Original research

Frequency and severity of respiratory infections prior to COPD diagnosis and risk of subsequent postdiagnosis COPD exacerbations and mortality: EXACOS-UK health care data study

Hannah Whittaker,[1] Clementine Nordon,[2] Annalisa Rubino,[2] Tamsin Morris,[3] Yang Xu,[3] Enrico De Nigris,[4] Hana Müllerová,[2] Jennifer K Quint [1]

ABSTRACT

Objective Little is known about how lower respiratory tract infections (LRTIs) before chronic obstructive pulmonary disease (COPD) are associated with future exacerbations and mortality. We investigated this association in patients with COPD in England.

Methods Clinical Practice Research Datalink Aurum, Hospital Episode Statistics and Office of National Statistics data were used. Start of follow-up was patient's first ever COPD diagnosis date and a 1-year baseline period prior to start of follow-up was used to find mild LRTIs (general practice (GP) events/no antibiotics), moderate LRTIs (GP events+antibiotics) and severe LRTIs (hospitalised). Patients were categorised as having: none, 1 mild only, 2+ mild only, 1 moderate, 2+ moderate and 1+ severe. Negative binomial regression modelled the association between baseline LRTIs and subsequent COPD exacerbations and Cox proportional hazard regression was used to investigate mortality.

Results In 215 234 patients with COPD, increasing frequency and severity of mild and moderate LRTIs were associated with increased rates of subsequent exacerbations compared with no recorded LRTIs (1 mild adjusted IRR 1.16, 95% CI 1.14 to 1.18, 2+ mild IRR 1.51, 95% CI 1.46 to 1.55, 1 moderate IRR 1.81, 95% CI 1.78 to 1.85, 2+ moderate IRR 2.55, 95% CI 2.48 to 2.63). Patients with 1+ severe LRTI (vs no baseline LRTIs) also showed an increased rate of future exacerbations (adjusted IRR 1.75, 95% CI, 1.70 to 1.80). This pattern of association was similar for risk of all-cause and COPD-related mortality; however, patients with 1+ severe LRTIs had the highest risk of all-cause and COPD mortality.

Conclusion Increasing frequency and severity of LRTIs prior to COPD diagnosis were associated with increasing rates of subsequent exacerbations, and increasing risk of all-cause and COPD-related mortality.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Exacerbations of chronic obstructive pulmonary disease (COPD) have been associated with future exacerbations and mortality, but little is known about how lower respiratory tract infections (LRTIs) before a COPD diagnosis are associated with exacerbations risk and mortality after diagnosis.

WHAT THIS STUDY ADDS

⇒ This study found that increasing frequency and severity of LRTIs prior to COPD diagnosis are associated with increased rate of future exacerbations of COPD and increased risk of mortality after COPD diagnosis.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This study emphasises the importance of identifying LRTIs early in order to treat and prevent subsequent COPD outcomes and improve the lives of people with COPD.

INTRODUCTION

Acute exacerbations of chronic obstructive pulmonary disease (COPD) are worsening of COPD symptoms (breathlessness, cough, and sputum volume and purulence) beyond normal day-to-day variation. Approximately 50% to 80% of patients with COPD experience at least one exacerbation in a given year, and even a single moderate exacerbation increases the risk of future multiple exacerbation events.[1] [2–4] Exacerbations have also been associated with other clinical outcomes including accelerated lung function decline, reduced quality of life, and mortality.[5–7]

Lower respiratory tract infections (LRTIs) are important to consider when thinking about a diagnosis of COPD. Following the National Institute for Health and Care Excellence (NICE) COPD guidelines, a diagnosis should be suspected in people who experience frequent winter bronchitis (ie, multiple LRTIs) or present with symptoms such as breathlessness, chronic cough, sputum production or wheeze.[8] It is therefore possible that individuals may have several LRTI events prior to COPD diagnosis. How frequently these events occur and how they are managed may relate to future COPD exacerbation events. This is important to understand, as it is recognised that current COPD exacerbations are related to future exacerbation events, and LRTIs prior to diagnosis may impact clinical burden, healthcare resource utilisation and the exacerbation pathway following diagnosis.[9] [10] To date, most published literature reports on prevalent

COPD cohorts or incident cohorts focusing on what happens after diagnosis, but not prior to diagnosis.

We aimed to quantify the association between frequency and severity of LRTIs prior to COPD diagnosis and future exacerbations and mortality in a newly diagnosed COPD cohort. The goal of the analysis was to provide an evidence-based resource to guide healthcare professionals as they endeavour to determine which patients may be at greatest risk of frequent exacerbations or exacerbations early on in their disease who could possibly be targeted with more aggressive COPD management, to prevent the downwards spiral of future exacerbations.

METHODS

An observational cohort study was conducted in the UK analysing deidentified routinely collected electronic healthcare data.

Study design and participants

This EXAcerbations of COPD and their OutcomeS (EXACOS-UK) study used Clinical Practice Research Datalink (CPRD) Aurum data, a primary care database of patients registered at general practices (GPs) in England. Linked mortality data from the Office for National Statistics (ONS), socioeconomic data from the Index of Multiple Deprivation and secondary care data from Hospital Episode Statistics (HES) were provided for this study by CPRD. People were included if they met the following criteria: (1) aged 40 or older; (2) current or ex-smokers; (3) any data recorded in CPRD Aurum from 1 January 2004 onwards. After all these inclusion criteria were satisfied, patients were included in the cohort if a first ever validated diagnostic code of COPD was identified ≥1 year later,[11] the date of this diagnosis being the index date. This was so that patients had at least 1 year of baseline data prior to COPD diagnosis but after all other criteria were satisfied. Follow-up was from index date until the 30 November 2019, death or the date at which the patient transferred to a non-CPRD contributing practice, whichever came first (figure 1).

Exposure

The exposure of interest was frequency and severity of LRTIs identified during baseline. LRTIs were deemed: (1) mild if the GP visit was not associated with any antibiotic prescription; (2) moderate if the GP visit was associated with a prescription of respiratory antibiotics for 5–14 days; (3) severe if they required hospitalisation with internationl classification of disease (ICD)-10 codes for LRTI or pneumonia. LRTIs are usually treated with antibiotics alone and therefore a prescription of oral corticosteroids (OCS) was not included in our definition. However, the proportion treated with both was described.

Events that were recorded within 2 weeks of each other were assumed to be the same event. Baseline LRTI frequency and severity was categorised into one of the six mutually exclusive categories: none, 1 mild event (and no moderate or severe), ≥2 mild events (and no moderate or severe), 1 moderate event (any mild event and no severe), ≥2 moderate events (any mild events and no severe) and ≥1 severe events (any mild or moderate events).

