NIHR | National Institute for Health Research

Journals Library

Check for updates

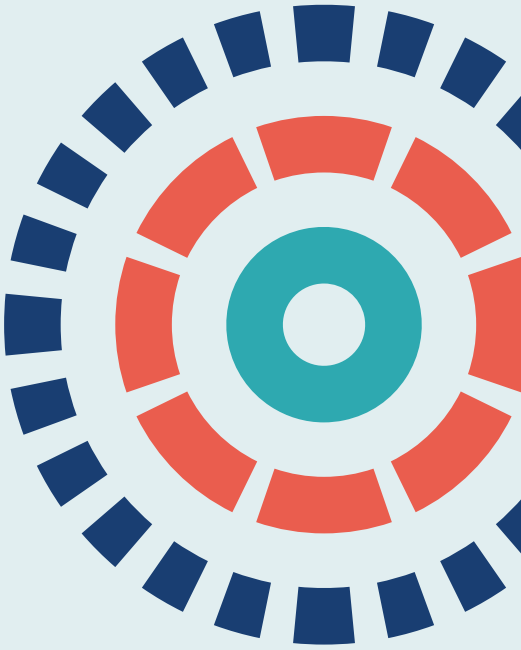
Health Technology Assessment

Volume 25 • Issue 51 • August 2021

ISSN 1366-5278

Medications for chronic obstructive pulmonary disease: a historical non-interventional cohort study with validation against RCT results

Kevin Wing, Elizabeth Williamson, James R Carpenter, Lesley Wise, Sebastian Schneeweiss, Liam Smeeth, Jennifer K Quint and Ian Douglas



DOI 10.3310/hta25510

Medications for chronic obstructive pulmonary disease: a historical non-interventional cohort study with validation against RCT results

Kevin Wing,[1*] Elizabeth Williamson,[2] James R Carpenter,[2] Lesley Wise,[1] Sebastian Schneeweiss,[3,4] Liam Smeeth,[1] Jennifer K Quint[5] and Ian Douglas[1]

[1]Department of Non-communicable Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, UK
[2]Department of Medical Statistics, Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, UK
[3]Department of Epidemiology, Harvard Medical School, Boston, MA, USA
[4]Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA
[5]National Heart and Lung Institute, Imperial College London, London, UK

*Corresponding author

Declared competing interests of authors: James R Carpenter is funded 50% by a grant from the Medical Research Council (MRC) (London, UK) via a secondment to the MRC Clinical Trials Unit (grants MC_UU_12023/21 and MC_UU_12023/29). He reports grants from the National Institute for Health Research (NIHR) (RP-PG-0614-20011) and the Economic and Social Research Council (ESRC) (Swindon, UK; MR/V020641/1 and ES/V005448/1), undertakes methodological consultancy work for Novartis Pharmaceuticals UK Ltd (London, UK) and GlaxoSmithKline plc (Brentford, UK), and has given missing data courses for GlaxoSmithKline plc, Bayer AG (Leverkusen, Germany) and Boehringer Ingelheim (Ingelheim am Rhein, Germany). Lesley Wise is an independent consultant to the pharmaceutical industry and is employed to provide advice by a number of different companies, none of which is involved in this therapeutic area. Sebastian Schneeweiss is a consultant to World Health Information Science Consultants, LLC (Dedham, MA, USA) and to Aetion, Inc. (New York, NY, USA), a software manufacturer of which he also owns equity. He is principal investigator of investigator-initiated grants to the Brigham and Women's Hospital (Boston, MA, USA) from Bayer, Genentech Inc. (South San Francisco, CA, USA) and Boehringer Ingelheim, unrelated to the topic of this study. He does not receive personal fees from biopharmaceutical companies. Sebastian Smeeth reports grants from Wellcome (London, UK), MRC, NIHR (14/49/159 – Statin Web-based Investigation of Side Effects Trial), the British Heart Foundation (London, UK), Diabetes UK (London, UK), ESRC and the European Union; grants and personal fees for advisory work from GlaxoSmithKline plc, and personal fees for advisory work from AstraZeneca plc (Cambridge, UK). He was a trustee of the British Heart Foundation (June 2014–May 2020). Jennifer K Quint's research group has received funding from MRC, Wellcome, the British Lung Foundation (London, UK), GlaxoSmithKline plc, Boehringer Ingelheim, AstraZeneca plc and Insmed, Inc. (Bridgewater Township, NJ, USA) for other projects, none of which relate to this work. Ian Douglas is funded by an unrestricted grant from GlaxoSmithKline plc and has consulted for, and holds stock in, GlaxoSmithKline plc.

Published August 2021
DOI: 10.3310/hta25510

This report should be referenced as follows:

Wing K, Williamson E, Carpenter JR, Wise L, Schneeweiss S, Smeeth L, *et al*. Medications for chronic obstructive pulmonary disease: a historical non-interventional cohort study with validation against RCT results. *Health Technol Assess* 2021;**25**(51).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded* (SciSearch®) and *Current Contents®*/ Clinical Medicine.

Health Technology Assessment

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.014

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, the Cochrane Library and Clarivate Analytics Science Citation Index.

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

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 15/80/28. The contractual start date was in April 2017. The draft report began editorial review in October 2020 and was accepted for publication in May 2021. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report 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 report.

This report presents independent research funded by the National Institute for Health 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, NETSCC, the HTA 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, NETSCC, the HTA programme or the Department of Health and Social Care.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

NIHR Journals Library Editor-in-Chief

NIHR Journals Library Editors

Please visit the website for a list of editors: www.journalslibrary.nihr.ac.uk/about/editors

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

Abstract

Medications for chronic obstructive pulmonary disease: a historical non-interventional cohort study with validation against RCT results

Kevin Wing[1*] Elizabeth Williamson[2] James R Carpenter[2] Lesley Wise[1] Sebastian Schneeweiss[3,4] Liam Smeeth[1] Jennifer K Quint[5] and Ian Douglas[1]

[1]Department of Non-communicable Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, UK
[2]Department of Medical Statistics, Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, UK
[3]Department of Epidemiology, Harvard Medical School, Boston, MA, USA
[4]Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA
[5]National Heart and Lung Institute, Imperial College London, London, UK

*Corresponding author kevin.wing@lshtm.ac.uk

Background: Chronic obstructive pulmonary disease treatment is informed by randomised controlled trial results, but it is unclear if these findings apply to people excluded from these trials. We used data from the TORCH (TOwards a Revolution in COPD Health) randomised controlled trial to validate non-interventional methods for assessing the clinical effectiveness of chronic obstructive pulmonary disease treatment in the UK Clinical Practice Research Datalink, before applying these methods to the analysis of people who would have been excluded from TORCH.

Objectives: To validate the use of non-interventional Clinical Practice Research Datalink data and methods for estimating chronic obstructive pulmonary disease treatment effects against trial results, and, using validated methods, to determine treatment effects in people who would have been excluded from the TORCH trial.

Design: A historical non-interventional cohort design, including validation against randomised controlled trial results.

Setting: The UK Clinical Practice Research Datalink.

Participants: People aged ≥ 18 years with chronic obstructive pulmonary disease registered in Clinical Practice Research Datalink GOLD between January 2000 and January 2017. For objective 1, we prepared a cohort that was analogous to the TORCH trial cohort by applying TORCH trial inclusion/exclusion criteria followed by individual matching to TORCH trial participants. For objectives 2 and 3, we prepared cohorts that were analogous to the TORCH trial that, nevertheless, would not have been eligible for the TORCH trial because of age, asthma, comorbidity or mild disease.

Interventions: The long-acting beta-2 agonist and inhaled corticosteroid combination product Seretide (GlaxoSmithKline plc) [i.e. fluticasone propionate plus salmeterol (FP-SAL)] compared with (1) no FP-SAL exposure or (2) exposure to salmeterol (i.e. the long-acting beta-2 agonist) only.

Main outcome measures: Exacerbations, mortality, pneumonia and time to treatment change.

Results: For objective 1, the exacerbation rate ratio was comparable to that in the TORCH trial for FP-SAL compared with salmeterol (0.85, 95% confidence interval 0.74 to 0.97, vs. TORCH trial 0.88, 95% confidence interval 0.81 to 0.95), but not for FP-SAL compared with no FP-SAL (1.30, 95% confidence interval 1.19 to 1.42, vs. TORCH trial 0.75, 95% confidence interval 0.69 to 0.81). Active comparator results were also consistent with the TORCH trial for mortality (hazard ratio 0.93, 95% confidence interval 0.65 to 1.32, vs. TORCH trial hazard ratio 0.93, 95% confidence interval 0.77 to 1.13) and pneumonia (risk ratio 1.39, 95% confidence interval 1.04 to 1.87, vs. TORCH trial risk ratio 1.47, 95% confidence interval 1.25 to 1.73). For objectives 2 and 3, active comparator results were consistent with the TORCH trial for exacerbations, with the exception of people with milder chronic obstructive pulmonary disease, in whom we observed a stronger protective association (risk ratio 0.56, 95% confidence interval 0.46 to 0.70, vs. TORCH trial risk ratio 0.85, 95% confidence interval 0.74 to 0.97). For the analysis of mortality, we saw a lack of association with being prescribed FP-SAL (vs. being prescribed salmeterol), with the exception of those with prior asthma, for whom we observed an increase in mortality (hazard ratio 1.49, 95% confidence interval 1.21 to 1.85, vs. TORCH trial-analogous HR 0.93, 95% confidence interval 0.64 to 1.32).

Conclusions: Routinely collected electronic health record data can be used to successfully measure chronic obstructive pulmonary disease treatment effects when comparing two treatments, but not for comparisons between active treatment and no treatment. Analyses involving patients who would have been excluded from trials mostly suggests that treatment effects for FP-SAL are similar to trial effects, although further work is needed to characterise a small increased risk of death in those with concomitant asthma.

Limitations: Some of our analyses had small numbers.

Future work: The differences in treatment effects that we found should be investigated further in other data sets. Currently recommended chronic obstructive pulmonary disease inhaled combination therapy (other than FP-SAL) should also be investigated using these methods.

Funding: This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 25, No. 51. See the NIHR Journals Library website for further project information.

Contents

List of tables

List of figures

List of boxes

List of abbreviations

BMI	body mass index	LABA	long-acting beta agonist
CI	confidence interval	LAMA	long-acting muscarinic antagonist
COPD	chronic obstructive pulmonary disease	LRT	log-likelihood ratio test
CPRD	Clinical Practice Research Datalink	NICE	National Institute for Health and Care Excellence
EHR	electronic health record	OCS	oral corticosteroid
FEV_1	forced expiratory volume in 1 second	RCT	randomised controlled trial
FP	fluticasone propionate	RCT-DUPLICATE	Randomized, Controlled Trials Duplicated Using Prospective Longitudinal Insurance Claims: Applying Techniques of Epidemiology
FP-SAL	fluticasone propionate plus salmeterol	RR	risk ratio
FVC	forced vital capacity	SAL	salmeterol
GP	general practitioner	TORCH	TOwards a Revolution in COPD Health
HES	Hospital Episode Statistics		
HR	hazard ratio		
ICS	inhaled corticosteroid		

Plain English summary

Chronic obstructive pulmonary disease affects 3 million people in the UK and is characterised by breathing difficulties that get worse over time, with sudden acute symptoms (exacerbations), possibly requiring hospitalisation. The evidence for use of medicines for treating chronic obstructive pulmonary disease comes from randomised controlled trial results. Randomised controlled trials generally include younger people with severe disease who do not have any other illnesses apart from chronic obstructive pulmonary disease, meaning that the effectiveness of these trials in all people with chronic obstructive pulmonary disease is unknown. Very large databases of anonymous electronic health records captured during NHS consultations can be used to study patients excluded from trials. However, confidence in results from studies using these data can be low because of fears of unaccounted bias, as patients are not randomised to treatment. In this project, we selected a group of patients from a very large electronic health record database called the Clinical Practice Research Datalink who were very similar to participants in a well-known large chronic obstructive pulmonary disease randomised controlled trial [the TORCH (TOwards a Revolution in COPD Health) trial]. When we analysed data from these patients, we found very similar results to the TORCH trial in relation to the reduction of exacerbations, development of pneumonia and time until death, when comparing one chronic obstructive pulmonary disease treatment with another. Having shown that our methods could be trusted to produce valid results when comparing one chronic obstructive pulmonary disease treatment with another, we then went on to analyse patients in the Clinical Practice Research Datalink who would have been excluded from the TORCH trial for the following reasons: aged > 80 years, having asthma as well as chronic obstructive pulmonary disease, or having only mild chronic obstructive pulmonary disease. For exacerbations, we found that, for people with milder chronic obstructive pulmonary disease, one of the treatments we studied seemed to work better than in the trial. For the analysis of mortality, we found that, for people with asthma as well as chronic obstructive pulmonary disease, one of the treatments seemed not to work so well, with more people dying. Future studies are needed in different populations (such as in a database from another country) to confirm these results.

Scientific summary

Text in this summary is reproduced with permission from Wing K, Williamson E, Carpenter JR, Wise L, Schneeweiss S, Smeeth L, *et al.* Real-world effects of medications for chronic obstructive pulmonary disease: protocol for a UK population-based non-interventional cohort study with validation against randomised trial results. *BMJ Open* 2018;**8**:e019475. This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Background

Chronic obstructive pulmonary disease affects 3 million people in the UK. The most common cause is smoking, and patients exhibit airflow obstruction that is not fully reversible. The disease is progressive, with declining lung function and worsening symptoms. Most troublesome are acute exacerbations manifested as a sudden worsening of symptoms (e.g. severe coughing, shortness of breath and chest congestion) that require urgent treatment and possibly hospitalisation. Although smoking cessation remains the most effective intervention, the rate of exacerbation can be reduced by regular medication, such as combination long-acting beta agonists and inhaled corticosteroids or long-acting muscarinic antagonists.

Chronic obstructive pulmonary disease treatment guidelines are largely informed by randomised controlled trial results, but it is not clear if these findings apply to large patient populations that are not studied in these trials. Fluticasone propionate plus salmeterol (FP-SAL) [Seretide (GlaxoSmithKline plc)] is a long-acting beta agonist/inhaled corticosteroid combination and is one of the most widely used chronic obstructive pulmonary disease treatments. It has been studied in large randomised trials [e.g. the TORCH (TOwards a Revolution in COPD Health) trial], but the effects of treatment in important patient groups who were not studied are unknown. Some patient groups were excluded from trials (e.g. those aged > 80 years, those with concomitant asthma or those with substantial comorbidity), whereas others are under-represented (e.g. people with mild chronic obstructive pulmonary disease), meaning that conclusions about these groups are difficult to make.

Although the conduct of non-interventional studies (sometimes also referred to as 'observational studies') to investigate possible drug harms is well established, the use of these studies to estimate treatment effectiveness is in its infancy. Issues of treatment channelling and indication bias mean that measuring the intended benefit of a treatment is beset with difficulties. Over the next few years, we believe that we will see more non-interventional studies of drug effectiveness emerging because of recent legislation that requires pharmaceutical companies to study the real-world effects of medications; however, rigorous, validated methodology is needed to translate these complex data into reliable evidence. For example, the availability of anonymised individual patient data from randomised controlled trials provides the potential for 'randomised controlled trial-analogous' cohorts to be selected from non-interventional data sources (by matching patient records from non-interventional data to the randomised controlled trial patient records on key characteristics). If subsequent analysis of a non-interventional randomised controlled trial-analogous cohort generates results that are similar to those generated by the reference randomised controlled trial, one could be confident in the validity of the results and in the non-interventional methods used to obtain these results in this setting.

In this study we used TORCH individual trial data to validate non-interventional methods for assessing chronic obstructive pulmonary disease treatment effectiveness, before going on to apply these methods to the analysis of treatment effectiveness within people excluded from, or under-represented in,

the TORCH trial. Non-interventional data were obtained from the UK Clinical Practice Research Datalink (linked to the Hospital Episodes Statistics database). The results generated could aid patients, prescribers and policy-makers in deciding the most appropriate treatment for chronic obstructive pulmonary disease for all types of patients. The approach used can also provide a template for treatment effectiveness research using non-interventional data with inbuilt validation against a randomised trial.

Aims and objectives

The aims of our study were as follows:

- to measure the association between treatments for chronic obstructive pulmonary disease and a number of chronic obstructive pulmonary disease outcomes, including exacerbation rate, mortality, pneumonia and time to treatment change, among patients not included in randomised clinical trials for chronic obstructive pulmonary disease treatments
- to develop a methodological framework with inbuilt validation against randomised controlled trial data for using non-interventional electronic health records to answer questions about drug treatment effects (i.e. both benefits and risks).

Specific objectives were to:

- validate methods for measuring chronic obstructive pulmonary disease medication effectiveness in electronic health record data by comparing with trial results
- use electronic health record data to measure chronic obstructive pulmonary disease medication effectiveness in patients excluded from trials (most importantly, those aged > 80 years or those with comorbidities)
- determine chronic obstructive pulmonary disease treatment effectiveness in an understudied disease stage (i.e. mild chronic obstructive pulmonary disease).

Methods

We performed a historical cohort study (2000–17) of chronic obstructive pulmonary disease drug treatment effects in the UK Clinical Practice Research Datalink. For objective 1 (i.e. validation of methods against the TORCH trial), two control groups were selected from the Clinical Practice Research Datalink by applying TORCH trial inclusion/exclusion criteria and 1 : 1 matching to individual TORCH trial participants. Control group 1 included people with chronic obstructive pulmonary disease not prescribed FP-SAL and control group 2 included people with chronic obstructive pulmonary disease who were prescribed salmeterol only. FP-SAL-exposed groups were then selected from Clinical Practice Research Datalink by propensity score matching to each control group. Outcomes studied were chronic obstructive pulmonary disease exacerbations, death from any cause and pneumonia. For objectives 2 and 3 (i.e. analyses of chronic obstructive pulmonary disease medication effectiveness in patients excluded from trials or with an understudied disease stage), the validated methods for patient selection and propensity score development from objective 1 were used to select and analyse the same outcomes in cohorts of people with chronic obstructive pulmonary disease in the Clinical Practice Research Datalink who would have been excluded from the TORCH trial because of age, comorbidities or having mild chronic obstructive pulmonary disease (but who would have otherwise met the TORCH trial criteria). For objectives 2 and 3, the control group was people with chronic obstructive pulmonary disease who were prescribed salmeterol.

Results

For the validation stage (i.e. objective 1), 2652 FP-SAL-exposed people were propensity score matched to 2652 unexposed people, and 991 FP-SAL-exposed people were propensity score matched to 991 salmeterol-exposed people. Exacerbation rate ratio was comparable to the TORCH trial for FP-SAL compared with salmeterol (0.85, 95% confidence interval 0.74 to 0.97, vs. TORCH trial 0.88, 95% confidence interval 0.81 to 0.95), but not for FP-SAL compared with no FP-SAL (1.30, 95% confidence interval 1.19 to 1.42, vs. TORCH trial 0.75, 95% confidence interval 0.69 to 0.81). Active comparator results were also consistent with the TORCH trial for mortality (hazard ratio 0.93, 95% confidence interval 0.65 to 1.32, vs. TORCH trial hazard ratio 0.93, 95% confidence interval 0.77 to 1.13) and pneumonia (risk ratio 1.39, 95% confidence interval 1.04 to 1.87, vs. TORCH trial risk ratio 1.47, 95% confidence interval 1.25 to 1.73). However, different result were obtained from the TORCH trial for the FP-SAL-exposed compared with FP-SAL-unexposed analysis of mortality and pneumonia (mortality hazard ratio 1.11, 95% confidence interval 0.95 to 1.26, vs. TORCH trial mortality hazard ratio 0.83, 95% confidence interval 0.68 to 1.00; pneumonia risk ratio 1.14, 95% confidence interval 0.96 to 1.34, vs. TORCH trial pneumonia risk ratio 1.59, 95% confidence interval 1.35 to 1.88). Time to treatment continuation differed from the TORCH trial for both the FP-SAL compared with salmeterol and FP-SAL compared with no FP-SAL analyses (e.g. FP-SAL vs. SAL hazard ratio 0.23, 95% confidence interval 0.20 to 0.27, vs. TORCH trial hazard ratio 0.89, 95% confidence interval 0.79 to 0.99).

For the over-80s cohort exacerbations analysis, we obtained a propensity score-matched rate ratio of 0.59 (95% confidence interval 0.36 to 0.95) and a propensity score-adjusted rate ratio of 0.83 (95% confidence interval 0.60 to 1.14), which is consistent with the association measured in the TORCH trial-analogous Clinical Practice Research Datalink population (0.85, 95% confidence interval 0.74 to 0.97). For the mortality outcome, we obtained a propensity score-matched hazard ratio of 0.99 (95% confidence interval 0.56 to 1.74) and a propensity score-adjusted hazards ratio of 1.29 (95% confidence interval 0.84 to 2.00). Again, this is consistent with the TORCH trial-analogous Clinical Practice Research Datalink result (0.93, 95% confidence interval 0.65 to 1.32). For the pneumonia analysis, we found no evidence for an increased risk associated with FP-SAL, with a propensity score-matched rate ratio of 0.82 (95% confidence interval 0.44 to 1.53) and a propensity score-adjusted rate ratio of 0.88 (95% confidence interval 0.54 to 1.42).

For the analysis of exacerbations in the cohort of people with concomitant asthma, we found a propensity score-matched rate ratio of 0.74 (95% confidence interval 0.62 to 0.89) and a propensity score-adjusted rate ratio of 0.67 (95% confidence interval 0.59 to 0.78), which is consistent with the association measured in the TORCH trial-analogous Clinical Practice Research Datalink population (0.85, 95% confidence interval 0.74 to 0.97). For the mortality outcome, we obtained a propensity score-matched hazard ratio of 1.49 (95% confidence interval 1.21 to 1.85) and propensity score-adjusted hazards ratio of 1.20 (95% confidence interval 1.04 to 1.40), contrary to the null findings with the TORCH trial-analogous Clinical Practice Research Datalink result (0.93, 95% confidence interval 0.65 to 1.32). For the pneumonia analysis, we found no evidence for an increased risk associated with FP-SAL, with a propensity score-matched rate ratio of 1.09 (95% confidence interval 0.74 to 1.63) and a propensity score-adjusted rate ratio of 1.04 (95% confidence interval 0.79 to 1.37).

