recurrent (≥2) CAP. Highest rates of both CAP and exacerbations were seen in winter. A greater proportion of frequent, compared to infrequent exacerbators experienced recurrent CAP (5.1% versus 2.0% respectively, p<0.001) 75.6% of CAP episodes were associated with hospital admission compared to 22.1% of exacerbations. Older age and increasing grade of airflow limitation were independently associated with increased odds of CAP and hospital admission at CAP. Other independent predictors of future CAP included lower body mass index, inhaled-corticosteroid use, prior frequent exacerbations and comorbidities including ischaemic heart disease and diabetes.
CAP in COPD demonstrates clear seasonal patterns, with patient characteristics predictive of the odds of future CAP and hospital admission at CAP. Highlighting this burden of COPD-associated CAP during the winter period, informs us of the likely triggers and the need for more effective preventive strategies.
Introduction

Community-acquired pneumonia (CAP) is a major cause of morbidity, mortality and healthcare burden globally.\textsuperscript{1,2} Certain population groups are more susceptible, particularly the elderly and those with chronic conditions such as chronic obstructive pulmonary disease (COPD).\textsuperscript{3,4} The prevalence of COPD increases with age and is estimated to affect around 210 million people globally and therefore this population contributes significantly to the overall disease burden of CAP.\textsuperscript{5,6}

The clinical course of COPD is interspersed by episodes of symptomatic deterioration, termed exacerbations. CAP and exacerbations can be difficult to distinguish clinically and there remains no clear consensus on whether they represent distinct events or are in fact a continuum. Exacerbations are highly seasonal, with increasing rates seen in the winter months, adding further to the pressures already experienced by overstretched health-care providers at this time of year.\textsuperscript{7,8} The winter peak coincides with the increased activity of viral infection and a causal relationship has been proposed.\textsuperscript{9} Although exacerbations are more frequent in winter it is not known whether COPD-associated CAP have a similar seasonal pattern. In Europe, hospitalisations due to CAP are increasing, with COPD a frequently encountered co-morbidity.\textsuperscript{10,11} Reports suggest that patients with COPD hospitalised with CAP, present with more severe symptoms, have longer hospital stays and increased need for ICU admission.\textsuperscript{12-14} CAP is common in the natural history of COPD, with a much higher incidence compared to that of the adult general population.\textsuperscript{4} Furthermore, in a UK-based audit, 18\% of COPD patients requiring hospitalisation for acute respiratory illness, had consolidation evident on chest X-ray, diagnostic of CAP.\textsuperscript{15} Recent reports have highlighted an increased risk of CAP associated with inhaled corticosteroid (ICS) use in COPD patients, underlining the need for stratification of patients for ICS use to
include an estimation of this risk.\textsuperscript{16,17} COPD-associated CAP is, therefore, particularly relevant considering the ageing global population, and will have important implications for future antibiotic usage and healthcare resources.

Despite the greater incidence of CAP in COPD compared to the adult general population, it remains a relatively infrequent event compared to exacerbations. Consequently, the true burden of CAP in COPD has been largely overlooked and limited to predominantly cross-sectional analyses of hospitalised patients.\textsuperscript{18,19} Analysis of large population databases can provide a ‘real-world’ perspective of incidence patterns, healthcare usage and risk-factors of CAP in COPD and thus, an understanding of the likely triggers, susceptibilities and socioeconomic burden in what still remains a poorly understood event in the natural history of many COPD patients.

This is key to guiding future, targeted research, focusing on preventative strategies in vulnerable patient groups.

Linked primary and secondary health-care records have the potential to improve the accuracy of disease incidence estimates compared to stand alone records.\textsuperscript{20} Therefore, using the Hampshire Health Record Analytical (HHRA) database, an anonymised database of linked health-care records covering approximately 75\% of primary care practices in Hampshire, southern England, we sought to describe the seasonal patterns, hospital admissions and risk-factors for CAP, amongst patients with COPD.

\textbf{Methods}

\textbf{Study design and population}

A retrospective, observational cohort design was used. Patients were defined as having COPD and eligible for inclusion, if they had $\geq1$ COPD-specific Read code
In their primary care record at any time prior to 1\(^{st}\) January 2010 and were aged ≥40 at the start of the study. Follow-up ran for 5-years from the 1\(^{st}\) January 2010 until the 31\(^{st}\) December 2014.

