



Health and Social Care Delivery Research

Volume 12 • Issue 43 • October 2024

ISSN 2755-0079

What happens between first symptoms and first acute exacerbation of COPD – observational study of routine data and patient survey

Alex Bottle, Alex Adamson, Xiubin Zhang, Benedict Hayhoe and Jennifer K Quint



What happens between first symptoms and first acute exacerbation of COPD – observational study of routine data and patient survey

Alex Bottle^{1*}, Alex Adamson¹, Xiubin Zhang¹,
Benedict Hayhoe¹ and Jennifer K Quint¹

¹School of Public Health, Imperial College London, London, UK

*Corresponding author

Published October 2024
DOI: 10.3310/CGTR6370

This report should be referenced as follows:

Bottle A, Adamson A, Zhang X, Hayhoe B, Quint JK. What happens between first symptoms and first acute exacerbation of COPD – observational study of routine data and patient survey. *Health Soc Care Deliv Res* 2024;**12**(43). <https://doi.org/10.3310/CGTR6370>

Health and Social Care Delivery Research

ISSN 2755-0079 (Online)

A list of Journals Library editors can be found on the [NIHR Journals Library website](#)

Health and Social Care Delivery Research (HSDR) was launched in 2013 and is indexed by Europe PMC, DOAJ, INAHTA, Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA), NCBI Bookshelf, Scopus and MEDLINE.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

This journal was previously published as *Health Services and Delivery Research* (Volumes 1–9); ISSN 2050-4349 (print), ISSN 2050-4357 (online)

The full HSDR archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hsdr.

Criteria for inclusion in the *Health and Social Care Delivery Research* journal

Manuscripts are published in *Health and Social Care Delivery Research* (HSDR) if (1) they have resulted from work for the HSDR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

HSDR programme

The HSDR programme funds research to produce evidence to impact on the quality, accessibility and organisation of health and social care services. This includes evaluations of how the NHS and social care might improve delivery of services.

For more information about the HSDR programme please visit the website at <https://www.nihr.ac.uk/explore-nihr/funding-programmes/health-and-social-care-delivery-research.htm>

This article

The research reported in this issue of the journal was funded by the HSDR programme or one of its preceding programmes as award number 17/99/72. The contractual start date was in June 2019. The draft manuscript began editorial review in March 2023 and was accepted for publication in December 2023. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HSDR editors and production house have tried to ensure the accuracy of the authors' manuscript and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this article.

This article presents independent research funded by the National Institute for Health and Care Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the HSDR programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the HSDR programme or the Department of Health and Social Care.

This article was published based on current knowledge at the time and date of publication. NIHR is committed to being inclusive and will continually monitor best practice and guidance in relation to terminology and language to ensure that we remain relevant to our stakeholders.

Copyright © 2024 Bottle *et al.* This work was produced by Bottle *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: <https://creativecommons.org/licenses/by/4.0/>. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Newgen Digitalworks Pvt Ltd, Chennai, India (www.newgen.co).

Abstract

What happens between first symptoms and first acute exacerbation of COPD – observational study of routine data and patient survey

Alex Bottle^{1*}, Alex Adamson¹, Xiubin Zhang¹,
Benedict Hayhoe¹ and Jennifer K Quint¹

¹School of Public Health, Imperial College London, London, UK

*Corresponding author robert.bottle@imperial.ac.uk

Background: Chronic obstructive pulmonary disease affects nearly 400 million worldwide – over a million in the United Kingdom – and is the third leading cause of death. However, there is limited understanding of what prompts a diagnosis, how long this takes from symptom onset and the different approaches to clinical management by primary care professionals.

Objectives: Map out the clinical management and National Health Service contacts from symptom presentation to chronic obstructive pulmonary disease diagnosis and first acute exacerbation of chronic obstructive pulmonary disease in three time periods; construct risk prediction for first acute exacerbation of chronic obstructive pulmonary disease.

Design: Retrospective cohort study and cross-sectional survey.

Setting: Primary care.

Participants: Patients with incident chronic obstructive pulmonary disease aged > 35 years in England.

Interventions: None.

Main outcome measures: First acute exacerbation of chronic obstructive pulmonary disease.

Data sources: Clinical Practice Research Datalink Aurum; new online survey.

Results: Forty thousand five hundred and seventy-seven patients were diagnosed between April 2006 and March 2007 (cohort 1), 48,249 between April 2016 and March 2017 (cohort 2) and 4752 between March and August 2020 (cohort 3). The mean (standard deviation) age was 68.3 years (12.0); 47.3% were female. Around three-quarters were diagnosed in primary care, with a slight fall in cohort 3. Compliance with National Institute for Health and Care Excellence diagnostic guidelines was slightly higher in cohorts 2 and 3 for all patients; 35.8% (10.0% in the year before diagnosis) had all four elements met for all cohorts combined. Multilevel modelling showed considerable between-practice variation in spirometry. The survey on the charity website had 156 responses by chronic obstructive pulmonary disease patients. Many respondents had not heard of the condition, hoped the symptoms would go away and identified various healthcare-related barriers to earlier diagnosis. Clinical Practice Research Datalink analysis showed notable changes in post-diagnosis prescribing from cohort 1 to 2, such as increases in long-acting muscarinic antagonist (21.7–46.3%). Triple therapy rose from 2.9% in cohort 2 to 11.1% in cohort 3. Documented pulmonary rehabilitation rose from just 0.8% in cohort 1 to 13.7% in cohort 2 and 20.9% in cohort 3. For all patients combined, the median time to first acute exacerbation of chronic obstructive pulmonary disease in patients who had one was 1.4 years in cohorts 1 and 2. Acute exacerbation of chronic obstructive pulmonary disease prediction models identified some

consistent predictors, such as age, deprivation, severity, comorbidities, post-diagnosis spirometry and annual review. Models without post-diagnosis general practitioner actions had a c-statistic of around 0.70; the highest c-statistic was 0.81, for cohort 2 with post-diagnosis general practitioner actions and 6-month follow-up. All models had good calibration. The three most important predictors in terms of their population attributable risks were being a current smoker and offered smoking cessation advice (32.8%), disease severity (30.6%) and deprivation (15.4%). The highest population attributable risks for variables with adjusted hazard ratios < 1 were chronic obstructive pulmonary disease review (-27.3%) and flu vaccination (-26.6%).

Limitations: Symptom recording and chronic obstructive pulmonary disease diagnosis vary between practice; predicted forced expiratory volume in 1 second had many missing values.

Conclusions: There has been some improvement over time in chronic obstructive pulmonary disease diagnosis and management, with large changes in prescribing, though patient and system barriers to further improvement exist. Data available to general practitioners cannot generate risk prediction models with sufficient accuracy.

Future work: It will be important to expand the COVID-era cohort with longer follow-up and augment general practitioner data for better prediction.

Study registration: This study is registered as Researchregistry.com: researchregistry4762.

Funding: This award was funded by the National Institute for Health and Care Research (NIHR) Health and Social Care Delivery Research programme (NIHR award ref: 17/99/72) and is published in full in *Health and Social Care Delivery Research*; Vol. 12, No. 43. See the NIHR Funding and Awards website for further award information.

Contents

List of tables	ix
List of figures	xi
List of abbreviations	xiii
Plain language summary	xv
Scientific summary	xvii
Chapter 1 Background and research objectives	1
Overall aims	2
Objectives	2
Chapter 2 Methods	3
Clinical Practice Research Datalink database	3
Air quality data set	3
Patient cohorts, definition of index date and acute exacerbation	3
Definition of patient characteristics and predictors	4
Statistical methods	4
<i>Route to diagnosis and first acute exacerbations of chronic obstructive pulmonary disease</i>	4
Modelling the risk of the first acute exacerbations of chronic obstructive pulmonary disease	6
Patient surveys and focus groups	7
<i>Focus groups</i>	8
<i>Analysis of the patient surveys</i>	8
<i>Analysis of the focus groups</i>	8
Chapter 3 Results	9
Cohort identification	9
Descriptive analysis	9
Route to diagnosis	9
Initial management following diagnosis in primary care	20
Time between diagnosis and first acute exacerbation	24
Chapter 4 Prediction of first acute exacerbation	33
Validation of risk prediction model	39
Potential cost savings for a population of 250,000	39
Chapter 5 Qualitative component	43
Twitter online survey results	43
British Lung Foundation online survey results	43
Clinical Practice Research Datalink online survey results	44
Patient focus group analysis	44
Chapter 6 Discussion	45
Main findings	45
Comparison with the literature	46
Strengths and limitations	48

CONTENTS

Equality, diversity and inclusion	49
<i>Participant representation</i>	49
<i>Reflections on our research team and wider involvement</i>	49
Patient and public involvement	49
Recommendations for future research	50
Implications for decision-makers	50
Chapter 7 Conclusions	53
Chapter 8 Issues faced by this study and lessons learnt	55
Additional information	57
References	61
Appendix 1 Survey of patient experience of getting diagnosed with chronic obstructive pulmonary disease	65
Appendix 2 Additional rows from the main tables	77

List of tables

TABLE 1 Patient characteristics of the three cohorts	10
TABLE 2 Patient characteristics by where diagnosed and cohort	13
TABLE 3 Compliance with NICE guidelines for COPD diagnosis: patients with no pre-existing asthma or HF	17
TABLE 4 General practitioner-recorded symptoms in the 5 years before diagnosis	19
TABLE 5 Prescribing in the 5 years before diagnosis by cohort	21
TABLE 6 Results of multilevel analysis for use of spirometry around the time of COPD diagnosis	22
TABLE 7 Funnel plot outlier practices for spirometry within 6 months prior to/after COPD diagnosis by cohort and diagnosis setting	24
TABLE 8 General practitioner actions in the year following the COPD diagnosis	25
TABLE 9 Characteristics of patients diagnosed via their first AE, by cohort	26
TABLE 10 General practitioner actions before diagnosis in patients diagnosed via their first AE, by cohort	29
TABLE 11 Time to first AECOPD in patients by asthma and HF status	31
TABLE 12 Hazard ratios for model 2 for each cohort for first acute exacerbation prediction	34
TABLE 13 Hazard ratios for model 3 for each cohort for first acute exacerbation prediction	36
TABLE 14 Cohort 2's PARs from the model including post-diagnosis GP actions	40
TABLE 15 Potential cost savings due to AE reduction per 5000 COPD patients	41
TABLE 16 Further GP-recorded symptoms before diagnosis	78

List of figures

FIGURE 1 Flow diagram for inclusion/exclusion of patients	5
FIGURE 2 Funnel plot of proportion of patients following the initial NICE pathways (performed spirometry within 6 months prior to/after COPD diagnosis) by number of COPD patients per practice: cohort 1, patients diagnosed in primary care	23
FIGURE 3 Funnel plot of proportion of patients following the initial NICE pathways (performed spirometry within 6 months prior to/after COPD diagnosis) by number of COPD patients per practice: cohort 2, patients diagnosed in primary care	23
FIGURE 4 Cumulative incidence of AECOPD (time 0 means that COPD and AECOPD were on the same day)	32
FIGURE 5 Cumulative incidence of all-cause mortality following COPD diagnosis by cohort	32
FIGURE 6 Calibration for the final Cox model for cohort 2 at 6 months since diagnosis	39

List of abbreviations

A&E	accident and emergency	IMD	Index of Multiple Deprivation
AE, AECOPD	acute exacerbation of chronic obstructive pulmonary disease	ISAC	Independent Scientific Advisory Committee
aHR	adjusted hazard ratio	LABA	long-acting beta agonist
AMI	acute myocardial infarction	LAMA	long-acting muscarinic antagonist
BMI	body mass index	MOR	median odds ratio
COPD	chronic obstructive pulmonary disease	OCS	oral corticosteroids
CPRD	Clinical Practice Research Datalink	ONS	Office for National Statistics
ED	emergency department	PAR	population attributable risk
GOLD	Global Initiative for Obstructive Lung Disease	PPI	patient and public involvement
GP	general practitioner	PR	pulmonary rehabilitation
HES	Hospital Episode Statistics	QOF	Quality and Outcomes Framework
HF	heart failure	RA	research assistant
ICC	intraclass correlation coefficient	SABA	short-acting beta agonist
ICS	inhaled corticosteroids	SAMA	short-acting muscarinic antagonist

Plain language summary

Chronic obstructive pulmonary disease is often caused by smoking and affects over 1 million people in the United Kingdom. While there are well-established treatments, less is known on where and when patients get the diagnosis, how general practitioners investigate their symptoms and to what extent the first major flare-up ('acute exacerbation') can be predicted and prevented.

Using a research database of general practitioner consultation records linked to hospital admissions and the national death register, we described patient characteristics, general practitioner actions before and following diagnosis, and, with statistical models, predictors of the first exacerbation. We looked at three time periods according to the date of diagnosis: April 2006–March 2007 (cohort 1), April 2016–March 2017 (cohort 2) and March–August 2020 (cohort 3). We sent patients a questionnaire asking about their experiences of developing symptoms, seeking medical help and getting diagnosed.

We analysed records of over 70,000 patients in total. The majority were diagnosed by their general practitioner. In cohorts 2 and 3, general practitioners did the recommended tests more than in cohort 1, though in the year before diagnosis, only 10% of patients had all four done. Our survey found that many people were unaware of chronic obstructive pulmonary disease and its symptoms before their diagnosis but also that some felt they were not taken seriously by the medical team and that their diagnosis was delayed. There were improvements over time in prescribing. Most patients were offered the flu jab. Older patients, current smokers and those with other conditions such as heart failure had higher risk of an acute exacerbation. The statistical models did not perform well enough to be used to guide decision-making.

Despite some improvements over time, there remain opportunities for better recognition of the condition among patients and general practitioners alike. Future work should more fully assess the impact of COVID-19.

Scientific summary

Background

Chronic obstructive pulmonary disease (COPD) affects nearly 400 million worldwide – over a million in the UK – and is the third leading cause of death. Despite this, there is limited understanding of what prompts a diagnosis, how long this takes from symptom onset and the different approaches to clinical management taken by primary care professionals. This is particularly true regarding people with comorbidities such as asthma and heart failure (HF) that can also cause breathlessness.

Objectives

Map out the clinical management and NHS contacts from symptom presentation to COPD diagnosis and first acute exacerbation, acute exacerbation of chronic obstructive pulmonary disease (AECOPD) (for some patients the latter two will be the same event); investigate whether and how this varied in three cohorts since 2006; rank predictors of the first AECOPD in importance and assess whether and how this changed over time; construct and validate risk prediction models for the first AECOPD.

Methods

The project involved the quantitative analysis of an existing database and a new survey. The main component used the Clinical Practice Research Datalink (CPRD), which collects deidentified patient electronic health records from participating general practitioner (GP) practices; its Aurum version includes healthcare records from GP practices using EMIS® software, representing around 13% of the population in England. It includes patient-level data on demographics, tests, symptoms, diagnoses, therapies, prescriptions and referrals to secondary care. Patient-level data from these practices were linked by CPRD staff to the Office for National Statistics death register, Hospital Episode Statistics and Index of Multiple Deprivation at small area level.

We included all individuals aged over 35 years with COPD diagnosed between 1 April 2006 and 31 March 2007 (cohort 1) and between 1 April 2016 and 31 March 2017 (cohort 2); a smaller COVID-era group for March–August 2020 made up cohort 3. For each patient, the index (diagnosis) date was defined as the first record of COPD, either in primary care records via SNOMED-CT codes or in hospital admission data via *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision* codes. Patients had to have at least 1 year's registration with the GP before the diagnosis date. Patient characteristics were described for the year up to and including the diagnosis date; GP actions were also described in the year before and since diagnosis. The first AECOPD was identified using our group's published algorithm and was restricted to hospital admissions to reliably capture the most serious ones. Much of the analysis was descriptive, including cumulative incidence plots for the time to first AECOPD by cohort. GP actions before diagnosis were compared with the National Institute for Health and Care Excellence (NICE) guidelines, which recommend spirometry, chest X-ray, full blood count (FBC) and the calculation of body mass index (BMI). These 'routes to diagnosis' analyses were stratified by pre-existing HF and asthma, conditions that share some symptoms and that could cause diagnostic confusion. Multilevel models assessed the variation between GP practices in the proportion of patients receiving spirometry in the 6 months prior to or after diagnosis; funnel plots were used to count statistical outliers at 2 and 3 standard deviations (SD) from the mean. Prescriptions for the main classes of medication were noted for the year following diagnosis. A set of Fine and Gray regression models quantified the association between patient characteristics, GP actions and first AECOPD in patients not diagnosed via an AECOPD, accounting for the competing risk of death from non-COPD

causes: model 1 contained only patient factors, model 2 additionally contained pre-diagnosis GP actions and model 3 also included post-diagnosis GP actions. Population attributable risks (PARs) were calculated for statistically significant risk factors. The focus for the reporting of the model outputs was on the model containing patient characteristics and pre-diagnosis GP actions.

We developed an online survey to investigate COPD patients' retrospective perceptions of their initial symptoms, what they did after developing those symptoms, what kind of professional advice was sought and year of diagnosis in order to distinguish between COVID and pre-COVID eras. It was designed jointly through a series of discussions by the project team at Imperial College London, which included researchers and patient representatives, and the teams at Asthma + Lung UK and the Taskforce for Lung Health, including its own patient advisory group. This was administered via the charity and GPs contributing to CPRD.

Results

Cohort 1 had 31,676 patients, cohort 2 had 37,393 and cohort 3 had 4752. Overall, the mean age was 68.3 years (SD 12.0), and 47.3% were female; the age-sex mix did not change over time with 82.7% being current or ex-smokers. Common comorbidities included hypertension, anxiety, depression, asthma, stroke, diabetes and renal disease, with an average of nearly four per patient. Around three-quarters were diagnosed in primary care, with a slight fall in this proportion in cohort 3. Those diagnosed this way were older, with lower blood pressure and had more comorbidities, with higher levels in recorded prevalence for anxiety, depression and diabetes being among the most notable. Nearly half of all patients had had a lower respiratory tract infection recorded in the 5 years before diagnosis, with presentations for other symptoms also common.

Compliance with NICE diagnostic guidelines was slightly higher in cohorts 2 and 3 for all patients; 35.8% (10.0% in the year before diagnosis) had all four elements met overall. Spirometry in the year before diagnosis rose from 55.7% in cohort 1 to 63.8% in cohort 2 but then fell to 37.4% in cohort 3; around a third had a chest X-ray, half had a FBC and half had a BMI measurement in this time. The use of pre-diagnosis echocardiography, cardiology referral and B-type natriuretic peptide (BNP) testing rose considerably over time, though BNP recording remained low despite its now widespread availability. In the 5 years before diagnosis, 36% were prescribed inhaled steroids (similar for each cohort and for oral steroids); 61% were prescribed short-acting beta agonist (SABA); 10% long-acting muscarinic antagonist (LAMA) (this rose to 20% for cohort 3). Compared with all patients combined, patients diagnosed through an emergency hospitalisation were less likely to have had pre-diagnosis spirometry in cohorts 1 and 2, but equally likely in cohort 3. They were much more likely to have had a chest X-ray and FBC in each cohort.

In cohort 1, practices had a median of 23 new COPD patients, with 18 diagnosed in primary care; these figures were 28 and 21 for cohort 2. There was considerable variation between practices in spirometry use, with median odds ratios of 1.5 or more. For patients diagnosed in primary care, 24.5% of practices in cohort 1 and 19.7% in cohort 2 were funnel plot outliers at 2 SD.

The survey on the charity website had 156 responses by COPD patients, of whom 124 (79.5%) were female and 82.7% were aged between 45 and 74 years. Many respondents had not heard of the condition, hoped the symptoms would go away, and identified healthcare-related barriers to earlier diagnosis such as difficulty in getting an appointment and the impression of not always being taken seriously by staff. The response rate from CPRD GPs was too low to allow analysis.

In the year following diagnosis, there were notable changes in prescribing from cohort 1 to 2, such as increases in LAMA (21.7–46.3%) and long-acting beta agonist (10.1–15.9%) and falls in short-acting muscarinic antagonist (12.4–2.1%) and inhaled corticosteroids (52.6–41.1%). SABA use changed little. These were maintained into the COVID era. Triple therapy rose from 2.9% in cohort 2 to 11.1% in cohort 3. Around four in five patients in each cohort were offered the influenza vaccine, with two-thirds

receiving it from the practice. Documented pulmonary rehabilitation (PR) rose from just 0.8% in cohort 1 to 13.7% in cohort 2 and 20.9% in cohort 3. Smoking cessation drug prescription fell over time, whereas advice fluctuated.

For all patients combined, the median time to first AECOPD in those who had one was 1.4 years in cohorts 1 and 2. This was generally slightly shorter in patients with HF and/or asthma. Those with HF, but not those with asthma, had much higher exacerbation rates. The all-cause death rate was notably higher for cohort 3 than the other two. The AECOPD prediction models consistently identified a number of predictors, including age (but not sex), deprivation, COPD severity, current smoking and various comorbidities such as osteoporosis, asthma and depression; other comorbidities, such as HF and diabetes with complications, were less consistently significant. Some medications, particularly LAMA, were associated with higher hazards. In the model including post-diagnosis GP actions, annual COPD review [cohort 2 adjusted hazard ratio (aHR) 0.63, 95% confidence interval (CI) 0.58 to 0.69], post-diagnosis spirometry (cohort 2 aHR 0.83, 95% CI 0.78 to 0.88) and flu vaccination (cohort 2 aHR 0.56, 95% CI 0.53 to 0.60) all had hazards < 1, but we did not find any significant association for specialist referral (cohort 2 aHR 1.10, 95% CI 0.99 to 1.20) or PR (cohort 2 aHR 1.04, 95% CI 0.96 to 1.13). Discrimination was moderate, with *c*-statistics of around 0.70 for models without post-diagnosis GP actions. The highest *c*-statistic was 0.81 and was obtained for model 3, with post-diagnosis GP actions, at 6 months of follow-up. Calibration was good for all models. For cohort 2, when the model with post-diagnosis GP actions was simplified to include only predictors with *p* < 0.05, the three most important predictors in terms of their PARs were being a current smoker (32.8%), Global Initiative for Obstructive Lung Disease severity (30.6%) and deprivation (15.4%). The highest PARs for variables with aHRs < 1 were COPD review (-27.3%) and flu vaccination (-26.6%). For a typical local population with 5000 COPD patients, these PARs translate into average potential annual cost savings of £193K for COPD review and £188K for flu vaccination.

Conclusions

There have been several improvements over time in NICE diagnostic guideline compliance, prescribing and referral for PR, but much more improvement is desirable, and there remains much variation between GP practices in spirometry use. There is also much unawareness of the condition among UK adults at risk of it. Data currently available in GP information technology systems are not enough to predict someone's first AECOPD with sufficient accuracy to guide shared decision-making.

Our recommendations for research are: (1) understand how and develop approaches to overcome NHS and patient barriers to earlier diagnosis; (2) seek strategies to reduce unwarranted variation in spirometry use between practices; (3) repeat the analysis on more data since March 2020 and with longer follow-up; (4) assess variations in prescribing between practices; and (5) evaluate the statistical reliability of practice-level spirometry as a potential quality indicator.

Study registration

This study is registered as Researchregistry.com: researchregistry4762.

Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Health and Social Care Delivery Research programme (NIHR award ref: 17/99/72) and is published in full in *Health and Social Care Delivery Research*; Vol. 12, No. 43. See the NIHR Funding and Awards website for further award information.

Chapter 1 Background and research objectives

Chronic obstructive pulmonary disease (COPD) affects nearly 400 million worldwide and is the third leading cause of death. According to the Global Burden of Disease project, COPD accounted for 3.3 million deaths and 74.4 million disability-adjusted life-years in 2019.¹ Affecting over a million people in the UK, it accounts for 30,000 deaths each year and puts the UK among the top 20 countries for COPD mortality worldwide.^{2,3} Over 110,000 people are diagnosed with COPD each year, and the number of people living with the disease is rising.⁴ However, there is limited understanding of what prompts a diagnosis, how long this takes from symptom onset and the different approaches to clinical management taken by primary care professionals. This is particularly true regarding people with comorbidities such as asthma and heart failure (HF) that can also cause breathlessness. A study using primary care records for 2000–9 found an improvement in COPD management and outcomes; however, they concluded that improvements in diagnosis were only modest during the period, and they were not able to look at exacerbations or hospitalisations.⁵ Another study using electronic patient records and patient-completed questionnaires concluded that COPD is not being treated in accordance with the Global Initiative for Chronic Obstructive Lung Disease (GOLD) or National Institute for Health and Care Excellence (NICE) guidelines in primary care.⁶ Local estimates of COPD prevalence are now published online by what was Public Health England, but overall, there is little information as to what extent the NHS is meeting the needs of current patients with COPD and how well it might meet those of future patients.