For the people with mild chronic obstructive pulmonary disease, we found a propensity score-matched rate ratio for exacerbations of 0.56 (95% confidence interval 0.46 to 0.70) and a propensity score-adjusted rate ratio of 0.52 (95% confidence interval 0.45 to 0.61), which suggests a stronger protective association than that measured in the TORCH trial-analogous Clinical Practice Research Datalink population (0.85, 95% confidence interval 0.74 to 0.97). Notably, however, the crude association in those with mild chronic obstructive pulmonary disease was also strongly protective, unlike in the TORCH trial-analogous population. For the mortality outcome, we obtained a propensity score-matched hazard ratio of 0.98 (95% confidence interval 0.67 to 1.45) and a propensity score-adjusted hazards ratio of 0.84 (95% confidence interval 0.66 to 1.08). Again, this is consistent with the TORCH trial-analogous Clinical

Practice Research Datalink result (0.93, 95% confidence interval 0.65 to 1.32). For the pneumonia analysis, we found no evidence for an increased risk associated with FP-SAL, with a propensity score-matched rate ratio of 0.78 (95% confidence interval 0.45 to 1.35) and a propensity score-adjusted rate ratio of 1.08 (95% confidence interval 0.74 to 1.57).

Conclusions

Our results suggest that routinely collected electronic health record data can be used to successfully identify the expected beneficial and harmful effects of treatments for chronic obstructive pulmonary disease when validating against results obtained from randomised trials. Importantly, successful replication was possible only when comparing between two active treatments, and could not be achieved for comparisons between active treatment and no treatment. These conclusions are specific to investigations of the effects of chronic obstructive pulmonary disease medication and cannot be assumed to replicate in other disease areas. In validating against the results of a large international multicentre randomised trial, it was also clear that, in some instances, some patient characteristics observed in a trial are not always observed in a single-country electronic health record setting. This raises questions of possible trial result heterogeneity by geographic region, which should be considered in future attempts to replicate trial findings in non-interventional data.

The step of directly comparing findings from non-interventional data with those from the trial provided a methodological validation and template, allowing further work to focus on the types of patients excluded from the original trials.

Analyses involving patients who would have been excluded from, or were under-represented in, chronic obstructive pulmonary disease treatment trials mostly suggest that treatment effects for FP-SAL are similar in patients aged > 80 years, those with mild chronic obstructive pulmonary disease and those with both asthma and chronic obstructive pulmonary disease. However, some potential differences were also suggested. For people with mild chronic obstructive pulmonary disease, the use of FP-SAL appears to be more beneficial with respect to exacerbations than was seen in the TORCH trial-analogous population. By contrast, we observed a small increased risk of mortality when comparing FP-SAL with salmeterol in the group with both chronic obstructive pulmonary disease and asthma. These associations should be interpreted with caution, and we recommend future studies to focus on further characterising these associations.

Overall, we have demonstrated the utility of non-interventional data to investigate the expected treatment effects of chronic obstructive pulmonary disease medications, in both trial-included and trial-excluded patient groups. Analyses largely suggest that chronic obstructive pulmonary disease treatment effects are consistent across different patient groups, but highlighted a small number of possible differences that should be investigated further in other data sets. Unanswered questions about the effectiveness of currently recommended chronic obstructive pulmonary disease inhaled combination therapy (other than FP-SAL) in patients excluded from trials should also be investigated using these methods, and further work on advanced technique (e.g. high-dimensional propensity scores) could be performed to investigate whether or not placebo-controlled randomised controlled trials can ever be replicated in this therapeutic area.

Funding

This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 25, No. 51. See the NIHR Journals Library website for further project information.

Chapter 1 Introduction

Background

Chronic obstructive pulmonary disease (COPD) affects 3 million people in the UK.[2] The most common cause is smoking, and patients exhibit airflow obstruction that is not fully reversible. The disease is progressive, with declining lung function and a worsening of symptoms. Most troublesome are acute exacerbations manifested as a sudden worsening of symptoms (e.g. severe coughing, shortness of breath and chest congestion) that require urgent treatment and possibly hospitalisation. Although smoking cessation remains the most effective intervention, the rate of exacerbation can be reduced by regular medication, such as combination long-acting beta agonists (LABAs) and inhaled corticosteroids (ICSs) or long-acting muscarinic antagonists (LAMAs).[3,4]

Chronic obstructive pulmonary disease treatment guidelines are largely informed by randomised controlled trial (RCT) results,[5] but it is not clear if these findings apply to the large patient populations who are not studied in these trials. Fluticasone propionate plus salmeterol (FP-SAL) [seretide (GlaxoSmithKline plc)] is a LABA/ICS combination and is one of the most widely used COPD treatments. It was studied in large randomised trials [e.g. the TORCH (TOwards a Revolution in COPD Health) trial],[3] but the effects of treatment in important patient groups who were not studied are unknown. Some patients were excluded from trials (e.g. those aged > 80 years, those with concomitant asthma or those with substantial comorbidity), whereas others are under-represented (e.g. people with mild COPD),[3,6] meaning that conclusions about these groups are difficult to make.

Although the conduct of non-interventional studies (sometimes also referred to as 'observational studies') to investigate possible drug harms is well established, the use of these studies to estimate treatment effectiveness is in its infancy. Issues of treatment channelling and indication bias mean that measuring the intended benefit of a treatment is beset with difficulties. Over the next few years, we believe that we will see more non-interventional studies of drug effectiveness emerging because of recent legislation that requires pharmaceutical companies to study the real-world effects of medications;[7,8] however, rigorous, validated methodology is needed to translate these complex data into reliable evidence.

For example, the availability of anonymised individual patient data from RCTs provides the potential for 'RCT-analogous' cohorts to be selected from non-interventional data sources (by first applying the trial inclusion and exclusion criteria to a non-interventional data source and then matching patient records from non-interventional data to the RCT patient records on key characteristics). Once a cohort of patients has been selected from non-interventional data with very similar characteristics to the original trial population, analysis can be performed of this cohort, looking at the same outcomes as the trial, but applying statistical methods for analysing non-interventional data. If the results of analysing a RCT-analogous cohort in this way are different from the trial results, this shows that issues with the validity of the analysis remain, even after creating a non-interventional cohort that is highly comparable to the trial. If, however, subsequent analysis of this non-interventional RCT-analogous cohort generates results that are similar to those generated by the reference RCT, one could be confident in the validity of the results and also in the non-interventional methods used to obtain these results in this setting.

This would then provide confidence that if one applies similar analysis approaches to cohorts of patients who were excluded from the trial but have been selected in a similar way to the trial (in terms of inclusion and exclusion criteria but not trial matching, as the trial did not include these patients by design so they would not be available for matching), then the results obtained are likely to be valid.

In this study, we used TORCH[3] individual trial data to validate non-interventional methods for assessing COPD treatment effectiveness, before going on to apply these methods to the analysis of treatment effectiveness within people excluded from, or under-represented in, the TORCH trial.[3] Non-interventional data were obtained from the UK Clinical Practice Research Datalink (CPRD) [linked to the Hospital Episode Statistics (HES) database].[9] The results generated could aid patients, prescribers and policy-makers in deciding the most appropriate treatment for COPD for all types of patients. The approach used can also provide a template for treatment effectiveness research using non-interventional data with inbuilt validation against a randomised trial.

Aims and objectives

The aims of our study were as follows:

- to measure the association between treatments for COPD and a number of COPD outcomes, including exacerbation rate, mortality, pneumonia and time to treatment change, among patients not included in randomised clinical trials for COPD treatments
- to develop a methodological framework with inbuilt validation against RCT data for using non-interventional electronic health records (EHRs) to answer questions about drug treatment effects (i.e. both benefits and risks).

Specific objectives were to:

- validate methods for measuring COPD medication effectiveness in EHR data by comparing with trial results
- use EHR data to measure COPD medication effectiveness in patients excluded from trials (most importantly, those aged > 80 years or those with comorbidities)
- determine COPD treatment effectiveness in an understudied disease stage (i.e. mild COPD).

Figure 1 provides a high-level overview of the study approach, detailing each objective and data source used. *Figure 1* illustrates how existing RCT data were used in objective 1 to validate methods for analysing COPD in routinely collected electronic data for application to unanswered questions in objectives 2 and 3.

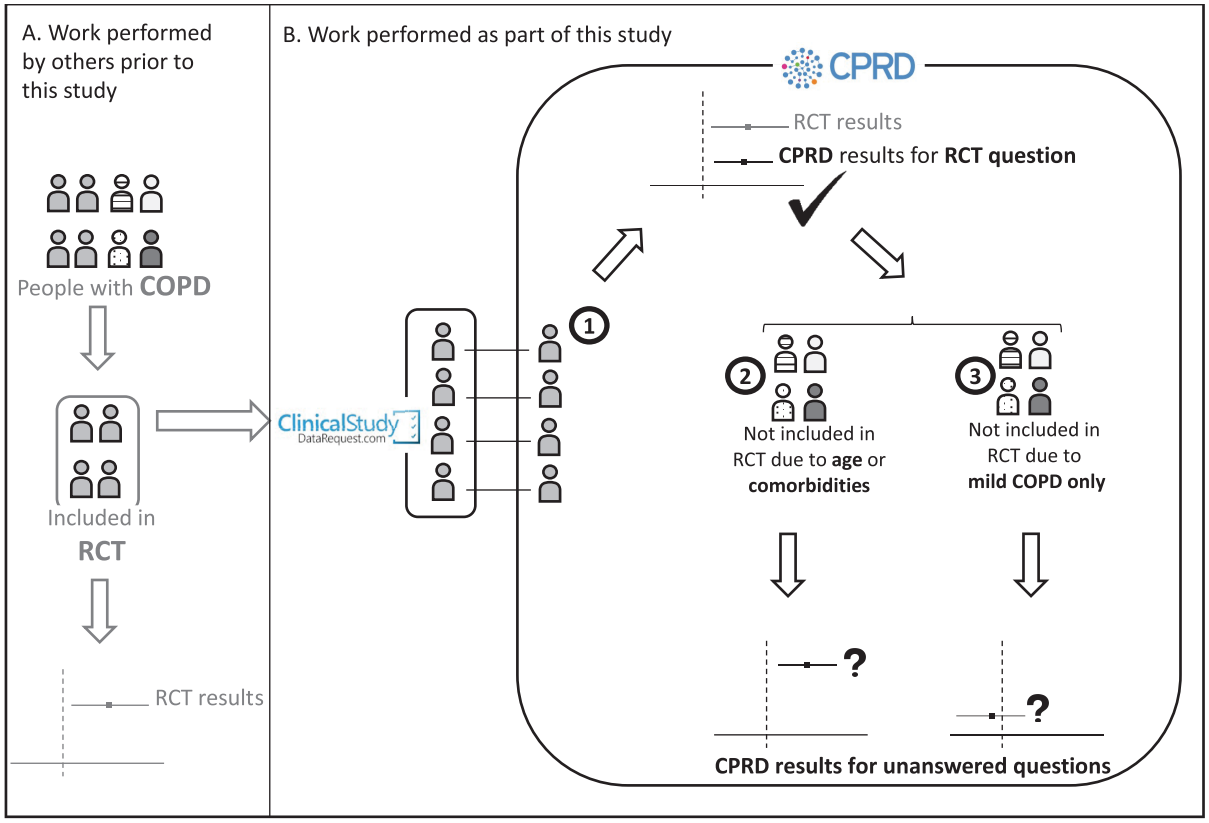


FIGURE 1 Overview of the COPD real-world medicines effects study. A, Work performed by others prior to this study. Of the total population of people with COPD, only a subset are included in RCTs of COPD treatments, based on the RCT inclusion/exclusion criteria. The RCT generates results that inform clinical practice, and the anonymised raw data for the study can be made available to other researchers via the Clinical Study Data Request website. For this study, the specific COPD treatment RCT of interest is the TORCH trial,[3] which investigated the effect of FP-SAL on COPD exacerbations. B, Work performed as part of this study. Objective 1: a cohort of TORCH trial-analogous[3] patients was selected from the UK CPRD by matching people with COPD within CPRD to the records of people included in the trial. Analyses of the effect of FP-SAL on COPD exacerbations were then performed on this TORCH-analogous[3] CPRD cohort. The result obtained were then compared with the TORCH trial[3] itself, serving as a validation step, with comparable results indicating that the data from the non-interventional ('real-world') CPRD source can reliably be used to study COPD treatment effects. Objective 2: the validated analysis techniques used for objective 1 were used to study people in CPRD who would not have been eligible for inclusion in a RCT because of their age and the presence of other comorbidities, and for whom the effect of FP-SAL is currently unknown. Objective 3: the validated analysis techniques were then used to study people with only mild COPD who have been under-represented in RCTs and for whom the effect of COPD treatments is unclear. Reproduced with permission from Wing *et al.*[1] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The figure includes minor additions and formatting changes to the original.

Chapter 2 Methods

Study design

A historical cohort study, with validation against RCT results.

Ethics approval and research governance

Scientific approval was provided by the London School of Hygiene & Tropical Medicine Research Ethics Committee (reference 11997) and the Independent Scientific Advisory Committee of the Medicines and Healthcare products Regulatory Agency (protocol number 17_114R). CPRD data are already approved via a National Research Ethics Committee for purely non-interventional research of this type. Approval for use of the TORCH trial[3] data was obtained from the Wellcome Trust (London, UK), the relevant sponsor (GlaxoSmithKline plc, Brentford, UK) and an independent review panel.

Setting/data sources

Patient data used in this study were obtained from two different sources: the TORCH trial[3] and the UK CPRD (linked to HES data).

The TORCH trial[3]

The TORCH trial[3] was a placebo-controlled randomised trial of combined inhaler FP-SAL for the treatment of COPD, published in 2007. Patients were randomised to receive FP-SAL, fluticasone propionate (FP) alone, salmeterol (SAL) alone or placebo, and the primary comparison of interest was between FP-SAL and placebo.[3] Key outcomes were expected benefits (with a primary outcome of decreased mortality and additional outcome of a decrease in the rate of COPD exacerbations) and an expected harm due to the immunosuppressive action of the corticosteroid FP (pneumonia). Although findings for the primary end point of mortality were null, this was thought to be because of poor statistical power as a result of a lower than anticipated mortality rate. Nonetheless, a lower rate of exacerbations was seen with FP-SAL and a higher rate of pneumonia was observed. As one of the largest trials in COPD, and with a 3-year follow-up, the TORCH trial[3] was a landmark study and provided a validation point for our study.
We obtained individual patient data from the TORCH trial[3] via www.clinicalstudydatarequest.com (accessed 28 May 2021) for use in objective 1 (see *Selection of participants*).

Clinical Practice Research Datalink

The CPRD is a very large database of prospectively collected, anonymised UK population-based EHRs. CPRD primary care records comprise ≈ 8–10% of the UK population and contain comprehensive information on clinical diagnoses, prescribing, referrals, tests and demographic/lifestyle factors.[9]
To contribute to the database, general practices and other health centres must meet prespecified standards for research-quality data (i.e. be 'up to standard'). Data quality/validity are, therefore, high and the data are nationally representative.[9,10] A patient starts contributing follow-up time to

the database at the date they join an 'up-to-standard' practice (or the date that their practice starts contributing up-to-standard data) and stop contributing follow-up time on the date of their death, their transfer out date (i.e. the date that they leave the database for reasons other than death) or the last collection date for their practice. Linkage between the primary care records in CPRD and HES is well established for > 60% of practices in the CPRD, providing a data set augmented with detailed secondary care diagnostic and procedural records. Algorithms have been established to identify COPD, COPD exacerbations, pneumonia (both hospital and primary care managed) and asthma in CPRD/HES-linked data (including validated algorithms for COPD and exacerbations).[11–13] A high-level overview of these algorithms is provided in *Table 1*, and all diagnostic and therapeutic codelist files used to search the CPRD and HES databases for exposure, outcome and covariate information described subsequently in this report are available for download at https://datacompass.lshtm.ac.uk/1655/ (accessed 28 May 2021). For body mass index (BMI) and smoking status, the algorithms we applied looked for the nearest status in

TABLE 1 Overview of algorithms used for detecting COPD, COPD exacerbations and pneumonia in CPRD and HES

Condition	Study	Algorithm description[a]	Validity[b]	Other notes
COPD	Quint *et al.*[11]	CPRD diagnostic (Read) code for COPD	PPV: 87% (95% CI 78% to 92%)	• Comparison with gold standard of respiratory physician review of information obtained by questionnaire from GPs • Eight algorithms presented in total; PPVs ranging from 12% to 89%
COPD exacerbation	Rothnie *et al.*[13]	CPRD diagnostic (Read) code for LRTI or AECOPD OR A prescription of a COPD-specific antibiotic combined with OCS for 5–14 days OR A record (Read code) of two or more respiratory symptoms of AECOPD with a prescription of COPD-specific antibiotics and/or OCS on the same day	PPV: 86% (95% CI 83% to 88%) Sensitivity: 63% (95% CI 55% to 70%)	• Comparison with gold standard of respiratory physician review of information obtained by questionnaire from GPs • Fifteen algorithms presented in total; PPVs ranging from 61% to 100% and sensitivities ranging from 1.6% to 63%
Pneumonia	Millet *et al.*[12]	CPRD diagnostic (Read) codes and HES diagnostic (ICD-10) codes for pneumonia (identified as a subset of an initial search for LRTI codes) Records in both databases within the 28 days considered the same illness episode	No validation performed	
Asthma	Nissen *et al.*[14]	CPRD diagnostic (Read) code indicating asthma	PPV: 86% (95% CI 77% to 95%)	

AECOPD, acute exacerbation of chronic obstructive pulmonary disease; CI, confidence interval; GP, general practitioner; ICD-10, *International Statistical Classification of Diseases and Related Health Problems*, Tenth Revision; LRTI, lower respiratory tract infection; OCS, oral corticosteroid; PPV, positive predictive value.

a Main algorithm applied in this real-world effects study (details on other algorithms presented in paper provided in the 'other notes' column where appropriate).

b Validity = measure of validity presented in article: result obtained (95% CI).

the period − 1 year to + 1 month from the index date (preferred). If this was not available, then the nearest in the period + 1 month to + 1 year after the index date was taken (second preferred). If this was not available, then the nearest before − 1 year from the index date was taken (third preferred) and if this not available, then we took the nearest after + 1 year from the index date (least preferred).

Selection of participants

Objective 1: validation of methods for measuring chronic obstructive pulmonary disease medication effectiveness in electronic health record data by comparing with trial results

For objective 1, two analyses were performed: (1) FP-SAL compared with no FP-SAL (for comparing with the TORCH trial[3] FP-SAL vs. placebo analysis) and (2) FP-SAL compared with SAL only (for comparing with the TORCH trial[3] FP-SAL vs. SAL analysis). The selection procedures for each of these analyses are detailed separately below.

FP-SAL exposed compared with unexposed analysis

Step 1: selection of all potentially eligible patients

An initial cohort was selected from all HES-linked patients actively registered in the CPRD between 1 January 2004 and 1 January 2017, who fulfilled the TORCH trial[3] inclusion criteria (*Box 1*).[3] The date that an individual met all inclusion criteria with at least 12 months prior registration in the CPRD was the 'eligible for the TORCH trial'[3] inclusion date.

Step 2: selection of pool of unexposed patients

Patients who had time periods in which they were unexposed to FP-SAL on or after the 'eligible for the TORCH trial'[3] inclusion date and who did not meet any of the TORCH trial[3] drug exposure exclusion criteria (see *Box 1*) were selected (*Figure 2*).[3] The start of follow-up date (i.e. the index date) for the unexposed time period was selected as a random date between the start and end of the unexposed period (see *Figure 2*). Individuals in CPRD were able to contribute more than one such unexposed time period to the total pool of unexposed time periods (see *Figure 2*) to avoid placing a restriction on a study entry that would not have existed if the potential participants were going to be recruited to a trial (i.e. they could have been recruited to a trial during any one of the eligible periods and we did not want to restrict to only one of these periods at this stage just because we were performing a study using data that had already been collected). Unexposed time periods were then removed from the cohort if the patient met any of the remaining TORCH trial[3] exclusion criteria prior to the index date.[3]

Step 3: selection of unexposed-to-FP-SAL people by 1 : 1 matching FP-SAL time periods to TORCH trial[3] participants

Each individual participant from the TORCH trial[3] [obtained via www.clinicalstudydatarequest.com (accessed 28 May 2021), as described *Setting/data sources*] was matched 1 : 1 with the closest available unexposed-to-FP-SAL time period on the following TORCH trial[3] baseline characteristics: age, sex, BMI, 1-year history of exacerbations requiring hospitalisation, history of cardiovascular disease and lung function [forced expiratory volume in 1 second (FEV_1)]. An individual could contribute only one unexposed period to the final TORCH trial-matched[3] unexposed cohort (see *Figure 2*) and, therefore, the output of this step was a cohort of unexposed-to-FP-SAL people. This trial-matching step was performed to obtain an unexposed cohort that was as similar as possible to that in the TORCH trial.[3]

Step 4: selection of exposed-to-FP-SAL time periods and application of TORCH trial[3] exclusion criteria

We identified all prescriptions for FP-SAL that started (1) on or after the initial 'eligible for the TORCH trial'[3] inclusion date (specified in step 1) and (2) at least 4 weeks after the end of a prescription for any of the TORCH trial[3] drugs. FP-SAL-exposed time periods were created with the index date assigned as the start of a FP-SAL prescription. The same exclusion criteria as applied to the unexposed FP-SAL

BOX 1 The TORCH trial[3] inclusion and exclusion criteria

TORCH[3] inclusion criteria applied to cohort

- A diagnosis of COPD.
- Aged 40–80 years.
- Smoking status of 'current' or 'ex'.
- Lung function criteria of $FEV_1 < 60\%$ predicted and a FEV_1/FVC ratio of $< 70\%$.

TORCH[3] exclusion criteria applied to cohort

Previous drug exposure criteria

- Any exposure to any of the TORCH trial[3] drugs (FP-SAL, SAL or FP) within the previous 4 weeks.
- Current use of a long-acting bronchodilator.[a]
- Current use of OCS therapy.[b]

Remaining exclusion criteria (after applying drug exposure criteria): all at any time prior to the index date unless specified

- A diagnosis of asthma (within the previous 5 years).[c]
- A diagnosis for any (non-COPD) respiratory disorder.
- A record of lung surgery.
- A diagnosis of alpha-1 antitrypsin deficiency.
- A record of having received long-term oxygen therapy.
- Diagnoses for conditions likely to interfere with the TORCH trial[3] or to cause death within the 3 years following the index date.
- Record of an exacerbation requiring OCS therapy or hospitalisation during the period equivalent to the trial 'run-in' period (i.e. the 2-week period following the index date).