The HHRA database was used for data collection (see online supplement for a detailed description). Briefly, it represents a large electronic database of individual patient anonymised primary and secondary care records, providing a single source of linked routine health data for patients living in Hampshire, southern England. At the time of the study, the HHRA contained records covering approximately 1.4 million patients from more than 140 primary care practices.

Use of the HHRA database for research is regulated by the Hampshire Health Record Information Governance Group (HHRIGG) and NHS South, Central and West Commissioning Support Unit. As HHRA data is anonymised it is not possible to identify patients, as such, and confirmed by local research governance, formal ethical approval was not required. However, patients have the option to opt-out of personal information being stored and used for analytical purposes.

**Demographics and clinical factors**

Data was collected for sex, age, weight, height, body mass index (BMI) and disease severity by GOLD stage at the most recently recorded time-point prior to the study start date, within a pre-assigned limit of 2 years. Comorbidities were identified using Read-code lists and included as present, if recorded at any time prior to the study start. Usage of concurrent therapies for COPD were included, if ≥1 prescription was evident in the 6-months prior to the study start, with the exception of oral corticosteroids (OCS), where frequent use was defined as ≥4 prescriptions in the year prior to the study start.
COPD-associated CAP

CAP episodes were identified from either a relevant ICD-10 code in the primary diagnostic position following an inpatient hospital stay or emergency department (ED) attendance, or from a physician Read-coded diagnosis of CAP (See on-line supplement, for coding lists). In order to avoid capturing multiple coded events for one illness as different episodes, single coded CAP events were grouped into clusters. Clusters were defined as the number of CAP events occurring within a continuous 70-day period, with the date of the first event regarded as the index date. A period of at least two event-free weeks following CAP episodes was required in order to avoid labelling prolonged recording of events for the same illness, as new episodes.

COPD exacerbations

Exacerbations were defined as either 1) a physician Read-coded diagnosis of exacerbation; 2) the prescription of specific antibiotics or OCS on the same day or up to 7 days before or after a COPD diagnostic or symptom Read Code; 3) a relevant respiratory ICD-10 code in the primary or secondary diagnostic positions, following an inpatient hospital stay or ED attendance (See on-line supplement, for coding lists). A continuous 21-day period defined an episode, with two event-free weeks required to separate exacerbations.

Seasonality

The month of the index event (CAP or exacerbation) date was taken as the month of episode occurrence. Winter was defined as December to February and summer as June to August.
Statistical analysis

Analyses were performed using SPSS version 22. CAP and exacerbation rates were calculated per month, adjusting for mortality, and displayed over time. Categorical variables are presented as number (percentage). Chi-square tests were used to compare data between groups with and without a diagnosis of CAP. Multivariate logistic regression was then used to identify adjusted odds ratios (OR) and 95% confidence intervals (CI) for having ≥1 CAP episode versus none, in patients with complete data, who either survived until the end of the observation period, or died following CAP. Covariates reaching significance from the bivariate analysis, were included in the multivariate model. To identify risk-factors for hospital admission during CAP we used STATA to conduct conditional logistic regression treating subject as a random effect. This accounted for the over-representation of some patients (i.e. those experiencing >1 CAP). A p-value of <0.05 was considered statistically significant.

Results

Population characteristics

The study cohort consisted of 14,513 patients with a diagnosis of COPD. The mean±SD age at the start of the study period was 70.3±10.8 years, 53.6% were male and 33.5% were current smokers (table 1). Of the cohort in which lung function data was available (n=10,358), mean±SD forced expiratory volume in 1 second (FEV1) % predicted was 59.5±19.8. Overall, exacerbations were more prevalent than CAP. The incidence rate for CAP was 37.6 per 1000 person years, with 13.4% (n=1938) experiencing at least one CAP and of these, 18.8% (n=364) had more than one.
Seasonality of CAP, exacerbations and hospital admissions

Figures 1A and 1B display the total monthly rates of CAP and exacerbations across 5 years, while figures 1C and 1D display the rates averaged into one calendar year. A clear seasonal pattern is evident with highest rates of CAP in the winter, especially December, compared to other periods of the year. Median monthly CAP rates were 4.0 per 1000 person months in winter versus 2.8 in summer (p=0.003). Similarly for exacerbations, higher rates occurred in winter, with lower rates in summer (83.6 versus 60.5 per 1000 person months respectively, p<0.001).