To improve outcomes, management of COPD should be better tailored to each patient, as recommended by the National Audit.⁷ One approach for personalising COPD treatment is to stratify patients according to the risk of acute exacerbations of chronic obstructive pulmonary disease (AECOPD) to prescribe treatments such as inhaled corticosteroids (ICS) or phosphodiesterase-4 inhibitors earlier. AECOPDs are responsible for the majority of the disease burden, contribute to the progressive decline in lung function and reduce patients' quality and quantity of life.^{8,9} It is also an important predictor of mortality and the second most common reason for emergency hospital admission in the UK.^{10,11} For many patients, AECOPD is diagnosed at the same time as COPD is. Early diagnosis and information on who is at higher risk of AECOPD would not only help shared decision-making between the practice clinicians (GP and practice nurses) but could also significantly reduce the burden borne by patients. A well-performing risk prediction model that uses information available to the clinician would inform clinical decision-making and timely management. This is particularly important to avoid the patient's first AECOPD, as each exacerbation damages the lungs and treatment is less effective thereafter, and the greatest predictor of subsequent exacerbations is having had one.^{9,12,13} Using trial data, a single moderate to severe exacerbation has been shown to cause a decline in post-bronchodilator lung function.¹⁴ Several studies tried to model the risk of AECOPD previously, though not successfully. One study using a number of different data sources found that one of the predictors for mortality for COPD patients was the number of GPs per 1000 patients.¹⁵ In addition, the authors found that the number of outpatient appointments attended and missed were strong predictors of mortality and readmissions. However, that study was limited to using practice-level information on primary care management. A recent systematic review of risk prediction models for AECOPD concluded that none of the existing models satisfied the requirements for risk-stratified treatment and personalised COPD care.¹⁶ Nine out of the 27 models included in the review included previous exacerbations as a predictor and are therefore of limited relevance to the prediction of the first one. The other main limitations of the models were problematic variable-selection procedures and the lack of external validation.

Since this project began in mid-2019, the COVID-19 pandemic has of course affected the nature and delivery of NHS healthcare provision and patient health-seeking behaviour, particularly during the early waves before the omicron variant became dominant. A systematic review of nine studies in nine countries up to May 2021 calculated a pooled rate ratio of hospital admissions for COPD exacerbations during the pandemic period compared with before it to be 0.50 (95% CI 0.44 to 0.57).¹⁷ The first 30 weeks in Scotland and Wales saw large falls in emergency department (ED) visits or hospital

admissions and also in AECOPDs in primary care, but with no rise in COPD deaths. This was followed by a gradual rise in ED visits and admissions after the end of the first lockdown in 2020.¹⁸ Also important was the lack of access to spirometry in primary care, documented in the most recent national audit, which used data for 314 general practices in Wales (80.7% of the country's practices) up to the end of July 2021 and noted that just 1.9% of patients had received post-bronchodilator spirometry in the previous 2 years.¹⁹ In view of these changes, we decided to add a third cohort to our analysis to cover the early COVID era.

Overall aims

The first aim was to describe and model the patient journey from symptom presentation to diagnosis and first acute exacerbation for COPD patients in England. The second aim was to investigate how patients obtain their COPD diagnosis (the 'route to diagnosis'), how they are managed in primary care and how they get their first AECOPD.

Objectives

The objectives of the project were:

1. Map out the clinical management and NHS contacts from symptom presentation to COPD diagnosis and first AECOPD (for some patients, the latter two will be the same event).
2. Investigate whether and how this varied in three cohorts since 2006.
3. Rank predictors of the first AECOPD in importance and assess whether and how this changed over time.
4. Construct and validate risk prediction models for the first AECOPD.

Chapter 2 Methods

The project principally involved the quantitative analysis of an existing database and a new survey.

Clinical Practice Research Datalink database

The Clinical Practice Research Datalink (CPRD) collects anonymised patient electronic health records from general practitioner (GP) practices using the Vision® or EMIS® software systems.²⁰⁻²³ CPRD GOLD contains data contributed by GP practices using Vision software. It covers approximately 7% of the UK population, with 674 participating GP practices and over 11.3 million patients (historical and current). CPRD Aurum includes healthcare records from GP practices using EMIS software, representing around 13% of the population in England. Both CPRD primary care databases include patient-level data on demographics, tests, symptoms, diagnoses, therapies, prescriptions and referrals to secondary care. All patients registered with the practices are included in the database unless they have requested to opt out of sharing their information for research purposes.²⁴ Data from patients from a subset of practices in England in CPRD GOLD and from all practices in CPRD Aurum can be linked to a range of other data sources. In particular, patient-level data from these practices can be linked to Office for National Statistics (ONS) death registration data, Hospital Episode Statistics (HES) data sets, small area-level data [Index of Multiple Deprivation (IMD)], cancer data and the Mental Health Services Data set. Other linkages are available on request (e.g. monitor identifiers for pollution/temperature monitoring stations).

The Independent Scientific Advisory Committee (ISAC) of the Medicine and Healthcare product Regulatory Agency database research approved this study (ISAC protocol no: 23_003056).

Air quality data set

Daily temperature and pollution data were obtained from the British Atmospheric Data Centre and the Department for Environment, Food and Rural Affairs. The latter source provides not only temperature data but also major pollutants that are believed to contribute to the risk of AECOPD: nitrogen dioxide, sulphur dioxide, ozone, carbon monoxide, particulate matter of size below 10 µm (PM10) and particulate matter of size below 2.5 µm (PM2.5). These data are updated daily and are available via R package *openair*. However, due to data extraction problems with CPRD, we used Aurum extracted for another project, which could not be linked to these temperature and pollution data. This and the other problems encountered during the project are discussed after the [Conclusions](#) section.

Patient cohorts, definition of index date and acute exacerbation

To make sure that we included patients with their first COPD record in either primary or secondary care, we obtained records for all patients aged over 35 years with a diagnosis of COPD in CPRD Aurum and also obtained records for patients who were diagnosed with COPD in HES who were linked to the Aurum data set. For each patient, the index (diagnosis) date was defined as the first record of COPD, either in the primary care record via SNOMED-CT codes or of AECOPD in the hospital admission data via *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision* (ICD-10) codes J44.9 (first position), J22 and J44.9 together in the first and second position or J44.0 or J44.1 in any position as per our published algorithms.^{25,26} We included all patients with an index date between 1 January 2006 and 31 December 2007 (cohort 1) and between 1 January 2016 and 31 December 2017

(cohort 2); a smaller COVID-era group for 1 March–31 August 2020 made up cohort 3. Patient-level data were extracted based on the following criteria:

- patients over 35 years who were flagged as having acceptable records in terms of data quality by standard CPRD criteria;
- patients who were registered at the current practice for at least 1 year pre diagnosis;
- practices with consent to linkage to HES;
- patients who were eligible for linkage with the following data sets:
 - HES admitted patient care (APC);
 - HES accident and emergency (A&E);
 - IMD data;
 - ONS death registration data.

Hospital Episode Statistics is an English data set, and so linkage with HES is only possible for English GP practices. In the last decade, there was a shift of GP practices within England moving from clinical computer system Vision (used for the CPRD GOLD data set) towards other available systems such as EMIS (used for the CPRD Aurum data set). By 2016, only 9% of GP practices used Vision software, which is 50% less compared with 2010–1,²¹ so the analysis was implemented using the Aurum data set. [Figure 1](#) shows the full flow chart for inclusion and exclusion of patients.

Definition of patient characteristics and predictors

Demographics were defined at diagnosis, whereas patient physiological characteristics and behaviours were defined using the most recent data at any time before diagnosis. Comorbidities were identified from the patient's records up to and including the date of diagnosis. COPD-related respiratory symptoms were sought up to 5 years before and including the date of diagnosis.

Statistical methods

Much of the initial analysis was descriptive, with patient characteristics summarised and compared between cohorts using standard basic tests. Times to diagnosis and times to first AECOPD were summarised using medians and interquartile ranges and by cumulative incidence plots.

Route to diagnosis and first acute exacerbations of chronic obstructive pulmonary disease

There have been several relevant NICE clinical guidelines. The first that is relevant to this study was published in 2004; relevant to the COVID-19 era, NG115 was published in December 2018.²⁷ They recognise that there is no single diagnostic test, though spirometry is essential to assess airway obstruction and that distinction from asthma is crucial. Both state that a COPD diagnosis should be considered in patients over the age of 35 years who have a risk factor (generally smoking) and who present with one or more of the five symptoms mentioned previously.²⁷ The guidelines recommend performing the following for patients presented with suspected COPD in primary care:

1. spirometry
2. chest radiograph to exclude other pathologies
3. full blood count to identify anaemia or polycythaemia
4. calculation of body mass index (BMI).

Additional investigations such as those related to HF, asthma medications or respiratory/COPD referrals are advised as optional. We characterised the route to COPD diagnosis in primary care through investigations, referrals and treatments, noting what proportion followed the NICE guidelines and how

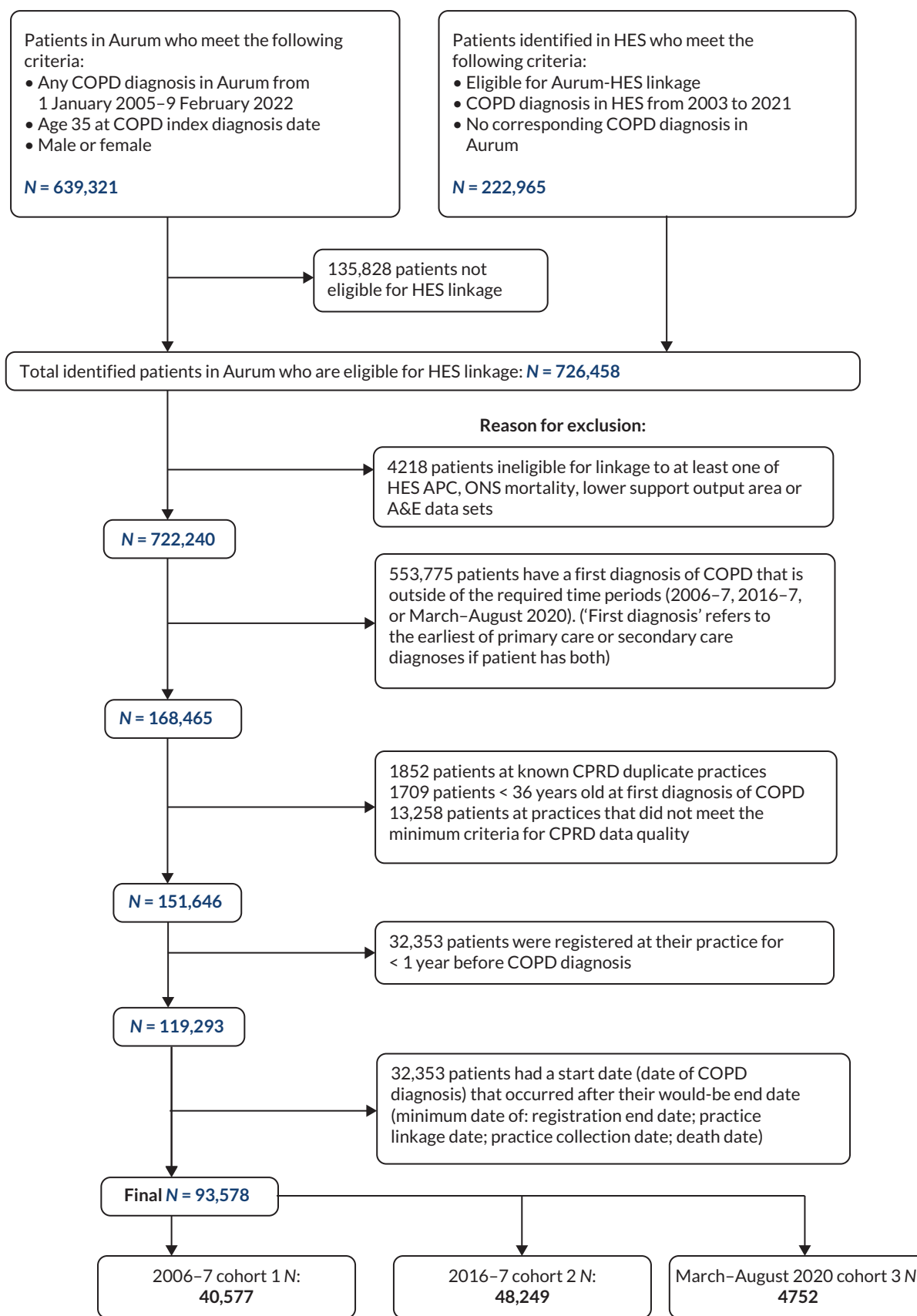


FIGURE 1 Flow diagram for inclusion/exclusion of patients.

long the pathway took. Given that COPD can present differently depending on what comorbidities are already present, we compared the route to diagnosis separately for people with asthma and HF. Time zero was defined as either the first recorded respiratory symptom [wheeze, lower respiratory tract infection (LRTI), cough, sputum, breathlessness] or the first respiratory symptom recorded up to a maximum of 5 years before COPD diagnosis, to try to protect against defining time zero on symptoms that may have been recorded for a disease other than COPD. We compared the time from first symptom to GP action, time from first symptom to COPD diagnosis and the number of symptoms recorded before diagnosis between cohorts 1, 2 and 3 and between patients with or without asthma or HF.

Multilevel logistic regression models were used to quantify the variation between GP practices in spirometry within 6 months prior to or after diagnosis. First, we fitted a 'null' model with only the practice-level random intercept to calculate the intraclass correlation coefficient (ICC). The ICC indicates how much of the total variation in the outcome is explained by between-practice variation. Next, the number of COPD patients per practice was added to the null model to give adjusted odds ratios (ORs). The median odds ratio (MOR) was reported.²⁸ The MOR is a median of the set of ORs that could be obtained by comparing two patients with identical characteristics from two randomly selected practices. It is therefore a summary measure of the amount of variation between practices. Funnel plots with 95% [2 standard deviation (SD)] and 99.7% (3 SD) control limits were used to graphically present the variation between practices, plotting the proportion of patients having spirometry against the number of patients with COPD per practice. This model was run for both the whole cohort and for just those who were diagnosed in primary care, to account for GPs being potentially unaware of the patient's diagnosis after hospitalisation. Cohort 3 was not included here due to its small sample size.

Modelling the risk of the first acute exacerbations of chronic obstructive pulmonary disease

To account for non-AECOPD mortality as a competing risk to AECOPD, Fine and Gray models were used instead of the Cox model.

Modelling was run only for those patients diagnosed in primary care, that is, not through hospitalisation for an AECOPD. Missing values for continuous variables were imputed using multiple imputation with 10 data sets, while those for categorical variables were included using a 'missing' category. Age was centred and modelled using a natural spline, with three knots at the 0.1, 0.5 and 0.9 quantiles. Use of splines increases the predictive power of the model but at the expense of model interpretability. The spline coefficients were positive between all knots and for all models, indicating a positive relationship between age and AECOPD risk.

We built the following models for each cohort separately:

1. Model 1: Patient baseline characteristics [age, sex, deprivation, smoking, BMI, blood pressure (BP), symptom history, COPD severity, comorbidities].
2. Model 2: As #1 plus GP actions prior to diagnosis [chest X-ray, echocardiogram, spirometry, specialist referral, HF referral and the following medications in the 5 years before diagnosis: long-acting muscarinic antagonist (LAMA) combined with long-acting beta agonist (LABA), ICS combined with LABA, triple therapy, LAMA, LABA, ICS, oral corticosteroids (OCS), short-acting beta agonist (SABA) and short-acting muscarinic antagonist (SAMA)]
3. Model 3: As #2 plus GP actions in the year since diagnosis but before any AECOPD [LAMA combined with LABA, ICS combined with LABA, triple therapy, LAMA, LABA, ICS, OCS, SABA, SAMA, oxygen therapy, COPD review, spirometry, flu vaccination, pulmonary rehabilitation (PR) and smoking status combined with whether cessation advice was given].

The ICC was calculated for the null model and found to be very small (0.021 for cohort 1 and 0.031 for cohort 2). Given the complexity of the models and the findings of previous work,²⁹ random effects for practices were not included.

Model fit statistics, discrimination (Harrell's c-statistic at intervals of 6 months) and calibration plots were reported. Validation of the models was done using k-fold cross-validation with 10 folds. The model was trained on nine of the folds (the 'training data') and tested on the remaining fold (the 'test data'), with the process repeated 10 times with each fold used as the test data once. The performance of the model was then averaged over the 10 iterations. Exponentiated coefficients [i.e. hazard ratios (HRs)] that had 95% confidence intervals that lay outside of 1 were determined to be statistically significant.

Patient surveys and focus groups

To further explore the patterns identified in the CPRD analysis, we created a survey and ran focus groups with the following objectives:

1. to describe the many potential routes by which patients obtain their COPD diagnosis;
2. to identify where patients think opportunities may have been missed in making the diagnosis;
3. to estimate the time between initial respiratory symptoms, first presentation to the health service with those symptoms and definitive COPD diagnosis;
4. to understand patients' approaches to managing an exacerbation;
5. to assess the relative merits of various digital formats and avenues when obtaining this information.

The online survey (see [Appendix 1](#)) was designed to investigate COPD patients' retrospective perceptions of their initial symptoms, what they did after developing those symptoms, what kind of professional advice was sought and year of diagnosis in order to distinguish between COVID and pre-COVID eras. It was designed jointly through a series of discussions by the project team at Imperial College London, which included researchers and patient representatives, and the teams at Asthma + Lung UK (formerly British Lung Foundation) and the Taskforce for Lung Health, including its own patient advisory group. Piloting with our local patients highlighted wording issues for correction and estimated the total time taken to complete the survey to ensure it was not too lengthy.

Online survey data were sought from three avenues:

- Direct messaging on Twitter for COPD patients who use Twitter. Direct messages were also sent to associated clinicians and patient groups, such as @pulsetoday, @gmcul, @LancetRespirMed.
- The same survey was administered by Asthma + Lung UK via their website. To fit in with the charity's wider goals, the survey was open to people with any of four lung conditions: COPD, interstitial lung disease, bronchiectasis and asthma. We only report the COPD results here.
- Primary care-based survey, administered by CPRD by sending out to participating practices to pass on to their existing COPD patients via CPRD's patient portal on the e-platform.

For the third route above, CPRD staff conducted a patient search on CPRD records based on the inclusion criteria (see [Patient cohorts, definition of index date and acute exacerbation](#) for inclusion criteria). CPRD sent a list of potential eligible patients to the relevant GP practices who agreed to be contacted for this study. After this screening by the GP practices, CPRD sent the study information pack (participant information sheet and invitation letter) with details on the link and log-in details for participants to access the patient survey. Patients were given their own ID number to enter so that their records could be identified.

The survey is included in [Appendix 1](#).

Focus groups

The aim was to convene up to six focus groups (six to eight patients per group, groups in urban and rural locations), for instance, at local British Lung Foundation Breathe Easy groups that welcome external speakers to their meetings, and PR groups. Due to COVID-19, these were run virtually using Microsoft Teams; the focus group discussions were audio-recorded and lasted up to 1 hour.

Analysis of the patient surveys

Data analysis was descriptive by reporting proportions, with chi-squared tests to make comparisons across the four diseases. For all tests, $p \leq 0.05$ was considered statistically significant. Questions answered in the free-text options were analysed simply using NVivo (QSR International, Warrington, UK) in word clouds.

Analysis of the focus groups

A thematic framework approach identified common themes and patterns relating to SABA and ICS use. The analysis followed the six steps described in the framework method: familiarisation, initial coding, generating themes, reviewing themes, defining themes and writing the final report.

Data were coded and organised using NVivo software version 12 to create new themes and topic nodes. These data were charted through the development of an analytical framework, rearranged according to themes, with a matrix developed with participants on the vertical axis, and main themes with subcategories across the horizontal axis.

Chapter 3 Results

Cohort identification

We included 40,577 patients with incident COPD between April 2006 and March 2007 (cohort 1) and 48,249 between April 2016 and March 2017 (cohort 2). The COVID-era group that included patients diagnosed between March and August 2020 (cohort 3) had 4752 patients. For three-quarters of the patients, the diagnosis was first recorded in primary care; this proportion was significantly lower in the COVID-era cohort.

[Table 1](#) describes the characteristics of the included patients in each cohort. Due to the similarity of respiratory symptoms between asthma and COPD, COPD is commonly misdiagnosed as asthma in the early stages of patient presentation. To ensure that patients with a diagnosis of asthma did truly have (current) asthma, we assumed that first asthma diagnoses in the 2 years prior to COPD diagnosis were a misdiagnosis. We therefore only defined a patient as having current asthma if they had a diagnosis of asthma 2–5 years prior to their COPD diagnosis or if they had a diagnosis 0–2 years prior to COPD diagnosis with another asthma diagnosis code following COPD diagnosis.

Descriptive analysis

The proportion in which the diagnosis was first recorded in secondary care increased from 21.9% in cohort 1 to 29.1% during the COVID-era cohort. [Table 2](#) compares the patient characteristics by cohort and diagnosis setting.

Patients diagnosed in hospital were older, less likely to have smoked and more likely to have comorbidities such as HF, diabetes and renal disease. These differences were maintained across the three cohorts. Comorbidities were all more likely to be recorded in cohort 2 than in cohort 1, and most were more likely to be recorded in cohort 3 than in cohort 2.

Route to diagnosis

Seventy-nine per cent of patients with a diagnosis of COPD had at least one respiratory symptom recorded in the previous 5 years, the first of which was used as the 'time zero' until diagnosis or GP actions. As by definition, all patients in our data set had a COPD diagnosis; we also looked backwards from first diagnosis to GP action to account for the fact that not all patients had symptoms recorded.

[Table 3](#) shows the compliance with NICE guidelines for diagnosing COPD, for patients without pre-existing asthma or HF only. The use of pre-diagnosis spirometry improved in cohort 2 on cohort 1 but fell back for the COVID group. In contrast, chest X-ray, full blood count (FBC) and BMI all improved after cohort 1 and were maintained for the COVID cohort, and almost all patients received one of these, but the improvements were smaller when the analysis was restricted to 1 year before diagnosis.

[Table 4](#) and [Appendix 2, Table 16](#) also show COPD symptoms recorded in the 5 years before the diagnosis date. In the 5 years before, fewer than 1 in 10 had wheeze, nearly half (42.1%) had a LRTI, 1 in 7 had abnormal sputum, half had breathlessness and half had a cough; 1 in 5 had none of these recorded. In those who did have a symptom recorded, the median time from symptom to diagnosis was 700 days in cohort 1, 830 days in cohort 2 and 957 days in cohort 3 when restricting to the 5-year pre-diagnosis period ($p < 0.001$). The median times between symptom and NICE-recommended test were also much longer for the COVID-era cohort.