FEV_1, forced expiratory volume in 1 second; FVC, forced vital capacity; OCS, oral corticosteroid.

a Current use of a long-acting bronchodilator defined in the CPRD population as any prescription for a long-acting bronchodilator occurring within the period that one of the study drugs was prescribed (or that ended within 7 days prior to the start of a prescription for one of the study drugs).

b Current use of OCS therapy in the TORCH trial[3] was defined as continuous use for > 6 weeks, with courses of OCSs separated by a period of < 7 days considered as continuous use. We applied the same approach to the CPRD population to define exclusion due to exposure to OCS.

c Asthma diagnosis based on a previously validated method for detecting cases of asthma in CPRD.[14]

time periods (step 3) were applied. If an individual contributed time periods to both the unexposed (step 2) and exposed (step 4) cohorts, they were contributing different periods of their person-time to each cohort (pre-FP-SAL treatment for step 2 vs. post-FP-SAL treatment for step 4) (see *Figure 2*).

Step 5: selection of comparable FP-SAL-exposed participants by matching FP-SAL-exposed time periods to FP-SAL-unexposed people

Using the index date baseline characteristics, propensity scores for receiving FP-SAL were calculated for the (TORCH trial-matched[3]) FP-SAL-unexposed people selected in step 3 and the FP-SAL-exposed time periods selected in step 4. Each FP-SAL-unexposed (TORCH trial-matched[3]) person selected in step 3 was matched 1 : 1 with the FP-SAL-exposed time period from step 4 with the closest propensity score. We applied a matching without replacement approach, which meant that an individual could appear only

FIGURE 2 Management of FP-SAL-exposed and FP-SAL-unexposed time periods in selection of people from the CPRD. Step 1: selection of all potentially eligible patients. Six example patients in CPRD. Green arrow = date at which individual met TORCH trial[3] inclusion criteria, grey time periods = FP-SAL exposed, white time periods = FP-SAL unexposed. Step 2: selection of pool of unexposed patients. Unexposed time periods selected and exclusion criteria relating to drug exposures applied. An unexposed record index date is then assigned as a random date within each unexposed period (indicated by diamond symbols) and further TORCH trial[3] exclusion criteria applied based on this date. In this example, one unexposed record from each of persons 1, 4 and 6 were excluded prior to step 3. Step 3: selection of unexposed-to-FP-SAL time periods by 1 : 1 matching to TORCH trial[3] participants. Dotted red lines indicate matching. Matching characteristics assessed on index date of specific unexposed time period and only one time period per person could be matched to the TORCH trial.[3] Step 4: selection of exposed-to-FP-SAL time periods and application of TORCH trial[3] exclusion criteria. In this example, one exposed time period from persons 3 and 6 was excluded based on TORCH trial[3] exclusion criteria. Step 5: selection of comparable FP-SAL-exposed participants. Pre matching there was one record per person for the FP-SA-unexposed cohort and one or more per person for the FP-SAL-exposed cohort. After matching there were an equal number of each. Exposed and unexposed records then followed up and analysed from index date onwards following an 'intention-to-treat' approach. Reproduced with permission from Wing *et al.*[15] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by/4.0/. The figure includes minor additions and formatting changes to the original.

once as an exposed participant in the final propensity score-matched cohort, meaning that this step selected FP-SAL-exposed participants from the initial pool of FP-SAL-exposed time periods. It was possible for the same person to be included in the FP-SAL-unexposed and FP-SAL-exposed cohorts, with different start of follow-up dates in each cohort. The matching of the (trial-matched) FP-SAL-unexposed cohort to the FP-SAL-exposed cohort was performed to obtain a FP-SAL-exposed cohort that was as comparable as possible to the (trial-matched) FP-SAL-unexposed cohort. Importantly, we did not apply matching to the TORCH trial[3] to select our FP-SAL-exposed group because we wanted to develop propensity score methodology for obtaining balanced groups that could then be applied to the study of groups of patients who were not included in the trial (i.e. groups that we would never be able to find to match to in a trial because they were excluded from the trial) (see *Objective 2: measurement of chronic obstructive pulmonary disease treatment effects in patients excluded from trials*).

Selection of participants: FP-SAL-exposed participants compared with salmeterol-exposed participants

The participant selection approach was analogous to the FP-SAL-exposed compared with the FP-SAL-unexposed participant selection, except that the comparator group selected was those exposed to SAL (rather than those unexposed to FP-SAL). The resulting differences in participant selection were as follows. For step 1, the study period was from 1 January 2000 to 1 January 2017 (increased to ensure sufficient numbers of eligible SAL-exposed individuals). For step 2, instead of selecting unexposed-to-FP-SAL time periods occurring on or after the 'eligible for the TORCH trial'[3] inclusion date, we selected periods of SAL exposure. Individuals in the CPRD who had more than one SAL-exposed eligibility period within their record were able to contribute more than once to the pool of SAL-exposed participants (with the covariates and person-time contributed unique to the specific SAL-exposed eligibility period). The index date for each SAL-exposed record was the first date of the eligible SAL exposure period (i.e. the first day of the SAL prescription). All other aspects of step 2 and steps 3–6 were then as described for the FP-SAL-exposed compared with FP-SAL-unexposed participant selection (with SAL-exposed records in place of FP-SAL-unexposed records wherever mentioned).

Objective 2: measurement of chronic obstructive pulmonary disease treatment effects in patients excluded from trials

We selected separate cohorts of patients with a valid COPD diagnosis in the CPRD who would not have been eligible for inclusion in the TORCH trial[3] (and, therefore, also not eligible for our objective 1) because of the following characteristics:

- aged > 80 years
- history of lung surgery
- history of long-term oxygen therapy
- evidence of drug/alcohol abuse
- an asthma diagnosis within the 5 years prior to study entry
- substantial comorbidity.

Separate cohorts were created and analysed for each of the specific characteristics listed, but in all other respects the people selected for each cohort met the TORCH trial[3] criteria (see *Box 1*).

In relation to substantial comorbidity, the TORCH trial[3] required people to be excluded from the trial if they had a serious uncontrolled disease with a likelihood of causing death within 3 years, and application of criteria based up TORCH trial[3] criteria in objective 1 allowed us to select these people (although we recognise this criterion is subjective). Participants for each of the objective 2 cohorts were selected in a similar fashion to the objective 1 cohort, with the amended eligibility criteria specified above applied (i.e. step 1 was modified for selection of each of the objective 2 cohorts).

As for objective 1, each participant was allowed to have multiple FP-SAL-exposed and FP-SAL-unexposed eligibility periods in their record (as described in *Figure 2*). In contrast to objective 1, there was no matching of unexposed patients to TORCH trial[3] patients, as we did not require a TORCH trial-analogous[3] cohort for this analysis (i.e. no step 3). Instead, we were specifically putting together cohorts of people who were not included in the TORCH trial[3] (and, therefore, would not be available or matching). All other selection steps were as applied for objective 1, including the use of propensity score matching to obtain comparable unexposed and exposed groups for analysis. This meant that our overall approach for objective 2 (and objective 3) was to apply the TORCH trial[3] inclusion and exclusion criteria to both the unexposed and exposed groups, but modify the criteria according to the specific trial exclusion criteria that we were interested in including (e.g. for those aged > 80 years we would include only those aged > 80 years, but would still apply the other criteria detailed in *Box 1*). We would then skip the TORCH trial-matching[3] step (as there were no people over the age of 80 years in the TORCH trial[3]), but would apply our propensity score matching approach to obtain comparable exposed and unexposed groups.

Objective 3: determination of treatment effects in an understudied disease stage

We selected separate cohorts of patients who had a valid COPD diagnosis in the CPRD and who would not have been eligible for inclusion in the TORCH trial[3] (or our objective 1) because of having milder COPD than those recruited, as determined by spirometry. This cohort, therefore, included periods of time from people who had a COPD diagnosis and whose spirometry measurements were a > 60% predicted FEV_1 (vs. the TORCH trial[3] requirement of a < 60% FEV_1) and/or a FEV_1/forced vital capacity (FVC) ratio of > 70% (vs. the TORCH trial[3] requirement of a FEV_1/FVC ratio of < 70%). We are aware that study protocols often require the presence of obstructive spirometry (i.e. a FEV_1/FVC ratio of < 0.7) for identification of patients with COPD; however, based on previous validation work[11] in the CPRD of the diagnosis of COPD and National Institute for Health and Care Excellence (NICE) guidance that recommends that clinicians should 'think about a diagnosis of COPD in younger people who have symptoms of COPD, even when their FEV_1/FVC ratio is above 0.7',[5] (© NICE 2010 Chronic obstructive pulmonary disease in over 16s: diagnosis and management. Available from www.nice.org.uk/guidance/cg101. All rights reserved. Subject to Notice of rights. NICE guidance is prepared for the National Health Service in England. All NICE guidance is subject to regular review and may be updated or withdrawn. NICE accepts no responsibility for the use of its content in this product/publication.) our criteria for mild COPD will allow individuals to be included who have a diagnostic code for COPD in the CPRD and a FEV_1/FVC ratio of > 0.7.[5,11]

Exposures, outcomes and covariates

Exposures

For all objectives, exposures were determined using CPRD prescribing records and codelists for COPD treatments [codelists are available from https://datacompass.lshtm.ac.uk/1655/ (accessed 28 May 2021)].

For all objectives, being prescribed FP-SAL was the primary exposure of interest and the comparison exposure groups were (1) people not being prescribed FP-SAL and (2) people being prescribed SAL only. In addition to FP-SAL and SAL, periods of exposure to oral corticosteroids (OCSs), ICS, FP, any LAMA or any LABA were identified to facilitate application of the inclusion and exclusion criteria described in *Selection of participants*.

For all drug exposures, duration of an exposure period was derived by multiplying the CPRD quantity variable by any relevant dose information stored in the packtype variable and then dividing by the value in the numeric daily dose CPRD variable. For example, for a prescription record with quantity = 1, packtype = '60 dose inhaler' and numeric daily dose = 2, the duration of the exposure period was (1 × 60)/2 = 30 days. For prescription records where it was not possible to calculate this exposure period (e.g. because of a missing quantity variable), the median value for that specific drug substance and packtype combination was imputed as the exposure duration. To attempt to account for any uncertainty in the end date of an exposure period (e.g. because of people not taking the medicine as

directed or relying on additional medication previously prescribed and kept at home), a grace period of half the median duration for the specific drug substance/pack type combination was added to the calculated exposure duration to estimate the end date of the exposure period.

Outcomes

Outcomes were COPD exacerbation, all-cause mortality, pneumonia and time to treatment discontinuation, and these were defined as follows:

- COPD exacerbation – defined using a CPRD-HES algorithm that was developed previously by one of the co-authors of this study.[11]
- All-cause mortality – recorded in Office for National Statistics mortality statistics (i.e. data that are linked to CPRD data).
- Pneumonia – defined using a CPRD-HES algorithm that was published previously by authors of this study.[12]
- Time to COPD treatment discontinuation – treatment discontinuation classified as a period of $\geq$ 90 days with no further prescription for the specific drug.

Covariates

Covariates available for inclusion in the propensity score models included lung function, age, sex, alcohol consumption, vascular disease, prescriptions for aspirin or statins, prior treatment with other COPD medication, type 2 diabetes, history of cancer, renal disease and health-care utilisation (i.e. rate of consultations, hospitalisations, hospital procedures and drug prescriptions).

Handling missing data

Complete records analysis was applied, given the small numbers of missing data (only socioeconomic status, alcohol consumption or BMI had any missing data, and all were < 5% missing).

Sample size considerations

Objective 1

Assuming a baseline conservative exacerbation rate of 0.5 per patient per year,[11] we required a sample of 408 patients per treatment group to detect a reduction in annual exacerbation rate of 0.4 per year, with 80% power and 5% significance. Our estimated sample size based on feasibility work assessing the number of people meeting TORCH trial[3] inclusion criteria was $\approx$ 12,000, providing ample power for the main outcomes of interest but also allowing stratification by patient characteristics to determine stratified results. For example, to detect a reduction from 0.5 to 0.4 exacerbations per year with 80% power and 1% significance we would have needed $\approx$ 600 people in each treatment group.

Objectives 2 and 3

We were also confident that we would have sufficient numbers to allow well-powered analyses for objectives 2 and 3. For example, a feasibility count looking at the number of people aged > 80 years eligible for inclusion in objective 2 estimated that there would be > 2000 people in each exposure group.

We were aware that further application of TORCH trial[3] exclusion criteria would lessen sample sizes further, but it was not possible to estimate the extent that this would happen from the data that were available to us prior to undertaking the study.

Blinding

Ascertainment of all outcomes was performed using pre-existing automated algorithms for detecting these outcomes in the CPRD (as detailed in *Outcomes*). Although the data management was performed separately for each drug exposure and, therefore, the person performing the analysis knew which exposure they were obtaining outcomes for at this stage, the code used was identical for each exposure and no edits to the code were permitted based on knowledge of exposure status.

Statistical analysis

Propensity score for addressing confounding

The propensity score for objective 1 was constructed using the principle that predictors of the exposure (i.e. FP-SAL) and outcome (i.e. exacerbations, mortality and pneumonia) or outcome only should be included. We considered a wide range of variables as the pool of initial variables for inclusion (as listed in *Covariates*) based on a priori knowledge of potential association with exposure or outcome, such as age, sex, BMI, alcohol consumption and a wide range of comorbidities (e.g. type 2 diabetes, coronary heart disease, cerebrovascular disease, peripheral vascular disease, heart failure, hypertension, renal disease and cancer). We also considered adjusting for health-care utilisation intensity (e.g. number of prior visits, hospitalisations, number of distinct medications used, number of procedures), as these are generic correlates of disease state and the likelihood of recording completeness. Our group has substantial prior experience of building propensity models.[16–19]

For us to then select variables for inclusion in the propensity score, we removed those variables from the pool of initial variables not associated with outcome in crude analysis before applying multivariable logistic regression (on drug exposure status) to generate propensity scores.[1] Variables were selected for inclusion in the final propensity score multivariable logistic regression model using log-likelihood ratio tests (LRTs) for goodness of fit. Starting from an initial fully adjusted model that included all initial variables found to be associated with outcome, goodness of fit was tested after removing variables sequentially from the logistic regression model (starting with the variable most weakly associated with exposure in the fully adjusted model). Variables with a LRT p-value > 0.1 were removed from the model until all variables remaining in the model had a LRT p-value < 0.1. These remaining variables were the final variables that we used to calculate the propensity score. Separate propensity scores were developed in this way for each outcome. Standardised differences were used to assess any residual imbalances after matching (with a standardised difference > 0.1 indicating substantial/important imbalance).[18]

The variable list used for the propensity score model obtained in objective 1 was the basis for propensity score modelling in objectives 2 and 3, but additional variables from the pool of initial variables were also considered, given the different nature of the patient populations being studied in these objectives. We also assessed the impact of adjusting for the propensity score (rather than matching) for these analyses.

Methods of analysis

For all objectives, comparisons were made according to FP-SAL (or other drugs being analysed, as specified in *Exposures*) status for rate of COPD exacerbation, pneumonia and mortality over 3 years. All analyses were performed according to the 'intention-to-treat' principle (as was carried out in the TORCH trial[3]), meaning that if a participant entered the study as either an exposed or an unexposed participant then they would remain assigned to that exposure category for the entire duration of their follow-up (irrespective as to whether or not their true exposure status changes). For exacerbations, a negative binomial model was used, accounting for variability between patients in the number and frequency of exacerbations, with the number of exacerbations as the outcome and the log of treated time as an offset variable. Time to mortality and treatment change was analysed using Cox proportional hazards regression. Risk of pneumonia was analysed using Poisson regression. This mirrors the TORCH trial[3] end points of major benefit and harm. We anticipated that the propensity matching process would allow us to assemble treated and untreated groups that were very similar with respect to baseline characteristics, except FP-SAL treatment status. However, this was tested by assessing standardised differences for each baseline variable. If substantial differences were noted for important variables, we made further adjustments to the statistical models. This could also include examining the effect of using a greedy matching approach (i.e. where once a match is made it is fixed) compared with an optimal matching approach (i.e. where the algorithm reconsiders all previously made matches before making a new match) to obtain the closest propensity score match and/or matching at a ratio other than 1 : 1.[19]

Validation of results against the TORCH trial[3]

Our findings were validated against the TORCH trial[3] as part of objective 1 by determining whether or not results of the CPRD FP-SAL compared with no FP-SAL treatment analysis were compatible with the TORCH trial[3] exacerbations rate ratio for FP-SAL compared with placebo [0.75, 95% confidence interval (CI) 0.69 to 0.81]. This outcome was selected because it is an outcome of key significance for people with COPD[5] and the result in TORCH trial[3] shows a clear benefit with 95% confidence limits of < 1. We set two criteria that needed to be met for us to conclude that results were consistent. First, the effect size needed to be clinically comparable with TORCH trial[3] findings (i.e. the rate ratio for exacerbations in the CPRD had to be between 0.65 and 0.9). This range was deliberately not symmetrical around the TORCH trial[3] estimate of 0.75, as we anticipated that the treatment effect in routine clinical care would be weaker than that seen in the optimised setting of a randomised trial. We recognised that this rule could be met with a poorly powered, inconclusive result, and so a second criterion was that the 95% CI for the rate ratio had to exclude 1. For the FP-SAL with SAL alone comparison (see *Exposures*), the 95% CI also needed to exclude 1 and the rate ratio had to be between 0.81 and 0.95 (compared with the TORCH trial[3] FP-SAL vs. SAL result of 0.88, 95% CI 0.81 to 0.95).

Sensitivity analyses

Handling measurement of adherence to medication

We considered that adherence to issued prescribing in general practice is likely to vary according to the treatment issued. For example, short-course antibiotic treatment is notoriously not well adhered to, whereas long-term life-saving treatment, such as antiretroviral medication, is more likely to be taken as prescribed. Although we were not aware of published figures for adherence for COPD medication in UK general practice, we reviewed the records for a random sample of 30 people with COPD starting treatment with FP-SAL to look at adherence patterns over the course of 1 year. Of 30 patients, 20 (67%) were still receiving seretide (FP-SAL) 1 year after starting treatment. Of the 20 patients who received seretide for a full year, 15 (75%) received sufficient prescriptions to suggest at least 50% adherence over the year and eight (40%) had sufficient prescriptions to suggest $\geq$ 80% adherence. As expected, we considered that this suggested that (1) adherence is likely to be poorer in routine clinical care than in the trial population (in the TORCH trial[3] 80% of participants were estimated to have adherence at $\geq$ 80%) and (2) there is a wide range of adherence in routine care. Although we acknowledge that prescribing can be only a proxy for used medication, we believed that it was not an unreasonable assumption that the amount of medication prescribed would be correlated with the amount consumed. We aimed to assess adherence for the cohort that we select for objective 1 beyond 1 year and report the findings. In the event that our analyses in objective 1 detected a null or poorer treatment effect than anticipated, we planned to conduct a sensitivity analysis restricted to people estimated to be covered by FP-SAL treatment for 80% of their follow-up.

Misclassification of (1) drug exposure periods and (2) outcome status

We considered that it would be possible that an individual may still be exposed to FP-SAL for some time after a prescription has finished (e.g. if they have medication at home that they have not used from a previous prescription). This would mean that people may become eligible for inclusion in the unexposed group while they are actually still exposed. We planned to conduct a sensitivity analysis if our results from objective 1 differed from the TORCH trial[3] results. In this analysis, we would include an additional (grace) exposed period that was equivalent to the length of a single prescription at the end of each actual exposed period and only classify individuals as eligible for inclusion as unexposed at the end of this additional period.

Misclassification of outcome can also have an impact on our results, given the routine nature of the data. Our initial approach for the detection of COPD exacerbations was to use a validated case definition from previous work that maximises positive predictive value while maintaining a relatively high sensitivity.[13] We therefore planned to perform a sensitivity analysis in which we assessed the impact of applying the alternative case definitions for COPD exacerbations from this publication if our objective 1 results differed from the TORCH trial.[3]

Safety reporting and disclosure

As this study was a non-interventional study that used stored electronic health data in the UK CPRD (with no recruitment of patients or intervention), there was no requirement for safety reporting or disclosure.

Deviations from original protocol

- In objective 3, when looking at people with milder COPD, as defined by spirometry, we originally specified that we would also look at people with no exacerbations at all in the year post COPD diagnosis. However, we decided not to analyse this outcome because we were already looking at a milder group, as defined by spirometry, and had limited time, as we effectively had to perform the main data management and analysis steps twice (i.e. once for the comparison of FP-SAL with no FP-SAL and then once for the comparison of FP-SAL with SAL because the FP-SAL and no FP-SAL comparison produced results that were different from the TORCH trial[3]).
- For objective 2, there were a number of subgroups we planned to analyse, but were unable to because of small numbers in these groups. These subgroups were people with a history of lung surgery, people with a history of long-term oxygen therapy, people with substantial comorbidity or people with evidence of drug/alcohol abuse. See *Chapter 3*, *Note on results presented in results and discussion part 2* for an overview of the actual numbers available for analysis from these groups.
- For objectives 1–3, in the protocol there were secondary analyses specified where we would repeat each analysis comparing COPD treatments other than FP-SAL with no treatment (e.g. no treatment vs. LABA, LAMA, LABA + LAMA, LABA + ICS other than FP-SAL and LABA + LAMA + ICS). We did not perform these secondary analyses comparing these exposures with no treatment because of the amount of work we had to do in repeating our primary analyses (i.e. FP-SAL vs. no FP-SAL followed by FP-SAL vs. SAL), but have specified comparing these treatments with SAL as future research (see *Chapter 4*, *Prioritised list of recommendations for future research*). However, the specific comparisons would need updated based on updated NICE guidance[20] published since the start of the project.

Patient and public involvement

We invited four patient and public representatives to be involved and advise on the project [via Breath Easy, www.blf.org.uk/support-for-you/breathe-easy (accessed 1 June 2021)], one of whom accepted our invitation and attended each of the four Steering Group meetings that took place during the project. During these meetings, the patient and public involvement representative provided feedback on whether or not the results we were presenting were clear and understandable, provided feedback on our plans on next steps based on results and provided insight on aspects of COPD treatment from a COPD patient's perspective (e.g. in relation to how a patient might adhere to/not adhere to COPD medication, and how a patient might typically go about managing COPD medications, which had been prescribed in UK primary care, at home).

Chapter 3 Results and discussion

Organisation of results and discussion section

The results and discussion section is organised into two parts, as illustrated in *Table 2*. We have organised the chapter in this way because the choice of analysis in the second part was determined by content in both the results and the discussion of analysis in the first part. Therefore, the most logical way to present the results and discussion was a single chapter that is split into two parts (see *Table 2*).