Outcomes

Study outcomes were measured between index date and the end of follow-up. These included postdiagnosis exacerbations of COPD (moderate or severe) and mortality. Moderate exacerbation events were defined as events recorded in the GP using the combination of exacerbation diagnosis codes, coprescription of antibiotics plus OCS for 5–14 days and recorded symptoms based on a previous validated definition.[11] Severe exacerbation events were defined as COPD and LRTI events leading to a hospitalisation as recorded in HES.[12] Events that were recorded within 2 weeks of each other were assumed to be the same event. All-cause mortality, COPD-related mortality and cardiovascular disease (CVD)-related mortality were extracted from the ONS mortality database and identified using the underlying ICD-10 cause of death code.

Statistical analysis

Analyses were performed in STATA V.17. Baseline characteristics were defined prior to COPD diagnosis. These included age, sex, smoking status (current or ex-smoker), body mass index, history of depression, anxiety disorder, gastro-oesophageal reflux disease, lung cancer, myocardial infarction, heart failure and stroke, current asthma, socioeconomic deprivation, Medical Research Council (MRC) dyspnoea score, Global Initiative for Chronic Obstructive Lung Disease (GOLD)-defined airflow obstruction and current elevated blood eosinophil level (using threshold of 300 cells/µL). See online supplemental material p.2 for further details on baseline characteristic definitions. These characteristics were described overall and by LRTI exposure
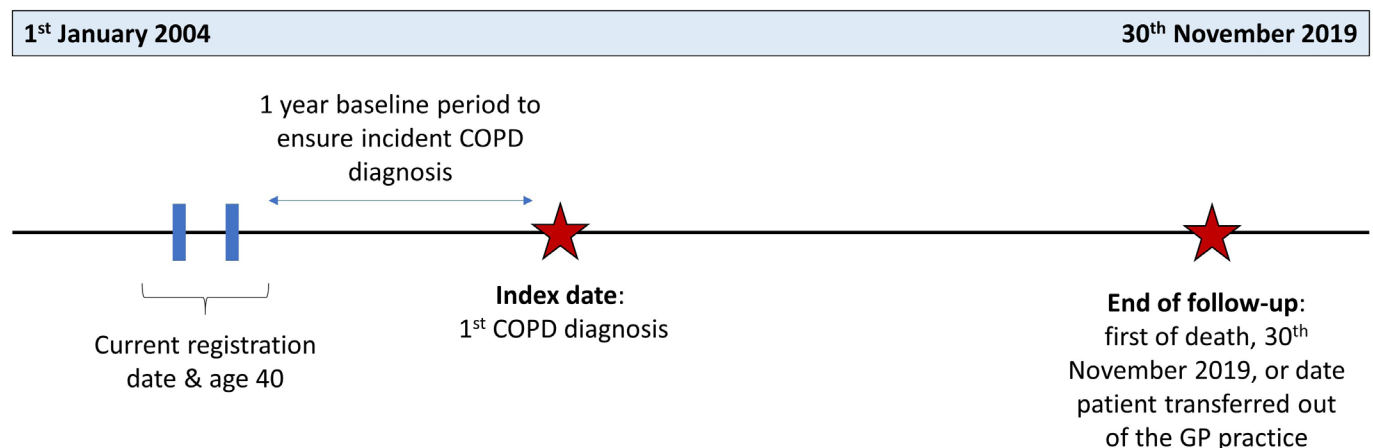

Figure 1 Study design for newly diagnosed COPD cohort. Patients were followed up from index date to the end of follow-up date. Outcome events were identified during this follow-up period. Baseline lower respiratory tract infections were identified in the baseline period prior to index date. COPD, chronic obstructive pulmonary disease; GP, general practice.

subgroup using mean (SD) for continuous measures and numbers (%) for categorical measures. COPD medication prescription in the year after COPD diagnosis (including combination of long and short-acting bronchodilators and inhaled corticosteroids) was described.

Negative binomial regression was used to investigate the association between baseline LRTI frequency and severity and the number of subsequent exacerbations expressed as incidence rate ratio (IRR) and its 95% CI (any type of exacerbation and separately for moderate and severe exacerbations). Person-time was defined as the time from start of follow-up to end of follow-up. Cox regression was used to investigate the association between baseline LRTI frequency and severity and all-cause, COPD-related, and CVD-related mortality. Adjusted models were adjusted for baseline characteristics listed. Missing values in baseline variables were coded as 'missing' and included in the final models ('missing' categories were not used as reference categories). No imputation was performed.

Sensitivity analyses

Sensitivity analyses were repeated using a different classification of baseline LRTIs in which mild and moderate LRTIs were grouped together, that is, irrespective of any antibiotics prescription as GPs may have different thresholds for choosing to prescribe antibiotics and antibiotics may not reflect the severity of the event.[13] The following baseline LRTI categories were used: none, 1 primary care recorded LRTI, ≥2 primary care recorded LRTIs, ≥3 primary care recorded LRTIs, ≥1 secondary care recorded LRTIs (ie, severe events). In addition, we repeated our main analysis using complete case analysis for variables that were not missing.

RESULTS

Overall, 215 234 patients newly diagnosed with COPD were included (online supplemental figure S1). In the baseline year, 158 074 (73.4%) patients had no LRTIs, 18 081 (8.4%) had 1 mild LRTI, 6133 (2.8%) had ≥2 mild LRTIs, 19 787 (9.2%) had 1 moderate LRTI, 6227 (2.9%) had ≥2 moderate LRTIs and 6932 (3.2%) had ≥1 severe LRTIs that required hospitalisation. Of those patients who had 1 moderate LRTI, 9.3% were also treated with OCS and 19.6% of patients with 2+ moderate were also treated with OCS.

Baseline characteristics of patients are detailed in table 1. Compared with patients with mild or moderate LRTIs, those with at least one severe LRTI in the year prior to COPD diagnosis tended to be older, were more likely to be ex-smokers, to have more breathlessness as per higher MRC scores ≥4 (although missing data concerned one third of the population on average) and more patients had a history of heart failure. Compared with patients with none, or only mild LRTIs, patients with moderate or severe LRTIs at baseline were more likely to have worse airflow obstruction on COPD diagnosis—with 20%–26% of patients with non-missing data having GOLD grade 3 or 4—and to have current asthma.

COPD medications prescribed in the first year after COPD diagnosis are detailed in table 2. Compared with patients with no LRTI, those with 2 or more mild LRTIs, 1 moderate LRTI, or 2 or more moderate LRTIs were more frequently prescribed COPD-related medications (67.6% vs between 72.8% and 82.1%) after initial diagnosis, whereas patients who had at least one severe LRTI (ie, required hospitalisation) were less frequently prescribed COPD-related medications (37.8%).