TABLE 1 Patient characteristics of the three cohorts

Variable		2006–7	2016–7	March–August 2020	Total	p-value
Diagnosis location	Primary care	31,676 (78.1)	37,393 (77.5)	3368 (70.9)	72,437 (77.4)	< 0.001
	Secondary care	8901 (21.9)	10,856 (22.5)	1384 (29.1)	21,141 (22.6)	
Sex	Female	19,194 (47.3)	22,764 (47.2)	2273 (47.8)	44,231 (47.3)	0.678
	Male	21,383 (52.7)	25,485 (52.8)	2479 (52.2)	49,347 (52.7)	
Age (years)	Mean (SD)	68.3 (11.7)	68.2 (12.2)	68.6 (12.5)	68.3 (12.0)	0.078
IMD quintile	1	5788 (14.3)	7068 (14.7)	663 (14.0)	13,519 (14.5)	0.114
	2	7048 (17.4)	8447 (17.5)	821 (17.3)	16,316 (17.4)	
	3	7540 (18.6)	8948 (18.6)	912 (19.2)	17,400 (18.6)	
	4	8714 (21.5)	10,567 (21.9)	1052 (22.1)	20,333 (21.7)	
	5	11,450 (28.2)	13,191 (27.4)	1302 (27.4)	25,943 (27.7)	
Smoking status	Missing	1442 (3.6)	2775 (5.8)	197 (4.1)	4414 (4.7)	< 0.001
	Never smoker	5015 (12.4)	6095 (12.6)	679 (14.3)	11,789 (12.6)	
	Ex-smoker	15,235 (37.5)	17,709 (36.7)	1832 (38.6)	34,776 (37.2)	
	Current smoker	18,885 (46.5)	21,670 (44.9)	2044 (43.0)	42,599 (45.5)	
BMI recorded	Yes	36,060 (88.9)	46,321 (96.0)	4555 (95.9)	86,936 (92.9)	< 0.001
	No	4517 (11.1)	1928 (4.0)	197 (4.1)	6642 (7.1)	
BMI category	Underweight	1861 (4.6)	2057 (4.3)	206 (4.3)	4124 (4.4)	< 0.001
	Normal	13,577 (33.5)	15,354 (31.8)	1527 (32.1)	30,458 (32.5)	
	Overweight	11,870 (29.3)	15,107 (31.3)	1393 (29.3)	28,370 (30.3)	
	Obese	8752 (21.6)	13,803 (28.6)	1429 (30.1)	23,984 (25.6)	
	Missing	4517 (11.1)	1928 (4.0)	197 (4.1)	6642 (7.1)	
Diastolic blood pressure taken	Yes	39,677 (97.8)	47,679 (98.8)	4692 (98.7)	92,048 (98.4)	< 0.001
	No	900 (2.2)	570 (1.2)	60 (1.3)	1530 (1.6)	

TABLE 1 Patient characteristics of the three cohorts (*continued*)

Variable		2006–7	2016–7	March–August 2020	Total	p-value
Diastolic blood pressure (mmHg)	Mean (SD)	77.6 (10.3)	76.0 (10.2)	76.7 (10.8)	76.8 (10.3)	< 0.001
Systolic blood pressure taken	Yes	39,666 (97.8)	47,682 (98.8)	4692 (98.7)	92,040 (98.4)	< 0.001
	No	911 (2.2)	567 (1.2)	60 (1.3)	1538 (1.6)	
Systolic blood pressure (mmHg)	Mean (SD)	135.8 (18.0)	131.8 (16.5)	131.7 (17.3)	133.5 (17.3)	< 0.001
Current asthma	Yes	10,139 (25.0)	9651 (20.0)	989 (20.8)	20,779 (22.2)	< 0.001
Any malignancy, including leukaemia and lymphoma (CCI)	Yes	5403 (13.3)	8867 (18.4)	997 (21.0)	15,267 (16.3)	< 0.001
CVD (CCI)	Yes	3759 (9.3)	5472 (11.3)	602 (12.7)	9833 (10.5)	< 0.001
Congestive HF (CCI)	Yes	3015 (7.4)	4064 (8.4)	557 (11.7)	7636 (8.2)	< 0.001
Dementia (CCI)	Yes	400 (1.0)	1720 (3.6)	212 (4.5)	2332 (2.5)	< 0.001
Diabetes without chronic complications (CCI)	Yes	4132 (10.2)	7342 (15.2)	932 (19.6)	12,406 (13.3)	< 0.001
Diabetes with chronic complications (CCI)	Yes	1418 (3.5)	4108 (8.5)	425 (8.9)	5951 (6.4)	< 0.001
AIDS/HIV (CCI)	Yes	32 (0.1)	205 (0.4)	35 (0.7)	272 (0.3)	< 0.001
Hemiplegia or paraplegia (CCI)	Yes	225 (0.6)	305 (0.6)	50 (1.1)	580 (0.6)	< 0.001
Metastatic solid tumour (CCI)	Yes	387 (1.0)	837 (1.7)	111 (2.3)	1335 (1.4)	< 0.001
Mild liver disease (CCI)	Yes	419 (1.0)	956 (2.0)	132 (2.8)	1507 (1.6)	< 0.001
Moderate or severe liver disease (CCI)	Yes	63 (0.2)	183 (0.4)	28 (0.6)	274 (0.3)	< 0.001
Myocardial infarction (CCI)	Yes	3422 (8.4)	3581 (7.4)	374 (7.9)	7377 (7.9)	< 0.001
Peptic ulcer disease (CCI)	Yes	2761 (6.8)	2929 (6.1)	299 (6.3)	5989 (6.4)	< 0.001
PVD (CCI)	Yes	3173 (7.8)	4087 (8.5)	418 (8.8)	7678 (8.2)	0.001
Renal disease (CCI)	Yes	3932 (9.7)	8073 (16.7)	843 (17.7)	12,848 (13.7)	< 0.001
Rheumatological disease (CCI)	Yes	2180 (5.4)	3180 (6.6)	397 (8.4)	5757 (6.2)	< 0.001
Hypertension	Yes	16,611 (40.9)	21,347 (44.2)	2186 (46.0)	40,144 (42.9)	< 0.001

continued

TABLE 1 Patient characteristics of the three cohorts (*continued*)

Variable		2006–7	2016–7	March–August 2020	Total	p-value
Anxiety	Yes	6494 (16.0)	12,099 (25.1)	1438 (30.3)	20,031 (21.4)	< 0.001
Depression	Yes	7837 (19.3)	14,160 (29.3)	1626 (34.2)	23,623 (25.2)	< 0.001
Osteoporosis	Yes	2194 (5.4)	3697 (7.7)	426 (9.0)	6317 (6.8)	< 0.001
Anaemia	Yes	3317 (8.2)	6560 (13.6)	825 (17.4)	10,702 (11.4)	< 0.001
Any LRTI diagnosis preceding COPD diagnosis	Yes	20,982 (51.7)	29,957 (62.1)	3111 (65.5)	54,050 (57.8)	< 0.001
Arrhythmia	Yes	857 (2.1)	1398 (2.9)	198 (4.2)	2453 (2.6)	< 0.001
Stroke	Yes	7991 (19.7)	11,867 (24.6)	1328 (27.9)	21,186 (22.6)	< 0.001
Atrial fibrillation	Yes	3052 (7.5)	4830 (10.0)	597 (12.6)	8479 (9.1)	< 0.001
Total number of comorbidities	Mean (SD)	3.5 (2.0)	4.3 (2.3)	4.5 (2.5)	3.9 (2.2)	< 0.001
GOLD status	GOLD stage 1: ≥ 80%	3097 (7.6)	9846 (20.4)	1070 (22.5)	14,013 (15.0)	< 0.001
	GOLD stage 2: 50–79%	12,269 (30.2)	17,168 (35.6)	1080 (22.7)	30,517 (32.6)	
	GOLD stage 3: 30–49%	5195 (12.8)	4401 (9.1)	256 (5.4)	9852 (10.5)	
	GOLD stage 4: < 30%	879 (2.2)	550 (1.1)	32 (0.7)	1461 (1.6)	
	Missing FEV1 %-pred measurement	19,137 (47.2)	16,284 (33.7)	2314 (48.7)	37,735 (40.3)	

CCI, Charlson Comorbidity Index; CVD, cerebrovascular disease; FEV1, forced expiratory volume in 1 second; LRTI, lower respiratory tract infection; PVD, peripheral vascular disease.

TABLE 2 Patient characteristics by where diagnosed and cohort

Variable		2006–7 and primary care	2006–7 and secondary care	2016–7 and primary care	2016–7 and secondary care	March–August 2020 and primary care	March–August 2020 and secondary care	Total	p-value
Sex	Female	14,583 (46.0)	4611 (51.8)	17,196 (46.0)	5568 (51.3)	1591 (47.2)	682 (49.3)	44,231 (47.3)	< 0.001
	Male	17,093 (54.0)	4290 (48.2)	20,197 (54.0)	5288 (48.7)	1777 (52.8)	702 (50.7)	49,347 (52.7)	
Age (years)	Mean (SD)	67.5 (11.3)	71.0 (13.0)	66.5 (11.4)	74.0 (12.9)	66.4 (11.7)	74.1 (12.7)	68.3 (12.0)	< 0.001
IMD quintile	1	4692 (14.8)	1096 (12.3)	5508 (14.7)	1560 (14.4)	468 (13.9)	195 (14.1)	13,519 (14.5)	< 0.001
	2	5684 (18.0)	1364 (15.3)	6544 (17.5)	1903 (17.5)	559 (16.6)	262 (18.9)	16,316 (17.4)	
	3	5968 (18.9)	1572 (17.7)	6932 (18.5)	2016 (18.6)	648 (19.3)	264 (19.1)	17,400 (18.6)	
	4	6683 (21.1)	2031 (22.8)	8181 (21.9)	2386 (22.0)	762 (22.6)	290 (21.0)	20,333 (21.7)	
	5	8621 (27.2)	2829 (31.8)	10,208 (27.3)	2983 (27.5)	929 (27.6)	373 (27.0)	25,943 (27.7)	
Smoking status	Missing	776 (2.4)	666 (7.5)	2129 (5.7)	646 (6.0)	115 (3.4)	82 (5.9)	4414 (4.7)	< 0.001
	Never smoker	3599 (11.4)	1416 (15.9)	3734 (10.0)	2361 (21.7)	349 (10.4)	330 (23.8)	11,789 (12.6)	
	Ex-smoker	12,306 (38.8)	2929 (32.9)	13,978 (37.4)	3731 (34.4)	1345 (39.9)	487 (35.2)	34,776 (37.2)	
	Current smoker	14,995 (47.3)	3890 (43.7)	17,552 (46.9)	4118 (37.9)	1559 (46.3)	485 (35.0)	42,599 (45.5)	
BMI measured	Yes	28,917 (91.3)	7143 (80.2)	36,430 (97.4)	9891 (91.1)	3272 (97.1)	1283 (92.7)	86,936 (92.9)	< 0.001
	No	2759 (8.7)	1758 (19.8)	963 (2.6)	965 (8.9)	96 (2.9)	101 (7.3)	6642 (7.1)	
BMI category	Underweight	1413 (4.5)	448 (5.0)	1473 (3.9)	584 (5.4)	130 (3.9)	76 (5.5)	4124 (4.4)	< 0.001
	Normal	10,905 (34.4)	2672 (30.0)	12,185 (32.6)	3169 (29.2)	1119 (33.2)	408 (29.5)	30,458 (32.5)	
	Overweight	9762 (30.8)	2108 (23.7)	12,223 (32.7)	2884 (26.6)	1010 (30.0)	383 (27.7)	28,370 (30.3)	
	Obese	6837 (21.6)	1915 (21.5)	10,549 (28.2)	3254 (30.0)	1013 (30.1)	416 (30.1)	23,984 (25.6)	
	Missing	2759 (8.7)	1758 (19.8)	963 (2.6)	965 (8.9)	96 (2.9)	101 (7.3)	6642 (7.1)	
Diastolic blood pressure taken	Yes	31,339 (98.9)	8338 (93.7)	37,256 (99.6)	10,423 (96.0)	3355 (99.6)	1337 (96.6)	92,048 (98.4)	< 0.001
	No	337 (1.1)	563 (6.3)	137 (0.4)	433 (4.0)	13 (0.4)	47 (3.4)	1530 (1.6)	

continued

TABLE 2 Patient characteristics by where diagnosed and cohort (*continued*)

Variable		2006–7 and primary care	2006–7 and secondary care	2016–7 and primary care	2016–7 and secondary care	March–August 2020 and primary care	March–August 2020 and secondary care	Total	p-value
Diastolic blood pressure (mmHg)	Mean (SD)	77.8 (10.1)	76.9 (11.2)	76.5 (9.9)	74.3 (11.3)	77.3 (10.2)	75.2 (12.2)	76.8 (10.3)	< 0.001
Systolic blood pressure taken	Yes	31,337 (98.9)	8329 (93.6)	37,257 (99.6)	10,425 (96.0)	3355 (99.6)	1337 (96.6)	92,040 (98.4)	< 0.001
	No	339 (1.1)	572 (6.4)	136 (0.4)	431 (4.0)	13 (0.4)	47 (3.4)	1538 (1.6)	
Systolic blood pressure (mmHg)	Mean (SD)	136.0 (17.7)	135.3 (19.2)	132.1 (15.9)	130.9 (18.6)	131.7 (16.2)	131.7 (19.7)	133.5 (17.3)	< 0.001
Current asthma	Yes	8170 (25.8)	1969 (22.1)	7356 (19.7)	2295 (21.1)	672 (20.0)	317 (22.9)	20,779 (22.2)	< 0.001
Any malignancy, including leukaemia and lymphoma (CCI)	Yes	4016 (12.7)	1387 (15.6)	6102 (16.3)	2765 (25.5)	605 (18.0)	392 (28.3)	15,267 (16.3)	< 0.001
CVD (CCI)	Yes	2611 (8.2)	1148 (12.9)	3494 (9.3)	1978 (18.2)	367 (10.9)	235 (17.0)	9833 (10.5)	< 0.001
Congestive HF (CCI)	Yes	2008 (6.3)	1007 (11.3)	2383 (6.4)	1681 (15.5)	286 (8.5)	271 (19.6)	7636 (8.2)	< 0.001
Dementia (CCI)	Yes	201 (0.6)	199 (2.2)	755 (2.0)	965 (8.9)	82 (2.4)	130 (9.4)	2332 (2.5)	< 0.001
Diabetes without chronic complications (CCI)	Yes	3081 (9.7)	1051 (11.8)	5375 (14.4)	1967 (18.1)	623 (18.5)	309 (22.3)	12,406 (13.3)	< 0.001
Diabetes with chronic complications (CCI)	Yes	996 (3.1)	422 (4.7)	2695 (7.2)	1413 (13.0)	239 (7.1)	186 (13.4)	5951 (6.4)	< 0.001
AIDS/HIV (CCI)	Yes	23 (0.1)	9 (0.1)	186 (0.5)	19 (0.2)	30 (0.9)	5 (0.4)	272 (0.3)	< 0.001
Hemiplegia or paraplegia (CCI)	Yes	157 (0.5)	68 (0.8)	180 (0.5)	125 (1.2)	28 (0.8)	22 (1.6)	580 (0.6)	< 0.001
Metastatic solid tumour (CCI)	Yes	95 (0.3)	66 (0.7)	173 (0.5)	194 (1.8)	32 (1.0)	29 (2.1)	589 (0.6)	< 0.001
Mild liver disease (CCI)	Yes	287 (0.9)	100 (1.1)	640 (1.7)	197 (1.8)	79 (2.3)	32 (2.3)	1335 (1.4)	< 0.001

TABLE 2 Patient characteristics by where diagnosed and cohort (continued)

Variable		2006–7 and primary care	2006–7 and secondary care	2016–7 and primary care	2016–7 and secondary care	March–August 2020 and primary care	March–August 2020 and secondary care	Total	p-value
Moderate or severe liver disease (CCI)	Yes	46 (0.1)	17 (0.2)	99 (0.3)	84 (0.8)	11 (0.3)	17 (1.2)	274 (0.3)	< 0.001
Myocardial infarction (CCI)	Yes	2499 (7.9)	923 (10.4)	2363 (6.3)	1218 (11.2)	222 (6.6)	152 (11.0)	7377 (7.9)	< 0.001
Peptic ulcer disease (CCI)	Yes	2170 (6.9)	591 (6.6)	2105 (5.6)	824 (7.6)	194 (5.8)	105 (7.6)	5989 (6.4)	< 0.001
PVD (CCI)	Yes	2371 (7.5)	802 (9.0)	2839 (7.6)	1248 (11.5)	273 (8.1)	145 (10.5)	7678 (8.2)	< 0.001
Renal disease (CCI)	Yes	2816 (8.9)	1116 (12.5)	5153 (13.8)	2920 (26.9)	456 (13.5)	387 (28.0)	12,848 (13.7)	< 0.001
Rheumatological disease (CCI)	Yes	1663 (5.3)	517 (5.8)	2205 (5.9)	975 (9.0)	251 (7.5)	146 (10.5)	5757 (6.2)	< 0.001
Hypertension	Yes	12,862 (40.6)	3749 (42.1)	15,536 (41.5)	5811 (53.5)	1421 (42.2)	765 (55.3)	40,144 (42.9)	< 0.001
Anxiety	Yes	5035 (15.9)	1459 (16.4)	9447 (25.3)	2652 (24.4)	1026 (30.5)	412 (29.8)	20,031 (21.4)	< 0.001
Depression	Yes	6086 (19.2)	1751 (19.7)	11,197 (29.9)	2963 (27.3)	1199 (35.6)	427 (30.9)	23,623 (25.2)	< 0.001
Osteoporosis	Yes	1605 (5.1)	589 (6.6)	2425 (6.5)	1272 (11.7)	243 (7.2)	183 (13.2)	6317 (6.8)	< 0.001
Anaemia	Yes	2377 (7.5)	940 (10.6)	4232 (11.3)	2328 (21.4)	479 (14.2)	346 (25.0)	10,702 (11.4)	< 0.001
Any LRTI diagnosis preceding COPD diagnosis	Yes	16,651 (52.6)	4331 (48.7)	23,076 (61.7)	6881 (63.4)	2184 (64.8)	927 (67.0)	54,050 (57.8)	< 0.001
Arrhythmia	Yes	632 (2.0)	225 (2.5)	954 (2.6)	444 (4.1)	115 (3.4)	83 (6.0)	2453 (2.6)	< 0.001
Stroke	Yes	6194 (19.6)	1797 (20.2)	8930 (23.9)	2937 (27.1)	919 (27.3)	409 (29.6)	21,186 (22.6)	< 0.001
Atrial fibrillation	Yes	2129 (6.7)	923 (10.4)	2907 (7.8)	1923 (17.7)	298 (8.8)	299 (21.6)	8479 (9.1)	< 0.001
Total number of comorbidities	Mean (SD)	3.5 (1.9)	3.3 (2.3)	4.2 (2.1)	4.8 (2.8)	4.2 (2.3)	5.2 (2.8)	3.9 (2.2)	< 0.001

continued

TABLE 2 Patient characteristics by where diagnosed and cohort (*continued*)

Variable		2006–7 and primary care	2006–7 and secondary care	2016–7 and primary care	2016–7 and secondary care	March–August 2020 and primary care	March–August 2020 and secondary care	Total	p-value
GOLD status	GOLD stage 1: ≥ 80%	2748 (8.7)	349 (3.9)	8516 (22.8)	1330 (12.3)	848 (25.2)	222 (16.0)	14,013 (15.0)	< 0.001
	GOLD stage 2: 50–79%	11,722 (37.0)	547 (6.1)	15,839 (42.4)	1329 (12.2)	913 (27.1)	167 (12.1)	30,517 (32.6)	
	GOLD stage 3: 30–49%	4927 (15.6)	268 (3.0)	4066 (10.9)	335 (3.1)	217 (6.4)	39 (2.8)	9852 (10.5)	
	GOLD stage 4: < 30%	819 (2.6)	60 (0.7)	502 (1.3)	48 (0.4)	26 (0.8)	6 (0.4)	1461 (1.6)	
	Missing FEV1 %-pred measurement	11,460 (36.2)	7677 (86.2)	8470 (22.7)	7814 (72.0)	1364 (40.5)	950 (68.6)	37,735 (40.3)	

CCI, Charlson Comorbidity Index; CVD, cerebrovascular disease; FEV1, forced expiratory volume in 1 second; LRTI, lower respiratory tract infection; PVD, peripheral vascular disease.

TABLE 3 Compliance with NICE guidelines for COPD diagnosis: patients with no pre-existing asthma or HF

Variable		2006–7	2016–7	March–August 2020	Total	p-value
Spirometry (pre diagnosis)	Yes	16,629 (59.3)	25,556 (72.4)	1905 (57.1)	44,090 (66.1)	< 0.001
	No	11,414 (40.7)	9761 (27.6)	1432 (42.9)	22,607 (33.9)	
Chest X-ray	Yes	13,334 (47.5)	22,141 (62.7)	2147 (64.3)	37,622 (56.4)	< 0.001
	No	14,709 (52.5)	13,176 (37.3)	1190 (35.7)	29,075 (43.6)	
FBC	Yes	20,589 (73.4)	30,876 (87.4)	3037 (91.0)	54,502 (81.7)	< 0.001
	No	7454 (26.6)	4441 (12.6)	300 (9.0)	12,195 (18.3)	
BMI	Yes	24,304 (86.7)	33,592 (95.1)	3157 (94.6)	61,053 (91.5)	< 0.001
	No	3739 (13.3)	1725 (4.9)	180 (5.4)	5644 (8.5)	
All of spirometry, FBC, CXR and BMI measurement before COPD diagnosis	Yes	6921 (24.7)	15,638 (44.3)	1297 (38.9)	23,856 (35.8)	< 0.001
	No	21,122 (75.3)	19,679 (55.7)	2040 (61.1)	42,841 (64.2)	
At least one of FBC, CXR and BMI measurement before COPD diagnosis	Yes	26,670 (95.1)	34,625 (98.0)	3269 (98.0)	64,564 (96.8)	< 0.001
	No	1373 (4.9)	692 (2.0)	68 (2.0)	2133 (3.2)	
Spirometry, CXR, BMI, FBC	Spirometry and one of CXR/BMI/FBC	16,347 (58.3)	25,482 (72.2)	1903 (57.0)	43,732 (65.6)	< 0.001
	Spirometry only	282 (1.0)	74 (0.2)	2 (0.1)	358 (0.5)	
	No spirometry and one of CXR/BMI/FBC	10,323 (36.8)	9143 (25.9)	1366 (40.9)	20,832 (31.2)	
	No spirometry, CXR, BMI or FBC	1091 (3.9)	618 (1.7)	66 (2.0)	1775 (2.7)	
Spirometry in the year before COPD diagnosis	Yes	15,609 (55.7)	22,528 (63.8)	1249 (37.4)	39,386 (59.1)	< 0.001
	No	12,434 (44.3)	12,789 (36.2)	2088 (62.6)	27,311 (40.9)	

continued

TABLE 3 Compliance with NICE guidelines for COPD diagnosis: patients with no pre-existing asthma or HF (continued)

Variable		2006–7	2016–7	March–August 2020	Total	p-value
Chest X-ray in the year before COPD diagnosis	Yes	7381 (26.3)	12,571 (35.6)	1116 (33.4)	21,068 (31.6)	< 0.001
	No	20,662 (73.7)	22,746 (64.4)	2221 (66.6)	45,629 (68.4)	
FBC in the year before COPD diagnosis	Yes	13,250 (47.2)	19,463 (55.1)	1813 (54.3)	34,526 (51.8)	< 0.001
	No	14,793 (52.8)	15,854 (44.9)	1524 (45.7)	32,171 (48.2)	
BMI measurement in the year before COPD diagnosis	Yes	14,878 (53.1)	20,600 (58.3)	1715 (51.4)	37,193 (55.8)	< 0.001
	No	13,165 (46.9)	14,717 (41.7)	1622 (48.6)	29,504 (44.2)	
All of spirometry, FBC, CXR and BMI measurement in the year before COPD diagnosis	Yes	2058 (7.3)	4354 (12.3)	273 (8.2)	6685 (10.0)	< 0.001
	No	25,985 (92.7)	30,963 (87.7)	3064 (91.8)	60,012 (90.0)	
At least one of FBC, CXR and BMI measurement in the year before COPD diagnosis	Yes	21,478 (76.6)	29,579 (83.8)	2663 (79.8)	53,720 (80.5)	< 0.001
	No	6565 (23.4)	5738 (16.2)	674 (20.2)	12,977 (19.5)	
Spirometry and one of CXR/BMI/FBC	Spirometry and one of CXR/BMI/FBC	13,410 (47.8)	20,518 (58.1)	1152 (34.5)	35,080 (52.6)	
	Spirometry only	2199 (7.8)	2010 (5.7)	97 (2.9)	4306 (6.5)	
	No spirometry and one of CXR/BMI/FBC	8068 (28.8)	9061 (25.7)	1511 (45.3)	18,640 (27.9)	
	No spirometry, CXR, BMI or FBC	4366 (15.6)	3728 (10.6)	577 (17.3)	8671 (13.0)	

FBC, full blood count.

TABLE 4 General practitioner-recorded symptoms in the 5 years before diagnosis

Variable		2006–7	2016–7	March–August 2020	Total	p-value
Any wheeze diagnosis	Yes	2056 (7.3)	3384 (9.6)	297 (8.9)	5737 (8.6)	< 0.001
Time from earliest wheeze diagnosis to COPD diagnosis (days)	Median (IQR)	225.5 (30.0 to 743.2)	136.0 (18.0 to 691.5)	331.0 (50.0 to 869.0)	177.0 (24.0 to 721.0)	< 0.001
Any LRTI diagnosis	Yes	11,550 (41.2)	15,049 (42.6)	1453 (43.5)	28,052 (42.1)	< 0.001
Time from earliest LRTI diagnosis to COPD diagnosis (days)	Median (IQR)	750.0 (265.0 to 1268.8)	787.0 (259.0 to 1364.0)	817.0 (282.0 to 1379.0)	773.0 (262.0 to 1327.0)	< 0.001
Any sputum diagnosis	Yes	3650 (13.0)	5382 (15.2)	597 (17.9)	9629 (14.4)	< 0.001
Time from earliest sputum diagnosis to COPD diagnosis (days)	Median (IQR)	450.0 (50.0 to 1057.8)	393.5 (55.0 to 1031.8)	440.0 (92.0 to 1055.0)	417.0 (55.0 to 1047.0)	0.361
Any breathlessness diagnosis	Yes	10,108 (36.0)	19,745 (55.9)	1667 (50.0)	31,520 (47.3)	< 0.001
Time from earliest breathlessness diagnosis to COPD diagnosis (days)	Median (IQR)	96.5 (6.0 to 637.0)	45.0 (0.0 to 522.0)	231.0 (14.0 to 898.0)	63.0 (0.0 to 586.0)	< 0.001
Any cough diagnosis	Yes	10,891 (38.8)	19,317 (54.7)	1758 (52.7)	31,966 (47.9)	< 0.001
Time from earliest cough diagnosis to COPD diagnosis (days)	Median (IQR)	574.0 (133.0 to 1131.0)	787.0 (190.0 to 1376.0)	847.0 (227.2 to 1425.5)	708.0 (167.0 to 1309.0)	< 0.001
Time from earliest COPD symptom to COPD diagnosis (days)	Median (IQR)	700.0 (161.0 to 1284.0)	830.0 (144.0 to 1449.0)	957.0 (273.0 to 1513.0)	780.0 (156.0 to 1393.0)	< 0.001
Time from first symptom to chest X-ray (days)	Median (IQR)	113.0 (–30.0 to 846.0)	117.0 (–52.0 to 985.5)	160.0 (–147.2 to 974.0)	118.0 (–49.0 to 932.8)	0.160
Time from first symptom to FBC (days)	Median (IQR)	359.0 (–35.0 to 967.0)	309.0 (–172.0 to 1093.0)	430.0 (–35.0 to 1122.0)	336.0 (–100.0 to 1050.0)	< 0.001
Time from first symptom to first recorded spirometry (days)	Median (IQR)	412.0 (28.0 to 1056.0)	89.0 (0.0 to 911.5)	30.0 (–1192.0 to 587.0)	188.0 (0.0 to 966.0)	< 0.001

FBC, full blood count; IQR, interquartile range; LRTI, lower respiratory tract infection.