Results and discussion part 1: analysis of chronic obstructive pulmonary disease treatment effects in a TORCH trial-analogous[3] cohort (objective 1)

Results

Participants

FP-SAL exposed compared with FP-SAL unexposed

Between 1 January 2004 and 1 January 2017 there were 125,671 people in the CPRD with a diagnosis of COPD, 73,889 (59%) of whom were from HES-linked CPRD practices (*Figure 3*). Application of TORCH trial[3] inclusion criteria reduced this to 18,715 people, contributing 35,746 unexposed-to-FP-SAL time periods and 26,390 exposed-to-FP-SAL time periods. After applying TORCH trial[3] exclusion criteria, dropping records with missing covariate data and matching the unexposed patients to TORCH trial[3] participants, there were 4196 unexposed patients available for propensity score matching to 10,463 FP-SAL-exposed time periods. The final propensity score-matched cohorts included 2652 patients in each exposure group for the exacerbations analysis, 2708 patients in each exposure group for mortality and 2779 patients in each exposure group for pneumonia.

FP-SAL exposed compared with SAL exposed

For the FP-SAL compared with SAL analysis, there were 154,785 people with a diagnosis of COPD in the CPRD between 1 January 2000 and 1 January 2017, 91,733 (59%) of whom were from HES-linked CPRD practices (*Figure 4*). A total of 1146 SAL-exposed patients were available for propensity score matching to 11,235 FP-SAL-exposed periods. The final propensity score-matched cohorts included 991 (exacerbations), 432 (mortality), 935 (pneumonia) and 996 (treatment discontinuation) patients per exposure group.

TABLE 2 Organisation of results and discussion section

Results and discussion part number	Population for analysis	Exposure	Outcome	Objective that section relates to
1	COPD treatment effects in a TORCH trial-analogous[3] cohort	FP-SAL vs. no FP-SAL FP-SAL vs. SAL	Exacerbations, mortality, pneumonia, time to treatment discontinuation	1
2	COPD treatment effects in (a) patients excluded from trials and (b) patients with milder COPD	FP-SAL vs. SAL	Exacerbations, mortality, pneumonia	2 and 3

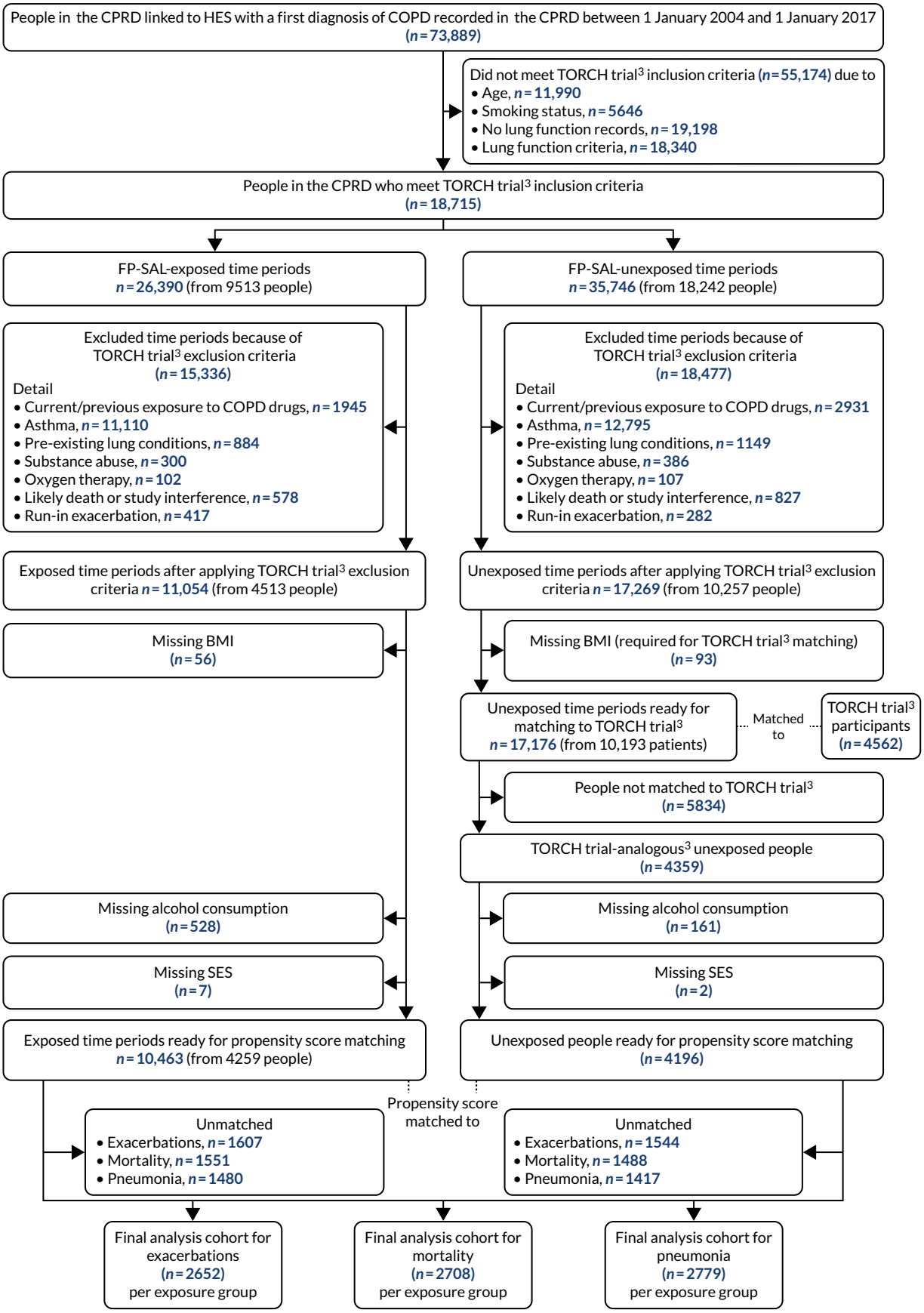


FIGURE 3 Flow of number of individuals included in the exposed to FP-SAL vs. unexposed to FP-SAL cohort analysis. Note that the current/previous use of COPD drugs relates to any of the drugs studied in the TORCH trial,[3] long-acting bronchodilators and OCSs (see *Box 1*). SES, socioeconomic status.

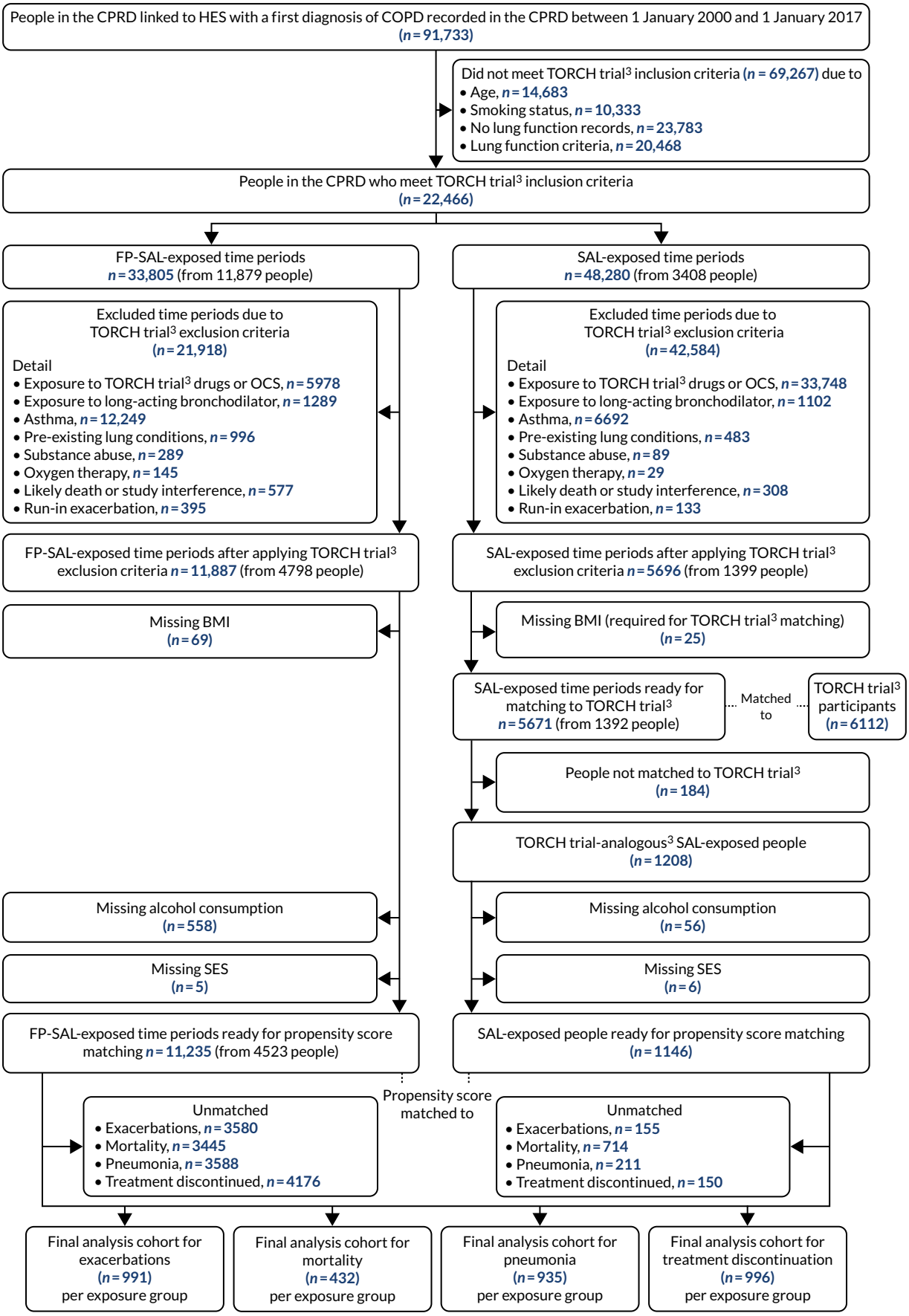


FIGURE 4 Flow of number of individuals included in the exposed to FP-SAL vs. exposed to SAL cohort analysis. SES, socioeconomic status.

Application of TORCH trial[3] inclusion/exclusion criteria and matching to the TORCH trial[3]

Applying the TORCH trial[3] inclusion/exclusion criteria and matching to the TORCH trial[3] resulted in cohorts that were much more similar to those recruited to the TORCH trial[3] (e.g. FEV_1% of predicted for the FP-SAL vs. unexposed to FP-SAL analysis was 66.3 in the CPRD before applying any criteria or matching, compared to 47.2 after these steps, compared to a TORCH[3] placebo group value of 44.2) (see *Table 3*). The largest residual difference to the TORCH trial[3] placebo group was for prior cardiovascular disease for both comparisons (*Tables 3* and *4*).

Propensity score matching of Clinical Practice Research Datalink cohorts

Details of the variables included in the final propensity score models are provided in *Table 5*.

TABLE 3 Characteristics of the CPRD non-interventional COPD cohort for the FP-SAL vs. no FP-SAL comparison, showing (1) the cohort of all HES-linked patients in the CPRD with a COPD diagnosis, (2) the cohort after applying TORCH trial[3] inclusion/exclusion criteria (unexposed to FP-SAL group only) and (3) the cohort after applying TORCH trial[3] inclusion/criteria and matching to TORCH trial[3] participants (unexposed to FP-SAL group only), in comparison with the baseline characteristics of the TORCH trial[3] placebo group

	CPRD non-interventional population			
		Unexposed to FP-SAL		
Variable	All [no TORCH trial[3] criteria or TORCH[3] matching applied[a] ($n = 45,939$ patients)]	After applying TORCH trial[3] inclusion/exclusion criteria[b] ($n = 17,176$ unexposed time periods from 10,193 people)	After matching[c] to individual TORCH trial[3] patients ($n = 4359$ unexposed people)	TORCH trial[3] placebo group ($n = 1524$ trial participants)
Age (years), median (IQR)	65 (58–74)	68.0 (61.0–73.0)	67.0 (61.0–73.0)	65 (59–71)
Sex (male), *n* (%)	24,182 (53)	10,671 (62)	3307 (76)	1163 (76)
BMI (kg/m^2), median (IQR)	26.7 (23.4–30.7)	26.3 (22.6–30.4)	25.5 (22.1–29.0)	25.0 (22.0–28.4)
Exacerbations requiring hospitalisation (mean ± SD)[d]	0.1 (0.9)	0.0 (0.3)	0.1 (0.3)	0.2 (0.7)
History of cardiovascular disease	11,564 (25)	4888 (28)	1987 (46)	784 (51)
Lung function: FEV_1 per cent of predicted, median (IQR)	66.3 (51.6–81.33)	51.7 (41.8–59.0)	47.2 (37.3–56.1)	44.2 (35.0–54.0)

IQR, interquartile range; SD, standard deviation.

a Includes all people in the CPRD with a diagnosis for COPD between 2004 and 2016 who have spirometry data recorded. All variables in this column measured at the date of COPD diagnosis recorded in the CPRD.

b Inclusion criteria: diagnosis of COPD, aged 40–80 years, all current or ex-smokers, lung function of $FEV_1 < 60\%$ predicted and a FEV_1/FVC ratio of $< 70\%$. Exclusion criteria: diagnosis of asthma within the previous 5 years, diagnosis for any non-COPD respiratory disorder, a record of lung surgery, a diagnosis of alpha-1 antitrypsin deficiency, evidence of drug or alcohol abuse, a record of having received long-term oxygen therapy, diagnoses likely to interfere with the TORCH trial[3] or cause death within 3 years, current use of OCS therapy and any exposure to FP-SAL within the previous 4 weeks. All variables in this column were measured at the earliest date that all inclusion criteria were met and all exclusion criteria were not met.

c Matched on all variables in this table.

d Within prior year.

TABLE 4 Characteristics of the CPRD non-interventional COPD cohort for the FP-SAL vs. SAL analysis showing (1) the cohort of patients in the CPRD with a COPD diagnosis (2) the cohort after applying TORCH[3] inclusion/exclusion criteria (exposed to SAL group) and (3) the cohort after applying TORCH[3] and inclusion/criteria and matching to TORCH[3] participants (exposed to SAL group), in comparison to the baseline characteristics of the TORCH trial[3] SAL group

	CPRD non-interventional population			
		Exposed to SAL		
Variable	All [no TORCH trial[3] criteria or TORCH trial[3] matching applied[a] (*n* = 53,099 people)]	After applying TORCH[3] inclusion/exclusion criteria[b] (*n* = 5671 SAL-exposed time periods from 1392 people)	After matching[c] to individual TORCH trial[3] patients (*n* = 1208 SAL-exposed people)	TORCH trial[3] SAL group (*n* = 1524 trial participants)
Age (years), median (IQR)	66.0 (58.0–74.0)	68.0 (63.0–74.0)	68.0 (62.0–73.0)	65.1 (60.0–71.0)
Sex (male), *n* (%)	35,045 (53)	3415 (60)	767 (63)	1160 (76)
BMI (kg/m^2), median (IQR)	25.8 (23.0–29.1)	26.9 (23.3–30.8)	26.2 (23.0–29.9)	24.8 (21.9–28.3)
Exacerbations requiring hospitalisation (mean ± SD)[d]	0.0 (0.3)	0.0 (0.1)	0.0 (0.2)	0.2 (0.6)
History of cardiovascular disease	13,274 (25)	1689 (30)	374 (31)	807 (53)
Lung function: FEV_1 per cent of predicted, median (IQR)	63.2 (49.1–76.8)	52.6 (43.4–61.1)	49.4 (40.5–57.1)	43.4 (33.8–53.4)

IQR, interquartile range; SD, standard deviation.
a Includes all people in the CPRD with a diagnosis for COPD between 2004 and 2016 who have spirometry data recorded. All variables in this column measured at the date of first SAL exposure.
b Inclusion criteria: diagnosis of COPD, aged 40–80 years, all current or ex-smokers, lung function of $FEV_1 < 60\%$ predicted and a FEV_1/FVC ratio of < 70%. Exclusion criteria: diagnosis of asthma within the previous 5 years, diagnosis for any non-COPD respiratory disorder, a record of lung surgery, a diagnosis of alpha-1 antitrypsin deficiency, evidence of drug or alcohol abuse, a record of having received long-term oxygen therapy, diagnoses likely to interfere with the TORCH trial[3] or cause death within 3 years, current use of OCS therapy and any exposure to any of the study drugs within the previous 4 weeks. All variables in this column were measured at the earliest date that all inclusion criteria were met and all exclusion criteria were not met.
c Matched on all variables in this table.
d Within prior year.

TABLE 5 Variables included in the final propensity score models

Analysis	Variables included in propensity score model	Matching
FP-SAL vs. unexposed to FP-SAL analysis		
Exacerbations	Sex, age, FEV_1, FEV_1/FVC, BMI, year of index date, previous diagnosis of cerebrovascular disease; having at least one prescription of (1) statin, (2) ICS, (3) LABA – ICS combination therapy or (4) LAMA in the previous year; and the frequency of consultations, prescriptions, hospitalisations, hospital procedures and exacerbations in the previous year	1 : 1 nearest neighbour, calliper[a] of 0.03
Mortality	Sex, age, FEV_1, FEV_1/FVC, BMI, SES, previous diagnosis of (1) coronary heart disease, (2) peripheral vascular disease or (3) cerebrovascular disease; having at least one prescription of (1) LAMA or (2) LABA – ICS combination therapy in the previous year; and the frequency of consultations, prescriptions, hospitalisations and exacerbations in the previous year	1 : 1 nearest neighbour, calliper of 0.03

continued

TABLE 5 Variables included in the final propensity score models (*continued*)

Analysis	Variables included in propensity score model	Matching
Pneumonia	Sex, age, FEV_1, FEV_1/FVC, BMI, alcohol consumption, previous diagnosis of (1) coronary heart disease, (2) peripheral vascular disease or (3) cerebrovascular disease; having at least one prescription of (1) LAMA or (2) aspirin in the previous year, and the frequency of prescriptions, hospitalisations and exacerbations in the previous year	1 : 1 nearest neighbour, calliper of 0.03
FP-SAL vs. SAL analysis		
Exacerbations	Sex, FEV_1, previous diagnoses for (1) type 2 diabetes or (2) chronic kidney disease, year of index date, having at least one prescription of an ICS in the previous year, and the frequency of consultations, hospitalisations and hospital procedures in the previous year	1 : 1 nearest neighbour, calliper of 0.03
Mortality	Sex, age, year of index date, BMI, SES, FEV_1, FEV_1/FVC, diagnoses for (1) peripheral vascular disease, (2) coronary heart disease, (3) cerebrovascular disease, (4) Type 2 diabetes, (5) cancer or (6) chronic kidney disease; having at least one prescription of (1) statin, (2) aspirin, (3) LAMA, (4) LABA or (5) LABA – ICS combination therapy in the previous year; and the frequency of consultations, exacerbations, prescriptions, hospitalisations and hospital procedures in the previous year	1 : 1 nearest neighbour, calliper of 0.03
Pneumonia	FEV_1, year of index date, SES, diagnoses for chronic kidney disease, and the frequency of consultations, prescriptions, hospitalisations and hospital procedures in the previous year	1 : 1 nearest neighbour, calliper of 0.03
Time to treatment discontinuation	FEV_1, FEV_1/FVC, alcohol intake, SES, year of index date, diagnoses for (1) peripheral vascular disease, (2) coronary heart disease, (3) cancer or (4) chronic kidney disease; having at least one prescription of (1) statin, (2) aspirin, (3) ICS or (4) LABA – ICS combination therapy in the previous year; and the frequency of consultations, exacerbations, prescriptions, hospitalisations and hospital procedures in the previous year	1 : 1 nearest neighbour, calliper of 0.03

SES, socioeconomic status.
a Calliper = maximum permitted difference in propensity score between matched pairs.

FP-SAL exposed compared with FP-SAL unexposed

Prior to propensity score matching, for the exacerbations, analysis differences by exposure status were noted for sex, FEV_1, BMI, prior exacerbations, coronary heart disease, peripheral vascular disease, cerebrovascular disease, prescriptions for aspirin, COPD medications, number of general practitioner (GP) consultations and number of distinct medications (*Table 6*). After propensity score matching, only the differences with respect to coronary heart disease, peripheral vascular disease and LABA persisted (see *Table 6*). Plots of propensity score distributions indicated close propensity score matching for exacerbations and all other outcomes under study (*Figure 5*).

FP-SAL exposed compared with salmeterol exposed

For the FP-SAL compared with SAL exacerbations analysis, after propensity score matching, there were notable imbalances in prior prescriptions for a LABA or an ICS and frequency of consultations, with smaller imbalances for lung function, BMI, coronary heart disease, statin prescription, aspirin prescription, LAMA, ICS plus LABA and prior GP consultations (*Table 7*). Plots of propensity score distribution indicated that, overall, groups were well matched on propensity score for each outcome (see *Figure 5*).