2413 CAP episodes were identified over 5-years. Trends in hospital admissions associated with CAP closely followed the seasonal distribution of total CAP episodes, with 75.6% of CAP associated with hospital admission, compared to only 22.1% of exacerbations.

Risk-factors for CAP

Using multivariate logistic regression modelling and controlling for potential confounders, factors independently predicting CAP risk were identified (table 2). The odds of experiencing CAP increased with age, with patients aged 60-79 and ≥80 years having higher odds than those aged 40-59 (OR 1.67, 95% CI 1.30-2.16; OR 4.10, 95% CI 3.05-5.49 respectively). BMI was also strongly associated with the odds of CAP, with those underweight having higher odds (OR 1.82, 95% CI 1.27-2.62) than those of normal weight. Conversely, the odds of CAP in overweight and obese patients was approximately 30% lower (OR 0.73, 95% CI 0.61-0.87; OR 0.70, 95% CI 0.58-0.85 respectively) than in patients of normal weight. Increasing grade of airflow limitation was associated with higher odds of CAP, with GOLD stage IV associated with almost three times the odds compared to GOLD stage I (OR 2.86, 95% CI 2.00-4.09).
Additional factors independently associated with increased odds of CAP were a prior history of CAP and comorbid diseases, including ischaemic heart disease (IHD), diabetes and chronic kidney disease (CKD).

**Exacerbation frequency and CAP risk**

A history of frequent exacerbations, defined as ≥2 exacerbations in the year prior to the study start date, appears to be predictive of future CAP. Frequent exacerbators had almost 30% higher odds of having CAP over the study period (OR 1.29, 95% CI 1.08-1.54), compared to infrequent exacerbators. In addition, a greater proportion of frequent exacerbators experienced recurrent CAP over 5-years, compared with infrequent exacerbators (5.1% versus 2.0%, p<0.001) (figure 2).

**Risk-factors for hospital admission associated with CAP**

Conditional logistic regression, accounting for subject effect, was used to examine associations between patient characteristics and the odds of hospital admission during CAP episodes (n patients=1368, n CAP=1721). Increasing age and grade of airflow limitation were independently associated with increased odds of hospital admission during CAP (table 3).

**All-cause mortality**

2877 deaths occurred over the 5-year study period. All-cause mortality followed a similar seasonal distribution to CAP and exacerbations (figure 3). 880 (30.6%) deaths occurred in winter compared to 602 (20.9%) in summer.

**Discussion**

This is the first population-based study showing the seasonal trends of CAP and exacerbations among people with COPD, using linked primary and secondary care
data sources. The incidence rate of 37.6 CAP cases per 1000 person years compares to that of 22.4 reported by Mullerova et al in COPD patients and 7.99 reported by Millett et al of adults aged ≥65 years, not limited to COPD.\textsuperscript{3,4}

13.4\% of our study population had CAP over 5 years, with almost a fifth of these having more than one episode. CAP occurrence is to some extent a predictable phenomenon and tracks with exacerbation frequency. Moreover, there are potentially modifiable factors such as BMI and concurrent medication use, which could form the basis of targeted risk-reducing strategies. Greater understanding of the underlying mechanisms is therefore required, to establish why some patients are more susceptible to CAP than others.

CAP occurred more commonly in the winter, compared to other times of the year. Similarly for exacerbations, and in line with findings from the Towards a Revolution in COPD Health (TORCH) study, a clear and consistent seasonal pattern from year-to-year was seen.\textsuperscript{7} Importantly, our study population differs from the TORCH study, as it benefits from providing a true insight into ‘real-world’ population trends across all severity of COPD, as compared to a highly selected patient cohort with more severe disease.