There were many statistically significant differences in the times from symptoms to diagnosis by cohort. For example, for those with recorded wheeze, the median time was 136 days for cohort 2 (shorter than for cohort 1) but 331 days for cohort 3. Similarly, for those with recorded breathlessness, the median time was 45 days for cohort 2 (again shorter than for cohort 1) but 231 days for cohort 3.

[Table 5](#) shows the pre-diagnosis prescribing, those who were referred to a respiratory/COPD specialist, and those who were given HF investigations. Of those listed, the most commonly prescribed medications were SABA and steroids (either inhaled or oral). The use of echocardiography, cardiology referral and B-type natriuretic peptide (BNP) testing rose considerably over time. BNP recording remained low despite its now widespread availability.

We ran multilevel models to describe the variation in compliance with NICE guidelines for diagnosis between practices using spirometry around the time of diagnosis as a proxy for this. The ICC and MORs are shown in [Table 6](#). There was most non-random variation between practices for cohort 1, but the MORs show considerable variation by practice for all three cohorts. At least 20% of practices were outliers on funnel plots at 2 SD ([Figures 2 and 3](#), [Table 7](#)). With purely random variation, we would expect 5% of practices to be outliers at 2 SD and just 0.3% at 3 SD. For patients diagnosed in primary care, the number of COPD patients at the practice was significantly and positively associated with spirometry for cohort 1 only, but the size of the effect was negligible.

Initial management following diagnosis in primary care

[Table 8](#) describes GP actions in the year after making the diagnosis. There were several notable changes in prescribing from cohort 1 to 2, such as increases in LAMA and LABA and falls in SAMA and ICS. These were maintained into the COVID era. There was a small fall in SABA from cohort 2 to 3. The year 2020 saw a big rise in triple therapy. Oral steroids were prescribed in around a quarter of patients in both cohorts 1 and 3. PR became more common in cohort 2 and even more so in cohort 3, though was still quite low at 21%. The prescription of smoking cessation drugs fell after cohort 1, with advice rising but then falling back in cohort 3. Around four in five patients in each cohort were offered the influenza vaccine, with two-thirds receiving it from the practice.

[Table 9](#) describes the patients diagnosed via their first AE. As shown earlier, these patients were the minority in each cohort, but it is useful to compare them over time. The sex balance did not change significantly, though the two most recent cohorts were older and living in less-deprived areas than the first one. There was a rise over time in the proportion recorded as never having smoked, with only a minor fall in the proportion with unknown smoking status. BMI recording improved after cohort 1; BMI categories had similar proportions in cohorts 2 and 3. Mean blood pressure was highest in cohort 1. Comorbidity patterns changed over time. Compared with cohort 1, later cohorts had more cancer, cerebrovascular disease, HF, dementia, diabetes, renal disease, hypertension (despite the lower mean BP), anxiety, depression and most of the conditions we considered. For many of these, cohort 3 had even higher prevalences than cohort 2, with a mean of 5.2 conditions compared with 4.8 in cohort 2 and 3.3 in cohort 1.

The use of NICE guideline pre-diagnosis tests varied by cohort in this subgroup in a similar way to the patterns for all patients combined ([Table 10](#)). However, pre-diagnosis spirometry was here commonest in cohort 3, as was the proportion having at least one and having all four tests. Restricting to the year before diagnosis, in the lower part of the table, the proportions having spirometry were low (< 10%) in each cohort, though the other three were commonest in cohort 3. Most of the 1-year pre-diagnosis period for the COVID-era cohort would have been before the pandemic.

TABLE 5 Prescribing in the 5 years before diagnosis by cohort

Variable	2006–7	2016–7	March–August 2020	Total	p-value
LAMA–LABA dual therapy prescribed in the 5 years preceding COPD diagnosis	158 (0.4)	787 (1.6)	233 (4.9)	1178 (1.3)	< 0.001
ICS–LABA dual therapy prescribed in the 5 years preceding COPD diagnosis	3535 (8.7)	939 (1.9)	37 (0.8)	4511 (4.8)	< 0.001
LAMA–LABA–ICS triple therapy prescribed in the 5 years preceding COPD diagnosis	269 (0.7)	219 (0.5)	129 (2.7)	617 (0.7)	< 0.001
LAMA therapy prescribed in the 5 years preceding COPD diagnosis	2321 (5.7)	6410 (13.3)	949 (20.0)	9680 (10.3)	< 0.001
LABA therapy prescribed in the 5 years preceding COPD diagnosis	4737 (11.7)	2273 (4.7)	413 (8.7)	7423 (7.9)	< 0.001
ICS therapy prescribed in the 5 years preceding COPD diagnosis	16,019 (39.5)	16,208 (33.6)	1684 (35.4)	33,911 (36.2)	< 0.001
Oral corticosteroids prescribed in the 5 years preceding COPD diagnosis	13,349 (32.9)	18,519 (38.4)	1889 (39.8)	33,757 (36.1)	< 0.001
SABA prescribed in the 5 years preceding COPD diagnosis	24,462 (60.3)	29,997 (62.2)	2926 (61.6)	57,385 (61.3)	< 0.001
SAMA prescribed in the 5 years preceding COPD diagnosis	3456 (8.5)	1194 (2.5)	62 (1.3)	4712 (5.0)	< 0.001
Respiratory referral before COPD diagnosis	58 (0.1)	4704 (9.7)	865 (18.2)	5627 (6.0)	< 0.001
Cardiology referral before COPD diagnosis	4084 (10.1)	8642 (17.9)	1133 (23.8)	13,859 (14.8)	< 0.001
Echocardiogram before COPD diagnosis	4995 (12.3)	10,615 (22.0)	1324 (27.9)	16,934 (18.1)	< 0.001
Natriuretic peptide test (BNP, NT-proBNP) before COPD diagnosis	121 (0.3)	2799 (5.8)	313 (6.6)	3233 (3.5)	< 0.001

TABLE 6 Results of multilevel analysis for use of spirometry around the time of COPD diagnosis

Statistic	2006–7 cohort, all	2006–7 cohort, diagnosed in primary care	2016–7 cohort, all	2016–7 cohort, diagnosed in primary care	March–August 2020 cohort, all	March–August 2020 cohort, diagnosed in primary care
N	40,577	31,676	48,249	37,393	4752	3368
Median (IQR) N per practice	23 (13–40)	18 (9–31)	28 (15–46)	21 (11–35)	3 (2–5)	2 (1–4)
ICC – null model	0.10	0.13	0.05	0.08	0.07	0.09
Median odds ratio – null model (bootstrap 95% CI)	1.76 (1.71 to 1.78)	1.95 (1.89 to 1.98)	1.49 (1.46 to 1.50)	1.62 (1.59 to 1.65)	1.62 (1.59 to 1.66)	1.72 (1.68 to 1.76)
ICC – COPD patient no. as predictor	0.09	0.12	0.05	0.08	0.07	0.09
Median odds ratio – COPD patient no. as predictor (bootstrap 95% CI)	1.74 (1.70 to 1.76)	1.92 (1.87 to 1.96)	1.48 (1.46 to 1.50)	1.67 (1.63 to 1.69)	1.62 (1.59 to 1.65)	1.72 (1.68 to 1.76)
Odds ratio (number of COPD patients at practice)	1.00 (1.00 to 1.01)	1.01 (1.00 to 1.01)	1.00 (1.00 to 1.00)	1.00 (1.00 to 1.00)	1.01 (1.00 to 1.03)	1.01 (0.98 to 1.03)
p-value (number of COPD patients at practice)	< 0.001	< 0.001	< 0.001	0.097	0.140	0.605

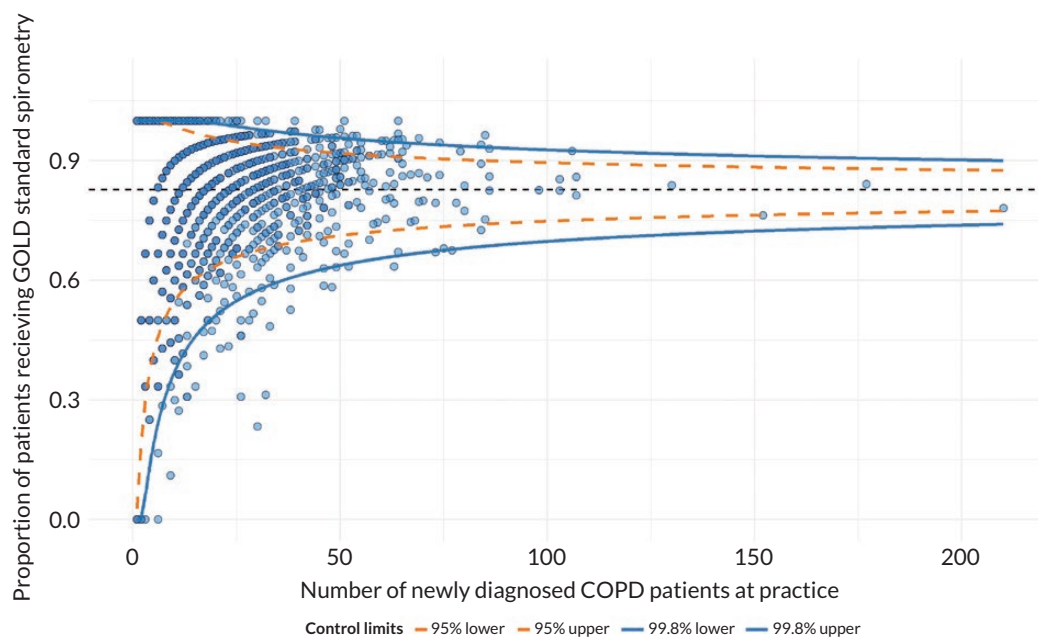


FIGURE 2 Funnel plot of proportion of patients following the initial NICE pathways (performed spirometry within 6 months prior to/after COPD diagnosis) by number of COPD patients per practice: cohort 1, patients diagnosed in primary care

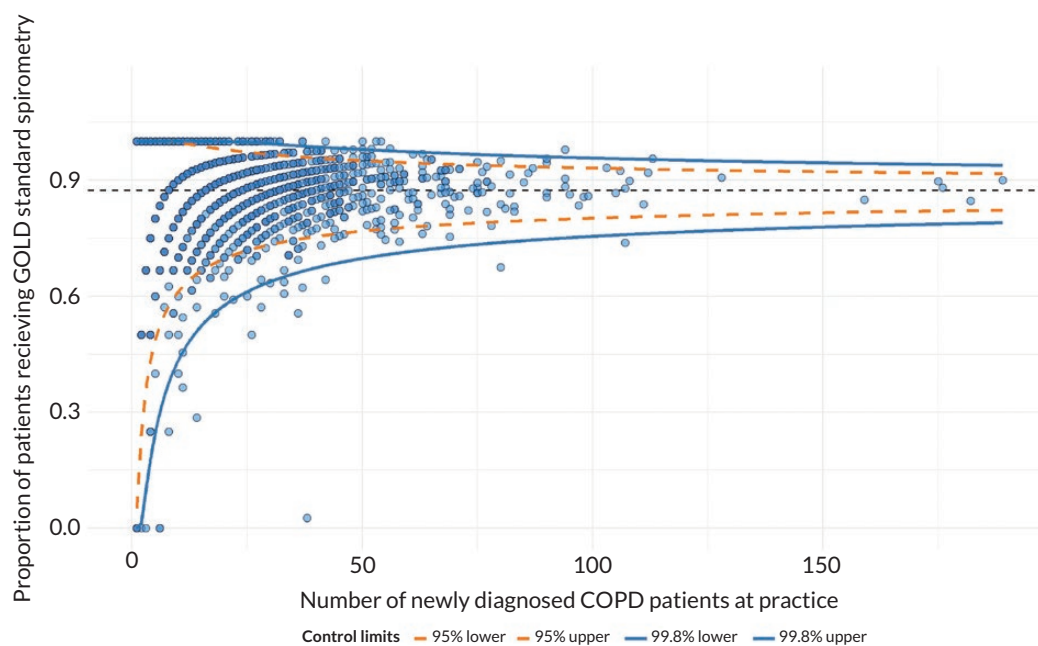


FIGURE 3 Funnel plot of proportion of patients following the initial NICE pathways (performed spirometry within 6 months prior to/after COPD diagnosis) by number of COPD patients per practice: cohort 2, patients diagnosed in primary care

TABLE 7 Funnel plot outlier practices for spirometry within 6 months prior to/after COPD diagnosis by cohort and diagnosis setting

Outlier status	2006–7 cohort, all	2006–7 cohort, just primary care	2016–7 cohort, all	2016–7 cohort, just primary care
Not an outlier	1044 (73.9)	1057 (75.5)	1162 (81.2)	1144 (80.3)
$x > +3$ SD	49 (3.5)	61 (4.4)	41 (2.9)	47 (3.3)
$+3$ SD $> x > +2$ SD	132 (9.3)	148 (10.6)	101 (7.1)	154 (10.8)
-3 SD $< x < -2$ SD	120 (8.5)	92 (6.6)	98 (6.8)	56 (3.9)
$x < -3$ SD	67 (4.7)	42 (3.0)	29 (2.0)	23 (1.6)

Time between diagnosis and first acute exacerbation

For some patients, their diagnosis and first acute exacerbation were the same event; as shown earlier, this proportion was highest in cohort 3. [Table 11](#) gives the AE rate by asthma and HF status, with the follow-up rate capped at 26 months after last eligibility date (i.e. 28 February 2010 and 28 February 2020 for cohorts 1 and 2, respectively) to allow fair comparison between the two pre-COVID cohorts and prevent overlap with the COVID cohort. This was generally slightly shorter in patients with HF and/or asthma. Those with HF, but not those with asthma, had much higher exacerbation rates. The median time to first AECOPD was shorter in cohort 3 due to the lower follow-up time of 6 months. Note that these are times only for patients with an AE, who are in the minority.

[Figure 4](#) summarises the time to first AE by cohort. Cohorts 1 and 2 had very similar AE rates over time, which were higher than that for cohort 3. This is for patients diagnosed in primary care only.

Mortality, however, showed noticeable differences by cohort, as seen in [Figure 5](#), with the COVID-era cohort having the highest risk.

TABLE 8 General practitioner actions in the year following the COPD diagnosis

Variable		2006–7	2016–7	March–August 2020	Total	p-value
LAMA–LABA dual therapy prescribed in the year after diagnosis	Yes	547 (1.3)	5445 (11.3)	805 (16.9)	6797 (7.3)	< 0.001
LABA–ICS dual therapy prescribed in the year after diagnosis	Yes	1904 (4.7)	418 (0.9)	21 (0.4)	2343 (2.5)	< 0.001
LAMA–LABA–ICS triple therapy prescribed in the year after diagnosis	Yes	562 (1.4)	1386 (2.9)	528 (11.1)	2476 (2.6)	< 0.001
LAMA therapy prescribed in the year after diagnosis	Yes	8802 (21.7)	22,327 (46.3)	2151 (45.3)	33,280 (35.6)	< 0.001
LABA therapy prescribed in the year after diagnosis	Yes	4092 (10.1)	7658 (15.9)	1295 (27.3)	13,045 (13.9)	< 0.001
ICS therapy prescribed in the year after diagnosis	Yes	21,348 (52.6)	19,821 (41.1)	1801 (37.9)	42,970 (45.9)	< 0.001
Oral corticosteroids prescribed in the year after COPD diagnosis	Yes	10,143 (25.0)	15,021 (31.1)	1176 (24.7)	26,340 (28.1)	< 0.001
SABA prescribed in the year after COPD diagnosis	Yes	28,038 (69.1)	32,873 (68.1)	2867 (60.3)	63,778 (68.2)	< 0.001
SAMA prescribed in the year after COPD diagnosis	Yes	5039 (12.4)	1004 (2.1)	46 (1.0)	6089 (6.5)	< 0.001
Oxygen prescription within 1 year of COPD diagnosis	Yes	158 (0.4)	4 (0.0)	0 (0.0)	162 (0.2)	< 0.001
Pulmonary rehabilitation within 1 year of COPD diagnosis	Yes	330 (0.8)	6604 (13.7)	993 (20.9)	7927 (8.5)	< 0.001
Smoking cessation drugs prescribed within 1 year of COPD diagnosis	Yes	4118 (21.8)	2597 (12.0)	206 (10.1)	6921 (16.2)	< 0.001
Smoking cessation advice within 1 year of COPD diagnosis	Yes	12,288 (65.1)	17,556 (81.0)	1135 (55.5)	30,979 (72.7)	< 0.001
Evidence of discussion of smoking cessation within 1 year of COPD diagnosis	Yes	12,816 (67.9)	17,802 (82.2)	1193 (58.4)	31,811 (74.7)	< 0.001
Influenza vaccine offered within 1 year of COPD diagnosis	Yes	31,347 (77.3)	41,190 (85.4)	3901 (82.1)	76,438 (81.7)	< 0.001
Influenza vaccine administered within 1 year of COPD diagnosis	Yes	27,580 (68.0)	32,218 (66.8)	3045 (64.1)	62,843 (67.2)	< 0.001

TABLE 9 Characteristics of patients diagnosed via their first AE, by cohort

Variable		2006–7	2016–7	March–August 2020	Total	p-value
Sex	Female	4611 (51.8)	5568 (51.3)	682 (49.3)	10,861 (51.4)	0.210
	Male	4290 (48.2)	5288 (48.7)	702 (50.7)	10,280 (48.6)	
Age (years)	Mean (SD)	71.0 (13.0)	74.0 (12.9)	74.1 (12.7)	72.8 (13.0)	< 0.001
IMD quintile	1	1096 (12.3)	1560 (14.4)	195 (14.1)	2851 (13.5)	< 0.001
	2	1364 (15.3)	1903 (17.5)	262 (18.9)	3529 (16.7)	
	3	1572 (17.7)	2016 (18.6)	264 (19.1)	3852 (18.2)	
	4	2031 (22.8)	2386 (22.0)	290 (21.0)	4707 (22.3)	
	5	2829 (31.8)	2983 (27.5)	373 (27.0)	6185 (29.3)	
Smoking status	Missing	666 (7.5)	646 (6.0)	82 (5.9)	1394 (6.6)	< 0.001
	Never smoker	1416 (15.9)	2361 (21.7)	330 (23.8)	4107 (19.4)	
	Ex-smoker	2929 (32.9)	3731 (34.4)	487 (35.2)	7147 (33.8)	
	Current smoker	3890 (43.7)	4118 (37.9)	485 (35.0)	8493 (40.2)	
BMI measured	Yes	7143 (80.2)	9891 (91.1)	1283 (92.7)	18,317 (86.6)	< 0.001
	No	1758 (19.8)	965 (8.9)	101 (7.3)	2824 (13.4)	
BMI category	Underweight	448 (5.0)	584 (5.4)	76 (5.5)	1108 (5.2)	< 0.001
	Normal	2672 (30.0)	3169 (29.2)	408 (29.5)	6249 (29.6)	
	Overweight	2108 (23.7)	2884 (26.6)	383 (27.7)	5375 (25.4)	
	Obese	1915 (21.5)	3254 (30.0)	416 (30.1)	5585 (26.4)	
	Missing	1758 (19.8)	965 (8.9)	101 (7.3)	2824 (13.4)	
Diastolic blood pressure taken	Yes	8338 (93.7)	10,423 (96.0)	1337 (96.6)	20,098 (95.1)	< 0.001
	No	563 (6.3)	433 (4.0)	47 (3.4)	1043 (4.9)	
Diastolic blood pressure (mmHg)	Mean (SD)	76.9 (11.2)	74.3 (11.3)	75.2 (12.2)	75.4 (11.4)	< 0.001

TABLE 9 Characteristics of patients diagnosed via their first AE, by cohort (*continued*)

Variable		2006–7	2016–7	March–August 2020	Total	p-value
Systolic blood pressure taken	Yes	8329 (93.6)	10,425 (96.0)	1337 (96.6)	20,091 (95.0)	< 0.001
	No	572 (6.4)	431 (4.0)	47 (3.4)	1050 (5.0)	
Systolic blood pressure (mmHg)	Mean (SD)	135.3 (19.2)	130.9 (18.6)	131.7 (19.7)	132.8 (19.0)	< 0.001
Current asthma	Yes	1969 (22.1)	2295 (21.1)	317 (22.9)	4581 (21.7)	0.128
Any malignancy, including leukaemia and lymphoma (CCI)	Yes	1387 (15.6)	2765 (25.5)	392 (28.3)	4544 (21.5)	< 0.001
CVD (CCI)	Yes	1148 (12.9)	1978 (18.2)	235 (17.0)	3361 (15.9)	< 0.001
Chronic pulmonary disease (CCI)	Yes	3526 (39.6)	4496 (41.4)	604 (43.6)	8626 (40.8)	0.003
Congestive HF (CCI)	Yes	1007 (11.3)	1681 (15.5)	271 (19.6)	2959 (14.0)	< 0.001
Dementia (CCI)	Yes	199 (2.2)	965 (8.9)	130 (9.4)	1294 (6.1)	< 0.001
Diabetes without chronic complications (CCI)	Yes	1051 (11.8)	1967 (18.1)	309 (22.3)	3327 (15.7)	< 0.001
Diabetes with chronic complications (CCI)	Yes	422 (4.7)	1413 (13.0)	186 (13.4)	2021 (9.6)	< 0.001
AIDS/HIV (CCI)	Yes	9 (0.1)	19 (0.2)	5 (0.4)	33 (0.2)	0.057
Hemiplegia or paraplegia (CCI)	Yes	68 (0.8)	125 (1.2)	22 (1.6)	215 (1.0)	0.002
Metastatic solid tumour (CCI)	Yes	66 (0.7)	194 (1.8)	29 (2.1)	289 (1.4)	< 0.001
Mild liver disease (CCI)	Yes	100 (1.1)	197 (1.8)	32 (2.3)	329 (1.6)	< 0.001
Moderate or severe liver disease (CCI)	Yes	17 (0.2)	84 (0.8)	17 (1.2)	118 (0.6)	< 0.001
Myocardial infarction (CCI)	Yes	923 (10.4)	1218 (11.2)	152 (11.0)	2293 (10.8)	0.159
Peptic ulcer disease (CCI)	Yes	591 (6.6)	824 (7.6)	105 (7.6)	1520 (7.2)	0.031
PVD (CCI)	Yes	802 (9.0)	1248 (11.5)	145 (10.5)	2195 (10.4)	< 0.001
Renal disease (CCI)	Yes	1116 (12.5)	2920 (26.9)	387 (28.0)	4423 (20.9)	< 0.001
Rheumatological disease (CCI)	Yes	517 (5.8)	975 (9.0)	146 (10.5)	1638 (7.7)	< 0.001
Hypertension	Yes	3749 (42.1)	5811 (53.5)	765 (55.3)	10,325 (48.8)	< 0.001

continued

TABLE 9 Characteristics of patients diagnosed via their first AE, by cohort (continued)

Variable		2006–7	2016–7	March–August 2020	Total	p-value
Anxiety	Yes	1459 (16.4)	2652 (24.4)	412 (29.8)	4523 (21.4)	< 0.001
Depression	Yes	1751 (19.7)	2963 (27.3)	427 (30.9)	5141 (24.3)	< 0.001
Osteoporosis	Yes	589 (6.6)	1272 (11.7)	183 (13.2)	2044 (9.7)	< 0.001
Anaemia	Yes	940 (10.6)	2328 (21.4)	346 (25.0)	3614 (17.1)	< 0.001
Any LRTI diagnosis preceding COPD diagnosis	Yes	4331 (48.7)	6881 (63.4)	927 (67.0)	12,139 (57.4)	< 0.001
Arrhythmia	Yes	225 (2.5)	444 (4.1)	83 (6.0)	752 (3.6)	< 0.001
Stroke	Yes	1797 (20.2)	2937 (27.1)	409 (29.6)	5143 (24.3)	< 0.001
Atrial fibrillation	Yes	923 (10.4)	1923 (17.7)	299 (21.6)	3145 (14.9)	< 0.001
Total number of comorbidities	Mean (SD)	3.3 (2.3)	4.8 (2.8)	5.2 (2.8)	4.2 (2.7)	< 0.001
GOLD status	GOLD stage 1: ≥ 80%	349 (3.9)	1330 (12.3)	222 (16.0)	1901 (9.0)	< 0.001
	GOLD stage 2: 50–79%	547 (6.1)	1329 (12.2)	167 (12.1)	2043 (9.7)	
	GOLD stage 3: 30–49%	268 (3.0)	335 (3.1)	39 (2.8)	642 (3.0)	
	GOLD stage 4: < 30%	60 (0.7)	48 (0.4)	6 (0.4)	114 (0.5)	
	Missing FEV1 %-pred measurement	7677 (86.2)	7814 (72.0)	950 (68.6)	16,441 (77.8)	

CCI, Charlson Comorbidity Index; CVD, cerebrovascular disease; FEV1, forced expiratory volume in 1 second; LRTI, lower respiratory tract infection; PVD, peripheral vascular disease.