Main results

FP-SAL exposed compared with FP-SAL unexposed

For the exacerbations analysis, the rate ratio in the propensity score-matched groups was 1.30 (95% CI 1.19 to 1.42) (*Table 8*). According to our prespecified protocol, this (harmful) association was not considered to be consistent with the (protective) TORCH trial[3] placebo-controlled result for the same outcome (0.75, 95% CI 0.69 to 0.81).[1] Similarly, our result for the mortality outcome [hazard ratio (HR) 1.11,

TABLE 6 Characteristics of the exposed to FP-SAL vs. unexposed to FP-SAL cohort before and after propensity score matching for the exacerbations analysis (CPRD non-interventional population)

Variable	Before propensity score matching			After propensity score[a] matching		
	Unexposed to FP-SAL[b] (*N* = 4196 people)	Exposed to FP-SAL[c] (*N* = 10,463 exposed time periods from 4259 people)	Standardised difference	Unexposed to FP-SAL (*N* = 2652 people)	Exposed to FP-SAL (*N* = 2652 people)	Standardised difference
Age (years), median (IQR)	67 (61–73)	68 (62–74)	0.103	68 (61–73)	68 (62–74)	0.083
Sex (male), *n* (%)	3175 (76)	6515 (62)	0.293	1868 (70)	1850 (70)	0.015
Lung function[d]						
FEV_1 per cent of predicted, median (IQR)	47 (38–56)	50 (40–60)	0.297	49 (39–57)	48 (38–56)	0.024
FEV_1 : FVC per cent, median (IQR)	53 (44–61)	53 (44–63)	0.073	53 (44–62)	52 (43–61)	0.045
BMI (kg/m^2),[d] median (IQR)	26 (22–29)	26 (23–31)	0.191	26 (23–30)	26 (22–30)	0.024
Prior exacerbations,[e] mean (SD)	0.51 (0.92)	0.66 (1.13)	0.148	0.56 (0.96)	0.62 (1.04)	0.060
Cardiovascular disease, *n* (%)[f]						
Coronary heart disease	1114 (27)	1783 (17)	0.232	720 (27)	441 (17)	0.257
Peripheral vascular disease	390 (9)	648 (6)	0.116	253 (10)	166 (6)	0.122
Cerebrovascular disease	434 (10)	714 (7)	0.126	212 (8)	222 (8)	0.014
Other atherosclerosis	11 (0)	20 (0)	0.015	7 (0)	7 (0)	0.008
Statin prescription, *n* (%)[g]	2066 (49)	4614 (44)	0.103	1227 (46)	1238 (47)	0.008
Aspirin prescription, *n* (%)[g]	1563 (37)	3129 (30)	0.156	954 (36)	828 (31)	0.101
Other COPD medication prescriptions, *n* (%)[g]						
LABA[h]	295 (7)	333 (3)	0.175	197 (7)	106 (4)	0.148
ICS[h]	530 (13)	842 (8)	0.151	280 (11)	333 (13)	0.063
LAMA[h]	1450 (35)	6284 (60)	0.528	1166 (44)	1177 (44)	0.008
ICS plus LABA[i]	526 (13)	488 (5)	0.284	196 (7)	258 (10)	0.084

continued

TABLE 6 Characteristics of the exposed to FP-SAL vs. unexposed to FP-SAL cohort before and after propensity score matching for the exacerbations analysis (CPRD non-interventional population) (*continued*)

Variable	Before propensity score matching			After propensity score[a] matching		
	Unexposed to FP-SAL[b] (*N* = 4196 people)	Exposed to FP-SAL[c] (*N* = 10,463 exposed time periods from 4259 people)	Standardised difference	Unexposed to FP-SAL (*N* = 2652 people)	Exposed to FP-SAL (*N* = 2652 people)	Standardised difference
Type 2 diabetes, *n* (%)[f]	543 (13)	1496 (14)	0.040	373 (14)	337 (13)	0.04
History of cancer, *n* (%)[f]	696 (17)	2105 (20)	0.091	486 (18)	451 (17)	0.035
Chronic kidney disease, *n* (%)[f]	540 (13)	1477 (14)	0.037	389 (15)	333 (13)	0.062
Health-care utilisation, median (IQR)[e]						
Number of GP consultations	21 (15–29)	16 (10–26)	0.409	18 (14–29)	16 (10–26)	0.143
Number of distinct medications	4 (2–7)	5 (3–8)	0.180	4 (2–7)	5 (3–8)	0.073
Number of hospitalisations	0 (0–1)	0 (0–1)	0.008	0 (0–1)	0 (0–1)	0.007
Number of hospital procedures	0 (0–0)	0 (0–1)	0.022	0 (0–0)	0 (0–0)	0.011

IQR, interquartile range; SD, standard deviation.
a See *Table 5* for list of variables included in final exacerbations propensity score model. Variables in this table that were included in the propensity score are in bold.
b TORCH trial[3] inclusion/exclusion criteria applied and matched to TORCH trial[3] individual patient data.
c TORCH trial[3] inclusion/exclusion criteria applied.
d Closest record prior to the index date.
e All counted within the year prior to the index date, includes exacerbations recorded in primary or secondary care.
f Any diagnosis for condition prior to the index date.
g Number of people who had at least one prescription within the previous year.
h Single product only.
i Combination product.

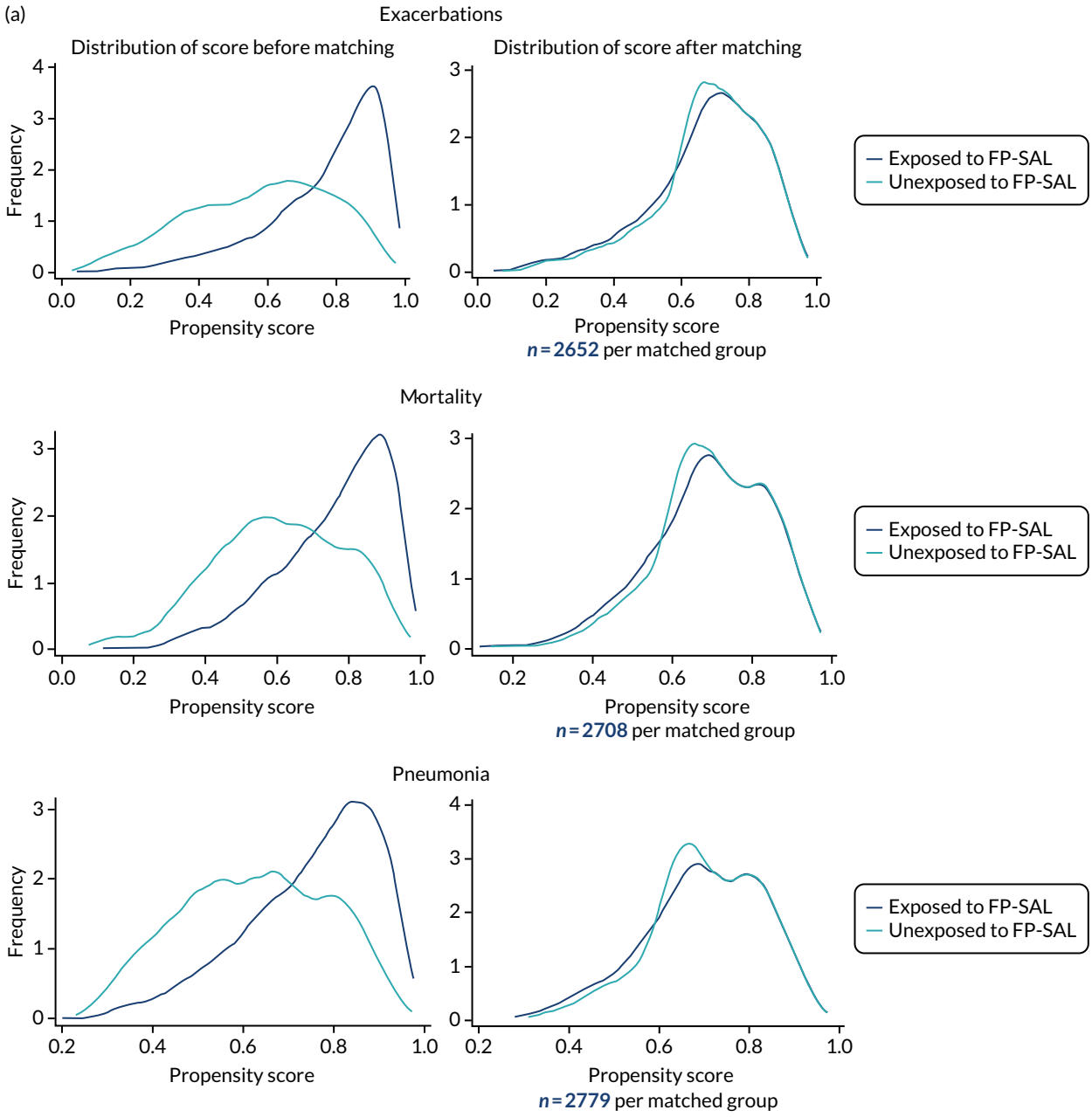


FIGURE 5 Propensity score distributions before and after matching. (a) FP-SAL exposed ($n = 10{,}926$ before matching) vs. FP-SAL unexposed ($n = 4391$ before matching); and (b) FP-SAL exposed ($n = 11{,}235$ before matching) vs. SAL exposed ($n = 1146$ before matching). Treatment discontinuation not included for these analysis as only one of the exposure groups was receiving treatment. (*continued*)

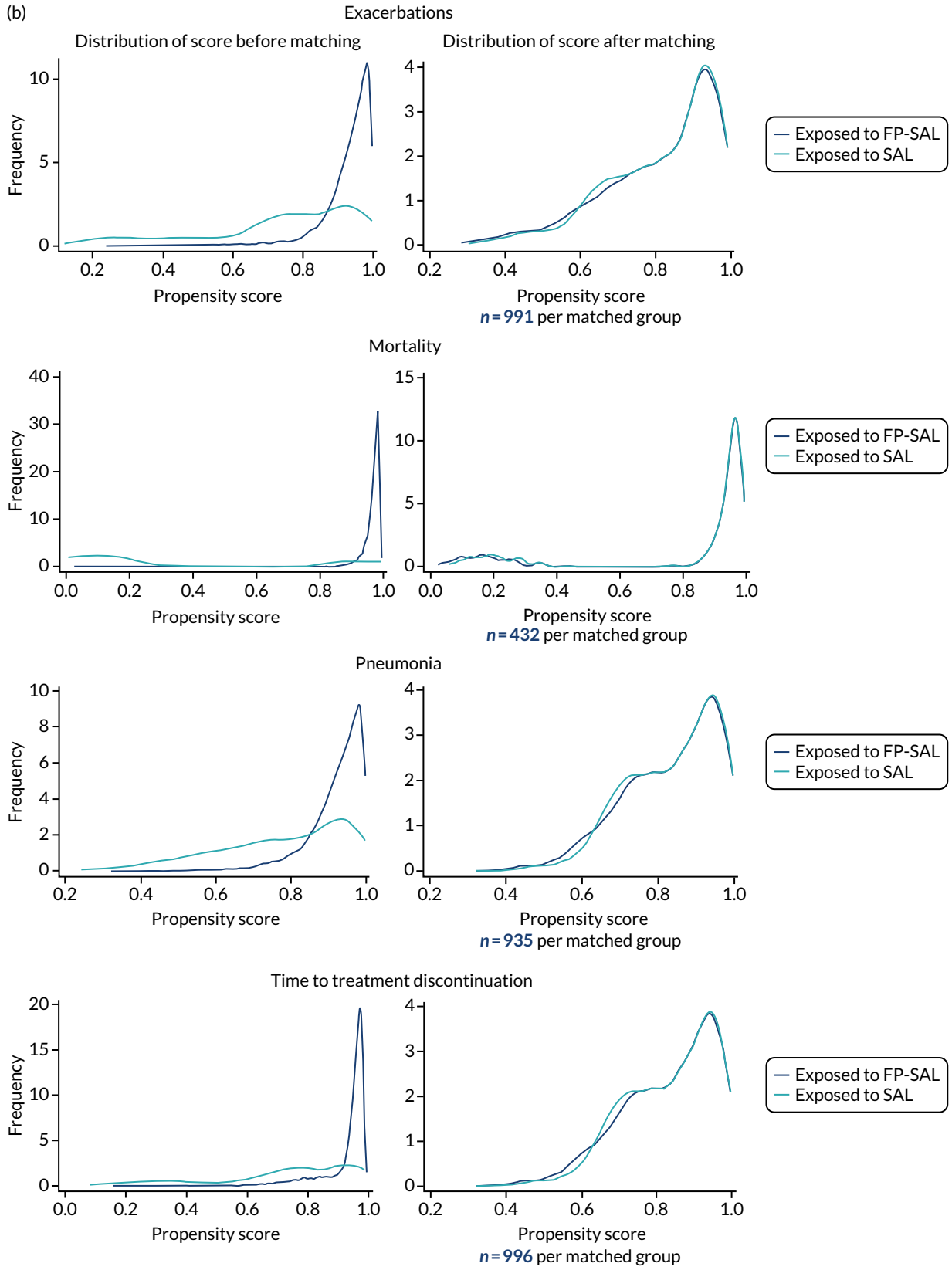


FIGURE 5 Propensity score distributions before and after matching. (a) FP-SAL exposed ($n = 10{,}926$ before matching) vs. FP-SAL unexposed ($n = 4391$ before matching); and (b) FP-SAL exposed ($n = 11{,}235$ before matching) vs. SAL exposed ($n = 1146$ before matching). Treatment discontinuation not included for these analysis as only one of the exposure groups was receiving treatment.

TABLE 7 Characteristics of the FP-SAL vs. SAL cohort before and after propensity score matching for the exacerbations analysis (CPRD non-interventional population)

Variable	Before propensity score matching			After propensity score[a] matching		
	SAL[b] (*N* = 1146 people)	FP-SAL[c] (*N* = 11,235 exposed time periods from 4523 people)	Standardised difference	SAL (*N* = 991 people)	FP-SAL (*N* = 991 people)	Standardised difference
Age (year), median (IQR)	68 (62–73)	68 (62–74)	0.051	68 (62–73)	67 (61–73)	0.038
Sex (male), *n* (%)	728 (64)	6960 (62)	0.033	628 (63)	637 (64)	0.019
Lung function[d]						
FEV_1 per cent of predicted, median (IQR)	49 (41–57)	50 (40–60)	0.272	50 (41–57)	49 (40–57)	0.107
FEV_1 : FVC per cent, median (IQR)	53 (44–61)	53 (44–62)	0.022	53 (45–62)	51 (42–60)	0.122
BMI (kg/m^2),[d] median (IQR)	26 (23–30)	26 (22–30)	0.057	26 (23–30)	26 (22–29)	0.123
Prior exacerbations,[e] mean (SD)	0.63 (1.02)	0.61 (1.07)	0.017	0.62 (1.01)	0.61 (1.03)	0.010
Cardiovascular disease, *n* (%)[f]						
Coronary heart disease	207 (18)	1958 (17)	0.017	175 (18)	129 (13)	0.129
Peripheral vascular disease	71 (6)	749 (7)	0.019	62 (6)	62 (6)	0.000
Cerebrovascular disease	87 (8)	792 (7)	0.021	81 (8)	64 (6)	0.066
Other atherosclerosis	1 (0)	21 (0)	0.027	1 (0)	1 (0)	0.026
Statin prescription, *n* (%)[g]	462 (40)	4906 (44)	0.068	411 (41)	344 (35)	0.140
Aspirin prescription, *n* (%)[g]	333 (29)	3376 (30)	0.022	297 (30)	246 (25)	0.116
Other COPD medication prescriptions, *n* (%)[g]						
LABA[h]	793 (69)	98 (1)	2.052	648 (65)	15 (2)	1.839
ICS[h]	419 (37)	862 (8)	0.742	275 (28)	387 (39)	0.241
LAMA[h]	477 (42)	6598 (59)	0.347	432 (44)	487 (49)	0.111
ICS plus LABA[i]	28 (2)	537 (5)	0.125	24 (2)	50 (5)	0.139

continued

TABLE 7 Characteristics of the FP-SAL vs. SAL cohort before and after propensity score matching for the exacerbations analysis (CPRD non-interventional population) (*continued*)

Variable	Before propensity score matching			After propensity score[a] matching		
	SAL[b] (*N* = 1146 people)	FP-SAL[c] (*N* = 11,235 exposed time periods from 4523 people)	Standardised difference	SAL (*N* = 991 people)	FP-SAL (*N* = 991 people)	Standardised difference
Type 2 diabetes, *n* (%)[f]	116 (10)	1549 (14)	0.113	101 (10)	100 (10)	0.003
History of cancer, *n* (%)[f]	200 (17)	2252 (20)	0.066	178 (18)	163 (16)	0.040
Chronic kidney disease, *n* (%)[f]	104 (9)	1535 (14)	0.145	89 (9)	85 (9)	0.014
Health-care utilisation, median (IQR)[e]						
Number of GP consultations	15 (9–23)	16 (9–26)	0.765	15 (9–23)	15 (9 –25)	0.021
Number of distinct medications	5 (3–8)	5 (3–8)	0.039	5 (3–8)	5 (3–8)	0.019
Number of hospitalisations	0 (0–1)	0 (0–1)	0.063	0 (0–1)	0 (0–1)	0.005
Number of hospital procedures	0 (0–0)	0 (0–1)	0.065	0 (0–0)	0 (0–0)	0.035

IQR, interquartile range; SD, standard deviation.
a See *Table 5* for list of variables included in final exacerbations propensity score model. Variables in this table that were included in the propensity score are in bold.
b TORCH trial[3] inclusion/exclusion criteria applied and matched to individual TORCH trial[3] patients.
c TORCH trial[3] inclusion/exclusion criteria applied.
d Closest record prior to the index date.
e All counted within the year prior to the index date.
f Any diagnosis for condition prior to the index date.
g Number of people who had at least one prescription within the previous year.
h Single product only.
i Combination product.

TABLE 8 Results for the analysis of exacerbations, mortality, pneumonia and time to treatment discontinuation for FP-SAL vs. no FP-SAL (compared with TORCH trial[3] results)

	CPRD non-interventional population		TORCH trial[3] population[a]	
Analysis	Unexposed to FP-SAL (*N* = 4196)	Exposed to FP-SAL (*N* = 10,463)	Placebo (*N* = 1524)	FP-SAL (*N* = 1533)
Exacerbations				
Person-years at risk	9330	22,054		
Events, *n*	4994	15,944		
Rate[b]	0.53	0.72	1.13	0.85
Crude rate ratio (95% CI)	1	1.35 (1.28 to 1.43)		
Propensity-matched rate ratio (95% CI)	1	1.30 (1.19 to 1.42)[c]	1	0.75 (0.69 to 0.81)
Mortality				
Person-years at risk	9330	22,054		
Events, *n*	543	1245		
Probability[d] at 3 years (%)	16.13	16.04	15.16	12.59
Crude HR (95% CI)	1	0.98 (0.88 to 1.08)		
Propensity-matched HR (95% CI)	1	1.11 (0.95 to 1.26)[e]	1	0.83 (0.68 to 1.00)
Pneumonia				
Events, *n*	350	998		
Per cent of total patients	8.34	9.54	12.31	19.60
Crude RR (95% CI)	1	1.14 (1.01 to 1.28)		
Propensity-matched RR (95% CI)	1	1.14 (0.96 to 1.34)[f]	1	1.59 (1.35 to 1.88)
Time to treatment discontinuation[g]				
Person-years at risk		20,402		
Events, *n*		2255		
Probability[d] at 3 years (%)		28.20	43.50	33.70
Crude HR				
Propensity-matched HR (95% CI)			1	0.69 (0.62 to 0.78)

a Only results reported in the TORCH trial[3] publication are shown.
b Per person per year.
c A total of 2652 people in each exposure group after propensity score matching. See *Table 5* for list of variables contributing to propensity score for exacerbations analysis.
d Probability of event at 3 years, calculated using a Cox proportional hazards model.
e A total of 2708 people in each exposure group after propensity score matching. See *Table 5* for list of variables contributing to propensity score for mortality analysis.
f A total of 2779 people in each exposure group after propensity score matching. See *Table 5* for list of variables contributing to propensity score for pneumonia analysis.
g Time to treatment discontinuation analysis not applicable to the unexposed to FP-SAL group.

95% CI 0.95 to 1.26] was in the opposite direction to the TORCH trial[3] placebo-controlled result (HR 0.83, 95% CI 0.68 to 1.00). For the pneumonia analysis, we found weak evidence for a 14% increased risk associated with FP-SAL [risk ratio (RR) 1.14, 95% CI 0.96 to 1.34], which was not consistent with the stronger harmful association found by the TORCH trial[3] placebo-controlled analysis (RR 1.59, 95% CI 1.35 to 1.88).

FP-SAL exposed compared with salmeterol exposed

For the exacerbations analysis, we obtained a propensity score-matched rate ratio of 0.85 (95% CI 0.74 to 0.97). According to our prespecified protocol, this (protective) effect was considered to be consistent with the TORCH trial[3] FP-SAL compared with SAL result for the same outcome (rate ratio 0.88, 95% CI 0.81 to 0.95) (*Table 9*).[1] Similarly, our result for the mortality outcome (HR 0.93, 95% CI 0.65 to 1.32) was consistent with the TORCH trial[3] FP-SAL compared with SAL result (HR 0.93, 95% CI 0.77 to 1.13). For the pneumonia analysis, we found evidence for a 39% increased risk associated with FP-SAL (RR 1.39, 95% CI

TABLE 9 Results for the analysis of exacerbations, mortality, pneumonia and time to treatment discontinuation for FP-SAL vs. SAL (compared with TORCH trial[3] results)

	CPRD non-interventional population		TORCH trial[3] population[a]	
Analysis	**SAL (*N* = 1146)**	**FP-SAL (*N* = 11,235)**	**SAL (*N* = 1521)**	**FP-SAL (*N* = 1533)**
Exacerbations				
Person-years at risk	2566	24,062		
Events, *n*	1515	14,034		
Rate[b]	0.73	0.59	0.97	0.85
Crude rate ratio (95% CI)	1	0.80 (0.72 to 0.88)		
Propensity-matched rate ratio (95% CI)	1	0.85 (0.74 to 0.97)[c]	1	0.88 (0.81 to 0.95)
Mortality				
Person-years at risk	2566	24,062		
Events, *n*	138	1445		
Probability[d] at 3 years (%)	15.09	16.84	13.48	12.59
Crude HR (95% CI)	1	1.12 (0.94 to 1.34)		
Propensity-matched HR (95% CI)	1	0.93 (0.65 to 1.32)[e]	1	0.93 (0.77 to 1.13)
Pneumonia				
Events, *n*	86	1137		
Per cent of total patients	7.50	10.12	13.29	19.60
Crude RR (95% CI)	1	1.35 (1.09 to 1.66)		
Propensity-matched RR (95% CI)	1	1.39 (1.04 to 1.87)[f]	1	1.47 (1.25 to 1.73)
Time to treatment discontinuation				
Person-years at risk	1251	21,587		
Events, *n*	740	2449		
Probability[d] at 3 years (%)	77.02	28.04	36.40	33.70
Crude HR (95% CI)	1	0.22 (0.20 to 0.23)		
Propensity-matched HR (95% CI)	1	0.23 (0.20 to 0.27)[g]	1	0.89 (0.79 to 0.99)

a Only results reported in the TORCH trial[3] publication are shown.
b Per person per year.
c A total of 991 people in each exposure group after propensity score matching. See *Table 5* for list of variables contributing to propensity score for exacerbations analysis.
d Probability of event at 3 years, calculated using a Cox proportional hazards model.
e A total of 443 people in each exposure group after propensity score matching. See *Table 5* for list of variables contributing to propensity score for mortality analysis.
f A total of 996 people in each exposure group after propensity score matching. See *Table 5* for list of variables contributing to propensity score for pneumonia analysis.
g A total of 935 people in each exposure group after propensity score matching. See *Table 5* for list of variables contributing to propensity score for time to treatment discontinuation analysis.

1.04 to 1.87), which was also consistent with the harmful association found by the TORCH trial[3] FP-SAL compared with SAL analysis (RR 1.47, 95% CI 1.25 to 1.73). For the time to treatment discontinuation analysis, the effect was apparently much stronger outside the trial setting (non-interventional HR 0.23, 95% CI 0.20 to 0.27 vs. TORCH trial[3] non-interventional HR 0.89, 95% CI 0.79 to 0.99).