Baseline LRTIs and future exacerbations of COPD

A total of 948 576 COPD exacerbation events were recorded over a median follow-up of 4.4 years (IQR 1.9–7.9) in the population. Compared with patients with no LRTI in the year prior to COPD diagnosis, patients with 1 moderate LRTI or ≥2 moderate LRTIs had higher incidence rates of future exacerbations of COPD of any severity: adjusted IRR 1.81 (95% CI, 1.78 to 1.85) and 2.55 (95% CI, 2.48 to 2.63), respectively (figure 2, online supplemental table S1). This higher incidence rate was also found for patients with severe LRTIs at baseline, however, with a smaller magnitude of association: adjusted IRR 1.75 95% CI, 1.70 to 1.801. Patients with 1 or ≥2 mild LRTIs had higher incidence rates compared with those with no LRTIs but also to a lower extent: adjusted IRR 1.16 95% CI, 1.14 to 1.18 and adjusted IRR 1.51, 95% CI, 1.46 to 1.55, respectively.

A similar pattern of association was seen with future moderate exacerbations of COPD (online supplemental table S2); however, the strongest association was seen between baseline severe LRTIs and future severe exacerbations (adjusted IRR 3.02 95% CI, 2.87 to 3.18; online supplemental table S3).

Baseline LRTIs and mortality

A total of 51 669 (24%) of 215 234 patients with COPD died during follow-up, the highest risk being for the baseline severe LRTI group (figure 3, online supplemental table S4) (adjusted HR: 1.92 95% CI, 1.84 to 1.99), followed by ≥2 moderate and 1 moderate baseline LRTI (adjusted IRR: 1.22 95% CI, 1.17 to 1.27 and 1.10 95% CI, 1.07 to 1.13, respectively). An association was seen between 1 or ≥2 mild baseline LRTIs but to a lesser extent: adjusted HR 1.11 95% CI, 1.07 to 1.15, and 1.14 95% CI, 1.08 to 1.22, respectively.

In terms of COPD and CVD-related mortality, compared with no baseline LRTIs, the strongest association was seen in patients with ≥1 severe LRTI (adjusted HR, 2.32 95% CI, 2.15 to 2.50 and IRR 1.68 95% CI, 1.55 to 1.82, respectively) (online supplemental tables S5 and S6).

Sensitivity analyses

When baseline LRTI events were grouped into the following categories: no LRTI, 1 primary care recorded LRTI, 2 primary care recorded LRTIs, ≥3 primary care recorded LRTIs, and ≥1 hospitalised LRTI, all categories were significantly associated with future exacerbations of COPD compared with no baseline LRTIs (online supplemental table S6). When analyses were repeated using complete case analysis for non-missing covariates, results remained consistent (online supplemental table S7).

DISCUSSION

We explored the association between LRTIs requiring antibiotics in primary care or not, as well as those requiring hospitalisation in the year prior to a COPD diagnosis and the rate of subsequent moderate or severe exacerbation in a population-based cohort.

In our cohort, over a quarter of people had evidence of an LRTI in the year prior to COPD diagnosis. All of these patients were at higher rate of having subsequent moderate, and/or severe exacerbations of COPD in the near term after diagnosis (median follow-up of 4 years). To our knowledge, no studies have investigated whether respiratory infection events prior to COPD diagnosis are associated with an increase in a patient's rate of subsequent exacerbations of COPD or mortality.[1 9 14–17] According to guidelines, COPD should be investigated in people with frequent bronchitis. These events in themselves may help to understand COPD progression and ultimately help to treat

Table 1 Baseline characteristics of newly diagnosed patients with COPD categorised by baseline LRTI frequency and severity

	Overall	None	1 mild	≥2 mild	1 moderate	≥2 moderate	≥1 severe
Patient, n (%)	215 234 (100)	158 074 (73.4)	18 081 (8.4)	6133 (2.8)	19 787 (9.2)	6227 (2.9)	6932 (3.2)
Median study follow-up in years (Q1–Q3)	4.4 (1.9–7.8)	4.5 (2.0–7.9)	3.1 (1.5–5.7)	2.9 (1.3–5.2)	6.4 (3.3–9.3)	6.5 (3.3–9.5)	2.4 (0.9–4.8)
Mean age (SD)	66.9 (11.4)	66.6 (11.3)	67.4 (11.5)	67.8 (11.6)	66.9 (11.2)	67.2 (11.4)	71.5 (12.0)
Sex							
Male	114 738 (53.3)	86 166 (54.5)	8998 (49.8)	2921 (47.6)	9991 (50.5)	2915 (46.8)	3747 (54.1)
Female	100 496 (46.7)	71 908 (45.5)	9083 (50.2)	3212 (52.4)	9796 (49.5)	3312 (53.2)	3185 (46.0)
Smoking status							
Ex-smoker	107 787 (50.1)	84 066 (53.2)	9407 (52.0)	3274 (53.4)	10 508 (53.1)	3324 (53.4)	4128 (59.6)
Current smoker	107 447 (49.9)	74 008 (46.8)	8674 (48.0)	2859 (46.6)	9279 (46.9)	2903 (46.6)	2804 (40.5)
GOLD grade of airflow limitation							
1	51 151 (23.8)	38 450 (24.3)	4370 (24.2)	1441 (23.5)	4627 (23.4)	1336 (21.5)	927 (13.4)
2	78 736 (36.6)	57 432 (36.3)	6757 (37.4)	2287 (37.3)	7793 (39.4)	2537 (40.7)	1930 (27.8)
3	24 985 (11.6)	17 764 (11.2)	2086 (11.5)	724 (11.8)	2705 (13.7)	859 (13.8)	847 (12.2)
4	3814 (1.8)	2781 (1.8)	282 (1.6)	99 (1.6)	361 (1.8)	138 (2.2)	153 (2.2)
Missing	56 548 (26.3)	41 647 (26.4)	4586 (25.4)	1582 (25.8)	4301 (21.7)	1357 (21.8)	3075 (44.4)
mMRC score							
1	34 191 (15.9)	25 687 (16.3)	3201 (17.7)	1000 (16.3)	2854 (14.4)	779 (12.5)	670 (9.7)
2	59 658 (27.7)	43 767 (27.7)	5580 (30.9)	1967 (32.1)	5297 (26.8)	1579 (25.4)	1468 (21.2)
3	32 330 (15.0)	22 858 (14.5)	3118 (17.2)	1161 (18.9)	2938 (14.9)	1001 (16.1)	1254 (18.1)
4	13 249 (6.2)	9050 (5.7)	1319 (7.3)	487 (7.9)	1188 (6.0)	434 (7.0)	771 (11.1)
5	2559 (1.2)	1652 (1.1)	279 (1.5)	94 (1.5)	235 (1.2)	78 (1.3)	221 (3.2)
Missing	73 247 (34.0)	55 060 (34.8)	4584 (25.4)	1424 (23.2)	7275 (36.8)	2356 (37.8)	2548 (36.8)
Myocardial infarction	15 062 (7.0)	10 599 (6.7)	1367 (7.6)	467 (7.6)	1408 (7.1)	491 (7.9)	730 (10.5)
Stroke	17 311 (8.0)	12 381 (7.8)	1553 (8.6)	588 (9.6)	1399 (7.1)	458 (7.4)	932 (13.4)
Heart failure	9136 (4.2)	6045 (3.8)	951 (5.3)	335 (5.5)	726 (3.7)	303 (4.9)	776 (11.2)
Lung cancer	412 (0.2)	285 (0.2)	42 (0.2)	<20	42 (0.2)	<20	<20
Asthma	45 751 (21.3)	33 480 (21.2)	3540 (19.6)	1285 (21.0)	4505 (22.8)	1745 (28.0)	1196 (17.3)
GORD	31 904 (14.8)	22 403 (14.2)	3154 (17.4)	1217 (19.8)	3005 (15.2)	1067 (17.1)	1058 (15.3)
Anxiety disorder	46 184 (21.5)	32 399 (20.5)	4625 (25.6)	1678 (27.4)	4409 (22.3)	1571 (25.2)	1502 (21.7)
Depression	55 323 (25.7)	39 000 (24.7)	5358 (29.6)	1990 (32.5)	5320 (26.9)	1839 (29.5)	1816 (26.2)
BMI group (in kg/m^2)							
<19	8395 (3.9)	5996 (3.8)	757 (4.2)	226 (3.7)	735 (3.7)	247 (4.0)	434 (6.3)
19–24	62 392 (29.0)	46 015 (29.1)	5126 (28.4)	1662 (27.1)	5919 (29.9)	1725 (27.7)	1945 (28.1)
25–29	64 885 (30.2)	48 313 (30.6)	5314 (29.4)	1802 (29.4)	5863 (29.6)	1877 (30.1)	1716 (24.8)
≥30	58 796 (27.3)	42 687 (27.0)	5251 (29.0)	1897 (30.9)	5325 (26.9)	1781 (28.6)	1855 (26.8)
Missing	20 766 (9.7)	15 063 (9.5)	1633 (9.0)	546 (8.9)	1945 (9.8)	597 (9.6)	982 (14.2)
IMD quintiles							
1 (least deprived)	34 469 (16.0)	25 774 (16.3)	2744 (15.2)	896 (14.6)	3148 (15.9)	909 (14.6)	998 (14.4)
2	38 946 (18.1)	28 881 (18.3)	3.122 (17.3)	1036 (16.9)	3589 (18.1)	1111 (17.8)	1207 (17.4)
3	42 090 (19.6)	31 054 (19.7)	3436 (19.0)	1206 (19.7)	3837 (19.4)	1202 (19.3)	1355 (19.6)
4	45 775 (21.3)	33 531 (21.2)	3913 (21.6)	1325 (21.6)	4120 (20.8)	1335 (21.4)	1551 (22.4)
5 (most deprived)	53 797 (25.0)	38 714 (24.5)	4852 (26.8)	1669 (27.2)	5076 (25.7)	1669 (26.8)	1817 (26.2)
Missing	157 (0.1)	120 (0.1)	14 (0.1)	<5	17 (0.1)	<5	<5
Blood eosinophil levels							
<300 cells/μL	83 244 (38.7)	60 574 (38.3)	7377 (40.8)	2659 (43.4)	7340 (37.1)	2329 (37.4)	2965 (42.8)
≥300 cells/μL	37 720 (17.5)	26 885 (17.0)	3408 (18.9)	1272 (20.7)	3611 (18.3)	1206 (19.4)	1338 (19.3)
Missing	94 270 (43.8)	70 615 (44.7)	7296 (40.4)	2202 (35.9)	8826 (44.7)	2692 (43.2)	2629 (37.9)