The majority of reports to date suggest that CAP in COPD is caused primarily by bacterial infection.\textsuperscript{18,22,23} However, the microbiological profile of CAP in adult patients without COPD has a proven seasonal basis with an important role attributed to viral infection.\textsuperscript{24,25} Holter et al demonstrated that CAP occurring in the winter and spring are associated with a greater frequency of bacterial-viral co-infection, supporting the results from an earlier study by Lieberman et al, showing greater frequencies of respiratory viral infection ascribed to CAP in these seasons.\textsuperscript{24,25} There is also a large
body of evidence supporting the role of influenza infection and secondary bacterial pneumonia, yet this finding is not apparent in reports for CAP in COPD. A separate analysis of our data (not shown) for each individual influenza season during the 5-year period, indicated no strong associations between reduced CAP rates and receipt of influenza vaccination. Certainly, our data shows a greater rate of CAP in the winter compared to other times of the year, suggesting that respiratory viral infection may have an important causative role. Indeed, environmental factors, notably lower temperatures and changes in relative humidity during winter, may favour the survival of respiratory viruses and increase host susceptibility to infection. It was not until relatively recently that viral infection and cold symptoms, a marker of putative viral infection, were established as important determinants of COPD exacerbations, leading to worsening symptoms and prolonged recovery. With the proven seasonality of exacerbations and now CAP in COPD, viral infection must be considered as a key early determinant in the sequence of events leading to both exacerbations and CAP in COPD patients.

Like other reports we found that increasing age, poorer lung function, lower BMI and co-morbid disease, particularly IHD, diabetes and CKD were all independently associated with future CAP risk. It is already known that the elderly, especially those aged >65 years, are particularly susceptible to developing CAP. The mechanisms are complex, but age-influenced immune system dysregulation, predisposing to infection risk is likely to be of importance. On a population-based level, physicians may be more likely to diagnose respiratory infection in older, frailer patients and yet we found no observed influence of age on the risk of exacerbation (data not shown). Like previous studies, we show the association between lower BMI and increased CAP risk and higher BMI with reduced risk, compared to normal weight patients. This
perceived protective effect of a higher BMI has also been reported previously in studies not limited to COPD patients.\textsuperscript{30} Yet, despite the underlying reasons remaining unclear, it could still provide a potentially useful target for risk-reducing interventions. A susceptibility to frequent exacerbations is now a well-established COPD phenotype and has been shown to be fairly consistent over time.\textsuperscript{31} Frequent exacerbators are more susceptible to viral infection and have higher levels of airway inflammation at stable-state, both proposed as possible underlying mechanisms.\textsuperscript{32} We have established prior exacerbation frequency as an independent risk-factor for CAP, with an increased proportion of frequent exacerbators experiencing recurrent CAP. However, a prior history of frequent exacerbations does not appear to influence the likelihood of more severe (i.e. hospitalised) CAP events. Furthermore, and in line with a previous report, we show that a prior history of CAP is a risk-factor for future CAP.\textsuperscript{4} Whether the mechanisms underlying exacerbation frequency and CAP susceptibility are the same, requires further research.

As anticipated, our study re-confirms ICS use as an independent risk-factor for future CAP, with a further trend evident between ICS use and increased odds of hospital admission associated with CAP. Interestingly, LAMA use was independently associated with increased odds of CAP. Published data has shown that in COPD patients with severe disease (FEV1 <50% predicted), LAMA use appears to confer less risk of CAP than ICS.\textsuperscript{17} However, data pertaining to CAP-risk with LAMA use in large population analyses of COPD patients and importantly across all disease severities is currently lacking.

The burden of respiratory-related healthcare utilisation and hospital attendances, especially for COPD and CAP is substantial, with the winter period being particularly
In line with an increased rate of CAP in winter, we have shown that the number of hospital admissions with CAP also increases substantially, proportionally much more so than with exacerbations. Patients who required hospital admission were older and had greater severity of underlying COPD, representing a vulnerable group and those most likely to impose a greater resource burden. More than 75% of CAP episodes were associated with hospital admission which is much higher than most reports for the adult general population and probably indicates a lower threshold for admitting COPD patients to hospital, or a manifestation of more severe symptoms. Regardless, with the ageing global population and increasing prevalence of COPD, CAP will become a growing future problem in this chronic disease group.