Analysis of impact of (1) TORCH trial[3] matching and (2) TORCH trial[3] criteria (post hoc analysis)

Repeating the FP-SAL compared with SAL analysis and omitting the TORCH trial-matching[3] step led to an exacerbations rate ratio of 0.87 (95% CI 0.81 to 0.94) (*Table 10*), which is very similar to both the main analysis and the TORCH trial[3] result. By contrast, neither applying the TORCH trial[3] criteria nor matching led to a completely different effect estimate (rate ratio 1.64, 95% CI 1.52 to 1.77).

Discussion

We have demonstrated that methods applied to non-interventional data can generate results comparable to active comparator trials for COPD treatment effects. By contrast, we found that the same methods were unable to replicate placebo-controlled trial results.

Comparison with previous studies

Previous studies applying similar 'trial-replication' approaches

Although a number of papers have compared the designs of observational studies with RCTs[21–26] and some studies have generated results similar to an earlier or subsequent trial,[27–29] to our knowledge there are very few non-interventional studies that have set out to explicitly replicate a specific trial cohort and its results.

Hernán *et al.*[30] replicated the design and result of the Women's Health Initiative[31] randomised trial on the effect of oestrogen/progestin therapy on coronary heart disease risk. Smeeth *et al.*[32] analysed the effect of statins on a range of health outcomes and replicated the Heart Protection Study[33] randomised trial. Fralick *et al.*[34] applied trial criteria and utilised propensity score matching to replicate cardiovascular results from ONTARGET[35] (ONgoing Telmisartan Alone and in combination with Ramipril Global Endpoint Trial).

Previous studies of chronic obstructive pulmonary disease drug treatment effects

Results of five (LABA/ICS vs. LABA) interventional studies (including the TORCH trial[3]) were summarised in a Cochrane review (rate ratio 0.76, 95% CI 0.68 to 0.84).[36] Three out of these five studies estimated effect sizes that were considerably greater than effect sizes reported in the TORCH trial.[3] As this study mirrored the TORCH trial,[3] our results aligned most closely to those of the TORCH trial.[3]

TABLE 10 Impact of choice of selection methods on ability to replicate trial results for the analysis of exacerbations in people exposed to FP-SAL vs. people exposed to SAL

	Rate ratio		
	SAL	FP-SAL (95% CI)	*n* per exposure group
TORCH trial[3]	1	0.88 (0.81 to 0.95)	1524
CPRD non-interventional selection method[a]			
TORCH trial[3] including/excluding criteria and matched to [b]TORCH trial[3]	1	0.85 (0.74 to 0.97)	991
TORCH trial[3] including/excluding criteria only	1	0.87 (0.81 to 0.94)	3225
No TORCH trial[3] criteria or matching	1	1.64 (1.52 to 1.77)	5951

a SAL and FP-SAL groups propensity score matched for all selection methods.
b As per the main analysis and presented in *Table 3*.

A number of studies have found strong survival benefits of ICS therapy after hospital discharge.[36–38] After accounting for likely time-related biases that have an impact on these studies, a null effect was obtained (rate ratio 0.94, 95% CI 0.81 to 1.09).[39] The methodology we applied obtained a mortality effect estimate comparable to the analysis designed to account for time-related biases (0.93, 95% CI 0.65 to 1.32).

In line with the TORCH trial,[3] previous studies have found an increased risk of pneumonia associated with ICS-containing treatments for COPD.[36,40,41] Our result (RR 1.39, 95% CI 1.04 to 1.87) was consistent with results of a meta-analysis of trials comparing LABA/ICS with LABA formulations (OR 1.55, 95% CI 1.20 to 2.01)[36] and very similar to a recent non-interventional study comparing LABA/ICS with LAMA formulations (HR 1.37, 95% CI 1.17 to 1.60).[42]

Our 3-year probability of treatment discontinuation for FP-SAL (28%) is comparable to non-adherence figures from previous non-interventional real-world data studies (49% and 43%).[43,44] The probability of discontinuation of SAL that we observed (77%) was higher than these two previous non-interventional studies, leading to the discrepancy with the TORCH trial.[3] We hypothesised that during our study period a large proportion of the patients who would have been initially prescribed SAL would have been likely to switch to FP-SAL because of prescribing decisions in primary care. A post hoc analysis found that 43% of people prescribed SAL switched to FP-SAL during follow-up (vs. only 2% switching from FP-SAL to SAL).

Implications and further work

When studying COPD treatment effects, if (1) the analysis is of active comparators, (2) trial exclusion and inclusion criteria are applied and (3) the propensity score models that we developed for each outcome are applied to balance exposure groups, then the results of studies carried out in routinely collected non-interventional data can be considered robust in the sense that they will be highly comparable to trial results. This now provides a methodological framework for being able to analyse COPD drug treatment effects in real-world data, focusing on groups that were either not included or under-represented in trials.[1]

Our inability to replicate placebo-controlled analyses suggests uncontrolled confounding by indication, a well-known bias in pharmacoepidemiology that is highly likely to be present when performing a comparison between people prescribed a drug and people not prescribed a drug.[45–48] An established design approach for addressing this bias is to perform an active comparator analysis (i.e. comparing the effects of one medicine with another, rather than one medicine with no treatment).[45–48] Based on the likelihood of confounding by indication having an impact on our results compared with the results of no treatment, we proceeded to perform the active comparator analysis and obtained results very similar to the trial, which indicates that confounding by indication is highly likely to be the reason for being unable to replicate placebo-controlled analyses in this setting.

One possibility for how this confounding by indication may be manifesting relates to an aspect of our study design that allowed people to be included in both the exposed and unexposed cohorts (i.e. the result we obtained could be strongly influenced by people initially in the unexposed group who are relatively healthy but then get more sick over time and require FP-SAL treatment and end up in the exposed group). However, in a post hoc analysis in which we dropped the 730 people (out of a total of 2652 per group) who appeared in both cohorts, our effect estimate was nearly identical (RR 1.33, 95% CI 1.20 to 1.47). We do consider, however, that because COPD treatment is based on a step-up approach, it is highly likely that patients not exposed to FP-SAL in routine primary care are generally likely to be those with milder COPD.

An additional point that further explains our inability to replicate the placebo-controlled analysis relates to the large difference in incidence rate between the TORCH trial[3] placebo group (1.13 exacerbations/person/year) and our FP-SAL-unexposed group (0.53 exacerbations/person/year). To investigate

underlying reasons for this discrepancy, we performed a post hoc analysis in which we compared the characteristics of the 1753 people from TORCH trial[3] who were not able to be matched to our unexposed-to-FP-SAL population in step 3 with those who were successfully matched. We found that those not matched were younger (mean age 60.7 vs. 65.8 years), more sick (history of cardiovascular disease: 93% vs. 46%), had worse lung function (FEV_1 34.9 vs. 45.9) and included a higher proportion of people recruited from Eastern European trial sites (27% vs. 17%). People with these characteristics may have been highly suitable for recruitment to the TORCH clinical trial,[3] but are very difficult to find in UK primary care, providing another reason why we were not able to replicate placebo-controlled analyses. Furthermore, both TORCH trial[3] placebo-assigned patients and patients in our own cohort were permitted to use other COPD treatments during follow-up, but given the time and setting differences between our FP-SAL-untreated group and the TORCH trial[3] placebo group, people in our placebo group are much more likely to have been prescribed, for example, an ICS during follow-up. More generally, it is also likely to be challenging to obtain comparable absolute rates in emulated cohorts within a single country based on historical international trials for reasons such as this.

Previous authors have gone as far to specifically recommend that when trying to emulate trial results it is important to choose an active comparator trial.[34] There are, however, examples in which placebo-controlled analyses have been successfully replicated.[30,32] One possibility is that replication of placebo-controlled results works better when the drug studied is (1) preventative and (2) used in a generally healthy cohort (e.g. the cited studies were of statins and of postmenopausal hormone therapy both prescribed, in some instances, to people without a specific underlying chronic disease, in contrast to the patients with COPD who received therapy in our study). We consider that further avenues of research could be followed to understand if there remains a possibility of replicating placebo-controlled studies within a non-interventional setting for COPD therapies. These could include the application of high-dimensional propensity scores or the use of instrumental variables. Our work also suggests that treatment discontinuation in the setting of non-interventional data may be driven by very different factors to those seen in trials and, at least in the setting of COPD, may not be a useful outcome to study. Interpretation of treatment discontinuation in routine data is challenging. For example, it is difficult to establish from routinely collected data whether a patient has truly stopped taking their medicine or is just taking the medicine differently than prescribed (e.g. is taking less than has been prescribed over a longer period), a point that was emphasised by our patient and public involvement representative.

Finally, in our post hoc analysis, we found that the application of the trial-matching step did not confer any advantage over the application of trial criteria alone in this setting. This suggests that treatment–covariate interactions are not as critical as we initially thought in this therapeutic area.

Limitations

Some of the TORCH trial[3] inclusion criteria were not fully assessable using CPRD data, meaning that the inclusion/exclusion criteria are analogous with TORCH trial[3] criteria, but we acknowledge that they are not identical. For example, TORCH trial[3] inclusion criteria included negative reversibility spirometry criteria, but it was not possible to replicate this in the CPRD data. In addition, at entry to the TORCH trial[3] 2-week run-in period all ICSs and inhaled long-acting bronchodilators were discontinued, but it was not possible to ensure that all patients selected from the CPRD had discontinued these at the start of the run-in period that we applied.

We originally planned to apply frequency of COPD therapy prescriptions in the previous year as a matching character/criterion. In practice, this was not feasible. However, it appears that matching at this level of detail was not required to be able to replicate trial results for active comparator analysis. Furthermore, it is clear from *Figures 3* and *4* that we had to drop around one-fifth of patients as they did not have spirometry measurements recorded in the CPRD. This does mean that our initial study population may be missing key groups of people who tend not to have their spirometry measured (e.g. people with COPD who have the least contact with the health services). This is a problem for any COPD research performed using routinely collected primary care data, and although we acknowledge the issue we think that in this work it is likely to

have minimal impact, as our aim was to create trial-analogous exposure groups that were highly comparable to each other, not to select a highly generalisable sample.

Finally, within the TORCH trial,[3] the dose of the fixed combination product FP-SAL was specified as 500 µg of FP and 50 µg of SAL (500/50) and the dose of SAL alone as 50 µg, whereas in our study we did not limit to a specific dose. The reason for this is that dosage information is incompletely captured in the CPRD. These are the only approved doses of FP-SAL and of SAL for COPD in the UK and so we consider the doses that people were prescribed in our study would have been generally similar to that administered in the TORCH trial,[3] although we do acknowledge that there are likely to be some differences in dosing between the TORCH trial[3] and our cohort because of the long-term management of patients with varying disease severity and varying concomitant conditions.

Conclusions

By replicating the COPD TORCH trial[3] selection procedures and inclusion/exclusion criteria in real-world data and developing propensity score models to account for any remaining differences between groups, we were able to obtain highly comparable relative effect estimates to the TORCH trial[3] active comparator analysis for exacerbations, mortality and pneumonia.

Replication of placebo-controlled analyses was not possible. This is a not entirely unexpected result because of the well-established fact that when comparing outcomes in people on treatment with people not on treatment using non-interventional data, confounding relating to the difference in the underlying health status of the two groups is usually too severe to address. We also found that trial placebo groups in this therapeutic area may be much less healthy than people attending UK primary care, having a further impact on the ability to replicate no-treatment comparisons. Further work to investigate whether or not confounding by indication for no treatment comparisons can ever be accounted for in this therapeutic area is warranted.

Performing active comparator analyses is a well-established design approach for minimising confounding by indication. In addition, obtaining such similar results to the TORCH trial[3] when comparing active comparator results provides confidence that our methods will provide valid results when used in other active comparator analyses within this therapeutic area. Application of the same selection procedures and propensity score models developed here to active comparator analyses of COPD drug treatment effects in groups under-represented or excluded from trials provides a practical way for key evidence gaps to be filled in relation to whether or not one treatment is more effective than another.

Results and discussion part 2: analysis of chronic obstructive pulmonary disease treatment effects in (1) patients excluded from trials and (2) patients with milder chronic obstructive pulmonary disease (objectives 2 and 3)

Note on results presented in results and discussion part 2

As presented and discussed in part 1 of this chapter (covering the primary analysis of objective 1), only active comparator analyses (FP-SAL vs. SAL) obtained valid results for the analysis of treatment effects (i.e. comparable to the TORCH trial[3]) using our methods. All of the primary analyses for objectives 2 and 3 (which are presented and discussed in this part of the results and discussion chapter) were, therefore, also performed using the FP-SAL compared with SAL comparison. Of the original analyses planned in the *Chapter 2*, *Objective 1: validation of methods for measuring chronic obstructive pulmonary disease medication effectiveness in electronic health record data by comparing with trial results* and *Objective 2: measurement of chronic obstructive pulmonary disease treatment effects in patients excluded from trials*, there were insufficient numbers of people exposed to SAL for performing analyses of cohorts that were excluded from the TORCH trial[3] because of a history of lung surgery ($n < 60$ prior to propensity score matching),

history of long-term oxygen therapy ($n < 29$), substantial comorbidity (i.e. a serious uncontrolled disease with high likelihood of causing death within 3 years) ($n < 60$) or evidence of drug/alcohol abuse ($n < 30$). Results are, therefore, presented for the analysis of the effect of FP-SAL compared with SAL on exacerbations, mortality and pneumonia for the following groups that would have been excluded from the TORCH trial:[3]

- people aged > 80 years
- people with an asthma diagnosis within the 5 years prior to inclusion
- people with milder COPD based on spirometry measurements (i.e. those who did not meet either or both of the TORCH[3] requirements of a FEV_1 of < 60% predicted or a FEV_1/FVC ratio of < 70%).

Results

Participants

After applying the TORCH trial[3] exclusion criteria, as outlined in part 1, to the 91,733 people from HES-linked CPRD practices with a diagnosis for COPD between 1 January 2000 and 1 January 2017 but with exclusion criteria altered relating to (1) aged > 80 years, (2) presence of asthma or (3) selection of people with milder COPD based on spirometry, the final cohort sizes before and after propensity score matching were as shown in *Table 11*.

TABLE 11 Number of people included in the analysis of CPRD COPD populations who would not have been eligible for the TORCH trial[3]

Population for analysis	SAL	FP-SAL
People aged > 80 years		
n per group before propensity score matching (*n* eligible time periods)[a]	194 (763)	670 (1645)
n per group after propensity score matching[b]		
Exacerbations	84	84
Mortality	94	94
Pneumonia	92	92
People with an asthma diagnosis within 5 years prior to inclusion		
n per group before propensity score matching (*n* eligible time periods)[a]	1175 (5585)	3577 (9719)
n per group after propensity score matching		
Exacerbations	544	544
Mortality	910	910
Pneumonia	573	573
People with milder COPD		
n per group before propensity score matching (*n* eligible time periods)[a]	877 (3876)	2013 (4783)
n per group after propensity score matching		
Exacerbations	434	434
Mortality	1362	1362
Pneumonia	425	425

a TORCH trial[3] exclusions applied as detailed in *Figure 4* (with the exception of the population characteristic under study in the cohort, i.e. aged > 80, previous asthma or milder COPD).
b Details of the variables included in the final propensity score models for each outcome in each cohort are provided in *Table 12*.

Propensity score matching of cohorts of people who would have been excluded from the TORCH trial[3]

Details of the variables included in the final propensity score models are provided in *Table 12*. As described in *Chapter 2*, *Propensity score for addressing confounding*, variables from the pool of initial variables that were not associated with the specific outcome under study were not included in the multivariable logistic regression model used to generate the propensity score. This, therefore, meant that the final propensity score models for each of the outcomes under study could contain a different set of variables, depending on the outcome being studied. For example, for people with milder COPD,

TABLE 12 Variables included in the final propensity score models for the cohorts of people who would have been excluded from the TORCH trial[3]

Outcome being analysed	Variables included in propensity score model	Matching
Aged > 80 years		
Exacerbations	Alcohol intake, year of index date, previous diagnosis of chronic kidney disease and having at least one prescription for (1) LABA or (2) ICS	1 : 1 nearest neighbour, calliper of 0.03
Mortality	Age, previous diagnosis of (1) chronic kidney disease or (2) cancer, year of index date, having at least one prescription for a LABA, and frequency of consultations and hospitalisations in the previous year	1 : 1 nearest neighbour, calliper of 0.03
Pneumonia	Age; FEV_1/FVC; year of index date; previous diagnosis of chronic kidney disease; having at least one prescription of a LABA; and frequency of prescriptions, consultations and hospitalisations in the previous year	1 : 1 nearest neighbour, calliper of 0.03
People with a prior asthma diagnosis		
Exacerbations	Sex; age; year of index date; FEV_1; FEV_1/FVC; BMI; previous diagnosis for coronary heart disease; having at least one prescription for (1) LABA, (2) ICS, (3) LAMA or (4) LABA – ICS combination therapy in the previous year; and frequency of consultations, prescriptions, hospitalisations and exacerbations in the previous year	1 : 1 nearest neighbour, calliper of 0.03
Mortality	Sex; age; index year; FEV_1/FVC; BMI; alcohol intake; previous diagnosis for (1) coronary heart disease, (2) peripheral vascular disease, (3) cancer or (4) chronic kidney disease; having at least one prescription for (1) statin, (2) aspirin, (3) ICS or (4) LAMA in the previous year; and frequency of consultations, hospitalisations, procedures, exacerbations in the previous year	1 : 1 nearest neighbour, calliper of 0.03
Pneumonia	Age; FEV_1; index year; previous diagnosis for (1) coronary heart disease or (2) peripheral vascular disease; having at least one prescription for (1) LABA, (2) LAMA or (3) aspirin; and frequency of consultations, hospitalisations and exacerbations in the previous year	1 : 1 nearest neighbour, calliper of 0.03
People with mild COPD based on lung function		
Exacerbations	Previous diagnosis for (1) coronary heart disease, (2) peripheral vascular disease or (3) cerebrovascular disease; having at least one prescription for (1) aspirin, (2) LABA or (3) ICS; and frequency of consultations and exacerbations in the previous year	1 : 1 nearest neighbour, calliper of 0.03
Mortality	Sex; age; BMI; previous diagnosis for (1) coronary heart disease, (2) peripheral vascular disease, (3) cerebrovascular disease or (4) cancer; having at least one prescription for (1) statin, (2) aspirin or (3) ICS; and the frequency of prescriptions, hospitalisations, procedures and exacerbations in the previous year	1 : 1 nearest neighbour, calliper of 0.03
Pneumonia	Index year; previous diagnosis for (1) coronary heart disease, (2) peripheral vascular disease or (3) cardiovascular disease; having at least one prescription for (1) aspirin or (2) LABA in the previous year; and the frequency of consultations, hospitalisations, procedures and exacerbations in the previous year	1 : 1 nearest neighbour, calliper of 0.03

SES, socioeconomic status.

the mortality propensity score included sex and age, but these variables did not end up in the final propensity score for exacerbations or pneumonia. Although sex and age were associated with each of these outcomes in crude analysis, exacerbations and pneumonia were dropped from the subsequent multivariable logistic regression models that had drug exposure as an outcome (specified in *Chapter 2*, *Propensity score for addressing confounding*) because of the lack of association with exposure after multivariable adjustments.

People aged over 80 years

For the analysis of COPD exacerbations, prior to propensity score matching, differences by exposure status were noted for sex, prior exacerbations, coronary heart disease, peripheral vascular disease, cerebrovascular disease, prescriptions for COPD medications, chronic kidney disease and all measures of health-care utilisation (*Table 13*). After propensity score matching, many differences persisted or became apparent. Variables with substantial imbalance were age, sex, BMI, lung function, coronary heart disease, peripheral vascular disease, statin use, prior LABA or LAMA use, type 2 diabetes, chronic kidney disease and most measures of health-care utilisation.

Tables 17 and *18* in *Appendix 1* show the balance of baseline variables for the analyses of mortality and pneumonia outcomes, respectively. For both outcomes, the pattern of imbalances before and after propensity score matching was very similar to that seen for COPD exacerbations.

People with an asthma diagnosis within 5 years prior to study entry

For the analysis of COPD exacerbations, prior to propensity score matching, differences by exposure status were noted for age, BMI, prescriptions for COPD medications, number of GP consultations and number of distinct medications prescribed (*Table 14*). After propensity score matching, many differences were minimised, although imbalances are notable for peripheral vascular disease, cerebrovascular disease, prior use of ICS and type 2 diabetes.

Tables 19 and *20* in *Appendix 1* show the balance of baseline variables for the analyses of mortality and pneumonia outcomes, respectively. After propensity score matching, imbalances were similarly reduced for both outcomes. Minor differences from the exacerbations analysis were noted. For the mortality analysis, post-matching imbalances were noted for FEV_1 and prior use of LABA or ICS plus LABA (see *Appendix 1*, *Table 19*). For pneumonia, post-matching imbalances were seen for LABA rather than ICS use, BMI, ICS plus LABA and number of hospital procedures (see *Appendix 1*, *Table 20*).

People with mild chronic obstructive pulmonary disease (based on spirometry)

For the analysis of COPD exacerbations, prior to propensity score matching, differences by exposure status were noted for FEV_1 : FVC, prior exacerbations, coronary heart disease, peripheral vascular disease, cerebrovascular disease, statin use, use of other COPD medications and most measures of health-care utilisation (*Table 15*). After propensity score matching, most differences were minimised, although small differences can be seen for number of distinct medications and number of hospitalisations.

Tables 21 and *22* in *Appendix 1* show the balance of baseline variables for the analyses of mortality and pneumonia outcomes, respectively. Similar imbalances to those seen for COPD exacerbations were seen before propensity score matching. After propensity score matching, for the mortality outcome, imbalances remained for FEV_1 : FVC, prior use of LABA or ICS, history of cancer, number of GP consultations and number of distinct medications (see *Appendix 1*, *Table 21*). For the pneumonia outcome, post-matching imbalances can be seen for FEV_1 : FVC, coronary heart disease and prior use of ICS (see *Appendix 1*, *Table 22*).