The following baseline variables were identified at or close to index date and were identified as 'current': age, smoking status, IMD, BMI, mMRC, GOLD, asthma, eosinophil level. The following baseline variables were recorded any time prior to index date: depression, anxiety, GORD, lung cancer, myocardial infarction, heart failure, stroke.
BMI, body mass index; COPD, chronic obstructive pulmonary disease; GOLD, Global Initiative for Chronic Obstructive Lung Disease; GORD, gastro-oesophageal reflux disease; IMD, Index of Multiple Deprivation; LRTI, lower respiratory tract infection; mMRC, modified Medical Research Council.

Table 2 COPD medications in the year after COPD diagnosis in patients categorised by baseline LRTI frequency and severity

	Overall N=2 15 234	None n=1 58 074	1 mild n=18 081	≥2 mild n=6133	1 moderate n=19 787	≥2 moderate n=6227	≥1 severe 6932
Any COPD medications	148 655 (69.1)	106 778 (67.6)	12 583 (69.6)	4467 (72.8)	15 403 (77.8)	5109 (82.1)	2617 (37.8)
Maintenance inhaled therapies							
Any COPD maintenance medications	125 540 (58.3)	89 637 (56.7)	10 690 (59.1)	3832 (62.5)	13 022 (65.8)	4485 (72.0)	3874 (55.9)
No COPD maintenance medications	89 694 (41.7)	68 437 (43.3)	7391 (40.9)	2301 (37.5)	6765 (34.2)	1742 (28.0)	3058 (44.1)
LABA monotherapy	4096 (1.9)	3062 (1.9)	285 (1.6)	93 (1.5)	446 (2.3)	146 (2.3)	64 (0.9)
LAMA monotherapy	20 328 (9.4)	14 896 (9.4)	1912 (10.6)	604 (9.9)	1802 (9.1)	497 (8.0)	617 (8.9)
ICS monotherapy	14 243 (6.6)	10 783 (6.8)	943 (5.2)	264 (4.3)	1563 (7.9)	491 (7.9)	199 (2.9)
LABA/ICS dual therapy	38 284 (17.8)	27 590 (17.5)	1073 (5.9)	1055 (17.2)	4214 (21.3)	1549 (24.9)	947 (13.7)
LABA/LAMA dual therapy	8837 (4.1)	6476 (4.1)	1073 (5.9)	388 (6.3)	438 (2.2)	104 (1.7)	358 (5.2)
LAMA/ICS dual therapy	3137 (1.5)	2234 (1.4)	242 (1.3)	89 (1.5)	360 (1.8)	131 (2.1)	81 (1.2)
Triple therapy	36 615 (17.0)	24 596 (15.6)	3306 (18.3)	1339 (21.8)	4199 (21.2)	1567 (25.2)	1608 (23.2)
Short-acting therapies							
SABA	114 354 (90.1)	81 630 (90.2)	10 029 (93.1)	3606 (93.3)	11 806 (87.5)	3942 (86.4)	3341 (91.6)
SAMA	2448 (1.9)	1770 (2.0)	157 (1.5)	55 (1.4)	320 (2.4)	107 (2.3)	39 (1.1)
SAMA/SABA	10 060 (7.9)	7122 (7.9)	583 (5.4)	206 (5.3)	1368 (10.1)	514 (11.3)	267 (7.3)

Any COPD medications includes any maintenance COPD medications or any short-acting therapies. COPD maintenance therapies include any combination of LABA, LAMA, ICS. Triple therapy=LABA/LAMA/ICS. Short-acting therapies include SABA, SAMA. Numbers are n(%).
COPD, chronic obstructive pulmonary disease; ICS, inhaled corticosteroids; LABA, long-acting beta agonist; LAMA, long-acting muscarinic antagonist; LRTI, lower respiratory tract infection; SABA, short-acting beta agonists; SAMA, short-acting muscarinic antagonists.

patients with COPD more effectively.[8] The results of our study illustrate the importance of LRTIs in the course of COPD.