TABLE 10 General practitioner actions before diagnosis in patients diagnosed via their first AE, by cohort

Variable	2006–7	2016–7	March–August 2020	Total	p-value	
Spirometry	Yes	1731 (19.4)	3805 (35.0)	541 (39.1)	6077 (28.7)	< 0.001
	No	7170 (80.6)	7051 (65.0)	843 (60.9)	15,064 (71.3)	
Chest X-ray	Yes	3592 (40.4)	6137 (56.5)	867 (62.6)	10,596 (50.1)	< 0.001
	No	5309 (59.6)	4719 (43.5)	517 (37.4)	10,545 (49.9)	
FBC	Yes	6452 (72.5)	9449 (87.0)	1261 (91.1)	17,162 (81.2)	< 0.001
	No	2449 (27.5)	1407 (13.0)	123 (8.9)	3979 (18.8)	
All of spirometry, FBC, CXR and BMI measurement	Yes	855 (9.6)	2695 (24.8)	414 (29.9)	3964 (18.8)	< 0.001
	No	8046 (90.4)	8161 (75.2)	970 (70.1)	17,177 (81.2)	
At least one of FBC, CXR and BMI measurement	Yes	8068 (90.6)	10,345 (95.3)	1336 (96.5)	19,749 (93.4)	< 0.001
	No	833 (9.4)	511 (4.7)	48 (3.5)	1392 (6.6)	
Spirometry, CXR, BMI, FBC	Spirometry and one of CXR/BMI/FBC	1714 (19.3)	3794 (34.9)	541 (39.1)	6049 (28.6)	< 0.001
	Spirometry only	17 (0.2)	11 (0.1)	0 (0.0)	28 (0.1)	
	No spirometry and one of CXR/BMI/FBC	6354 (71.4)	6551 (60.3)	795 (57.4)	13,700 (64.8)	
	No spirometry, CXR, BMI or FBC	816 (9.2)	500 (4.6)	48 (3.5)	1364 (6.5)	
Spirometry in the year before COPD diagnosis	Yes	886 (10.0)	977 (9.0)	98 (7.1)	1961 (9.3)	0.001
Chest X-ray in the year before COPD diagnosis	Yes	1282 (14.4)	2252 (20.7)	343 (24.8)	3877 (18.3)	< 0.001

continued

TABLE 10 General practitioner actions before diagnosis in patients diagnosed via their first AE, by cohort (*continued*)

Variable	2006–7	2016–7	March–August 2020	Total	p-value	
FBC in the year before COPD diagnosis	Yes	4170 (46.8)	6088 (56.1)	790 (57.1)	11,048 (52.3)	< 0.001
BMI measurement in the year before COPD diagnosis	Yes	3561 (40.0)	4413 (40.7)	622 (44.9)	8596 (40.7)	0.002
All of spirometry, FBC, CXR and BMI measurement in the year before COPD diagnosis	Yes	117 (1.3)	203 (1.9)	35 (2.5)	355 (1.7)	< 0.001
At least one of FBC, CXR, and BMI measurement in the year before COPD diagnosis	Yes	5871 (66.0)	7985 (73.6)	1055 (76.2)	14,911 (70.5)	< 0.001
	Spirometry and one of CXR/BMI/FBC	757 (8.5)	882 (8.1)	96 (6.9)	1735 (8.2)	
	Spirometry only	129 (1.4)	95 (0.9)	2 (0.1)	226 (1.1)	
	No spirometry and one of CXR/BMI/FBC	5114 (57.5)	7103 (65.4)	959 (69.3)	13,176 (62.3)	
No spirometry, CXR, BMI or FBC	2901 (32.6)	2776 (25.6)	327 (23.6)	6004 (28.4)		

FBC, full blood count.

TABLE 11 Time to first AECOPD in patients by asthma and HF status

Variable		2006–7	2016–7	March–August 2020	Total	p-value
No asthma or HF: First AECOPD	Yes	2517 (11.7)	3298 (11.9)	76 (3.1)	5891 (11.4)	< 0.001
	No	19,031 (88.3)	24,399 (88.1)	2345 (96.9)	45,775 (88.6)	
First AECOPD per person-years rate	Rate (95% CI)	0.047 (0.045 to 0.049)	0.046 (0.045 to 0.048)	0.039 (0.030 to 0.048)	0.046 (0.045 to 0.048)	0.050
Asthma, no HF: First AECOPD	Yes	941 (12.3)	756 (11.1)	17 (2.8)	1714 (11.4)	< 0.001
	No	6705 (87.7)	6077 (88.9)	581 (97.2)	13,363 (88.6)	
First AECOPD per person-years rate	Rate (95% CI)	0.047 (0.044 to 0.050)	0.042 (0.039 to 0.045)	0.034 (0.018 to 0.050)	0.044 (0.042 to 0.046)	0.019
HF, no asthma: First AECOPD	Yes	313 (20.4)	430 (22.8)	15 (7.4)	758 (20.9)	< 0.001
	No	1222 (79.6)	1458 (77.2)	189 (92.6)	2869 (79.1)	
First AECOPD per person-years rate	Rate (95% CI)	0.103 (0.092 to 0.114)	0.106 (0.096 to 0.116)	0.104 (0.053 to 0.155)	0.105 (0.097 to 0.112)	0.724
Asthma and HF: First AECOPD	Yes	88 (21.9)	94 (22.7)	5 (7.7)	187 (21.2)	0.021
	No	314 (78.1)	320 (77.3)	60 (92.3)	694 (78.8)	
First AECOPD per person-years rate	Rate (95% CI)	0.109 (0.087 to 0.130)	0.100 (0.080 to 0.120)	0.096 (0.012 to 0.180)	0.104 (0.090 to 0.118)	0.780
All patients: First AECOPD	Yes	4145 (13.3)	4778 (13.0)	117 (3.6)	9040 (12.7)	< 0.001
	No	26,986 (86.7)	32,054 (87.0)	3171 (96.4)	62,211 (87.3)	
First AECOPD per person-years rate	Rate (95% CI)	0.050 (0.048 to 0.052)	0.048 (0.047 to 0.050)	0.043 (0.035 to 0.050)	0.049 (0.048 to 0.050)	0.013

RESULTS

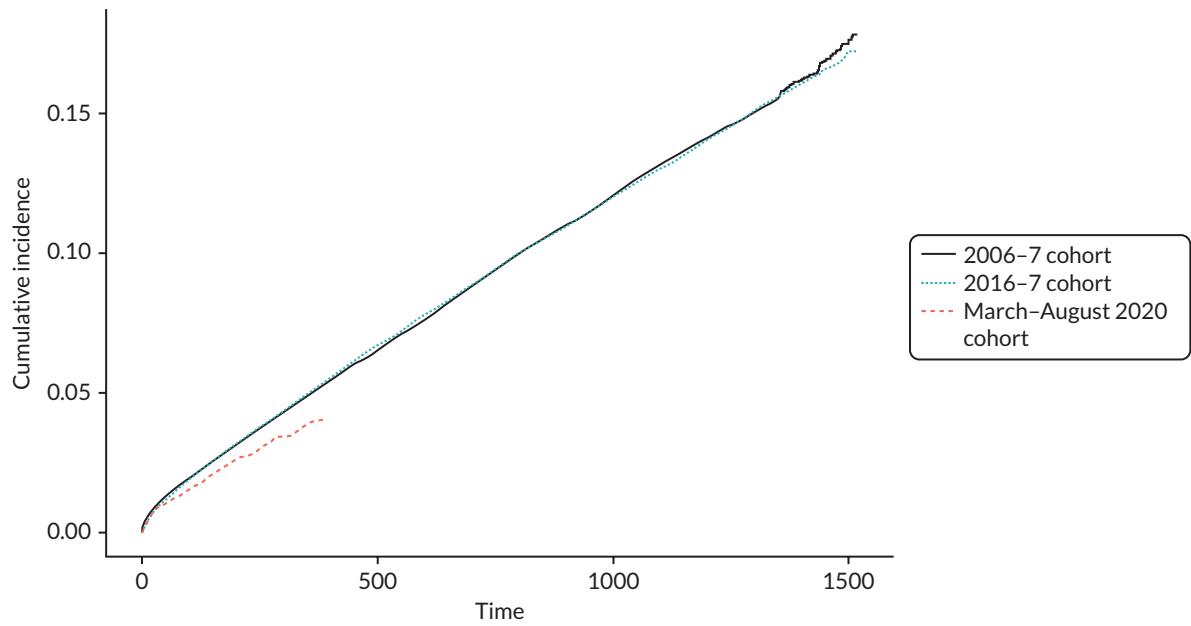


FIGURE 4 Cumulative incidence of AECOPD (time 0 means that COPD and AECOPD were on the same day)

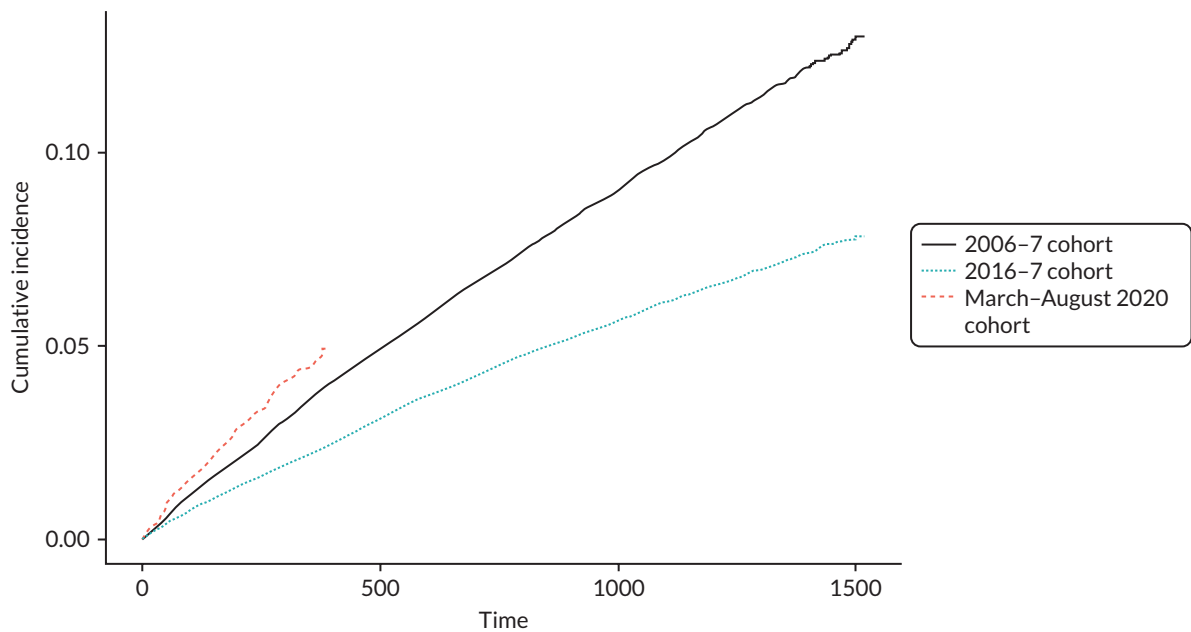


FIGURE 5 Cumulative incidence of all-cause mortality following COPD diagnosis by cohort

Chapter 4 Prediction of first acute exacerbation

Beginning at time of diagnosis, we used Fine–Gray regression to model the time to first AE as competing risks were deemed to be a concern, with 10% and 6% of cohort 1 and 2 patients, respectively, dying from a non-COPD cause. Inspection of residuals revealed that the usual assumptions were met; splines were included due to non-linear associations and to enhance prediction ability.

We fitted three models for each cohort: patient characteristics (model 1), patient characteristics plus GP actions only before diagnosis (model 2) and patient characteristics plus GP actions in the year since diagnosis (model 3).

For cohort 1, model 1, the significant predictors were age, IMD, GOLD stage, asthma, HF, prior acute myocardial infarction (AMI), chronic kidney disease (CKD), depression, osteoporosis, being a current smoker, having unknown smoking status, LRTI in previous 5 years and BMI (negative association). For cohort 2, for this model, the significant predictors were age, IMD, anaemia, GOLD stage, asthma, AF, any malignancy, diabetes with complications, mild liver disease, moderate/severe liver disease, peptic ulcer, peripheral vascular disease (PVD), CKD, rheumatic diseases, depression, osteoporosis, current smoker, LRTI in previous 5 years, sputum disease in previous 5 years and BMI (negative association).

For cohort 1, model 2, the significant predictors were age, IMD, GOLD stage, asthma, HF, prior AMI, CKD, depression, osteoporosis, current smoker, unknown smoking status, LRTI in previous 5 years, spirometry (negative association) and prescription of LAMA, ICS, OCS, SABA and SAMA (all in previous 5 years). The largest HR among the medications was LAMA (1.51, 95% CI 1.34 to 1.70). For cohort 2, model 2, the significant predictors were age, IMD, anaemia, GOLD stage, malignancy, HF, dementia, diabetes with complications, liver disease (any severity), peptic ulcer, PVD, rheumatic disease, depression, osteoporosis, current smoker, never smoker (negative association), LRTI in previous 5 years, BMI (negative association), echo, spirometry (negative association), specialist referral and LAMA, ICS and OCS (all in previous 5 years). The largest HR among the medications was LAMA (1.46, 95% CI 1.35 to 1.58). Regarding the HRs, most of the predictors have overlapping CIs and hence probably similar effects in cohorts 1 and 2.

For cohort 1, model 3, the significant predictors were age, IMD, GOLD stage, HF, prior AMI, depression, hypertension, osteoporosis, BMI (negative association but small effect), spirometry (negative association), LAMA, ICS, OCS, SAMA (all medications 1 year post diagnosis) and LAMA, SAMA and SABA (all up to 5 years post diagnosis), COPD annual review (negative association), spirometry in the year since diagnosis and flu vaccination in the year since diagnosis. As with model 2, the largest HRs among the medications were for LAMA. We did not find an association for PR (HR 1.16, 95% CI 0.87 to 1.55). For cohort 2, model 3, the significant predictors were age, IMD, GOLD stage, male sex (small effect), anaemia, asthma (negative association), HF, dementia, diabetes with complications, liver disease (any severity), PVD, CKD, depression, hypertension (small effect), LRTI in previous 5 years, echo, spirometry (negative association), LAMA up to 5 years post diagnosis, OCS up to 5 years post diagnosis (small effect), LAMA in the year post diagnosis, ICS in the year post diagnosis, OCS in the year post diagnosis, COPD annual review, spirometry, flu vaccination (these three all in the year since diagnosis and with a negative association) and BMI (negative association but small effect). We did not find an association for PR (HR 1.04, 95% CI 0.96 to 1.13).

[Table 12](#) gives the coefficients for model 2 for each cohort.

[Table 13](#) gives the coefficients for model 3 for each cohort.

TABLE 12 Hazard ratios for model 2 for each cohort for first acute exacerbation prediction

Variable	Variable value	Cohort 1 hazard ratio (95% CI)	Cohort 2 hazard ratio (95% CI)
IMD quintile	1	Reference	Reference
	2	1.08 (0.96 to 1.22)	1.19 (1.07 to 1.32)
	3	1.18 (1.05 to 1.32)	1.20 (1.08 to 1.33)
	4	1.32 (1.19 to 1.47)	1.28 (1.16 to 1.42)
	5	1.45 (1.31 to 1.61)	1.31 (1.19 to 1.45)
	Missing IMD quintile	0.32 (0.03 to 3.24)	1.05 (0.27 to 4.02)
Male sex		1.05 (0.98 to 1.12)	1.06 (0.99 to 1.13)
Anaemia		0.99 (0.88 to 1.11)	1.21 (1.12 to 1.31)
GOLD status	GOLD stage 1: ≥ 80%	Reference	Reference
	GOLD stage 2: 50–79%	1.22 (1.05 to 1.41)	1.30 (1.20 to 1.42)
	GOLD stage 3: 30–49%	2.20 (1.89 to 2.56)	2.23 (2.01 to 2.47)
	GOLD stage 4: < 30%	3.80 (3.13 to 4.61)	3.06 (2.52 to 3.72)
	Missing FEV1 %-pred measurement	1.93 (1.64 to 2.27)	1.63 (1.45 to 1.84)
Anxiety		1.04 (0.95 to 1.14)	1.03 (0.96 to 1.11)
Arrhythmia		0.99 (0.80 to 1.22)	0.92 (0.77 to 1.10)
Current asthma		0.93 (0.86 to 1.02)	0.92 (0.84 to 1.01)
Atrial fibrillation		0.98 (0.87 to 1.11)	1.08 (0.98 to 1.20)
Any malignancy, including leukaemia and lymphoma (CCI)		0.91 (0.83 to 1.00)	1.08 (1.00 to 1.16)
CVD (CCI)		1.07 (0.96 to 1.18)	1.06 (0.96 to 1.16)
Congestive HF (CCI)		1.44 (1.28 to 1.63)	1.20 (1.08 to 1.35)
Dementia (CCI)		1.28 (0.95 to 1.73)	1.31 (1.11 to 1.54)
Diabetes without chronic complications (CCI)		1.09 (0.98 to 1.21)	1.04 (0.96 to 1.12)
Diabetes with chronic complications (CCI)		1.02 (0.86 to 1.21)	1.20 (1.08 to 1.33)
AIDS/HIV (CCI)		1.27 (0.43 to 3.73)	0.81 (0.50 to 1.32)
Hemiplegia or paraplegia (CCI)		1.04 (0.69 to 1.57)	1.27 (0.89 to 1.83)
Metastatic solid tumour (CCI)		1.03 (0.61 to 1.73)	1.13 (0.78 to 1.64)
Mild liver disease (CCI)		1.28 (0.96 to 1.71)	1.40 (1.14 to 1.71)
Moderate or severe liver disease (CCI)		0.99 (0.42 to 2.30)	1.86 (1.26 to 2.74)
Myocardial infarction (CCI)		1.15 (1.03 to 1.28)	1.03 (0.93 to 1.15)
Peptic ulcer disease (CCI)		1.09 (0.97 to 1.22)	1.13 (1.01 to 1.25)
PVD (CCI)		1.08 (0.97 to 1.21)	1.28 (1.17 to 1.40)
Renal disease (CCI)		1.12 (1.01 to 1.25)	1.10 (1.02 to 1.19)
Rheumatological disease (CCI)		1.06 (0.93 to 1.21)	1.12 (1.00 to 1.25)
Depression		1.11 (1.02 to 1.21)	1.12 (1.05 to 1.21)
Hypertension		1.05 (0.99 to 1.13)	1.05 (0.99 to 1.12)

TABLE 12 Hazard ratios for model 2 for each cohort for first acute exacerbation prediction (*continued*)

Variable	Variable value	Cohort 1 hazard ratio (95% CI)	Cohort 2 hazard ratio (95% CI)
Osteoporosis		1.27 (1.12 to 1.43)	1.15 (1.04 to 1.28)
Smoking status	Ex-smoker	Reference	Reference
	Current smoker	1.39 (1.30 to 1.50)	1.41 (1.32 to 1.51)
	Never smoker	0.67 (0.60 to 0.76)	0.71 (0.63 to 0.79)
	Missing	1.35 (1.11 to 1.63)	1.07 (0.93 to 1.23)
Any breathlessness diagnosis in 5 years preceding COPD diagnosis		1.05 (0.99 to 1.12)	1.01 (0.95 to 1.08)
Any cough diagnosis in 5 years preceding COPD diagnosis		0.85 (0.70 to 1.03)	0.88 (0.77 to 1.00)
Any LRTI diagnosis in 5 years preceding COPD diagnosis		1.07 (1.01 to 1.14)	1.17 (1.10 to 1.24)
Any sputum diagnosis in 5 years preceding COPD diagnosis		1.05 (0.96 to 1.14)	1.06 (0.98 to 1.14)
Any wheeze diagnosis in 5 years preceding COPD diagnosis		0.91 (0.82 to 1.01)	0.99 (0.90 to 1.08)
BMI		0.98 (0.97 to 0.99)	0.99 (0.98 to 0.99)
BMI ²		1.001 (1.001 to 1.001)	1.001 (1.001 to 1.001)
Systolic blood pressure taken		0.999 (0.997 to 1.001)	0.999 (0.998 to 1.001)
Chest X-ray		0.995 (0.94 to 1.06)	0.998 (0.94 to 1.06)
Echocardiogram before COPD diagnosis		1.05 (0.95 to 1.17)	1.20 (1.11 to 1.30)
Spirometry (pre diagnosis)		0.84 (0.76 to 0.94)	0.80 (0.71 to 0.90)
Respiratory referral before COPD diagnosis		1.55 (0.81 to 2.96)	1.12 (1.02 to 1.23)
Cardiology referral before COPD diagnosis		0.95 (0.86 to 1.06)	0.95 (0.87 to 1.03)
LAMA-LABA dual therapy prescribed in the 5 years preceding COPD diagnosis		0.98 (0.66 to 1.46)	0.89 (0.66 to 1.21)
ICS-LABA dual therapy prescribed in the 5 years preceding COPD diagnosis		1.12 (0.93 to 1.35)	0.92 (0.68 to 1.25)
LAMA-LABA-ICS triple therapy prescribed in the 5 years preceding COPD diagnosis		0.87 (0.63 to 1.19)	0.79 (0.54 to 1.17)
LAMA therapy prescribed in the 5 years preceding COPD diagnosis		1.51 (1.34 to 1.70)	1.46 (1.35 to 1.58)
LABA therapy prescribed in the 5 years preceding COPD diagnosis		1.17 (0.99 to 1.39)	1.18 (0.93 to 1.51)
ICS therapy prescribed in the 5 years preceding COPD diagnosis		1.09 (1.00 to 1.19)	1.09 (1.00 to 1.18)
Oral corticosteroids prescribed in the 5 years preceding COPD diagnosis		1.17 (1.09 to 1.26)	1.16 (1.08 to 1.24)
SABA prescribed in the 5 years preceding COPD diagnosis		1.14 (1.05 to 1.24)	1.08 (1.00 to 1.15)
SAMA prescribed in the 5 years preceding COPD diagnosis		1.24 (1.13 to 1.37)	1.11 (0.93 to 1.32)

CCI, Charlson Comorbidity Index; CVD, cerebrovascular disease; FEV1, forced expiratory volume in 1 second; LRTI, lower respiratory tract infection; PVD, peripheral vascular disease.