TABLE 13 Characteristics of the aged > 80 years cohort before and after propensity score matching for the exacerbations analysis (FP-SAL vs. SAL) (CPRD non-interventional population)

Variable	Before propensity score matching			After propensity score matching		
	Exposed to SAL (*N* = 763 time periods from 194 people)	Exposed to FP-SAL (*N* = 1645 exposed time periods from 670 people)	Standardised difference	Exposed to SAL (*N* = 84 people)	Exposed to FP-SAL (*N* = 84 people)	Standardised difference
Age (years), median (IQR)	85 (83–88)	85 (83–88)	0.038	85 (82–87)	83 (82–85)	0.365
Sex (male), *n* (%)	517 (68)	976 (59)	0.176	57 (68)	50 (60)	0.174
Lung function, median (IQR)						
FEV_1 per cent of predicted	52 (46–59)	52 (41–59)	0.085	53 (43–59)	46 (37–56)	0.185
FEV_1 : FVC per cent	55 (47–63)	55 (45–64)	0.092	52 (47–61)	49 (41–59)	0.326
BMI (kg/m^2), median (IQR)	25 (22–28)	25 (22–28)	0.016	26 (22–29)	25 (21–28)	0.136
Prior exacerbations, mean (SD)	0.60 (0.97)	0.73 (1.16)	0.124	0.77 (1.17)	0.61 (0.84)	0.164
Cardiovascular disease, *n* (%)						
Coronary heart disease	55 (7)	391 (24)	0.47	12 (14)	20 (24)	0.244
Peripheral vascular disease	21 (3)	158 (10)	0.288	3 (4)	9 (11)	0.28
Cerebrovascular disease	18 (2)	184 (11)	0.357	5 (6)	4 (5)	0.053
Other atherosclerosis	0 (0)	3 (0)	0.06	0 (0)	0 (0)	0 (0)
Statin prescription, *n* (%)	328 (43)	694 (42)	0.016	33 (39)	39 (46)	0.145
Aspirin prescription, *n* (%)	308 (40)	700 (43)	0.044	36 (43)	33 (39)	0.073
Other COPD medication prescriptions, *n* (%)						
LABA	675 (88)	14 (1)	3.728	8 (10)	11 (13)	0.113
ICS	297 (39)	108 (7)	0.837	34 (40)	36 (43)	0.048
LAMA	244 (32)	901 (55)	0.473	35 (42)	44 (52)	0.216
ICS plus LABA	19 (2)	71 (4)	0.101	7 (8)	5 (6)	0.093

Variable	Before propensity score matching			After propensity score matching		
	Exposed to SAL (*N* = 763 time periods from 194 people)	Exposed to FP-SAL (*N* = 1645 exposed time periods from 670 people)	Standardised difference	Exposed to SAL (*N* = 84 people)	Exposed to FP-SAL (*N* = 84 people)	Standardised difference
Type 2 diabetes, *n* (%)	103 (13)	195 (12)	0.049	10 (12)	16 (19)	0.198
History of cancer, *n* (%)	282 (37)	545 (33)	0.08	28 (33)	26 (31)	0.051
Chronic kidney disease, *n* (%)	223 (29)	675 (41)	0.249	24 (29)	16 (19)	0.225
Health-care utilisation, median (IQR)						
Number of GP consultations	49 (35–72)	19 (12–31)	1.161	52 (38–86)	19 (13–30)	1.453
Number of distinct medications	6 (3–9)	6 (3–10)	0.108	7 (4–9)	7 (5–10)	0.158
Number of hospitalisations	0 (0–1)	0 (0–1)	0.106	0 (0–1)	0 (0–1)	0.021
Number of hospital procedures	0 (0–1)	0 (0–2)	0.143	0 (0–1)	0 (0–2)	0.132

IQR, interquartile range; SD, standard deviation.

TABLE 14 Characteristics of the cohort with an asthma diagnosis (during the previous 5 years) before and after propensity score matching for the exacerbations analysis (FP-SAL vs. SAL) (CPRD non-interventional population)

	Before propensity score matching			After propensity score matching		
Variable	Exposed to SAL (*N* = 5585 exposed time periods from 1175 people)	Exposed to FP-SAL (*N* = 9719 exposed time periods from 3577 people)	Standardised difference	Exposed to SAL (*N* = 544 people)	Exposed to FP-SAL (*N* = 544 people)	Standardised difference
Age (years), median (IQR)	68 (61–73)	66 (59–72)	0.192	68 (62–73)	68 (61–74)	0.011
Sex (male), *n* (%)	2923 (52)	5409 (56)	0.067	276 (50)	282 (51)	0.022
Lung function, median (IQR)						
FEV_1 per cent of predicted	53 (42–62)	52 (41–61)	0.031	51 (42–60)	51 (41–60)	0.026
FEV_1 : FVC per cent	55 (45–63)	55 (46–63)	0.022	53 (44–62)	53 (44–63)	0.029
BMI (kg/m^2), median (IQR)	27 (24–31)	27 (24–32)	0.106	27 (23–31)	26 (23–30)	0.067
Prior exacerbations, mean (SD)	0.62 (1.02)	0.48 (1.01)	0.14	0.65 (1.02)	0.75 (1.19)	0.083
Cardiovascular disease, *n* (%)						
Coronary heart disease	190 (3)	268 (3)	0.037	22 (4)	20 (4)	0.019
Peripheral vascular disease	60 (1)	60 (1)	0.05	16 (3)	5 (1)	0.146
Cerebrovascular disease	73 (1)	163 (2)	0.031	13 (2)	4 (1)	0.132
Other atherosclerosis	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Statin prescription, *n* (%)	2141 (38)	3742 (39)	0.003	189 (34)	186 (34)	0.011
Aspirin prescription, *n* (%)	1281 (23)	2401 (25)	0.042	134 (24)	131 (24)	0.013
Other COPD medication prescriptions, *n* (%)						
LABA	5123 (92)	172 (2)	4.166	131 (24)	133 (24)	0.008
ICS	4177 (75)	1071 (11)	1.685	287 (52)	357 (64)	0.258
LAMA	1630 (29)	5193 (53)	0.508	173 (31)	188 (34)	0.058
ICS plus LABA	145 (3)	507 (5)	0.136	38 (7)	39 (7)	0.007

Variable	Before propensity score matching			After propensity score matching		
	Exposed to SAL (*N* = 5585 exposed time periods from 1175 people)	Exposed to FP-SAL (*N* = 9719 exposed time periods from 3577 people)	Standardised difference	Exposed to SAL (*N* = 544 people)	Exposed to FP-SAL (*N* = 544 people)	Standardised difference
Type 2 diabetes, *n* (%)	718 (13)	1429 (15)	0.054	70 (13)	51 (9)	0.110
History of cancer, *n* (%)	980 (18)	1629 (17)	0.021	98 (18)	88 (16)	0.048
Chronic kidney disease, *n* (%)	579 (10)	1142 (12)	0.044	49 (9)	56 (10)	0.043
Health-care utilisation, median (IQR)						
Number of GP consultations	13 (8–21)	16 (9–26)	0.259	15 (9–23)	15 (9–22)	0.040
Number of distinct medications	4 (2–7)	5 (3–8)	0.201	5 (3–8)	5 (3–7)	0.027
Number of hospitalisations	0 (0–1)	0 (0–1)	0.069	0 (0–1)	0 (0–1)	0.059
Number of hospital procedures	0 (0–0)	0 (0–1)	0.093	0 (0–0)	0 (0–0)	0.075

IQR, interquartile range; SD, standard deviation.

TABLE 15 Characteristics of the mild COPD cohort before and after propensity score matching for the exacerbations analysis (FP-SAL vs. SAL) (CPRD non-interventional population)

Variable	Before propensity score matching			After propensity score matching		
	Exposed to SAL (*N* = 3876 exposed time periods from 877 people)	Exposed to FP-SAL (*N* = 4783 exposed time periods from 2013 people)	Standardised difference	Exposed to SAL (*N* = 417 people)	Exposed to FP-SAL (*N* = 417 people)	Standardised difference
Age (years), median (IQR)	66 (60–72)	66 (59–73)	0.003	66 (59–72)	66 (59–72)	0.04
Sex (male), *n* (%)	2104 (54)	2483 (52)	0.047	222 (53)	213 (51)	0.043
Lung function, median (IQR)						
FEV_1 per cent of predicted	76 (69–87)	76 (67–87)	0.041	74 (67–84)	74 (66–85)	0.043
FEV_1 : FVC per cent	68 (61–75)	70 (63–77)	0.143	67 (61–75)	71 (62–78)	0.099
BMI (kg/m^2), median (IQR)	28 (25–32)	28 (25–32)	0.055	28 (24–32)	28 (25–32)	0.051
Prior exacerbations, mean (SD)	0.50 (0.87)	0.29 (0.72)	0.263	0.59 (0.98)	0.53 (0.85)	0.055
Cardiovascular disease, *n* (%)						
Coronary heart disease	173 (4)	925 (19)	0.472	52 (12)	43 (10)	0.068
Peripheral vascular disease	52 (1)	256 (5)	0.224	13 (3)	10 (2)	0.044
Cerebrovascular disease	52 (1)	271 (6)	0.237	16 (4)	11 (3)	0.068
Other atherosclerosis	0 (0)	1 (0)	0.02	0 (0)	0 (0)	0.069
Statin prescription, *n* (%)	1589 (41)	2202 (46)	0.102	169 (41)	151 (36)	0.089
Aspirin prescription, *n* (%)	1172 (30)	1391 (29)	0.025	119 (29)	112 (27)	0.038
Other COPD medication prescriptions, *n* (%)						
LABA	3455 (89)	36 (1)	3.871	31 (7)	33 (8)	0.018
ICS	1177 (30)	544 (11)	0.481	167 (40)	179 (43)	0.058
LAMA	965 (25)	1795 (38)	0.275	138 (33)	144 (35)	0.03
ICS plus LABA	37 (1)	152 (3)	0.157	15 (4)	18 (4)	0.037

Variable	Before propensity score matching			After propensity score matching		
	Exposed to SAL (*N* = 3876 exposed time periods from 877 people)	Exposed to FP-SAL (*N* = 4783 exposed time periods from 2013 people)	Standardised difference	Exposed to SAL (*N* = 417 people)	Exposed to FP-SAL (*N* = 417 people)	Standardised difference
Type 2 diabetes, *n* (%)	401 (10)	598 (13)	0.068	46 (11)	41 (10)	0.039
History of cancer, *n* (%)	634 (16)	768 (16)	0.008	69 (17)	60 (14)	0.06
Chronic kidney disease, *n* (%)	426 (11)	602 (13)	0.049	48 (12)	39 (9)	0.071
Health-care utilisation, median (IQR)						
Number of GP consultations	14 (9–21)	16 (10–25)	0.230	15 (10–24)	16 (10–25)	0.059
Number of distinct medications	4 (2–7)	5 (3–8)	0.21	6 (3–8)	6 (3–8)	0.106
Number of hospitalisations	0 (0–1)	0 (0–1)	0.109	0 (0–1)	0 (0–1)	0.125
Number of hospital procedures	0 (0–0)	0 (0–1)	0.093	0 (0–0)	0 (0–0)	0.041

IQR, interquartile range; SD, standard deviation.

Main results

Table 16 shows the results for the association between FP-SAL and COPD exacerbations, mortality and pneumonia in the patient populations aged > 80 years, those with asthma and those with mild COPD. Crude associations and associations from propensity score-matched and propensity score-adjusted analyses are shown.

People aged over 80 years

For the exacerbations analysis, we obtained a propensity score-matched rate ratio of 0.59 (95% CI 0.36 to 0.95) and a propensity score-adjusted rate ratio of 0.83 (95% CI 0.60 to 1.14), which is consistent with the association measured in the TORCH trial-analogous[3] CPRD population (0.85, 95% CI 0.74 to 0.97). For the mortality outcome, we obtained a propensity score-matched HR of 0.99 (95% CI 0.56 to 1.74) and a propensity score-adjusted HR of 1.29 (95% CI 0.84 to 2.00), which, again, is consistent with the

TABLE 16 Results for the analysis of exacerbations, mortality and pneumonia comparing FP-SAL with SAL in people in the CPRD who would not have been eligible for the TORCH trial[3] because of (1) being aged > 80 years, (2) having a diagnosis of asthma within 5 years prior to their start of follow-up or (3) having mild COPD based on lung function (compared with results for the CPRD TORCH trial-analogous[3] cohort)

	Cohorts of people with COPD in CPRD			
Analysis	**Over 80 years old[a]**	**With asthma[b]**	**With mild COPD[c]**	**TORCH trial[3] analogous**
Exacerbations				
Crude rate ratio (95% CI)	1.19 (1.04 to 1.35)	0.71 (0.67 to 0.75)	0.47 (0.44 to 0.51)	0.80 (0.72 to 0.88)
Propensity score[d]-matched rate ratio (95% CI)	0.59 (0.36 to 0.95)[e]	0.74 (0.62 to 0.89)[f]	0.56 (0.46 to 0.70)[g]	0.85 (0.74 to 0.97)
Propensity score-adjusted rate ratio (95% CI)	0.83 (0.60 to 1.14)	0.67 (0.59 to 0.78)	0.52 (0.45 to 0.61)	NA
Mortality				
Crude HR (95% CI)	0.89 (0.76 to 1.03)	1.07 (0.96 to 1.19)	1.32 (1.06 to 1.64)	1.12 (0.94 to 1.34)
Propensity score-matched HR (95% CI)	0.99 (0.56 to 1.74)[h]	1.49 (1.21 to 1.85)[i]	0.98 (0.67 to 1.45)[j]	0.93 (0.65 to 1.32)
Propensity score-adjusted HR (95% CI)	1.29 (0.84 to 2.00)	1.20 (1.04 to 1.40)	0.84 (0.66 to 1.08)	NA
Pneumonia				
Crude RR (95% CI)	0.97 (0.82 to 1.15)	1.18 (1.05 to 1.32)	1.62 (1.35 to 1.94)	1.35 (1.09 to 1.66)
Propensity score-matched RR (95% CI)	0.82 (0.44 to 1.53)[k]	1.09 (0.74 to 1.63)[l]	0.78 (0.45 to 1.35)[m]	1.39 (1.04 to 1.87)
Propensity score-adjusted RR (95% CI)	0.88 (0.54 to 1.42)	1.04 (0.79 to 1.37)	1.08 (0.74 to 1.57)	NA

NA, not applicable.
a SAL group, 763 time periods from 194 people; FP-SAL group, 1645 time periods from 670 people.
b SAL group, 5585 time periods from 1175 people; FP-SAL group, 9719 time periods from 3577 people.
c SAL group, 3876 time periods from 877 people; FP-SAL group, 4783 time periods from 2013 people.
d See *Table 11* for details of variables included in each of the propensity score models included in this table.
e A total of 84 people per exposure group after matching.
f A total of 554 people per exposure group after matching.
g A total of 434 people per exposure group after matching.
h A total of 94 people per exposure group after matching.
i A total of 910 people per exposure group after matching.
j A total of 1362 people per exposure group after matching.
k A total of 92 people per exposure group after matching.
l A total of 573 people per exposure group after matching.
m A total of 425 people per exposure group after matching.

TORCH trial-analogous[3] CPRD result (0.93, 95% CI 0.65 to 1.32). For the pneumonia analysis, we found no evidence for an increased risk associated with FP-SAL, with a propensity score-matched rate ratio of 0.82 (95% CI 0.44 to 1.53) and a propensity score-adjusted rate ratio of 0.88 (95% CI 0.54 to 1.42).

People with an asthma diagnosis within 5 years prior to study entry

For the exacerbations analysis, we found a propensity score-matched rate ratio of 0.74 (95% CI 0.62 to 0.89) and a propensity score-adjusted rate ratio of 0.67 (95% CI 0.59 to 0.78), which is consistent with the association measured in the TORCH trial-analogous[3] CPRD population (0.85, 95% CI 0.74 to 0.97). For the mortality outcome, we obtained a propensity score-matched HR of 1.49 (95% CI 1.21 to 1.85) and propensity score-adjusted HR of 1.20 (95% CI 1.04 to 1.40), contrary to the null findings with the TORCH trial-analogous[3] CPRD result (0.93, 95% CI 0.65 to 1.32). For the pneumonia analysis, we found no evidence for an increased risk associated with FP-SAL, with a propensity score-matched rate ratio of 1.09 (95% CI 0.74 to 1.63) and a propensity score-adjusted rate ratio of 1.04 (95% CI 0.79 to 1.37).

People with mild chronic obstructive pulmonary disease (based on spirometry)

For the exacerbations analysis, we found a propensity score-matched rate ratio of 0.56 (95% CI 0.46 to 0.70) and a propensity score-adjusted rate ratio of 0.52 (95% CI 0.45 to 0.61), which suggests a stronger protective association than that measured in the TORCH trial-analogous[3] CPRD population (0.85, 95% CI 0.74 to 0.97). Notably, however, the crude association in those with mild COPD was also strongly protective, unlike in the TORCH trial-analogous[3] population. For the mortality outcome, we obtained a propensity score-matched HR of 0.98 (95% CI 0.67 to 1.45) and a propensity score-adjusted HR of 0.84 (95% CI 0.66 to 1.08). Again, this is consistent with the TORCH trial-analogous[3] CPRD result (0.93, 95% CI 0.65 to 1.32). For the pneumonia analysis, we found no evidence for an increased risk associated with FP-SAL, with a propensity score-matched rate ratio of 0.78 (95% CI 0.45 to 1.35) and a propensity score-adjusted rate ratio of 1.08 (95% CI 0.74 to 1.57).

Discussion

In the second part of the project, we applied the observational methods that we previously validated against the TORCH trial[3] to the study of three groups of people who would have been excluded from the TORCH trial:[3] (1) people aged > 80 years, (2) people with a diagnosis of asthma in the preceding 5 years and (3) people with milder COPD, as determined by lung function. When analysing exacerbations, our results were consistent with our analysis of the TORCH trial-analogous[3] CPRD cohort (and, therefore, with the TORCH trial[3] itself), with the exception of people with milder COPD, for whom we observed a stronger protective association with FP-SAL. For the analysis of mortality, as for the TORCH trial-analogous[3] cohort, we saw a lack of association between being prescribed FP-SAL (vs. being prescribed SAL), with the exception of those with prior asthma, for whom we observed an increase in mortality. Finally, for the analysis of pneumonia, although we found an increase in those prescribed FP-SAL in our TORCH trial-analogous[3] cohort, we found no evidence of an association in any of our TORCH trial-excluded[3] cohorts.

Comparison with previous studies

Previous studies that have extended 'trial-replication' approaches to understudied groups

In *Comparison with previous studies*, we highlighted a number of other studies in which the aim was to replicate specific trials (and their results).[30,32,34] None of these studies specified that further work was to include the use of the methodology that had been applied to specific study groups excluded or under-represented in trials. The authors of one of these studies[34] (one of whom is a co-investigator on this grant) are working on the RCT-DUPLICATE (Randomized, Controlled Trials Duplicated Using Prospective Longitudinal Insurance Claims: Applying Techniques of Epidemiology) project,[49] which aims to use similar trial-replication approaches to those we have used here to determine which types of clinical questions and real-world data analyses can be conducted with confidence with real-world data. The work we have performed here contrasts and compliments the RCT-DUPLICATE project,[49] as it is specifically concerned with answering questions about populations that would not be included in trials.

Previous studies of chronic obstructive pulmonary disease drug treatment effects in groups excluded from trials

Although it is recognised that treatment of COPD in the elderly population is a particular challenge for both clinicians and patients,[50] there is a lack of studies of treatment effects. Pooled analyses of trial data have been performed to address this knowledge gap and have demonstrated that effects obtained are similar to those obtained in overall trial populations.[51] In contrast to our methodology, these analyses may require access to trial data from a large number of trials to achieve sufficient power, rely on the trials being well performed to produce valid results and can be performed only when these data exist already.[52]

There is a lack of studies focusing on the effects of COPD treatments on people who have both asthma and COPD. The 'recommendations for research' section of the current NICE guidance on COPD treatment includes 'inhaled therapies for people with COPD and asthma'.[20] Emerging evidence suggests that this group of people is a group that is more sick, with higher levels of inflammation and a worse prognosis.[53,54]

In a recent non-systematic review[55] of the evidence for COPD treatment effects in people with mild COPD, the authors commented on the concerning lack of evidence for commonly used maintenance treatments (e.g. LABA/ICS combinations) in this population. This review highlights only two studies[56,57] that focused on any of the outcomes we studied in mild COPD patients, both focusing on exacerbations and both only studying monotherapy. The first study was a systematic review and meta-analysis of ICS compared with placebo and found a lack of effect on exacerbations, although the size of the analysis was small ($n = 191$ from three RCTs) and the result had wide CIs (RR 0.92, 95% CI 0.55 to 1.53).[56] In the second study (a RCT), there was no evidence of a difference between exacerbation rate in 391 patients with mild COPD taking the LABA FP compared with placebo.[57]

Implications and further work

Our propensity score-matched results for the > 80-year-old cohort suggest that the effect of FP-SAL (vs. SAL) on exacerbations is of a similar protective magnitude as it is in a cohort of people with similar characteristics to those in the TORCH trial[3] (RR 0.59, 95% CI 0.36 to 0.95 in the > 80-year-old group compared with RR 0.85, 95% CI 0.74 to 0.97 in our TORCH trial-analogous[3] cohort). This provides some reassurance of the effectiveness of this treatment in this population in a UK primary care setting. Further work could be to extend this analysis to other drug treatments (e.g. looking at inhaled triple combination therapy LAMA + LABA + ICS, as recently added to NICE guidance[5]) and also assessing whether or not the same effect is seen when using the same methods applied in a different setting (e.g. an EHR database from another country). For the mortality and pneumonia effects, although both our propensity score-matched and our propensity score-adjusted associations were consistent with those found in our TORCH trial-analogous[3] analysis, there were wide CIs. One possibility could be to repeat these analyses and include the additional EHRs now available in CPRD Aurum.[58]

For our asthma cohort, the increased mortality observed in the FP-SAL group (in contrast with the null effect found in our TORCH trial[3] analysis) could be a spurious result, possibly due to imbalances in the cohort that were introduced by our propensity score method (although from *Appendix 1*, *Table 19*, the only observed imbalance introduced to matching is a small difference in FEV_1). Further work is, therefore, needed to investigate this finding. However, as mentioned previously, people with both asthma and COPD have a poorer prognosis than those with just COPD, and how this has a selective impact on those on FP-SAL compared with those on SAL while having a protective effect on exacerbations and a null impact on pneumonia needs further consideration. Future work (in both CPRD GOLD and other data sources) should focus on identifying any patterns in cause-specific mortality and whether or not this represents any possible causal association with treatment or unadjusted confounding due to underlying differences between treatment groups.

Finally, for the cohort of people with mild COPD, our results indicate that there is a stronger protective effect on exacerbations than those people in the TORCH trial-analogous[3] cohort. This is a potentially reassuring result; however, as it has not been seen before, it would be advisable to try to

replicate this finding in a completely separate setting using this methodology (e.g. an EHR database containing a completely distinct population to the CPRD).