Patients with one mild LRTI in 12 months prior to incident COPD diagnosis had the lowest rate of subsequent exacerbations compared with patients with increased frequency and severity at baseline. It is possible that one mild LRTI event in primary care that is not treated with antibiotics is extremely mild and misclassification may exist between mild LRTIs and events occurring but not recorded by the GP. We also saw a difference in associations between mild and moderate primary care events in terms of subsequent moderate exacerbations of COPD, whereby LRTI events treated in primary care were more strongly associated with future moderate COPD exacerbation events. This was not seen with future severe exacerbations of COPD, whereby severe LRTI events were most strongly associated with future severe exacerbations followed by two or more mild or moderate LRTI. This suggests that other than severe LRTI, increased frequency

of any type of baseline LRTI is also important in relation to future severe exacerbations of COPD.

The association between exacerbations of COPD and increased risk of future exacerbations is well known.[4 9] In addition, recent studies suggest that the presence of moderate exacerbations of COPD during a year, even a single one, are associated with future risk of exacerbations.[2–4] Our study expands on this result by highlighting the importance of even a single LRTI requiring antibiotics prior to first COPD diagnosis, in the estimation of the future rate of COPD exacerbations.

When LRTIs were categorised into levels of frequency and severity, we observed that the rate of subsequent moderate COPD exacerbations (ie, managed in primary care settings) was the highest for patients with moderate LRTIs prior to COPD diagnosis, while the rate of subsequent severe exacerbations (ie, managed in the hospital) was the highest for patients with severe LRTIs, which may reflect the underlying severity of COPD on diagnosis. Indeed, the group of patients who were diagnosed after severe LRTIs or multiple LRTIs had worse airflow limitation, dyspnoea, and were older suggesting that COPD had existed
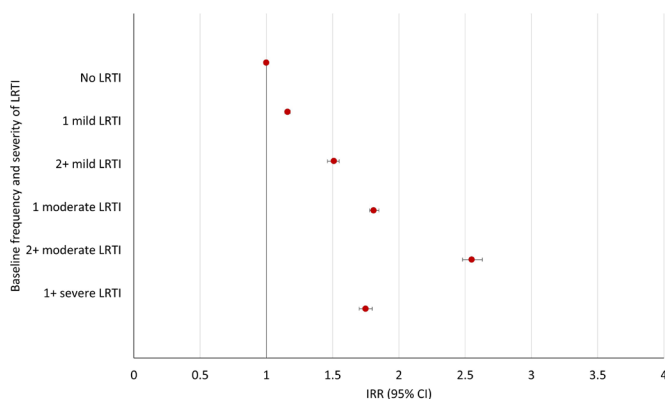

Figure 2 Incidence rate ratios (IRRs) of at least one acute exacerbation of chronic obstructive pulmonary disease (COPD) following COPD diagnosis by baseline frequency and severity of lower respiratory tract infection (LRTI).
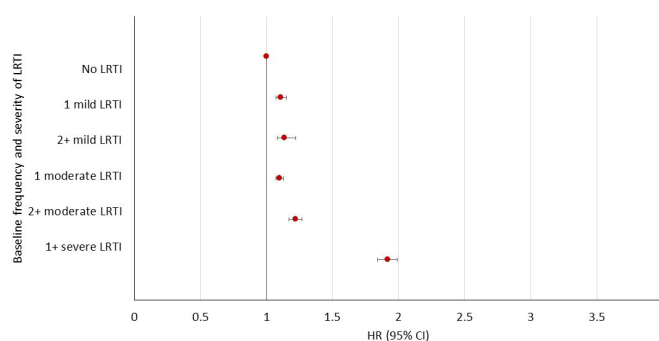

Figure 3 HRs for all-cause mortality following chronic obstructive pulmonary disease diagnosis by baseline frequency and severity of lower respiratory tract infection (LRTI).

for a longer period and that the formal diagnosis of COPD was made at a later stage. This suggests a missed opportunity for an earlier diagnosis in people experiencing frequent or severe LRTI. Another plausible explanation could be residual confounding related to healthcare utilisation and quality of care, as patients with severe LRTIs at baseline had the lowest proportion of any COPD therapies. Finally, it is possible that some patients who have LRTIs, or COPD exacerbations, go to the emergency room instead of the GP.[18]

In terms of mortality, we found that increased frequency and severity of baseline LRTI were associated with increased rate of all-cause and COPD-related mortality but only multiple moderate LRTIs and severe LRTIs were associated with CVD-related mortality.

A previous study investigated exacerbations of COPD in the year following diagnosis and future risk of exacerbations and mortality, namely in incident COPD.[9] This study defined GP related exacerbations as exacerbation of COPD events or LRTIs, a prescription of antibiotics, OCS or both, for 5–14 days, or exacerbation-related symptoms combined with prescription for antibiotics and/or OCS. Authors found that increasing frequency of GP-related exacerbations in the year following COPD diagnosis was associated with increased risk of future exacerbations of COPD after COPD diagnosis, which is similar to the results of our sensitivity analysis. The main difference, however, was that we included events using a different definition for LRTIs. Our definition did not include the use of OCS nor did it include events recorded as exacerbations of COPD in order to identify LRTI events which general practitioners would have recorded prior to COPD diagnosis. Despite this, we found that the vast proportion of events treated with a respiratory antibiotic were also treated with a course of OCS before a clinical COPD diagnosis was given. This suggests that GPs may suspect COPD; however, they do not formally diagnose these patients until a short while after. In addition, we expanded on previous work by showing that increasing frequency and severity of LRTIs are associated with future exacerbations of COPD, all-cause mortality and COPD-related mortality and that in patients with LRTIs, multiple moderate or any severe LRTIs had similar CVD-related risk of death than those who had no LRTIs at baseline.