TABLE 13 Hazard ratios for model 3 for each cohort for first acute exacerbation prediction

Variable	Variable value	Cohort 1 hazard ratio (95% CI)	Cohort 2 hazard ratio (95% CI)
IMD quintile	1	Reference	Reference
	2	1.07 (0.95 to 1.20)	1.20 (1.08 to 1.33)
	3	1.17 (1.05 to 1.31)	1.19 (1.07 to 1.32)
	4	1.31 (1.17 to 1.46)	1.27 (1.15 to 1.41)
	5	1.39 (1.25 to 1.55)	1.25 (1.13 to 1.38)
	Missing IMD quintile	0.28 (0.02 to 3.31)	1.12 (0.29 to 4.28)
Male sex		1.07 (1.00 to 1.14)	1.08 (1.01 to 1.15)
Anaemia		0.98 (0.87 to 1.09)	1.19 (1.10 to 1.29)
GOLD status	GOLD stage 1: ≥ 80%	Reference	Reference
	GOLD stage 2: 50–79%	1.17 (1.01 to 1.36)	1.28 (1.18 to 1.40)
	GOLD stage 3: 30–49%	2.00 (1.71 to 2.33)	2.08 (1.88 to 2.31)
	GOLD stage 4: < 30%	3.24 (2.66 to 3.94)	2.79 (2.28 to 3.41)
	Missing FEV1 %-pred measurement	1.81 (1.54 to 2.14)	1.53 (1.35 to 1.73)
Anxiety		1.04 (0.95 to 1.14)	1.04 (0.96 to 1.12)
Arrhythmia		0.99 (0.79 to 1.23)	0.93 (0.77 to 1.11)
Current asthma		0.94 (0.86 to 1.02)	0.90 (0.82 to 0.99)
Atrial fibrillation		1.01 (0.89 to 1.14)	1.03 (0.93 to 1.15)
Any malignancy, including leukaemia and lymphoma (CCI)		0.91 (0.83 to 1.00)	1.07 (1.00 to 1.16)
CVD (CCI)		1.05 (0.95 to 1.17)	1.06 (0.96 to 1.16)
Congestive HF (CCI)		1.40 (1.23 to 1.58)	1.18 (1.06 to 1.33)
Dementia (CCI)		1.23 (0.91 to 1.66)	1.25 (1.06 to 1.48)
Diabetes without chronic complications (CCI)		1.11 (1.00 to 1.23)	1.04 (0.96 to 1.13)
Diabetes with chronic complications (CCI)		1.06 (0.89 to 1.26)	1.21 (1.09 to 1.35)
AIDS/HIV (CCI)		1.20 (0.41 to 3.51)	0.82 (0.50 to 1.34)
Hemiplegia or paraplegia (CCI)		1.03 (0.67 to 1.58)	1.19 (0.82 to 1.71)
Metastatic solid tumour (CCI)		0.93 (0.56 to 1.54)	0.99 (0.67 to 1.46)
Mild liver disease (CCI)		1.26 (0.94 to 1.70)	1.42 (1.16 to 1.74)
Moderate or severe liver disease (CCI)		1.00 (0.42 to 2.36)	1.94 (1.32 to 2.86)
Myocardial infarction (CCI)		1.15 (1.03 to 1.29)	1.02 (0.92 to 1.14)
Peptic ulcer disease (CCI)		1.10 (0.98 to 1.24)	1.11 (0.99 to 1.24)
PVD (CCI)		1.09 (0.97 to 1.21)	1.27 (1.16 to 1.40)
Renal disease (CCI)		1.11 (0.99 to 1.23)	1.10 (1.01 to 1.19)
Rheumatological disease (CCI)		1.02 (0.90 to 1.17)	1.11 (0.90 to 1.24)
Depression		1.12 (1.03 to 1.23)	1.12 (1.04 to 1.20)

TABLE 13 Hazard ratios for model 3 for each cohort for first acute exacerbation prediction (*continued*)

Variable	Variable value	Cohort 1 hazard ratio (95% CI)	Cohort 2 hazard ratio (95% CI)
Hypertension		1.09 (1.02 to 1.17)	1.07 (1.00 to 1.14)
Osteoporosis		1.25 (1.11 to 1.42)	1.15 (1.03 to 1.27)
Any breathlessness diagnosis in 5 years preceding COPD diagnosis		1.04 (0.97 to 1.11)	1.04 (0.98 to 1.12)
Any cough diagnosis in 5 years preceding COPD diagnosis		0.85 (0.70 to 1.03)	0.89 (0.78 to 1.02)
Any LRTI diagnosis in 5 years preceding COPD diagnosis		1.06 (1.00 to 1.14)	1.14 (1.07 to 1.22)
Any sputum diagnosis in 5 years preceding COPD diagnosis		1.05 (0.97 to 1.15)	1.06 (0.98 to 1.14)
Any wheeze diagnosis in 5 years preceding COPD diagnosis		0.92 (0.82 to 1.02)	0.98 (0.89 to 1.07)
Chest X-ray		1.00 (0.94 to 1.07)	0.99 (0.93 to 1.06)
Echocardiogram before COPD diagnosis		1.09 (0.98 to 1.21)	1.22 (1.12 to 1.32)
Spirometry (pre diagnosis)		0.86 (0.77 to 0.96)	0.87 (0.77 to 0.97)
Respiratory referral before COPD diagnosis		1.33 (0.68 to 2.60)	1.10 (1.00 to 1.20)
Cardiology referral before COPD diagnosis		0.96 (0.86 to 1.07)	0.95 (0.88 to 1.03)
LAMA-LABA dual therapy prescribed in the 5 years preceding COPD diagnosis		1.08 (0.72 to 1.61)	0.85 (0.62 to 1.16)
ICS-LABA dual therapy prescribed in the 5 years preceding COPD diagnosis		1.14 (0.94 to 1.38)	0.98 (0.72 to 1.34)
LAMA-LABA-ICS triple therapy prescribed in the 5 years preceding COPD diagnosis		0.85 (0.61 to 1.19)	0.83 (0.55 to 1.23)
LAMA therapy prescribed in the 5 years preceding COPD diagnosis		1.27 (1.12 to 1.44)	1.30 (1.19 to 1.42)
LABA therapy prescribed in the 5 years preceding COPD diagnosis		1.15 (0.97 to 1.37)	1.14 (0.89 to 1.47)
ICS therapy prescribed in the 5 years preceding COPD diagnosis		1.06 (0.97 to 1.16)	1.05 (0.96 to 1.15)
Oral corticosteroids prescribed in the 5 years preceding COPD diagnosis		1.07 (0.99 to 1.16)	1.07 (1.00 to 1.15)
SAMA prescribed in the 5 years preceding COPD diagnosis		1.21 (1.09 to 1.35)	1.06 (0.88 to 1.27)
SABA prescribed in the 5 years preceding COPD diagnosis		1.11 (1.02 to 1.21)	1.04 (0.96 to 1.12)
LAMA-LABA dual therapy prescribed in the year after diagnosis		0.87 (0.67 to 1.14)	1.06 (0.89 to 1.27)
LABA-ICS dual therapy prescribed in the year after diagnosis		0.91 (0.74 to 1.11)	0.88 (0.62 to 1.24)
LAMA-LABA-ICS triple therapy prescribed in the year after diagnosis		0.91 (0.70 to 1.18)	0.74 (0.60 to 0.91)
LAMA therapy prescribed in the year after diagnosis		1.41 (1.30 to 1.52)	1.30 (1.21 to 1.39)

continued

TABLE 13 Hazard ratios for model 3 for each cohort for first acute exacerbation prediction (continued)

Variable	Variable value	Cohort 1 hazard ratio (95% CI)	Cohort 2 hazard ratio (95% CI)
LABA therapy prescribed in the year after diagnosis		1.09 (0.92 to 1.28)	1.17 (0.98 to 1.39)
ICS therapy prescribed in the year after diagnosis		1.10 (1.01 to 1.19)	1.12 (1.03 to 1.21)
Oral corticosteroids prescribed in the year after COPD diagnosis		1.39 (1.29 to 1.49)	1.30 (1.22 to 1.39)
SAMA prescribed in the year after COPD diagnosis		1.11 (1.01 to 1.22)	1.13 (0.93 to 1.37)
SABA prescribed in the year after COPD diagnosis		1.04 (0.96 to 1.13)	1.04 (0.97 to 1.12)
Oxygen prescription within 1 year of COPD diagnosis		1.33 (0.85 to 2.07)	1.64 (0.42 to 6.35)
COPD review within 1 year of diagnosis		0.92 (0.86 to 0.98)	0.63 (0.58 to 0.69)
Spirometry taken within 1 year of diagnosis		0.83 (0.78 to 0.89)	0.83 (0.78 to 0.88)
Influenza vaccine administered within 1 year of COPD diagnosis		0.51 (0.48 to 0.55)	0.56 (0.53 to 0.60)
Pulmonary rehabilitation within 1 year of COPD diagnosis		1.16 (0.87 to 1.55)	1.04 (0.96 to 1.13)
Smoking and smoking cessation composite variable	Ex-smoker	Reference	Reference
	Current smoker, cessation evidence	1.42 (1.32 to 1.54)	1.38 (1.29 to 1.48)
	Current smoker, no cessation evidence	1.20 (1.08 to 1.33)	1.24 (1.09 to 1.41)
	Never smoker	0.67 (0.60 to 0.76)	0.69 (0.62 to 0.78)
	Missing smoking status	1.27 (1.05 to 1.54)	1.03 (0.89 to 1.19)
BMI		0.98 (0.97 to 0.99)	0.99 (0.98 to 0.99)
BMI ²		1.001 (1.001 to 1.001)	1.001 (1.001 to 1.001)
Systolic blood pressure taken		0.999 (0.997 to 1.001)	0.999 (0.997 to 1.001)

CCI, Charlson Comorbidity Index; CVD, cerebrovascular disease; FEV1, forced expiratory volume in 1 second; LRTI, lower respiratory tract infection; PVD, peripheral vascular disease.

Hazard ratios for many factors were similar in the two cohorts. Hazards were consistently higher in more deprived areas, at older ages, for males, current smokers, those with higher GOLD stage and for some comorbidities (HF, depression, hypertension and osteoporosis), though a few comorbidities, such as liver disease, were statistically significant only in cohort 2. The HRs for GP actions were remarkably consistent between the cohorts. The coefficients for age are not shown due to the use of splines, with knots at the 0.1, 0.5 and 0.9 quantiles, making interpretation more difficult.

Discrimination (c-statistic) varied by cohort and time since diagnosis but was generally moderate at 0.68–0.70 for the sets of models with just patient characteristics with or without pre-diagnosis GP actions (models 1 and 2). For model 3, which included post-diagnosis GP actions, the c-statistics were

higher, especially after the first 6 months since diagnosis, at 0.79 for cohort 1 (falling gradually to 0.70 after 4 years) and 0.81 for cohort 2 (falling gradually to 0.72 after 4 years). Calibration was good for all models except for some underestimation in the highest-risk decile. [Figure 6](#) shows the calibration plot for cohort 2's model 3.

[Table 14](#) gives the population attributable risks (PARs) for the statistically significant variables from model 3 for cohort 2. The most important ones for variables with a positive association with the outcome [i.e. adjusted hazard ratios (aHRs) > 1] were current smoking, GOLD stage and IMD quintile; the most important ones for variables with a negative association with the outcome (i.e. aHRs < 1) were an annual COPD review and flu vaccination.

Validation of risk prediction model

We ran internal cross-validation of the above model and split the data into testing and training portions. This showed good consistency.

Potential cost savings for a population of 250,000

We did not have CCG in our data extract, so in [Table 15](#) we illustrate the potential costs and savings to an average CCG serving a population of 250,000 people that includes 5000 COPD patients, based on an NHS Innovation Accelerator called 'myCOPD'.³⁰ The rate of first AECOPD was calculated with data from cohort 2 by dividing the total number of first AECOPD events by the total person-years at

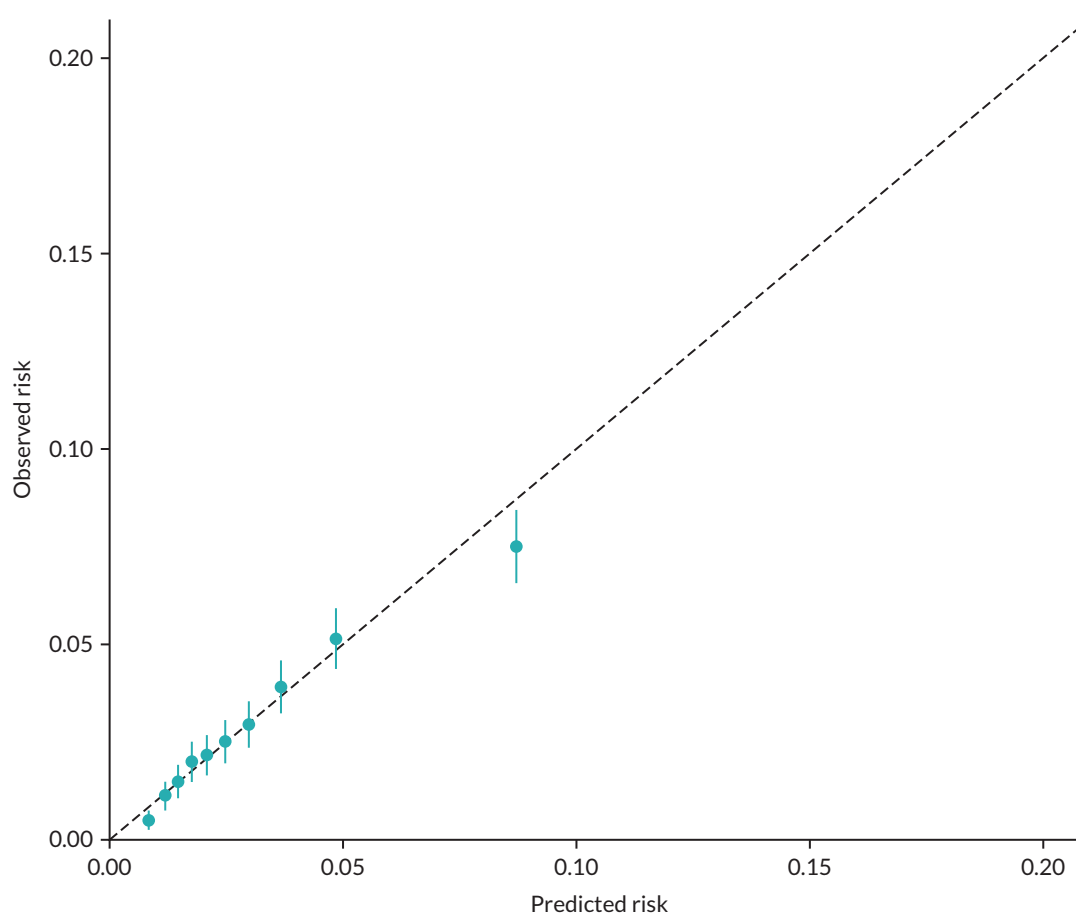


FIGURE 6 Calibration for the final Cox model for cohort 2 at 6 months since diagnosis

TABLE 14 Cohort 2's PARs from the model including post-diagnosis GP actions

Variable	PAR
BMI	0.022
Age	0.012
IMD quintile	0.154
GOLD stage	0.306
Current smokers given cessation advice	0.328
Sex	0.041
Anaemia	0.042
Asthma	-0.013
CCI congestive HF	0.035
CCI dementia	0.018
CCI diabetes with end-organ damage	0.026
CCI liver disease, mild	0.007
CCI liver disease, moderate or severe	0.005
CCI PVD	0.041
CCI CKD	0.029
Depression	0.024
Hypertension	0.039
Osteoporosis	0.015
LRTI diagnosed in 5 years before diagnosis	0.066
Echocardiogram	0.071
Spirometry before diagnosis	-0.083
LAMA 5 years pre dx	0.045
OCS 5 years pre dx	0.028
Triple therapy 1 year since dx	-0.006
LAMA 1 year since dx	0.108
ICS 1 year since dx	0.037
OCS 1 year since dx	0.079
COPD review 1 year since dx	-0.273
Spirometry 1 year since dx	-0.052
Flu vaccination given 1 year since dx	-0.266

CCI, Charlson Comorbidity Index; CKD, chronic kidney disease; LRTI, lower respiratory tract infection; PVD, peripheral vascular disease.

TABLE 15 Potential cost savings due to AE reduction per 5000 COPD patients

Variable	PAR	Savings (£)
BMI	0.022	15,600
Smoking status and smoking cessation advice composite variable	0.328	232,000
Anaemia	0.042	29,700
CCI congestive HF	0.035	24,800
CCI dementia	0.018	12,700
CCI diabetes with end-organ damage	0.026	18,400
CCI liver disease, mild	0.007	5000
CCI liver disease, moderate or severe	0.005	3500
CCI PVD	0.041	29,000
CCI CKD	0.029	20,500
Depression	0.024	17,000
Hypertension	0.039	27,600
Osteoporosis	0.015	10,600
Spirometry before diagnosis	-0.083	58,700
COPD review 1 year since dx	-0.273	193,100
Spirometry 1 year since dx	-0.052	36,800
Flu vaccination given 1 year since dx	-0.266	188,200

Dx, diagnosis; CCI, Charlson Comorbidity Index; CKD, chronic kidney disease; PVD, peripheral vascular disease.

Note

Savings calculated based upon a total cost of £707,435 for 5000 COPD patients over the course of 1 year.

risk (restricted to February 2020 at the latest). Four thousand seven hundred and seventy-eight events occurred during 98,596 patient-years, to give a rate of 0.048 first AECOPD per patient-year. The rate was multiplied by the weighted average cost of a COPD hospitalisation (£2919.65) for the most recent published year (FY 2020–1) using cost data obtained from www.england.nhs.uk/national-cost-collection/. This in turn was multiplied by the expected number of 5000 COPD patients at an average CCG, to give a total yearly cost of £707,435. Potential cost savings were calculated by multiplying the total cost by the PAR. The table suggests that the biggest savings could potentially be made by patients stopping smoking, COPD annual reviews and flu vaccination.

Chapter 5 Qualitative component

Twitter online survey results

The questionnaire was published on 1 February 2022, and access to the questionnaire was closed on 22 April 2022. Seventy-six participants responded to this questionnaire. However, there were only four completed surveys, with the rest being partially completed ($N = 9$) or left totally blank ($N = 63$). There were 9 participants who responded to question 1; 7 participants responded to questions 2 and 18; 6 participants responded to questions 3 and 5; 5 participants responded to question 4, question 9 and questions 6–17; and only 4 participants responded to questions 20–23.

British Lung Foundation online survey results

Full results, including a comparison between the four diseases covered, are given in a forthcoming publication. In brief, there were 156 responses by COPD patients, of whom 124 (79.5%) were female. Ninety-one per cent were white; 82.7% were aged between 45 and 74; 29.5% said they had another heart or lung condition; 9.0% ($n = 14$) were diagnosed since the COVID era began, too few to allow more detailed analysis by time period.

The results suggested that a combination of patient and healthcare system factors was important regarding the time to diagnosis. The most common symptoms reported before seeking professional help were increased breathlessness, chest infections, wheezing and tiredness, with smaller numbers reporting low mood, weight loss and coughing up blood. Also, 39.1% ($n = 61$) said that these symptoms were restricting daily activities. Thirty-four per cent of respondents had not heard of the condition before their diagnosis; 37.2% reported that they did not know the signs of potential lung disease; and 58.3% said they did not appreciate the severity or urgency of the situation, which is likely why nearly half said that their first act on realising that there was something wrong with them was to do nothing, hoping the symptoms would disappear, whereas 39.1% went to see their GP.

Just over half said they were diagnosed by their GP, 19.2% in an outpatient department and 1 in 10 in A&E or as an inpatient. Two-thirds reported having had spirometry and nearly half having had a chest X-ray. One in five had a computed tomography (CT) scan, and the same number had blood tests and an oxygen saturation test (finger probe).

When asked what the causes of their COPD were, respondents were allowed to select multiple options. Tobacco exposure was selected by everyone – 78.2% for their own smoking, 21.8% for passive smoking – but genetics (29.5%), air pollution (21.2%) and previous infections (20.5%) were all reported by more than one in five. Poor housing, poor health and exposure at work were also ticked.

Waiting time prior to a diagnosis varied across COPD respondents, from less than a week (12.8% of respondents) to 10 or more years ($n = 4$, 2.6%), though a time of a few months was the commonest experience. Nearly half said their symptoms worsened between first seeking professional help and getting the diagnosis; 39.1% believed that their diagnosis was delayed, with a variety of barriers to getting a diagnosis proposed; 19.2% believed their diagnosis was delayed due to the health professionals' lack of expertise or knowledge. Nearly one in five thought the healthcare professional did not take enough time to investigate their case or they reported being initially treated for another lung condition, and one in nine respondents felt they had to fight for their care. One in 10 thought they had been misdiagnosed. Other options commonly ticked were difficulty in getting appointments and lack of follow-up to discuss test results.

Clinical Practice Research Datalink online survey results

The survey was open from 26 October 2022 to 6 February 2023, and there were 21 responses but only 17 valid ones. We therefore limited the analysis to some comparisons with the charity website survey, which shows that there were consistent responses. For example, the most common symptoms reported before seeking professional help were increased breathlessness (61.5%), cough (47.1%) and chest infections (41.2%). Also, 58.5% ($n = 10$) of respondents had heard of the condition before their diagnosis; 47.1% reported that they did not know the signs of potential lung disease; 41.2% said they were diagnosed by their GP and 23.5% in an outpatient department; and 70.6% responded that they smoked.

Patient focus group analysis

As only one focus group was run in time for submission of this report, we do not present the findings here.

Chapter 6 Discussion

This mixed-methods project used routine electronic health record (EHR) data from CPRD with new data from surveys to describe what happens between the first symptoms, diagnosis and first acute exacerbation for COPD in an unselected population in England. The quantitative data compared three cohorts: two of them 10 years apart as per the original protocol and a third, smaller one during the COVID era.

Main findings

The age and sex distribution changed little by cohort, but there was more obesity but lower BP in cohorts 2 and 3. Recorded comorbidity rose over time, especially renal disease, anxiety, depression, anaemia, stroke and atrial fibrillation.

Around three-quarters were diagnosed in primary care, with a slight fall in this proportion in cohort 3. Patients diagnosed in hospital were older and more likely to have comorbidities, such as HF, diabetes and renal disease.

From CPRD, the time from earliest COPD symptom to COPD diagnosis had a median of 700 days in cohort 1, 830 days in cohort 2 and 957 days in cohort 3, with wide interquartile ranges. Subanalyses of the time between different recorded symptoms and diagnosis also showed statistically significant differences between cohorts, with cohort 3 having the longest times to diagnosis. From our self-selected survey respondents, the time to diagnosis varied from less than a week to several years, though a few months was the commonest experience. Their perception of where they were diagnosed broadly aligned with what we found in CPRD data. Thirty-nine per cent believed that their diagnosis was delayed, with a variety of system barriers to getting a diagnosis proposed, such as difficulties getting an appointment and not always being taken seriously by the doctors. The survey highlighted patient factors too, such as that the majority did not know the symptoms and the admission by many respondents that they had not taken their symptoms seriously enough, hoping they would go away.

Compliance with NICE guidelines for diagnosis rose a little over time, with much non-random variation in spirometry use between practices. Only a small minority received all four recommended tests and measurements. Eighty per cent were offered the flu vaccination by the GP practice, which is important as a systematic review concluded that this prevented some AEs and hospital admissions.³¹

Although it has improved over time, PR referral remains low at 20% in cohort 3, despite NICE NG115 published in December 2018. However, conventional face-to-face PR programmes were widely suspended during the early waves of COVID-19 to protect vulnerable groups, with many staff redeployed to prioritise those acutely unwell.

Prescribing changed in several ways over time. A systematic review found LABA and LAMA to be effective in lowering the risk of hospital admission,³² and both medications, either given singly or in combination, rose considerably over time in our analysis. The use of SABA changed little and is still given to 60% of patients in the year following diagnosis. The use of inhaled steroids fell from 53% to 38% but remains common. One study of 511 patients at 10 practices found that ICS was often inappropriately given to those with only mild to moderate disease.³³ In the year since diagnosis, around one in four were given oral steroids in our cohorts 1 and 3 (31% in cohort 2). This is not normally recommended in maintenance therapy, being the treatment for AECOPDs, so this rate is likely higher than it should be. Triple therapy rose over time and was 11% for cohort 3. A 2021 meta-analysis of six randomised controlled trials (RCTs) found that triple therapy led to lower rates of AECOPDs and mortality plus less dyspnoea compared with dual LAMA-LABA treatment, though at the cost of 52% higher odds of

pneumonia.³⁴ However, these results relate to those with symptomatic, moderate to severe COPD and a history of exacerbations. We cannot tell from our data how many minor AECOPDs our patients had.

We identified a number of consistent predictors of the first AECOPD. Patient factors included age, deprivation, GOLD stage, being a current smoker, various comorbidities, such as osteoporosis, asthma and depression. Other comorbidities, such as diabetes with complications, HF and CKD, were often significant but not in every model. Pre- and post-diagnosis medications were often associated with higher hazards, particularly LAMA. Annual COPD review, post-diagnosis spirometry and flu vaccination all reduced the hazard, but we did not find any significant association for specialist referral or PR.

The performance of the risk prediction models that did not include post-diagnosis GP actions was moderate, rather than poor or good, with a *c*-statistic of around 0.70. This is not good enough to be used for guiding individual patients and their medical team. Adding post-diagnosis GP actions improved the discrimination considerably, especially over the first 6–12 months since diagnosis.

In terms of PARs, the most important variables with a positive association with the outcome (i.e. aHRs > 1) were current smoking, GOLD stage and IMD quintile; the most important variables with a negative association with the outcome (i.e. aHRs < 1) were an annual COPD review and flu vaccination. Converting the modifiable risk factors for the first AECOPD into healthcare costs for a population of 5000 patients with COPD, the biggest savings could potentially be made with smoking cessation advice (£232K), COPD annual reviews (£193K) and flu vaccination (£188K). It should be remembered that the PARs all assume causal relations with the outcome and are meant as a guide.

Comparison with the literature

For clinical guidelines to be effective, medical staff need to know about them and be able and competent to act on them; patients then need to comply with the management plan. For example, it is not enough for spirometry to be recommended; it needs to be available and accessible to clinicians and patients. GPs need to be able to interpret the findings and act on them accordingly. In previous work, our group examined 218 spirometry traces conducted in primary care that were of sufficient quality and for which a GP diagnosis of COPD had been made; only three further traces were poor quality. Two chest physicians reviewed these 218 traces and found that 72.5% of them showed evidence of obstruction, suggesting an unmet training in primary care.³⁵ A Dutch study showed greater disagreement between clinicians. Fifteen GPs and 2 pulmonologists assessed 149 spirometry tests and questionnaires on clinical usefulness and formulated a diagnosis. Low agreements were found on diagnosis between GPs and pulmonologists 1 ($\kappa = 0.39$) and 2 ($\kappa = 0.44$).³⁶

Following diagnosis, prescribing is another core primary care activity and is known to vary between clinicians, though not a feature that we assessed in this project. A Dutch study among 219 GPs and 25 GPs in training asked about prednisolone prescription for AECOPD. They found good uniformity within the scope of Dutch guidelines between exacerbations but substantial variation in treatment duration in response to exacerbation severity, disease severity and in patients with diabetes, where steroids can cause hyperglycaemia.³⁷ Despite a fall over time, we found high rates of ICS prescribed alone in the year since diagnosis. ICS with LABA or in triple therapy is intended to reduce AECOPDs but carries with it an increased risk of complications such as osteoporosis and pneumonia and so should not be given to patients with GOLD A/B (i.e. not severe disease). A CPRD study on data for 2005–15 found various predictors of ICS in GOLD A/B and a fall over time in use.³⁸ The authors showed regional variation and stated that GP practice was a significant predictor, but they did not appear to use multilevel modelling or otherwise appropriately handle the clustering.

Even if a medication has been prescribed appropriately, the patient needs to take it properly. Inhaler technique is of such importance that its assessment by the medical team is commonly used as a

performance indicator. The most recent national COPD audit found that only 28.2% of those prescribed an inhaler had their technique checked in the previous year, down from 44.4% in the previous, pre-COVID audit.¹⁹ This is important, as many patients' technique is not good enough to fully benefit from the medication. The commonest devices are the pressurised metered-dose inhaler, where patients need to inhale correctly and co-ordinate breathing and dose delivery, and the dry powder inhaler, with most devices relying on a rapid and powerful inhalation manoeuvre for drug delivery, which can be very difficult for patients who struggle to inhale forcefully. Study definitions of errors and results vary, but estimates of inhaler error rates range up to 90% of patients, irrespective of the device type used.³⁹⁻⁴¹ We were unable to assess inhaler technique in this study.