We acknowledge that this study has limitations. We were unable to validate CAP diagnoses with independent assessment of chest X-ray findings, as this was a large population database containing anonymised data in which radiological features were assessed by the attending physician. However, Skull and colleagues have reported that ICD-10 codes can be a valid method of retrospectively identifying CAP from medical records. A large proportion of CAP in our study were associated with hospital admission and based on hospital discharge diagnoses, and therefore it is highly likely that a chest X-ray was performed in these cases, confirming the diagnosis.

With regards our study population, we included patients with a Read coded diagnosis of COPD and did not require spirometric confirmation for study inclusion. However, this method has been shown to accurately identify COPD patients from UK primary care records with spirometry only marginally improving the accuracy.

In conclusion, like exacerbations, COPD-associated CAP is highly seasonal, with an increased burden observed over winter. Despite occurring less frequently in the
natural history of COPD than exacerbations, their relative impact on health-care utilisation is considerable, with the most vulnerable patients often at greatest risk of both events. In addition, establishing common risk-factors, may allow for a prospective estimation of a patients risk of future CAP, aiding clinicians in treatment and management decisions. Research, focusing on preventative strategies such as interventions to reduce risk and improve outcomes including nutritional interventions, treatment stratification and new and improved vaccination programmes, are key to tackling the increasing future burden of CAP in at-risk patients, such as those with COPD.

**Acknowledgements**

The authors thank NHS South, Central and West Commissioning Support Unit and the Hampshire Health Record Information Governance Group for their support, and for provision of access to HHRA data.
References


Figure 1. Total monthly rates and rate of hospital admission for CAP (A) and exacerbations (B) in COPD over 5 years. Figures C and D represent data for CAP and exacerbations respectively, averaged over one calendar year and displayed as the median monthly rate.
Figure 2. Associations between CAP frequency and the frequent exacerbator phenotype over 5 years. P-values from Chi-square test. *p<0.001
Figure 3. All-cause mortality rate over 5 years, averaged into one calendar year and displayed as the median monthly rate.
Supplementary data

Methods - Data Source

Hampshire Health Record (HHR) and electronic health care datasets

The HHR is an electronic shared care record for people living in Hampshire that provides a single source of linked routine health data for individual patients, collected in “real-time”, and links data from GP practices, hospitals and recently includes some data from social care. It is able to combine data from many different IT systems. It is not part of the National Summary Care Record but rather is an arrangement local to Hampshire, storing data from a wide range of practices that have “opted in” to the data sharing scheme. The HHR is used by clinicians to share information and aid clinical decision-making and is currently supported with data supplied by many organisations in the south of England, including local primary care practices and providers of secondary, community and social care.

Hampshire Health Record Analytical database (HHRA)

The Hampshire Health Record Analytical database (HHRA) is a separate electronic database created for research and analysis to support health improvement and planning. At the time of the present study it included data from more than 140 practices across Hampshire, covering about 1.4 million patients. It shares some of the health data contained within the HHR but in anonymised form to protect patients’ identity. The HHR and the HHRA are hosted and maintained by NHS South, Central and West Commissioning Support Unit (CSU). The Governance body is the HHR Information Governance Group (HHRIGG), which ensures the security and confidentiality of the HHR and HHRA and considers issues of data integration or data sharing. Linked data contained within HHRA includes:
• Primary Care data: coded clinical entries made during routine patient care

• Secondary Care data:
  • Radiology and pathology data
  • Inpatient, outpatient and Emergency Department (ED) Secondary Uses Service (SUS) data

Primary care coding within the HHRA uses Read codes, which are the standard clinical terminology system used in primary care in the UK. The Read code system comprises a detailed hierarchy of codes describing diagnoses, symptoms, signs, processes of care, investigations, administrative items, procedures and medication. 90% of UK practices use Read version 2 (V2) codes, which were used for the purpose of data capture for this study.

The HHRA allows full access to coded hospital activity data from SUS, thereby allowing quantification of respiratory-related hospital-based healthcare use to a much greater extent than is possible through databases accessing primary care data alone. The hospital data is the same as that used in contracting, accounting and NHS reporting processes, so high quality is assured.

The secondary care information used was Payment by Results data drawn from the SUS, which involves many quality control/validation processes applied at national level. For primary care information, basic data cleaning had already been performed within the NHS environment before data were imported into the HHR and HHRA. For this study, data cleaning involved excluding from the dataset, duplicated records and assessments with null and zero values and values deemed incompatible with life.