Studies of this type have a number of limitations. First, while severe exacerbations were identified by hospitalised events, we cannot be sure that these events are severe events in the clinical sense. It is possible that some patients who exacerbate go to the emergency room instead of the GP and are in turn hospitalised in the absence of outpatient management. Indeed, access to GPs in the UK is becoming increasingly difficult.[18] In addition, we included LRTI events as LRTI/pneumonia recorded events with or without antibiotic prescriptions; however, it is possible that events might have been missed if they were coded as exacerbations of COPD. Similarly, antibiotic prescriptions for LRTI events will depend on whether the infection is viral or bacterial based on the general practitioner's judgement. In the same way that prior exacerbations are associated with subsequent exacerbation events, it could be argued that LRTI events are associated with subsequent exacerbation events after a diagnosis of COPD. Alternatively, it could be that in fact the LRTI events were themselves exacerbations of COPD (just unknown as a diagnosis of COPD had not been made), as those with LRTI events had lower lung function at COPD diagnosis. However, as these events occurred prior to a COPD diagnosis, they will have been managed differently to an exacerbation in a diagnosed patient with COPD, for example, prednisolone will not have been prescribed. Despite this, bacterial infections or bacterial coinfections are associated with increased symptoms and can be more severe in patients with COPD.[19] In addition, residual confounding is likely to exist as we were unable to account for variables such as environmental factors and time-varying confounders such as COPD medication over follow-up. Lastly, this study investigated the association between events that occurred prior to a COPD diagnosis and COPD-related events that occurred after COPD diagnosis and does not imply causation. While LRTI events might have occurred earlier than in the year prior to COPD diagnosis, the events that occur close to a diagnosis are important in order to understand how better management of LRTI events or earlier diagnosis could improve COPD-related outcomes in the future.

CONCLUSION

The frequency and severity of LRTI events prior to first COPD diagnosis are associated with future increased rates of exacerbations of COPD, all-cause mortality and COPD-related mortality. Overall, this study suggests that LRTIs prior to COPD diagnosis may be indicative of subsequent increased rate of exacerbations and increased risk of mortality and should be taken seriously by healthcare professionals.

Patient consent for publication Not applicable.

Ethics approval This work is based in part on data from the CPRD obtained under license from the UK Medicines and Healthcare products Regulatory Agency. The data is provided by patients and collected by the National Health Service as part of their care and support. The interpretation and conclusions contained in this study are those of the authors alone. A protocol for this research was approved by the Independent Scientific Advisory Committee (ISAC) for Medicines and Healthcare products Regulatory Agency Database Research (protocol number 20_103R) and the approved protocol was made available to the journal in which this research is published, and to the reviewers during peer review. Generic ethical approval for observational research using the CPRD with approval from ISAC has been granted by a Health Research Authority Research Ethics Committee (East Midlands—Derby, REC reference number 05/MRE04/87). The interpretation and conclusions contained in this study are those of the authors alone.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. Data are available on request from the CPRD. Their provision requires the purchase of a license, and this license does not permit the authors to make them publicly available to all.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

ORCID iD
Jennifer K Quint http://orcid.org/0000-0003-0149-4869

REFERENCES

1 Hurst JR, Vestbo J, Anzueto A, *et al*. Susceptibility to exacerbation in chronic obstructive pulmonary disease. *N Engl J Med* 2010;363:1128–38.

2 Haughney J*et al*. The long-term clinical and economic impact of COPD exacerbations: an observational study (SHERLOCK). *Europ Res J* 2020;56:4910.

3 Vogelmeier CF, Diesing J, Kossack N, *et al*. COPD exacerbation history and impact on future exacerbations - 8-year retrospective observational database cohort study from Germany.

4 Whittaker H*et al*. Frequency and severity of exacerbations of COPD associated with future risk of exacerbations and mortality: a UK routine health care data study.

5 Miravitlles M, Ferrer M, Pont A, *et al*. Effect of exacerbations on quality of life in patients with chronic obstructive pulmonary disease: a 2 year follow up study. *Thorax* 2004;59:387–95.

6 Donaldson GC, Seemungal TAR, Bhowmik A, *et al*. Relationship between exacerbation frequency and lung function decline in chronic obstructive pulmonary disease. *Thorax* 2002;57:847–52.

7 Soler-Cataluña JJ, Martínez-García MA, Román Sánchez P, *et al*. Severe acute exacerbations and mortality in patients with chronic obstructive pulmonary disease. *Thorax* 2005;60:925–31.

8 NICE. *Chronic obstructive pulmonary disease in over 16s: diagnosis and management (ed Excellence, N.I.f.H.a.C.)*, 2019.

9 Rothnie KJ, Müllerová H, Smeeth L, *et al*. Natural history of chronic obstructive pulmonary disease exacerbations in a general practice-based population with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2018;198:464–71.

10 Pasquale MK, Sun SX, Song F, *et al*. Impact of exacerbations on health care cost and resource utilization in chronic obstructive pulmonary disease patients with chronic bronchitis from a predominantly Medicare population. *Int J Chron Obstruct Pulmon Dis* 2012;7:757–64.

11 Rothnie KJ, Müllerová H, Hurst JR, *et al*. Validation of the recording of acute exacerbations of COPD in UK primary care electronic healthcare records. *PLoS One* 2016;11:e0151357.

12 Rothnie KJ, Müllerová H, Thomas SL, *et al*. Recording of hospitalizations for acute exacerbations of COPD in UK electronic health care records. *Clin Epidemiol* 2016;8:771–82.

13 Palin V, Mölter A, Belmonte M, *et al*. Antibiotic prescribing for common infections in UK general practice: variability and drivers. *J Antimicrob Chemother* 2019;74:2440–50.

14 Sadatsafavi M, Xie H, Etminan M, *et al*. The association between previous and future severe exacerbations of chronic obstructive pulmonary disease: updating the literature using robust statistical methodology. *PLoS One* 2018;13:e0191243.

15 Hartley BF, Barnes NC, Lettis S, *et al*. Risk factors for exacerbations and pneumonia in patients with chronic obstructive pulmonary disease: a pooled analysis. *Respir Res* 2020;21:5.

16 Müllerová H, Shukla A, Hawkins A, *et al*. Risk factors for acute exacerbations of COPD in a primary care population: a retrospective observational cohort study. *BMJ Open* 2014;4:e006171.

17 Margüello MS, Garrastazu R, Ruiz-Nuñez M, *et al*. Independent effect of prior exacerbation frequency and disease severity on the risk of future exacerbations of COPD: a retrospective cohort study. *NPJ Prim Care Respir Med* 2016;26:16046.

18 Baker M, Jeffers H. *Continuity of care in modern day general practice*. London: Royal College of General Practitioners, 2016.

19 Fau FH, Jungblut S. The influence of virus infections on the course of COPD.

Frequency and severity of respiratory infections prior to COPD diagnosis and risk of future exacerbations and mortality: a UK routine health care data study

Supplementary material

Contents

Additional methodology

Analyses were performed in STATA17. Baseline characteristics were defined prior to COPD diagnosis. These included age, sex, smoking status (current or ex-smoker), BMI, history of depression, anxiety disorder, gastro-oesophageal reflux disease (GORD), lung cancer, myocardial infarction, heart failure and stroke, current asthma, socioeconomic deprivation, Medical Research Council (MRC) dyspnoea score, GOLD-defined airflow obstruction, and current elevated blood eosinophil level (using threshold of 300 cells/µl). These characteristics were described overall and by LRTI exposure sub-group using mean (SD) for continuous measures and numbers (%) for categorical measures. They included age at index date, sex, closest smoking status to index date (current or ex-smoker), current BMI within 5 years of index date (underweight <18.5kg/m2, normal 18.5-24.9kg/m2, overweight 25.0-29.9kg/m2, and obese ≤30.0kg/m2), history of depression, anxiety disorder, gastro-oesophageal reflux disease (GORD), lung cancer, myocardial infarction, heart failure and stroke ever recorded prior to index date, current asthma defined in a period of 3 years 2 years prior to index date, current socioeconomic deprivation using the Index of Multiple Deprivation (categorized into quintiles), current Medical Research Council (MRC) dyspnoea score identified within 2 years of index date (1-5, a higher score indicating more severe dyspnoea), current GOLD-defined airflow obstruction defined within 2 years of index date (GOLD 1: FEV1 % predicted ≥80%; GOLD 2: FEV1 % predicted 50-79%; GOLD 3: FEV1 % predicted 30-49%; GOLD 4: FEV1 % predicted <30%), and current elevated blood eosinophil level defined within 1 years of index date (using threshold of 300 cells/µl).