The Quality and Outcomes Framework (QOF) has facilitated studies on the association between quality of primary care and outcomes, particularly hospitalisation. A London study using data for 2006–10 used these QOF indicators for COPD: receipt of the influenza vaccination in the preceding 12 months; confirmation of diagnosis using spirometry; recording of forced expiratory volume in 1 second (FEV1) in the previous 15 months; recording of good inhaler technique; and review by a health professional in the preceding 15 months. Emergency hospital admission rates were found to be related to COPD prevalence and small-area deprivation but not to GP QOF performance, GP supply or nurse supply.⁴² Another QOF study looked at the influence of primary care nursing levels on chronic disease performance, including that for COPD. This made use not only of the usual clinical indicators but also those describing organisational performance. The many organisational variables were reduced by factor analysis. Although the relation with staffing levels was statistically significant, the strongest predictors of quality of care were found to be the organisational factors of clinical recording, education and training and use of patient experience surveys.⁴³ An interaction term suggested that the adverse effects on quality of the lowest levels of nurse staffing were mitigated by good organisational quality.

The most direct comparison with this study is the national audit, though it relates to practices in Wales rather than in England, as in this study. The most recent report is based on patients coded with the disease between April 2020 and July 2021, which subsumes the time frame for our cohort 3. Its key findings for COPD were: just 1.9% got post-bronchodilator spirometry in the previous 2 years, 3.5% with MRC 3–5 had a PR referral in the previous 3 years (this was 39.8% with any MRC score), 28.2% of those prescribed an inhaler had their technique checked in the previous year (this was 44.4% in the previous, pre-COVID audit) and, overall, no improvement since the previous audit. We were unable to assess post-bronchodilator spirometry or inhaler technique, but we found a rise in PR referral to 20.9% within a year of diagnosis in cohort 3, up from 13.7% in 2015–7. This was despite the fact that conventional face-to-face PR programmes were widely suspended to protect vulnerable groups, with many staff redeployed to support the care of those acutely unwell. However, there is emerging evidence from RCTs that virtual PR can be as effective as face-to-face options without disadvantaging people who are less web-literate.⁴⁴

The systematic review of risk prediction models cited earlier¹⁶ found nine studies of models that did not separate first from subsequent AECOPDs as the outcome of interest. In these, the main predictor was the number of previous AECOPDs, whereas this study focused on the first. The main statistical limitations in the 27 studies in total were problematic variable-selection procedures and the lack of external validation. We chose our variables and models a priori, preventing the well-known problems of stepwise-type selection, though many studies in the review also did this. Sample sizes ranged from 109 to 8020, whereas our 2 main cohorts were over 30,000 each. Follow-up was usually limited to up to 1 year, shorter than in our case. Twelve out of 27 models were deemed to have an easily available set of predictors across non-specialised and specialised healthcare setting; our set would also fit this description. The number of initial predictors in the review was not always given but ranged from 3 to 60; the number of final predictors ranged from just 1 to 18. Since that review, further models have been published. The ACCEPT study combined data from three RCTs on patients who already had had one or more AECOPDs.⁴⁵ Predictors of AECOPDs within a year of enrolment were history of exacerbations, age, sex, BMI, smoking status, domiciliary oxygen therapy, lung function, symptom burden and baseline

medication use. Its *c*-statistic had a maximum of 0.74. It was later recalibrated and externally validated on observational study data for 1803 patients, achieving a *c*-statistic of 0.76.⁴⁶ Singh *et al.*⁴⁷ combined complete data on 19,194 patients with moderate to severe disease from 7 RCTs to predict moderate or severe AECOPDs within a year. Machine learning methods such as gradient boosting and elastic nets were used to take a set of predictors forward to regression modelling, which then used backward elimination for parsimony. Prior exacerbations, eosinophil count, FEV1 per cent predicted, prior maintenance treatments, reliever medication use, sex, COPD Assessment Test score, smoking status and region were significant predictors. The *c*-statistic was 0.70, with reasonable calibration.

Strengths and limitations

Clinical Practice Research Datalink is much used in research. Its main strengths include easy access, large sample size, broad representativeness of the UK population, the almost universal coverage by primary care in the UK and longevity which means that many groups have experience with it and have conducted validation studies on the accuracy of the coding of diseases and dates. The introduction of QOF in 2004 led to the improved measurement and recording of important factors such as smoking status, blood pressure, body mass index and diseases (once known) rather than symptoms. CPRD capture of medications prescribed by GPs is complete, and linkage of English practices to HES data is now very high. Linkage to the ONS deaths register ensures a high capture of deaths, wherever they occur in the country. We believe that CPRD data – and GP EHRs in general – are good enough for our primary purpose of mapping the route to diagnosis and first AE, capturing important actions such as spirometry, PR referral and prescribing.

Mild and moderate AECOPDs are common but are not always recorded in CPRD and/or reported by patients to their GP. We therefore chose to focus on the severe end, that is, those requiring hospitalisation. This ensures a high level of data capture but is subject to the effects of patient behaviour (though likely less so than with more minor AECOPDs) and hospital admission thresholds.

Clinical Practice Research Datalink has non-zero rates of missing values for some items such as blood pressure, smoking status and BMI, and some tests that are done are not coded and recorded by practices. The coding system means that any social history detail could be captured but relies on the GP asking the question of the patient and recording it electronically. Records rarely contain genetic information, which may be useful in risk prediction as our understanding grows. The recording by GPs of symptoms is known to vary by practice.⁴⁸ Medications given in hospital (but not then continued by the GP) or bought over the counter are typically not captured; the utility of these in risk prediction is likely to be low but is uncertain. Patient preferences regarding the management of their condition(s) are not always recorded. These will affect to some extent the rates of some prescriptions and referrals.

We were unable to assess inhaler technique in this study, which is known to be important regarding outcomes. This technique has been assessed in various countries and found to be frequently suboptimal. According to the World Health Organization, patient adherence to long-term therapy in chronic diseases averages 50% in developed countries,⁴⁹ a figure that is typical for COPD. CPRD-derived measures such as the medication possession ratio are useful to estimate whether the patient has been given the drug and been to the pharmacist to collect it, but not whether they have taken it properly or at all.

The performance of our risk prediction models without post-diagnosis GP actions was moderate. Due to data problems explained in the [Issues faced by this study and lessons learnt](#), we were unable to externally validate the models, a necessary step before implementation in practice. However, given the modest performance of the model without post-diagnosis GP actions, such external validation would not have been worthwhile.

Our COVID-era cohort was small, with limited follow-up time. We were therefore unable to use it in the risk prediction models and to assess between-practice variations in management. Due to COVID

disruptions, we would in any case not have the numbers to examine differences in spirometry use. For similar reasons, there may have been more misdiagnosis for cohort 3 than before. This slightly impairs the comparisons between cohort 3 and the earlier two.

Our survey was developed by clinicians, researchers and patients and pilot-tested by an external patient group over several rounds of feedback. After the failure with Twitter, the sample size from the charity website was sufficient for analysis, though we did not have the numbers to stratify by COVID era. Unfortunately, we had neither sufficient numbers of focus groups nor CPRD survey respondents to allow analysis.

Equality, diversity and inclusion

In this section, we have aimed to follow the National Institute for Health and Care Research (NIHR)-INCLUDE guidance.

Participant representation

Around 99% of the UK population is registered with a GP, but use of primary care services varies considerably for many reasons. The population covered by CPRD practices is known to be broadly representative of the UK population as a whole. The respondents to our survey on the charity website were more likely to be female and less likely to be non-white than the overall COPD population in the UK; the survey was only in English as we lacked the resources to translate it, and we acknowledge that patients using that website could be more engaged in their care than, and different in other ways from, the average person with COPD. The Twitter and GP-administered survey aimed to reach different subpopulations, but recruitment was minimal. There was no financial incentive to take part, though we tried to limit the length of the survey to encourage participation. No paper version of the survey was made available. A number of further analyses are possible in the future, such as the description of the routes to diagnosis by age, sex and other factors. Ethnic groups are not adequately captured in the data currently, and there are recognised limitations on this even in HES data.

The primary outcome – severe acute exacerbation leading to hospitalisation – is expected to be of great interest to most patients. This study did not target a specific underserved group but included everyone registered with a CPRD practice. We described our sample and presented estimated risks of an AECOPD by key patient characteristics such as age, sex, deprivation, BMI and various comorbidities (including anxiety and depression).

Reflections on our research team and wider involvement

As it was formed from our existing network at Imperial College, our research team is fairly gender-balanced but all-white apart from our professor of primary care; at the time of grant submission, the academics comprised two professors, two readers and a lecturer. The junior team members were incorporated into our existing groups, which are very diverse in terms of skillsets, experience, gender and ethnicity. We had two patient and public involvement (PPI) reps, one male and one female, with one contracting COPD because of smoking and the other because of a common genetic disorder. They are both retired.

Patient and public involvement

Our two PPI representatives reviewed the Integrated Research Approval System application for the qualitative component. They advised on the set of Twitter users with large followings who we direct-messaged to distribute the survey and on the themes and final wording of the survey posted on the charity website and sent out via CPRD practices. One of them also inputted into the questions for the focus groups and made comments on the draft publication of the charity website survey results. He

also reviewed the final draft report to NIHR but had no major comments to make. Another PPI member attended all the Project Oversight Committee meetings. Our questionnaire also had considerable input from, and pilot testing by, the patient group at the Asthma + Lung UK charity, which ensured that we had no substantive queries from participants about the purpose or wording.

Recommendations for future research

COVID changed much that is related to personal behaviour, health and health care, but our COVID-era sample was limited in size and follow-up length. Our survey had too few responses in people diagnosed in 2020 and afterwards for more analysis. Future research using more data since early 2020 is therefore warranted. Many AECOPDs are caused by respiratory infections, and these fell with the government-imposed restriction of movement in 2020 and the winter of 2021. On the other hand, the lack of spirometry during the early waves noted earlier would have led to delays in diagnosis for new patients, and it will be important to assess this impact in future.

We assessed prescribing for COPD medications but not antibiotics or other management of the AECOPD itself. Many, but not all, AECOPDs are caused by bacteria, and antibiotic prescribing for AECOPDs has been found to vary considerably by GP practice.⁵⁰ Also, the way in which GP practices usually prescribe antibiotics for patients with COPD (as 'rescue packs' in advance) will make it difficult to associate prescriptions with AECOPDs using routinely coded data. We assessed variations by practice in spirometry, but, by combining further years of data, one could quantify variations in prescribing and other important features of management.

We aimed to predict each patient's first AECOPD, given the subsequent lung damage, but another approach would be to try to predict those at risk of multiple AECOPDs. A Medicare study found that the number of AECOPDs is associated with a multiplicative increase in healthcare costs.⁵¹ Any future risk prediction model, whether for single or multiple acute exacerbations, would benefit from richer data, such as lifestyle and omics. These could be exploited by machine learning approaches alongside conventional statistical methods.

We measured variation by GP practice in spirometry rates, with numbers of patients for cohorts 1 and 2 that were sufficiently large for many but not all practices. This could be a quality indicator but should be subject to reliability analysis as in a Dutch study using administrative data that assessed the ICC, reliability and sample size requirements at common reliability thresholds for different indicators for diabetes and COPD.⁵²

In summary, we make the following research recommendations:

- continue analysis of the quality of primary care and patient experience since March 2020;
- use social media better to obtain patient experience data from more representative samples;
- determine the impact of the disruption due to COVID-19 on the diagnosis and management of COPD patients in terms of their outcomes;
- quantify variations by GP practice in prescribing;
- assess whether spirometry use is suitable as a quality of primary care indicator;
- seek expansion of data available to researchers and GPs to derive better risk prediction.

Implications for decision-makers

Patients, especially smokers, need to be more aware of COPD symptoms and take them seriously. When they do, GPs need more help with recognising COPD symptoms and following NICE guidelines on diagnosis. We found considerable variation by GP practice in the use of spirometry. Pulmonary

rehabilitation referral has improved but is still low, so finding ways to improve both referral and uptake are needed. There may be a role for virtual sessions.

Risk prediction for the first AECOPD, when prevention of lung damage is most effective, is still not very powerful with primary care data, so augmentation is needed, when one could make use of machine learning methods as well as statistical approaches.

The impact of the COVID-19 pandemic is still to be seen, though we found some evidence of diagnostic delay compared with earlier cohorts. It will be important to monitor longer-term management and outcomes.

Chapter 7 Conclusions

Around three-quarters of COPD patients were diagnosed in primary care. Comorbidity seems to be increasing over time and associated with higher odds of diagnosis through emergency hospitalisation. Findings from the survey and CPRD indicate a variety of both patient and system factors hindering timely diagnosis. While flu vaccination rates are high, the use of NICE-recommended diagnosis tests and PR warrant further improvement. However, LABA and LAMA prescribing increased considerably over time. Prediction of the first AECOPD with current primary care data shows only moderate performance, and our models would need further data enhancements before field testing and roll-out are warranted. Given the changes caused by the COVID-19 pandemic, it will be important to continue to monitor the diagnosis and management of COPD in primary care and longer-term outcomes.

Chapter 8 Issues faced by this study and lessons learnt

The study had three research associates (RAs) during its course, incurring some minor delays while we looked for replacements. More importantly, the temperature and pollution data and R code to identify the nearest GP practice were obtained by someone external to the project team who went on leave unexpectedly for several months without us knowing. This delayed the data extraction, greatly limiting what the first RA could do, as CPRD were unwilling to give us an extract without the linkage while we waited for it. When the pandemic began, CPRD prioritised data extracts that were directly COVID related, unlike ours. After our second RA began processing the data, always a lengthy task with such complex databases, during several small project team meetings, we noticed that an unusually high number of patients had COPD recorded in HES but not in CPRD, particularly for the second cohort. As we began work with the first cohort, which appeared plausible in its patterns, this took time to become apparent. We identified one potential reason as data artefact, namely that the gradual changeover of IT systems in UK primary care during the 2010–9 decade (which has continued) meant that CPRD GOLD was diminishing and CPRD Aurum increasing in size. We compared several algorithms to take account of this, but it was our third RA who ran other checks and made us conclude that CPRD staff had made an error in the extraction, as there were too many patients with GP consultations before and/or after a HES record containing a COPD code in whom those GP consultations contained no mention of COPD or its treatment. We were fortunate to be able to bid for and win internal NIHR Biomedical Research Centres (BRC) funding to cover the third RA's salary until the project extension date of mid-March 2023. With time running out for analysis, we therefore decided to use a different extract, from another project, which was CPRD Aurum. The alternative would have been to try to get CPRD to correct their error, but, given the growing pressures at CPRD and increasing timescales, we have noticed for other projects and also for the survey component of this one (see below), we decided this would not be feasible. Our chosen option had the advantage of being able to include a third, albeit relatively small, COVID-era cohort.

The qualitative component of the project also faced numerous difficulties. Ideally, we would have run enough quantitative analysis in Year 1 to inform it, but that was not possible as detailed above. However, due to internal problems at CPRD, it took around 16 months between our initial meeting with them and the roll-out of our survey to CPRD practices in December 2022; recruitment from this route proved to be extremely low. The final wording and posting of the survey on the charity website was delayed by staff changes at their end, but we were still able to analyse the results well before the CPRD practices were recruited. The focus groups could not be started until this analysis had been done in order to devise the questions to pose to participants, which was around November 2022. A research physiotherapist from our local Trust helped us convene one, but it was near Christmas, and some potential participants declined due to busy schedules or poorer health due to cold weather. This reduced the number of focus groups that we were able to run in time to submit this report. We were fortunate to obtain non-NIHR funding for a qualitative researcher for 12 months to work on the non-CPRD component of the project; we were asked by the NIHR Board at funding to ask patients directly about their experience of diagnosis and acute exacerbation, and we greatly undercosted this, unsure whether NIHR would have agreed to fund the necessary extra budget. The third route for the survey, Twitter, was extremely disappointing. Most people who attempted it made very little progress, leaving too few complete answers for analysis. We suspect that the low attempt rate was due to our inexperience with using Twitter for this purpose and poor selection of targets for our direct messaging, though it is also possible that Twitter is not a good medium for this in any case. We did not have a project manager, unfortunately, which added administrative and various other tasks to the principal investigator's (PI's) workload.

The project's oversight committee were kept informed, were sympathetic regarding our problems, and were very supportive of our applications to NIHR for two no-cost extensions, which were granted. We submitted our 6-monthly progress reports to NIHR, but it sometimes took several months for these to be reviewed. More positively, we appreciated the opportunity to discuss the project's difficulties and ways forward with NIHR in January 2022.

We believe that this considerable set of issues could have been mitigated with:

- More appropriate costing of the project as a whole by allowing for more time for things to go wrong and by better appreciating how much resource was required to do justice to the qualitative components.
- Costing for a part-time project manager to lessen the administrative burden on the PI.
- Greater flexibility by CPRD for non-standard data linkages – they were unwilling to give us the data extract without the linkage to temperature and pollution estimates, preventing us from doing any meaningful data preparation and analysis for many months.

Reconsideration of timescales when working with CPRD: they are in high demand, more so since COVID and their new internal processes do not seem to help.

Additional information

Acknowledgements

We are very grateful to our two PPI representatives, Roger Williams and Patricia Craik, for their help with this study. Roger's input into the scope and wording of the survey was particularly valuable, and we appreciate the time he spent reviewing this final report. Data management was provided by the Big Data and Analytical Unit (BDAU) at the Institute of Global Health Innovation at Imperial College London. The Dr Foster Unit is an academic unit in the Department of Primary Care and Public Health, within the School of Public Health, Imperial College London. The unit received research funding from Dr Foster Intelligence, an independent health service research organisation (a wholly owned subsidiary of Telstra), until September 2021. The Dr Foster Unit at Imperial is affiliated with the National Institute of Health and Care Research (NIHR) Imperial Patient Safety Translational Research Centre. The NIHR Imperial Patient Safety Translational Centre is a partnership between the Imperial College Healthcare NHS Trust and Imperial College London. The Department of Primary Care and Public Health at Imperial College London is grateful for support from the NW London NIHR Applied Research Collaboration and the Imperial NIHR Biomedical Research Centre. Finally, we are grateful to the anonymous reviewers who took time and effort to make constructive comments that improved this report. We also thank the Project Oversight Committee members for their support.

Contributions of authors

Alex Bottle (<https://orcid.org/0000-0001-9978-2011>) (Professor of Medical Statistics and Principal Investigator) conceived the study, obtained funding and drafted this report.

Alex Adamson (<https://orcid.org/0000-0003-0265-5900>) (Research Associate) analysed the quantitative data.

Xiubin Zhang (<https://orcid.org/0000-0003-2660-3937>) (Research Associate) collected and analysed the survey data.

Benedict Hayhoe (<https://orcid.org/0000-0002-2645-6191>) (GP and researcher) conceived the study and obtained funding.

Jennifer K Quint (<https://orcid.org/0000-0003-0149-4869>) (Clinical Professor of Respiratory Epidemiology) conceived the study and obtained funding.

All authors reviewed and critically edited the report content. All authors approved the version to be published and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Disclosure of interests

Full disclosure of interests: Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page at <https://doi.org/10.3310/CGTR6370>.

Primary conflicts of interest: This study was funded by NIHR HS&DR (award number 17/99/72) and NIHR BRC; the qualitative component was part-funded by the Taskforce through Asthma and Lung UK. Alex Bottle's Dr Foster Unit declares grant income from Medtronic and Dr Foster (Telstra Health) during the project but outside this work. Alex Bottle declares income for advisory work from AstraZeneca and Lilly during the project but outside this work; he was on the NIHR HS&DR Researcher Led and Prioritisation Panel from 2016 to 2020. Alex Adamson declares no conflicts of interest. Xiubin Zhang declares no conflicts of interest. Benedict Hayhoe declares other academic work funded under the NIHR ARC programme for NW London; he also works for eConsult, provider of electronic consultations for NHS primary, secondary and urgent/emergency care (outside of this work). Jenni Quint declares grants from MRC, HDR UK, GSK, BI, Asthma + Lung UK, AstraZeneca and personal fees for advisory board participation or speaking fees from GlaxoSmithKline, AstraZeneca and Insmmed outside of this work.

Patient data statement

This work uses data provided by patients and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people's patient records, to understand more about disease, develop new treatments, monitor safety and plan NHS services. Patient data should be kept safe and secure, to protect everyone's privacy, and it is important that there are safeguards to make sure that they are stored and used responsibly. Everyone should be able to find out about how patient data are used. #datasaveslives You can find out more about the background to this citation here: <https://understandingpatientdata.org.uk/data-citation>.

Data-sharing statement

The anonymised patient data from CPRD that were used for this study can be accessed by contacting the Clinical Practice Research Datalink at enquiries@cpdr.com. Access to these data is subject to a data-sharing agreement (DSA) containing detailed terms and conditions of use following protocol approval from CPRD's Independent Scientific Advisory Committee (ISAC).

Ethics statement

We have approval from the Secretary of State and the Health Research Authority under Regulation 5 of the Health Service (Control of Patient Information) Regulations 2002 to hold confidential data and analyse them for research purposes (CAG ref 15/CAG/0005). We have approval to use them for research and measuring quality of delivery of healthcare, from the London - South East Ethics Committee (REC ref 20/LO/0611). The favourable opinion letter is dated 27 May 2020. Linked pseudonymised data were provided for this study by CPRD. Data are linked by NHS Digital, the statutory trusted third party for linking data, using identifiable data held only by NHS Digital. Select general practices consent to this process at a practice level with individual patients having the right to opt out. We obtained CPRD data for protocol number 19_116 ('What happens between first symptoms and first acute exacerbation of COPD? Mapping and prediction study') for the quantitative work; surveys were sent out by CPRD-participating GPs as per protocol number 20_000074 ('Diagnosing COPD: Survey of Patient Experience across the UK') and with HRA approval dated 7 April 2021 (REC ref 21/PR/0098).

Information governance statement

Imperial College London is committed to handling all personal information in line with the UK Data Protection Act (2018) and the General Data Protection Regulation (EU GDPR) 2016/679. Under the Data Protection legislation, we (Imperial College London) are the Data Processor, CPRD is the Data Controller, and we process personal data in accordance with their instructions. You can find out more about how we handle personal data, including how to exercise your individual rights and the contact details for our Data Protection Officer and policy here: www.imperial.ac.uk/admin-services/secretariat/information-governance/data-protection/our-policy/

Publications

Poster on CPRD GOLD data processing challenges presented at Public Health Science in 2021.

Oral presentation at ISQua annual conference on routes to COPD diagnosis in October 2022.

Zhang X, Ellis A, Quint JK, Bottle A. Survey-identified experiences of prediagnosis and diagnosis process among patients with COPD, asthma, interstitial lung disease and bronchiectasis. *BMJ Open Respir Res* 2023;**10**(1):e001588.

References

1. Safiri S, Carson-Chahhoud K, Noori M, Nejadghaderi SA, Sullman MJM, Ahmadian Heris J, *et al.* Burden of chronic obstructive pulmonary disease and its attributable risk factors in 204 countries and territories, 1990–2019: results from the Global Burden of Disease Study 2019. *BMJ* 2022;**378**:e069679.
2. Adeloye D, Chua S, Lee C, Basquill C, Papan A, Theodoratou E, *et al.*; Global Health Epidemiology Reference Group (GHERG). Global and regional estimates of COPD prevalence: systematic review and meta-analysis. *J Glob Health* 2015;**5**:020415.
3. Snell N, Strachan D, Hubbard R, Gibson J, Gruffydd-Jones K, Jarrold I. S32 Epidemiology of chronic obstructive pulmonary disease (COPD) in the UK: findings from the British Lung Foundation's 'respiratory health of the nation' project. *Thorax* 2016;**71**:A20.
4. British Lung Foundation. *Chronic Obstructive Pulmonary Disease (COPD) Statistics*. 2022. URL: <https://statistics.blf.org.uk/copd> (accessed 21 December 2022).
5. James GD, Donaldson GC, Wedzicha JA, Nazareth I. Trends in management and outcomes of COPD patients in primary care, 2000–2009: a retrospective cohort study. *NPJ Prim Care Respir Med* 2014;**24**:1–7.
6. Price D, West D, Brusselle G, Gruffydd-Jones K, Jones R, Miravittles M, *et al.* Management of COPD in the UK primary-care setting: an analysis of real-life prescribing patterns. *Int J Chron Obstruct Pulmon Dis* 2014;**9**:889–904.
7. *National COPD Audit Programme: Primary Care Audit (Wales) 2015–17 – Planning for Every Breath*. URL: www.rcplondon.ac.uk/projects/outputs/primary-care-audit-wales-2015-17-planning-every-breath (accessed 19 December 2022).
8. Donaldson GC, Seemungal TA, Bhowmik A, Wedzicha JA. Relationship between exacerbation frequency and lung function decline in chronic obstructive pulmonary disease. *Thorax* 2002;**57**:847–52.
9. Seemungal TA, Donaldson GC, Paul EA, Bestall JC, Jeffries DJ, Wedzicha JA. Effect of exacerbation on quality of life in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998;**157**:1418–22.
10. Marcos PJ, Huerta A, Enzler MJ. Using standardized care bundles in the emergency department to decrease mortality in patients presenting with community-acquired pneumonia (CAP) and acute exacerbation of chronic obstructive pulmonary disease (AECOPD). *Curr Infect Dis Rep* 2015;**17**:1–6.
11. Nguyen PL, Uddin MM, Mir T, Khalil A, Regmi N, Pervaiz A, *et al.* Trends in incidence, and mortality of acute exacerbation of chronic obstructive pulmonary disease in the United States Emergency Department (2010–2018). *COPD* 2021;**18**:567–75.
12. Hurst JR, Vestbo J, Anzueto A, Locantore N, Müllerova H, Tal-Singer R, *et al.*; Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints (ECLIPSE) Investigators. Susceptibility to exacerbation in chronic obstructive pulmonary disease. *N Engl J Med* 2010;**363**:1128–38.
13. Soler-Cataluña JJ, Martínez-García MA, Román Sánchez P, Salcedo E, Navarro M, Ochando R. Severe acute exacerbations and mortality in patients with chronic obstructive pulmonary disease. *Thorax* 2005;**60**:925–31.
14. Halpin DMG, Decramer M, Celli BR, Mueller A, Metzdorf N, Tashkin DP. Effect of a single exacerbation on decline in lung function in COPD. *Respir Med* 2017;**128**:85–91.

15. Bottle A, Honeyford K, Chowdhury F, Bell D, Aylin P. *Factors Associated with Hospital Emergency Readmission and Mortality Rates in Patients with Heart Failure or Chronic Obstructive Pulmonary Disease: A National Observational Study*. Southampton: NIHR Journals Library; 2018.
16. Guerra B, Gaveikaite V, Bianchi C, Puhan MA. Prediction models for exacerbations in patients with COPD. *Eur Respir Rev* 2017;**26**:160061.
17. Alqahtani JS, Oyelade T, Aldhahir AM, Mendes RG, Alghamdi SM, Miravittles M, *et al*. Reduction in hospitalised COPD exacerbations during COVID-19: a systematic review and meta-analysis. *PLOS ONE* 2021;**16**:e0255659.
18. Alsallakh MA, Sivakumaran S, Kennedy S, Vasileiou E, Lyons RA, Robertson C, *et al.*; EAVE II Collaborators. Impact of COVID-19 lockdown on the incidence and mortality of acute exacerbations of chronic obstructive pulmonary disease: national interrupted time series analyses for Scotland and Wales. *BMC Med* 2021;**19**:124.
19. National COPD Audit Programme. *Wales Primary Care Clinical Audit Report 2021*. URL: www.rcp.ac.uk/projects/outputs/wales-primary-care-clinical-audit-report-2021 (accessed 21 December 2022).
20. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, Smeeth L. Data resource profile: clinical practice research datalink (CPRD). *Int J Epidemiol* 2015;**44**:827–36.
21. Kontopantelis E, Stevens RJ, Helms PJ, Edwards D, Doran T, Ashcroft DM. Spatial distribution of clinical computer systems in primary care in England in 2016 and implications for primary care electronic medical record databases: a cross-sectional population study. *BMJ Open* 2018;**8**:e020738.
22. Williams T, van Staa T, Puri S, Eaton S. Recent advances in the utility and use of the General Practice Research Database as an example of a UK Primary Care Data resource. *Ther Adv Drug Saf* 2012;**3**:89–99.
23. Wolf A, Dedman D, Campbell J, Booth H, Lunn D, Chapman J, Myles P. Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. *Int J Epidemiol* 2019;**48**:1740–1740g.
24. NHS Digital. *National Data Opt-out*. 2020. URL: <https://digital.nhs.uk/data-and-information/publications/statistical/national-data-opt-out/ode-2020> (accessed 14 March 2023).
25. Rothnie KJ, Müllerová H, Hurst JR, Smeeth L, Davis K, Thomas SL, Quint JK. Validation of the recording of acute exacerbations of COPD in UK primary care electronic healthcare records. *PLOS ONE* 2016;**11**:e0151357.
26. Rothnie KJ, Müllerová H, Thomas SL, Chandan JS, Smeeth L, Hurst JR, *et al*. Recording of hospitalizations for acute exacerbations of COPD in UK electronic health care records. *Clin Epidemiol* 2016;**8**:771–82.
27. National Institute for Health and Care Excellence. *Chronic Obstructive Pulmonary Disease in Over 16s: Diagnosis and Management*. 2018. URL: www.nice.org.uk/guidance/ng115/resources/chronic-obstructive-pulmonary-disease-in-over-16s-diagnosis-and-management-pdf-66141600098245 (accessed 14 March 2023).
28. Sanagou M, Wolfe R, Forbes A, Reid CM. Hospital-level associations with 30-day patient mortality after cardiac surgery: a tutorial on the application and interpretation of marginal and multilevel logistic regression. *BMC Med Res Methodol* 2012;**12**:28.
29. Bouwmeester W, Twisk JW, Kappen TH, van Klei WA, Moons KG, Vergouwe Y. Prediction models for clustered data: comparison of a random intercept and standard regression model. *BMC Med Res Methodol* 2013;**13**:19.
30. NHS Innovation Accelerator. *Implementation Toolkit*. myCOPD. URL: https://wessexahsn.org.uk/img/projects/Implementation%20Toolkit_myCOPD.pdf (accessed 14 March 2023).

31. Bao W, Li Y, Wang T, Li X, He J, Wang Y, *et al.* Effects of influenza vaccination on clinical outcomes of chronic obstructive pulmonary disease: a systematic review and meta-analysis. *Ageing Res Rev* 2021;**68**:101337.
32. Bobrovitz N, Heneghan C, Onakpoya I, Fletcher B, Collins D, Tompson A, *et al.* Medications that reduce emergency hospital admissions: an overview of systematic reviews and prioritisation of treatments. *BMC Med* 2018;**16**:115.
33. Thomas M, Radwan A, Stonham C, Marshall S. COPD exacerbation frequency, pharmacotherapy and resource use: an observational study in UK primary care. *COPD* 2014;**11**:300–9.
34. Koarai A, Yamada M, Ichikawa T, Fujino N, Kawayama T, Sugiura H. Triple versus LAMA/LABA combination therapy for patients with COPD: a systematic review and meta-analysis. *Respir Res* 2021;**22**:183.
35. Rothnie KJ, Chandan JS, Goss HG, Müllerová H, Quint JK. Validity and interpretation of spirometric recordings to diagnose COPD in UK primary care. *Int J Chron Obstruct Pulmon Dis* 2017;**12**:1663–8.
36. van de Hei SJ, Flokstra-de Blok BMJ, Baretta HJ, Doornewaard NE, van der Molen T, Patberg KW, *et al.* Quality of spirometry and related diagnosis in primary care with a focus on clinical use. *NPJ Prim Care Respir Med* 2020;**30**:22.
37. de Vries M, Berendsen AJ, Bosveld HE, Kerstjens HA, van der Molen T. COPD exacerbations in general practice: variability in oral prednisolone courses. *BMC Fam Pract* 2012;**13**:3.
38. Chalmers JD, Tebboth A, Gayle A, Ternouth A, Ramscar N. Determinants of initial inhaled corticosteroid use in patients with GOLD A/B COPD: a retrospective study of UK general practice. *NPJ Prim Care Respir Med* 2017;**27**:43.
39. Al-Showair RA, Tarsin WY, Assi KH, Pearson SB, Chrystyn H. Can all patients with COPD use the correct inhalation flow with all inhalers and does training help? *Respir Med* 2007;**101**:2395–401.
40. Lavorini F, Magnan A, Dubus JC, Voshaar T, Corbetta L, Broeders M, *et al.* Effect of incorrect use of dry powder inhalers on management of patients with asthma and COPD. *Respir Med* 2008;**102**:593–604.
41. Usmani OS, Lavorini F, Marshall J, Dunlop WCN, Heron L, Farrington E, Dekhuijzen R. Critical inhaler errors in asthma and COPD: a systematic review of impact on health outcomes. *Respir Res* 2018;**19**:10.
42. Jamieson AL, Harries TH, Thornton H, Crichton S, White P. Emergency admissions for COPD in an urban population: the role of population and primary care factors. *COPD* 2015;**12**:606–12.
43. Griffiths P, Maben J, Murrells T. Organisational quality, nurse staffing and the quality of chronic disease management in primary care: observational study using routinely collected data. *Int J Nurs Stud* 2011;**48**:1199–210.
44. Houchen-Wolloff L, Steiner MC. Pulmonary rehabilitation at a time of social distancing: prime time for tele-rehabilitation? *Thorax* 2020;**75**:446–7.
45. Adibi A, Sin DD, Safari A, Johnson KM, Aaron SD, FitzGerald JM, Sadatsafavi M. The Acute COPD Exacerbation Prediction Tool (ACCEPT): a modelling study. *Lancet Respir Med* 2020;**8**:1013–21.
46. Safari A, Adibi A, Sin DD, Lee TY, Khoa Ho J, Sadatsafavi M. ACCEPT 2 0: recalibrating and externally validating the Acute COPD exacerbation prediction tool (ACCEPT). *eClinicalMedicine* 2022;**51**:101574.

REFERENCES

47. Singh D, Hurst JR, Martinez FJ, Rabe KF, Bafadhel M, Jenkins M, *et al.* Predictive modelling of COPD exacerbation rates using baseline risk factors. *Thorax* 2022;**75**:1073-1081.
48. Jordan K, Porcheret M, Croft P. Quality of morbidity coding in general practice computerized medical records: a systematic review. *Fam Pract* 2004;**21**:396-412.
49. World Health Organization. *Adherence to Long-term Therapies: Evidence for Action*. World Health Organization; 2003. URL: <https://apps.who.int/iris/handle/10665/42682> (accessed 23 December 2022).
50. Boggon R, Hubbard R, Smeeth L, Gulliford M, Cassell J, Eaton S, *et al.* Variability of antibiotic prescribing in patients with chronic obstructive pulmonary disease exacerbations: a cohort study. *BMC Pulm Med* 2013;**13**:32.
51. Dhamane AD, Moretz C, Zhou Y, Burslem K, Saverno K, Jain G, *et al.* COPD exacerbation frequency and its association with health care resource utilization and costs. *Int J Chron Obstruct Pulmon Dis* 2015;**10**:2609-18.
52. Eijkenaar F, van Vliet RCJA. Profiling individual physicians using administrative data from a single insurer: variance components, reliability, and implications for performance improvement efforts. *Med Care* 2013;**51**:731-9.

Appendix 1 Survey of patient experience of getting diagnosed with chronic obstructive pulmonary disease

1. Have you been diagnosed with chronic obstructive pulmonary disease (COPD)?

COPD includes conditions such as chronic bronchitis and emphysema.

Yes

No

2. Which country in the UK do you currently live in?

England

Scotland

Wales

Northern Ireland

I don't live in the UK

3. Thinking just about your diagnosis of COPD, which of the following symptoms did you experience before seeking professional help (e.g. visiting a GP or A&E)? Please tick all that apply.

Increased breathlessness

Cough

Wheezing

Unusual phlegm / sputum

Chest infections

Tiredness

Low mood

Pain

Coughing up blood

Unusual weight changes

Other (please specify)

APPENDIX 1

None of the above / I don't know

4. What caused you to notice something was wrong? Please tick all that apply.

- Nothing in particular, I just felt odd or generally unwell/rundown
- I was finding it more difficult to play sports or exercise
- Daily activities such as housework were more difficult to do
- Symptoms like cough or chest infections weren't going away as I would expect or they kept recurring
- It was taking me longer to recover from illness or I was not recovering fully
- A healthcare professional told me something was wrong
- Someone close to me noticed a change in me
- Other (please specify)

None of the above / I don't know

5. What did you do after you noticed you were experiencing symptoms?

This question is asking about what you did when you first noticed your symptoms, up to and including when you **first** sought professional help (e.g. visiting a GP or A&E). So **does not include** referrals or specialist appointments that may have happened later on.

Please rank these in the order that you did them. It may be best to read through all the options before answering. For options that don't apply to you, select 'N/A'.

If you did multiple things at once, please give them the same rank. If you can't remember the order, please try to pick at least one thing you did first and order all the other things you did equally. (e.g. all as 2nd or all as 3rd)

	1st	2nd	3rd	4th	5th	6th	7th	8th	N/A
Did nothing, hoping symptoms would go away	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Used at home									

remedies or alternative medicine	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Searched the internet	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Spoke to family & friends	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Spoke to a pharmacist	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sought telephone advice from NHS 111	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Went to the GP	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Felt concerned and rang 999	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Went to A&E	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

6. How long did you live with your COPD symptoms before seeking professional help (e.g. visiting a GP or A&E)?

Common symptoms include increasing breathlessness, a persistent chesty cough with phlegm that does not go away, frequent chest infections and persistent wheezing.

- Less than 1 month
- 1–2 months
- 3–6 months
- 7–12 months
- 1–2 years
- 3–5 years
- 6–10 years
- 10 + years
- I don't know / not applicable

7. What motivated you to seek professional help? Please tick all that apply.

- I was worried about my symptoms
- I was no longer able to live my life in the same way
- My partner encouraged me to seek help
- There's a history of respiratory disease in my family so I was aware of the risks
- I was aware my job or lifestyle factors put me at risk of health issues
- I read or saw something on the internet that motivated me to take action
- I read or saw something on the TV that motivated me to take action
- I had a scare such as collapsing
- Other (please specify)
- None of the above / I don't know

8. What, if anything, prevented you from seeking professional help earlier? Please tick all that apply.

- Busy at work
- Busy with home life
- Not being registered with GP
- Not knowing the signs of potential lung disease
- Not appreciating the severity or urgency of the situation
- Not wanting to know if something was wrong
- Hoping things would go away on their own
- Concern about catching COVID
- Not being able to get an appointment at a time that suited me
- Other (please specify)

9. How long did you have to wait between seeking / receiving professional

help for your COPD (e.g. visiting a GP or A&E) and receiving a formal diagnosis? Please estimate if you're not sure.

- Less than 1 week
- 1–2 weeks
- 2–4 weeks
- 1–2 months
- 3–6 months
- 7–12 months
- 1–2 years
- 2–5 years
- 5–10 years
- 10+ years
- I don't remember

10. What year were you first diagnosed with COPD? Please give your best guess if you can't remember.

11. Where were you when you were given your COPD diagnosis?

- GP practice
- A&E
- While admitted as a hospital inpatient
- Hospital outpatient department (respiratory clinic)
- Hospital outpatient department (other clinic)
- At home (told of my diagnosis over the phone)
- Other (please specify)

- None of the above / I don't know

12. How were you diagnosed? Please tick all that apply.

- Discussion with my doctor
- Spirometry (breathing test where you blow hard into a mouthpiece on a small machine)
- 6-minute walking test
- Chest X-ray
- Oxygen saturation (finger probe) test
- Phlegm (sputum) test
- Sweat test
- CT scan
- Other scan
- Bronchoscopy (when a flexible tube is put into your nose or mouth and into your lungs)
- Lung biopsy
- Bronchoalveolar lavage (when liquid is put through the bronchoscope to get a sample)
- Blood tests
- Feno testing (test where you breathe out slowly through a filter into a portable machine)
- Pulmonary Exercise Stress Test
- Histamine testing (skin prick test to look for allergies)
- Other (please specify)
- None of the above / I don't know

13. Had you heard of COPD before you were diagnosed?

- Yes
- No
- I don't know

14. Did your symptoms worsen between first seeking professional help and getting your diagnosis?

- Yes
- No
- I don't know

15. Do you feel your diagnosis was delayed?

- Yes
- No
- I don't know

16. If you feel that your diagnosis was delayed, how did this affect you? Please tick all that apply.

- I worried for longer than I needed to
- I felt my condition worsened more than it had to
- I became demotivated
- My mental health suffered
- I felt like I wasn't being taken seriously
- I didn't get medication quickly enough
- I didn't get advice on managing my condition quickly enough
- Other (please specify)
-
- None of the above / I don't know

17. What were the main barriers, if any, to getting a diagnosis? Please tick all that apply.

- Being treated for another lung condition
- Being treated for another non-lung condition (e.g. heart condition)

APPENDIX 1

- Symptoms were attributed to a pre-existing condition that I had
- Lack of expertise or knowledge, e.g. healthcare professional not recognising my symptoms
- Lack of effort or motivation, e.g. healthcare professional not taking the time to investigate
- Feeling like I had to fight for my care, e.g. being turned away by my GP
- Difficulty getting appointments
- Long waiting times or delays
- Lack of follow-up to discuss test results
- COVID-19
- I was misdiagnosed
- Other (please specify)
- I do not recall there being any barriers

18. What factors do you think contributed to your COPD? Please tick all that apply.

- Smoking
- Passive smoking
- Genetic factors
- Poor housing (e.g. damp, mould, living near sewage)
- Air pollution
- Exposure at work
- Previous infections
- Poor health
- Other (please specify)
- None of the above / I don't know

19. Had you already been diagnosed with heart or lung problems prior to your COPD diagnosis?

- Yes
- No
- I don't know

If yes, please specify the name of these

20. In the year after being diagnosed, were you offered the following vaccines?

	I received this vaccine	I was offered this vaccine but did not receive it	I was not offered this vaccine	I don't know
Annual flu jab	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pneumonia jab	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

APPENDIX 1

Questions about you help us understand if health care is being delivered fairly to all people no matter their location, age, gender, ethnicity or socio-economic background.

21. What was your postcode leading up to your diagnosis? Please provide your full postcode.

By survey participants providing their postcodes it allows us to explore how the delivery of health care varies across country, between local areas and between different groups of people. Using this we can understand if health care is being delivered fairly to all people or if certain socioeconomic factors affect your care.

22. What is your age?

- 18–24
- 25–34
- 35–44
- 45–54
- 55–64
- 65–74
- 75–84
- 85–94
- 95+
- Prefer not to say

23. What is your gender?

- Male
- Female
- I prefer to self-describe

I would prefer not to say

24. Which of these best describes your ethnic group?

- White
- Mixed or multiple ethnic groups
- Asian or Asian British
- Black, African, Caribbean or Black British
- Arab
- Prefer not to say
- Other ethnic group (please specify)

25. Do you belong to or follow any patient advisory groups? Please tick all that apply.

- Patient support group
- British Lung Foundation Breathe Easy group
- Facebook groups/pages/accounts relating to COPD
- Instagram accounts relating to COPD
- Twitter accounts relating to COPD
- Other social media relating to COPD
- Online patient forum/group
- WhatsApp group relating to COPD
- Other (please specify)

I don't belong to or follow any patient advisory groups

Appendix 2 Additional rows from the main tables

TABLE 16 Further GP-recorded symptoms before diagnosis (see [Table 4](#))

Variable		2006–7	2016–7	March–August 2020	Total	p-value
Any wheeze diagnosis preceding COPD diagnosis	Yes	2629 (9.4)	4581 (13.0)	424 (12.7)	7634 (11.4)	< 0.001
	No	25,414 (90.6)	30,736 (87.0)	2913 (87.3)	59,063 (88.6)	
Time from earliest wheeze diagnosis to COPD diagnosis (days)	Median (IQR)	569.0 (62.0–2081.0)	613.0 (44.0–2727.0)	926.0 (190.5–3341.0)	619.5 (54.0–2512.8)	< 0.001
Any LRTI diagnosis preceding COPD diagnosis	Yes	13,631 (48.6)	20,529 (58.1)	2030 (60.8)	36,190 (54.3)	< 0.001
	No	14,412 (51.4)	14,788 (41.9)	1307 (39.2)	30,507 (45.7)	
Time from earliest LRTI diagnosis to COPD diagnosis (days)	Median (IQR)	1345.0 (478.0–3174.5)	2847.0 (995.0–5099.0)	3442.0 (1175.2–6041.0)	2198.5 (711.0–4553.0)	< 0.001
Any sputum diagnosis preceding COPD diagnosis	Yes	4662 (16.6)	7960 (22.5)	875 (26.2)	13497 (20.2)	< 0.001
	No	23,381 (83.4)	27,357 (77.5)	2462 (73.8)	53,200 (79.8)	
Time from earliest sputum diagnosis to COPD diagnosis (days)	Median (IQR)	971.5 (160.2–2536.2)	1575.0 (252.0–4005.0)	1613.0 (296.5–4226.5)	1288.0 (218.0–3456.0)	< 0.001
Any breathlessness diagnosis preceding COPD diagnosis	Yes	10,640 (37.9)	21,157 (59.9)	1876 (56.2)	33,673 (50.5)	< 0.001
	No	17,403 (62.1)	14,160 (40.1)	1461 (43.8)	33,024 (49.5)	
Time from earliest breathlessness diagnosis to COPD diagnosis (days)	Median (IQR)	161.0 (9.0–974.0)	114.0 (0.0–1478.0)	588.0 (45.0–2335.0)	148.0 (2.0–1310.0)	< 0.001
Any cough diagnosis preceding COPD diagnosis	Yes	12,304 (43.9)	23,186 (65.7)	2250 (67.4)	37,740 (56.6)	< 0.001
	No	15,739 (56.1)	12,131 (34.3)	1087 (32.6)	28,957 (43.4)	
Time from earliest cough diagnosis to COPD diagnosis (days)	Median (IQR)	940.0 (245.8–2249.0)	2107.0 (643.2–4099.0)	2793.5 (1049.5–4895.8)	1662.0 (456.0–3657.2)	< 0.001
Time from earliest COPD symptom to COPD diagnosis (days)	Median (IQR)	1205.0 (335.0–2951.5)	2764.0 (790.0–4976.0)	3667.0 (1401.0–6074.0)	2047.0 (543.0–4435.5)	< 0.001
Time from first symptom to chest X-ray (days)	Median (IQR)	748.0 (9.0–2386.8)	1986.0 (84.0–4400.0)	2809.5 (355.2–5370.2)	1484.0 (34.0–3909.0)	< 0.001

TABLE 16 Further GP-recorded symptoms before diagnosis (see [Table 4](#)) (continued)

Variable		2006–7	2016–7	March–August 2020	Total	p-value
Time from first symptom to FBC (days)	Median (IQR)	922.0 (36.0–2607.0)	2171.0 (286.0–4540.0)	3005.0 (819.5–5581.5)	1646.0 (164.0–4052.8)	< 0.001
Time from first symptom to first recorded spirometry (days)	Median (IQR)	950.0 (96.0–2651.0)	1740.0 (92.0–4173.0)	1946.0 (109.0–4569.0)	1384.0 (94.0–3709.0)	< 0.001
Number of wheeze diagnoses in 5 years preceding COPD diagnosis	Median (IQR)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	< 0.001
Number of LRTI diagnoses preceding COPD diagnosis	Median (IQR)	0.0 (0.0–2.0)	1.0 (0.0–3.0)	1.0 (0.0–3.0)	1.0 (0.0–2.0)	< 0.001
Number of LRTI diagnoses in 5 years preceding COPD diagnosis	Median (IQR)	0.0 (0.0–1.0)	0.0 (0.0–1.0)	0.0 (0.0–1.0)	0.0 (0.0–1.0)	0.017
Number of sputum diagnoses preceding COPD diagnosis	Median (IQR)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–1.0)	0.0 (0.0–0.0)	< 0.001
Number of sputum diagnoses in 5 years preceding COPD diagnosis	Median (IQR)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	< 0.001
Number of breathlessness diagnoses preceding COPD diagnosis	Median (IQR)	0.0 (0.0–1.0)	1.0 (0.0–2.0)	1.0 (0.0–2.0)	1.0 (0.0–1.0)	< 0.001
Number of breathlessness diagnoses within 5 years preceding COPD diagnosis	Median (IQR)	0.0 (0.0–1.0)	1.0 (0.0–1.0)	0.0 (0.0–2.0)	0.0 (0.0–1.0)	< 0.001
Number of cough diagnoses preceding COPD diagnosis	Median (IQR)	0.0 (0.0–2.0)	1.0 (0.0–3.0)	1.0 (0.0–4.0)	1.0 (0.0–3.0)	< 0.001
Number of cough diagnoses within 5 years preceding COPD diagnosis	Median (IQR)	0.0 (0.0–1.0)	1.0 (0.0–2.0)	1.0 (0.0–2.0)	0.0 (0.0–2.0)	< 0.001
Number of presentations with any COPD symptoms preceding COPD diagnosis	0	6804 (24.3)	3635 (10.3)	400 (12.0)	10,839 (16.3)	< 0.001
	1	7938 (28.3)	7394 (20.9)	646 (19.4)	15,978 (24.0)	
	2	6691 (23.9)	9672 (27.4)	807 (24.2)	17,170 (25.7)	
	3	4286 (15.3)	8868 (25.1)	850 (25.5)	14,004 (21.0)	
	4	1932 (6.9)	4669 (13.2)	525 (15.7)	7126 (10.7)	

continued

TABLE 16 Further GP-recorded symptoms before diagnosis (see Table 4) (continued)

Variable		2006–7	2016–7	March–August 2020	Total	p-value
	5	392 (1.4)	1079 (3.1)	109 (3.3)	1580 (2.4)	
Any COPD symptoms in 5 years preceding COPD diagnosis	0	7869 (28.1)	5474 (15.5)	674 (20.2)	14,017 (21.0)	< 0.001
	1	8784 (31.3)	10,176 (28.8)	877 (26.3)	19,837 (29.7)	
	2	6423 (22.9)	9961 (28.2)	832 (24.9)	17,216 (25.8)	
	3	3456 (12.3)	6537 (18.5)	639 (19.1)	10,632 (15.9)	
	4	1298 (4.6)	2677 (7.6)	261 (7.8)	4236 (6.4)	
	5	213 (0.8)	492 (1.4)	54 (1.6)	759 (1.1)	

FBC, full blood count.

EME
HSDR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

*This report presents independent research funded by the National Institute for Health and Care Research (NIHR).
The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the
Department of Health and Social Care*

Published by the NIHR Journals Library