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Thorax

Table S1: Incidence rates of at least one AECOPD (moderate or severe) following COPD diagnosis in different categories of baseline frequency and severity of LRTI

Baseline LRTI category	n (%) of events N=948,576	Rate 100py^{-1} (95% CI)	Crude IRR (95% CI)	Adjusted* IRR (95% CI)
No LRTI n=158,074	512,658	61.9 (61.7 – 62.1)	Reference	Reference
1 mild LRTI n=18,081	52,232	72.0 (71.3 – 72.6)	1.18 (1.15-1.20)	1.16 (1.14-1.18)
2+ mild LRTI n=6,133	21,717	95.7 (94.5 – 97.0)	1.56 (1.51-1.61)	1.51 (1.46-1.55)
1 moderate LRTI n=19,787	145,161	113.2 (112.6 – 113.8)	1.92 (1.88-1.95)	1.81 (1.78-1.85)
2+ moderate LRTI n=6,227	68,822	168.2 (167.0 – 169.5)	2.88 (2.80-2.97)	2.55 (2.48-2.63)
1+ severe LRTI n=6,932	22,280	96.1 (94.8 – 97.4)	2.05 (1.99-2.12)	1.75 (1.70-1.80)

Legend: AECOPD (Acute Exacerbations of Chronic Obstructive Pulmonary Disease), LRTI (Lower Respiratory Tract Infection), IRR (Incidence Rate Ratio). PY (Person years). *Adjusted for all baseline covariates listed in Table 1. Number of patients included crude and fully adjusted models N= 215,234.

3

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance
placed on this supplemental material which has been supplied by the author(s)

Thorax

Table S2: Incidence rates of at least one moderate AECOPD following COPD diagnosis in different categories of baseline frequency and severity of LRTI

Baseline LRTI category	n (%) of events N=948,576	Rate 100py^{-1} (95% CI)	Crude IRR (95% CI)	Adjusted* IRR (95% CI)
No LRTI n=158,074	512,658	61.9 (61.7 – 62.1)	Ref	Ref
1 mild LRTI n=18,081	52,232	72.0 (71.3 – 72.6)	1.14 (1.12-1.16)	1.14 (1.12-1.16)
2+ mild LRTI n=6,133	21,717	95.7 (94.5 – 97.0)	1.54 (1.49-1.59)	1.50 (1.45-1.55)
1 moderate LRTI n=19,787	145,161	113.2 (112.6 – 113.8)	2.02 (1.98-2.06)	1.90 (1.87-1.94)
2+ moderate LRTI n=6,227	68,822	168.2 (167.0 – 169.5)	3.08 (2.98-3.17)	2.71 (2.64-2.79)
1+ severe LRTI n=6,932	22,280	96.1 (94.8 – 97.4)	1.60 (1.55-1.66)	1.44 (1.39-1.49)

Legend: AECOPD (Acute Exacerbations of Chronic Obstructive Pulmonary Disease), LRTI (Lower Respiratory Tract Infection), IRR (Incidence Rate Ratio). PY (Person years). *Adjusted for all baseline covariates listed in Table 1. Number of patients included crude and fully adjusted models N= 215,234.

4

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Thorax

Table S3: Incidence rates of at least one severe AECOPD following COPD diagnosis in different categories of baseline frequency and severity of LRTI

Baseline LRTI category	n (%) of events N=948,576	Rate 100py^{-1} (95% CI)	Crude IRR (95% CI)	Adjusted* IRR (95% CI)
No LRTI n=158,074	81,259	9.8 (9.7 – 9.9)	Ref	Ref
1 mild LRTI n=18,081	9,539	13.1 (12.9 – 13.4)	1.35 (1.29-1.40)	1.25 (1.20-1.29)
2+ mild LRTI n=6,133	3,629	16.0 (15.5 – 16.5)	1.65 (1.55-1.76)	1.48 (1.40-1.57)
1 moderate LRTI n=19,787	16,466	12.8 (12.6 – 13.0)	1.37 (1.32-1.42)	1.30 (1.26-1.34)
2+ moderate LRTI n=6,227	6,760	16.5 (16.1 – 16.9)	1.82 (1.72-1.93)	1.62 (1.54-1.70)
1+ severe LRTI n=6,932	8,053	34.7 (34.0 – 35.5)	4.46 (4.22-4.72)	3.03 (2.87-3.18)

Legend: AECOPD (Acute Exacerbations of Chronic Obstructive Pulmonary Disease), LRTI (Lower Respiratory Tract Infection), IRR (Incidence Rate Ratio). PY (Person years). *Adjusted for all baseline covariates listed in Table 1. Number of patients included crude and fully adjusted models N= 215,234.

5

Table S4: Hazard ratios for all-cause mortality following COPD diagnosis in different categories of baseline frequency and severity of LRTI.

Baseline LRTI category	n (%) of events N=51,669	Crude HR (95% CI)	Adjusted* HR (95% CI)
No LRTI n=158,074	36,151 (22.9)	Ref	Ref
1 mild LRTI n=18,081	3,491 (19.3)	1.13 (1.09-1.17)	1.11 (1.07-1.15)
2+ mild LRTI n=6,133	1,096 (17.9)	1.15 (1.08-1.22)	1.14 (1.08-1.22)
1 moderate LRTI n=19,787	6,046 (30.6)	1.07 (1.04-1.09)	1.10 (1.07-1.13)
2+ moderate LRTI n=6,227	2,183 (35.1)	1.20 (1.15-1.26)	1.22 (1.17-1.27)
1+ severe LRTI n=6,932	2,702 (39.0)	2.77 (2.66-2.88)	1.92 (1.84-1.99)

Legend: AECOPD (Acute Exacerbations of Chronic Obstructive Pulmonary Disease), LRTI (Lower Respiratory Tract Infection), HR (Hazard Ratio). *Adjusted for all baseline covariates listed in Table 1. Number of patients included crude and fully adjusted models N= 215,234.

6

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Thorax

Table S5: Hazard ratios for COPD-related mortality following COPD diagnosis in different categories of baseline frequency and severity of LRTI

Baseline LRTI category	n (%) of events N=11,770	Crude HR (95% CI)	Adjusted* HR (95% CI)
No LRTI n=158,074	7,886 (67.0)	Ref	Ref
1 mild LRTI n=18,081	823 (7.0)	1.27 (1.18-1.36)	1.19 (1.11-1.28)
2+ mild LRTI n=6,133	277 (2.4)	1.40 (1.24-1.58)	1.29 (1.14-1.45)
1 moderate LRTI n=19,787	1,425 (12.1)	1.13 (1.07-1.20)	1.13 (1.06-1.19)
2+ moderate LRTI n=6,227	546 (4.6)	1.35 (1.24-1.47)	1.27 (1.16-1.39)
1+ severe LRTI n=6,932	813 (6.9)	4.01 (3.73-4.31)	2.32 (2.15-2.50)

Legend: AECOPD (Acute Exacerbations of Chronic Obstructive Pulmonary Disease), LRTI (Lower Respiratory Tract Infection), HR (Hazard ratio). *Adjusted for all

baseline covariates listed in Table 1. Number of patients included crude and fully adjusted models N= 215,234.

7

Table S6: Hazard ratios for CVD-related mortality following COPD diagnosis in different categories of baseline frequency and severity of LRTI

Baseline LRTI category	n (%) of events N=13,116	Crude HR (95% CI)	Adjusted* HR (95% CI)
No LRTI n=158,074	9,325 (71.1)	Ref	Ref
1 mild LRTI n=18,081	875 (6.7)	1.08 (1.01-1.16)	1.06 (0.99-1.14)
2+ mild LRTI n=6,133	291 (2.2)	1.16 (1.03-1.30)	1.17 (1.04-1.31)
1 moderate LRTI n=19,787	1,432 (10.9)	0.99 (0.93-1.04)	1.03 (0.98-1.09)
2+ moderate LRTI n=6,227	544 (4.1)	1.17 (1.08-1.28)	1.20 (1.10-1.31)
1+ severe LRTI n=6,932	649 (4.9)	2.52 (2.33-2.73)	1.68 (1.55-1.82)

Legend: AECOPD (Acute Exacerbations of Chronic Obstructive Pulmonary Disease), LRTI (Lower Respiratory Tract Infection), HR (Hazard Ratio). *Adjusted for all

baseline covariates listed in Table 1. Number of patients included crude and fully adjusted models N= 215,234.

8

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance
placed on this supplemental material which has been supplied by the author(s)

Thorax

Table S7: Incidence rates of at least one AECOPD (moderate or severe) following COPD diagnosis in different categories of baseline frequency and severity of LRTI (mild and moderate GP events combined)

Baseline LRTI category	n (%) of events N=948,576	Rate 100py^{-1} (95% CI)	Crude IRR (95% CI)	Adjusted* IRR (95% CI)
No LRTI n=158,074	593,917	71.7 (71.5 – 71.9)	Ref	Ref
1 GP LRTI n=34,849	197,612	108.2 (107.7 – 108.7)	1.52 (1.20-1.55)	1.48 (1.46-1.50)
2+ GP LRTI n=10,315	75,068	136.7 (135.7 – 137.7)	1.95 (1.91-2.00)	1.83 (1.79-1.88)
3+ GP LRTI n=5,064	51,646	192.4 (190.7 – 194.1)	2.88 (2.78-2.97)	2.53 (2.45-2.61)
1+ HES LRTI n=6,932	30,33	130.8 (129.4 – 132.3)	2.05 (1.99-2.12)	1.76 (1.70-1.81)

Legend: AECOPD (Acute Exacerbations of Chronic Obstructive Pulmonary Disease), LRTI (Lower Respiratory Tract Infection), IRR (Incidence Rate Ratio). PY (Person years). *Adjusted for all baseline covariates listed in Table 1. Number of patients included crude and fully adjusted models N= 215,234.

9

Figure S1: Inclusion of COPD patients

COPD patients aged over 40 years old with CPRD Aurum data between 2004 and 2019
N=431,917

↓

COPD patients with HES linkage
N=355,696

↓

COPD patients with current or ex-smoking status
N=340,515

↓

Patients with a COPD diagnosis at least one year after the 1st January 2004 and the data at which they turned 40
N=215,234
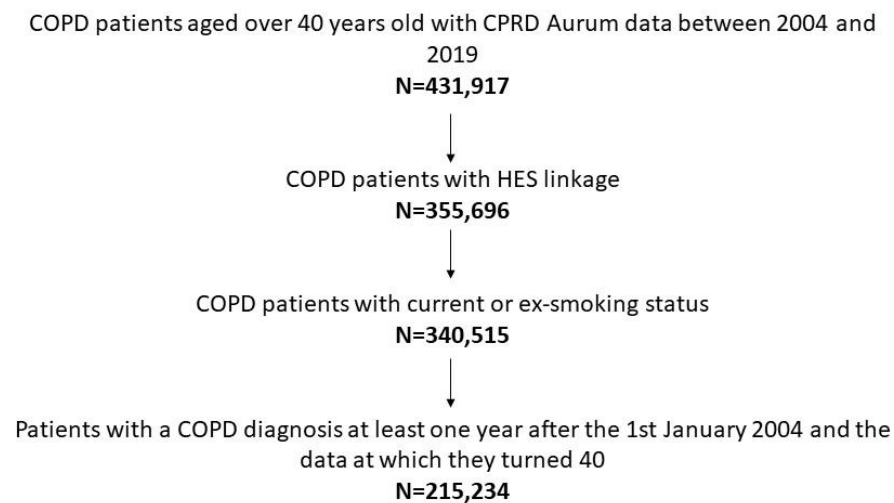
Legend: CPRD (Clinical Practice Research Datalink); HES (Hospital Episode Statistics)

10

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Thorax

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No.	Recommendation	Page No.
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4-5
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5-6
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	5-6
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	n/a
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5-6
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	*5-6*
Bias	9	Describe any efforts to address potential sources of bias	5-6
Study size	10	Explain how the study size was arrived at	Figure S1

Continued on next page

1

Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	7-8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	7-8
		(*b*) Describe any methods used to examine subgroups and interactions	7-8
		(*c*) Explain how missing data were addressed	7-8
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	7-8
		(*e*) Describe any sensitivity analyses	8
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	9 & fig S1
		(b) Give reasons for non-participation at each stage	Fig S1
		(c) Consider use of a flow diagram	Fig S1
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Table 1 & 2
		(b) Indicate number of participants with missing data for each variable of interest	Table 1 & 2
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	Table 1
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	*Table3 & 4*
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Table 3 & 4
		(*b*) Report category boundaries when continuous variables were categorized	All tables
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Table 3& 4

2

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	Tables in supplementary material
Discussion			
Key results	18	Summarise key results with reference to study objectives	11
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	14
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	11-13
Generalisability	21	Discuss the generalisability (external validity) of the study results	11-13
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	22

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

3