Limitations

Some of our analyses had very small numbers. In particular, the propensity score-matched groups for those aged > 80 years were all < 100 (see *Table 11*), despite using one of the largest data sets available for these kinds of analyses. To help address this, we also presented results that were adjusted for the propensity score rather than matched. Although some of the point estimates differed when comparing matched with adjusted estimates, none of our conclusions changed for this group when considering the propensity score-matched cohort with the propensity score-adjusted cohort. Nonetheless, the more recent availability of even larger quantities of EHR data (i.e. CPRD Aurum) will allow for more precise estimates in future and would likely mean that any residual discrepancy between matched and adjusted estimates would be accounted for.

A key limitation of this part of the analysis (that also effects all observational studies) is that we cannot have the same level of certainty regarding causal associations as would be possible in a well-conducted randomised trial. Owing to the nature of the research questions in this part of the project, we cannot know if our results are valid and would replicate those of a reference trial, as for this stage we were not matching our cohorts to actual trial participants or comparing the results to actual clinical trial results. However, as discussed in *Analysis of impact of (1) TORCH trial[3] matching and (2) TORCH trial[3] criteria (post hoc analysis)*, it was the application of the trial inclusion/exclusion criteria when selecting cohorts that led to trial-comparable results not matching to the individual trial participants, and we did apply the same rigorous trial selection-type approach (described in *Chapter 2*, *Selection of participants*) to the selection of people for inclusion in all of the cohorts in this section. Furthermore, our methods for preparing the propensity score models were identical to those used in the TORCH trial-analogous[3] analysis (described in *Chapter 2, Propensity score for addressing confounding*).

Conclusions

Analyses of the association between FP-SAL and COPD exacerbations, mortality and pneumonia compared with SAL are largely consistent with the results from the TORCH trial[3] when we extend to patients excluded from or under-represented in the TORCH trial.[3] This is largely reassuring, but two potential differences were noted. First, we observed a larger protective association than expected for COPD exacerbations in people with mild COPD. Second, and more concerningly, we observed a small increased risk of death in patients with evidence of both COPD and asthma. Both these signals warrant further investigation in CPRD GOLD and in other data sets.

Chapter 4 Overall conclusions

The aims of this project were to measure the association between treatments for COPD and COPD outcomes among patients not included in RCTs for COPD treatments, and to develop a methodological framework for using non-interventional EHRs. We have met these aims by completing each of the three specific primary objectives outlined in *Chapter 1, Aims and objectives*: (1) validate methods for measuring COPD medication effectiveness in EHR data by comparing with trial results, (2) use EHR data to measure COPD medication effectiveness in patients excluded from trials (most importantly, those aged > 80 years or those with comorbidities) and (3) determine COPD treatment effectiveness in an understudied disease stage (i.e. mild COPD).

Our results show that routinely collected EHR data can be used to successfully identify the expected beneficial and harmful effects of treatments for COPD when validating against results obtained from randomised trials. Importantly, successful replication was possible when comparing between two active treatments only and could not be achieved for comparisons between active treatment and no treatment. This was because of (1) confounding by indication (an established challenge when trying to compare with no treatment) and (2) the fact that the TORCH trial[3] placebo-controlled group were such an unhealthy group in comparison with people with COPD routinely treated in UK primary care (overall much more recently than the TORCH trial[3]). These conclusions are specific to investigations of the effects of COPD medication and cannot be assumed to replicate in other disease areas. In validating against the results of a large international multicentre randomised trial, it was also clear that, in some instances, some patient characteristics observed in a trial are not always observed in a single-country EHR setting. This raises questions of possible trial result heterogeneity by geographic region, which should be considered in future attempts to replicate trial findings in non-interventional data.

The step of directly comparing findings from non-interventional data with those from the trial provided a methodological validation and template, allowing further work to focus on the types of patients excluded from the original trials.

Analyses involving patients who would have been excluded from, or were under-represented in, COPD treatment trials mostly suggests that treatment effects for FP-SAL are similar in patients aged > 80 years, those with mild COPD and those with both asthma and COPD. However, some potential differences were also suggested. For people with mild COPD, the use of FP-SAL appears to be more beneficial with respect to exacerbations than was seen in the TORCH trial-analogous[3] population. By contrast, we observed a small increased risk of mortality when comparing FP-SAL with SAL in the group with both COPD and asthma. These associations should be interpreted with caution, and we recommend that future studies in CPRD GOLD and in other data sets focus on further characterising these associations.

Overall, we have demonstrated the utility of non-interventional data to investigate the expected treatment effects of COPD medications, in both trial-included and trial-excluded patient groups. Analyses largely suggest that COPD treatment effects are consistent across different patient groups, but highlighted a small number of possible differences that need to be investigated further in CPRD GOLD and other data sets before any conclusions that might have an impact on COPD treatment decisions can be reached. Unanswered questions about the effectiveness of currently recommended COPD inhaled combination therapy (other than FP-SAL) in patients excluded from trials should also be investigated using these methods. In addition, further work on advanced technique (e.g. high-dimensional propensity scores) could be performed to investigate whether or not placebo-controlled RCTs can ever be replicated in this therapeutic area (particularly when the target trial is a historical trial that has recruited patients in the placebo group who are much more sick than patients under the care of clinicians in a much more recent primary care setting).

Prioritised list of recommendations for future research

- To further investigate the harmful effect of FP-SAL (vs. SAL) observed with respect to mortality in people with asthma within the CPRD GOLD data set.
- To update the cohort so that (1) patients can be selected up to January 2020 and (2) patients can be selected from both CPRD GOLD and CPRD Aurum databases to increase the power of the subgroup analyses (i.e. for those aged > 80 years, those with mild COPD and those with concomitant asthma) and to facilitate analysis in those groups that we were unable to analyse because of small numbers (i.e. people with a history of lung surgery, people with a history of long-term oxygen therapy, people with substantial comorbidity or people with evidence of drug/alcohol abuse).
- To apply the methodology developed here to the analysis of other COPD treatments that are of current interest and relevance to prescribing practice and now specified in NICE guidance, such as the use of new triple therapy formulations (e.g. LAMA–ICS–LABA) compared with ICS–LABA or LABA–LAMA formulations.[5]
- To investigate advanced techniques (e.g. high-dimensional propensity scores) to assess whether or not it is possible to replicate placebo-controlled outcomes analyses in this therapeutic area when comparing with historical trials.

Acknowledgements

Contributions of authors

Kevin Wing (https://orcid.org/0000-0003-2335-9641) (Assistant Professor of Epidemiology) applied for access to the clinical trial data used for matching and liaised with the study sponsor of the trial, did the formal analysis, contributed to the study methods, contributed to disease category conceptualisation and codelists, acquired ethics approval for the study, wrote the original manuscript draft, contributed to reviewing and editing of the manuscript, and approved the final manuscript.

Elizabeth Williamson (https://orcid.org/0000-0001-6905-876X) (Professor of Medical Statistics) contributed to the conceptualisation of the study, contributed to the study methods, contributed to reviewing and editing of the manuscript, was involved in design and conceptual development, and reviewed and approved the final manuscript.

James R Carpenter (https://orcid.org/0000-0003-3890-6206) (Professor of Medical Statistics) contributed to the conceptualisation of the study, contributed to reviewing and editing of the manuscript, was involved in design and conceptual development, and reviewed and approved the final manuscript.

Lesley Wise (https://orcid.org/0000-0002-5105-5509) (Honorary Associate Professor of Pharmacoepidemiology) contributed to the conceptualisation of the study, contributed to reviewing and editing of the manuscript, was involved in design and conceptual development, and reviewed and approved the final manuscript.

Sebastian Schneeweiss (https://orcid.org/0000-0003-2575-467X) (Professor of Medicine and Epidemiology) contributed to the conceptualisation of the study, contributed to reviewing and editing of the manuscript, was involved in design and conceptual development, and reviewed and approved the final manuscript.

Liam Smeeth (https://orcid.org/0000-0002-9168-6022) (Professor of Clinical Epidemiology) contributed to the conceptualisation of the study, contributed to reviewing and editing of the manuscript, was involved in design and conceptual development, and reviewed and approved the final manuscript.

Jennifer K Quint (https://orcid.org/0000-0003-0149-4869) (Professor of Respiratory Epidemiology) contributed to the conceptualisation of the study, contributed to disease category conceptualisation and codelists, contributed to reviewing and editing of the manuscript, was involved in design and conceptual development, and reviewed and approved the final manuscript.

Ian Douglas (https://orcid.org/0000-0002-8970-1406) (Professor of Pharmacoepidemiology) contributed to the conceptualisation of the study, contributed to the study methods, contributed to disease category conceptualisation and codelists, acquired ethics approval for the study, contributed to project administration, supervised the study, was involved in design and conceptual development, and reviewed and approved the final manuscript.

Publications

Wing K, Williamson E, Carpenter JR, Wise L, Schneeweiss S, Smeeth L, *et al.* Real-world effects of medications for chronic obstructive pulmonary disease: protocol for a UK population-based non-interventional cohort study with validation against randomised trial results. *BMJ Open* 2018;**8**:e019475.

Wing K, Williamson E, Carpenter JR, Wise L, Schneeweiss S, Smeeth L, *et al.* Real world effects of COPD medications: a cohort study with validation against RCT results. *Eur Respir J* 2021;**57**:2001586.

Data-sharing statement

All data analysed in this project was obtained from the UK CPRD under a licence agreement with the London School of Hygiene & Tropical Medicine and cannot be shared. Applications for access to the same data can be made to the CPRD. For further information please contact the corresponding author.

References

1. Wing K, Williamson E, Carpenter JR, Wise L, Schneeweiss S, Smeeth L, *et al.* Real-world effects of medications for chronic obstructive pulmonary disease: protocol for a UK population-based non-interventional cohort study with validation against randomised trial results. *BMJ Open* 2018;**8**:e019475. https://doi.org/10.1136/bmjopen-2017-019475
2. The Healthcare Commission. *Clearing the Air: A National Study Of Chronic Pulmonary Disease*. London: The Healthcare Commission; 2006.
3. Calverley PM, Anderson JA, Celli B, Ferguson GT, Jenkins C, Jones PW, *et al.* Salmeterol and fluticasone propionate and survival in chronic obstructive pulmonary disease. *N Engl J Med* 2007;**356**:775–89. https://doi.org/10.1056/NEJMoa063070
4. Tashkin DP, Celli B, Senn S, Burkhart D, Kesten S, Menjoge S, Decramer M, UPLIFT Study Investigators. A 4-year trial of tiotropium in chronic obstructive pulmonary disease. *N Engl J Med* 2008;**359**:1543–54. https://doi.org/10.1056/NEJMoa0805800
5. National Institute for Health and Care Excellence. *Chronic Obstructive Pulmonary Disease in Over 16s: Diagnosis and Management*. 2010. URL: www.nice.org.uk/guidance/cg101 (accessed 26 May 2017).
6. Ferguson GT, Anzueto A, Fei R, Emmett A, Knobil K, Kalberg C, *et al.* Maintenance pharmacotherapy of mild and moderate COPD: what is the evidence? *Respir Med* 2011;**105**:1268–74. https://doi.org/10.1016/j.rmed.2011.02.005
7. European Union. *Regulation (EU) No 1235/2010 of the European Parliament and of the Council*. 2010. URL: http://ec.europa.eu/health//sites/health/files/files/eudralex/vol-1/reg_2010_1235/reg_2010_1235_en.pdf (accessed 2 Jaune 2021).
8. Spring S. *Guidance for Industry Postmarketing Studies and Clinical Trials – Implementation of Section 505(o)(3) of the Federal Food, Drug, and Cosmetic Act*. URL: www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm (accessed 13 June 2017).
9. 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. https://doi.org/10.1093/ije/dyv098
10. Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. *Br J Clin Pharmacol* 2010;**69**:4–14. https://doi.org/10.1111/j.1365-2125.2009.03537.x
11. Quint JK, Müllerova H, DiSantostefano RL, Forbes H, Eaton S, Hurst JR, *et al.* Validation of chronic obstructive pulmonary disease recording in the Clinical Practice Research Datalink (CPRD-GOLD). *BMJ Open* 2014;**4**:e005540. https://doi.org/10.1136/bmjopen-2014-005540
12. Millett ER, Quint JK, Smeeth L, Daniel RM, Thomas SL. Incidence of community-acquired lower respiratory tract infections and pneumonia among older adults in the United Kingdom: a population-based study. *PLOS ONE* 2013;**8**:e75131. https://doi.org/10.1371/journal.pone.0075131
13. 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. https://doi.org/10.1371/journal.pone.0151357

14. Nissen F, Morales DR, Mullerova H, Smeeth L, Douglas IJ, Quint JK. Validation of asthma recording in the Clinical Practice Research Datalink (CPRD). *BMJ Open* 2017;**7**:e017474. https://doi.org/10.1136/bmjopen-2017-017474

15. Wing K, Williamson E, Carpenter JR, Wise L, Schneeweiss S, Smeeth L, *et al.* Real world effects of COPD medications: a cohort study with validation against RCT results. *Eur Respir J* 2021;**57**:2001586. https://doi.org/10.1183/13993003.01586-2020

16. Douglas I, Evans S, Smeeth L. Effect of statin treatment on short term mortality after pneumonia episode: cohort study. *BMJ* 2011;**342**:d1642. https://doi.org/10.1136/bmj.d1642

17. Williamson E, Morley R, Lucas A, Carpenter J. Propensity scores: from naive enthusiasm to intuitive understanding. *Stat Methods Med Res* 2012;**21**:273–93. https://doi.org/10.1177/0962280210394483

18. Williamson EJ, Forbes A. Introduction to propensity scores. *Respirology* 2014;**19**:625–35. https://doi.org/10.1111/resp.12312

19. Root AA, Wong AY, Ghebremichael-Weldeselassie Y, Smeeth L, Bhaskaran K, Evans SJ, *et al.* Evaluation of the risk of cardiovascular events with clarithromycin using both propensity score and self-controlled study designs. *Br J Clin Pharmacol* 2016;**82**:512–21. https://doi.org/10.1111/bcp.12983

20. National Institute for Health and Care Excellence. *Chronic Obstructive Pulmonary Disease in Over 16s: Diagnosis and Management*. URL: www.nice.org.uk/guidance/ng115 (accessed 3 June 2019).

21. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med* 2000;**342**:1887–92. https://doi.org/10.1056/NEJM200006223422507

22. Golder S, Loke YK, Bland M. Meta-analyses of adverse effects data derived from randomised controlled trials as compared to observational studies: methodological overview. *PLOS Med* 2011;**8**:e1001026. https://doi.org/10.1371/journal.pmed.1001026

23. Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovitch C, Song F, *et al.* Evaluating non-randomised intervention studies. *Health Technol Assess* 2003;**7**(27). https://doi.org/10.3310/hta7270

24. Lawlor DA, Davey Smith G, Kundu D, Bruckdorfer KR, Ebrahim S. Those confounded vitamins: what can we learn from the differences between observational versus randomised trial evidence? *Lancet* 2004;**363**:1724–7. https://doi.org/10.1016/S0140-6736(04)16260-0

25. Vandenbroucke JP. The HRT controversy: observational studies and RCTs fall in line. *Lancet* 2009;**373**:1233–5. https://doi.org/10.1016/S0140-6736(09)60708-X

26. Danaei G, Rodríguez LA, Cantero OF, Logan R, Hernán MA. Observational data for comparative effectiveness research: an emulation of randomised trials of statins and primary prevention of coronary heart disease. *Stat Methods Med Res* 2013;**22**:70–96. https://doi.org/10.1177/0962280211403603

27. Douglas IJ, Evans SJ, Pocock S, Smeeth L. The risk of fractures associated with thiazolidinediones: a self-controlled case-series study. *PLOS Med* 2009;**6**:e1000154. https://doi.org/10.1371/journal.pmed.1000154

28. Langan SM, Smeeth L, Margolis DJ, Thomas SL. Herpes zoster vaccine effectiveness against incident herpes zoster and post-herpetic neuralgia in an older US population: a cohort study. *PLOS Med* 2013;**10**:e1001420. https://doi.org/10.1371/journal.pmed.1001420

29. Schneeweiss S, Seeger JD, Landon J, Walker AM. Aprotinin during coronary-artery bypass grafting and risk of death. *N Engl J Med* 2008;**358**:771–83. https://doi.org/10.1056/NEJMoa0707571

30. Hernán MA, Alonso A, Logan R, Grodstein F, Michels KB, Willett WC, *et al.* Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology* 2008;**19**:766–79. https://doi.org/10.1097/EDE.0b013e3181875e61

31. Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, *et al.* Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *JAMA* 2002;**288**:321–33. https://doi.org/10.1001/jama.288.3.321

32. Smeeth L, Douglas I, Hall AJ, Hubbard R, Evans S. Effect of statins on a wide range of health outcomes: a cohort study validated by comparison with randomized trials. *Br J Clin Pharmacol* 2009;**67**:99–109. https://doi.org/10.1111/j.1365-2125.2008.03308.x

33. Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20,536 high-risk individuals: a randomised placebo-controlled trial. *Lancet* 2002;**360**:7–22. https://doi.org/10.1016/S0140-6736(02)09327-3

34. Fralick M, Kesselheim AS, Avorn J, Schneeweiss S. Use of health care databases to support supplemental indications of approved medications. *JAMA Intern Med* 2018;**178**:55–63. https://doi.org/10.1001/jamainternmed.2017.3919

35. The ONTARGET Investigators. Telmisartan, ramipril, or both in patients at high risk for vascular events. *N Engl J Med* 2008;**358**:1547–59. https://doi.org/10.1056/NEJMoa0801317

36. Nannini LJ, Lasserson TJ, Poole P. Combined corticosteroid and long-acting beta2-agonist in one inhaler versus long-acting beta2-agonists for chronic obstructive pulmonary disease. *Cochrane Database Syst Rev* 2012;**9**:CD006829. https://doi.org/10.1002/14651858.CD006829.pub2

37. Soriano JB, Vestbo J, Pride NB, Kiri V, Maden C, Maier WC. Survival in COPD patients after regular use of fluticasone propionate and salmeterol in general practice. *Eur Respir J* 2002;**20**:819–25. https://doi.org/10.1183/09031936.02.00301302

38. Mapel DW, Hurley JS, Roblin D, Roberts M, Davis KJ, Schreiner R, Frost FJ. Survival of COPD patients using inhaled corticosteroids and long-acting beta agonists. *Respir Med* 2006;**100**:595–609. https://doi.org/10.1016/j.rmed.2005.08.006

39. Suissa S. Inhaled steroids and mortality in COPD: bias from unaccounted immortal time. *Eur Respir J* 2004;**23**:391–5. https://doi.org/10.1183/09031936.04.00062504

40. Suissa S, Patenaude V, Lapi F, Ernst P. Inhaled corticosteroids in COPD and the risk of serious pneumonia. *Thorax* 2013;**68**:1029–36. https://doi.org/10.1136/thoraxjnl-2012-202872

41. Ernst P, Gonzalez AV, Brassard P, Suissa S. Inhaled corticosteroid use in chronic obstructive pulmonary disease and the risk of hospitalization for pneumonia. *Am J Respir Crit Care Med* 2007;**176**:162–6. https://doi.org/10.1164/rccm.200611-1630OC

42. Suissa S, Dell'Aniello S, Ernst P. Comparative effectiveness of LABA-ICS versus LAMA as initial treatment in COPD targeted by blood eosinophils: a population-based cohort study. *Lancet Respir Med* 2018;**6**:855–62. https://doi.org/10.1016/S2213-2600(18)30368-0

43. Bogart M, Stanford RH, Laliberté F, Germain G, Wu JW, Duh MS. Medication adherence and persistence in chronic obstructive pulmonary disease patients receiving triple therapy in a USA commercially insured population. *Int J Chron Obstruct Pulmon Dis* 2019;**14**:343–52. https://doi.org/10.2147/COPD.S184653

44. Mueller S, Wilke T, Bechtel B, Punekar YS, Mitzner K, Virchow JC. Non-persistence and non-adherence to long-acting COPD medication therapy: a retrospective cohort study based on a large German claims dataset. *Respir Med* 2017;**122**:1–11. https://doi.org/10.1016/j.rmed.2016.11.008

45. Kyriacou DN, Lewis RJ. Confounding by indication in clinical research. *JAMA* 2016;**316**:1818–19. https://doi.org/10.1001/jama.2016.16435

46. Yoshida K, Solomon DH, Kim SC. Active-comparator design and new-user design in observational studies. *Nat Rev Rheumatol* 2015;**11**:437–41. https://doi.org/10.1038/nrrheum.2015.30

47. Gagne JJ, Fireman B, Ryan PB, Maclure M, Gerhard T, Toh S, *et al.* Design considerations in an active medical product safety monitoring system. *Pharmacoepidemiol Drug Saf* 2012;**21**:32–40. https://doi.org/10.1002/pds.2316

48. Nørgaard M, Ehrenstein V, Vandenbroucke JP. Confounding in observational studies based on large health care databases: problems and potential solutions – a primer for the clinician. *Clin Epidemiol* 2017;**9**:185–93. https://doi.org/10.2147/CLEP.S129879

49. *RCT DUPLICATE – Home.* URL: https://aetion.com/company/news/first-rct-duplicate-findings-published-in-circulation/ (accessed 15 October 2021).

50. Fried TR, Vaz Fragoso CA, Rabow MW. Caring for the older person with chronic obstructive pulmonary disease. *JAMA* 2012;**308**:1254–63. https://doi.org/10.1001/jama.2012.12422

51. Ferguson GT, Buhl R, Bothner U, Hoz A, Voß F, Anzueto A, Calverley PMA. Safety of tiotropium/olodaterol in chronic obstructive pulmonary disease: pooled analysis of three large, 52-week, randomized clinical trials. *Respir Med* 2018;**143**:67–73. https://doi.org/10.1016/j.rmed.2018.08.012

52. Ray R, Tombs L, Naya I, Compton C, Lipson DA, Boucot I. Efficacy and safety of the dual bronchodilator combination umeclidinium/vilanterol in COPD by age and airflow limitation severity: a pooled post hoc analysis of seven clinical trials. *Pulm Pharmacol Ther* 2019;**57**:101802. https://doi.org/10.1016/j.pupt.2019.101802

53. Lee H, Ryu J, Chung SJ, Park DW, Sohn JW, Yoon HJ, Kim SH. Coexisting COPD increases mortality in patients with corticosteroid-dependent asthma: a nationwide population-based study. *Allergy Asthma Immunol Res* 2020;**12**:821–31. https://doi.org/10.4168/aair.2020.12.5.821

54. Harada T, Yamasaki A, Fukushima T, Hashimoto K, Takata M, Kodani M, *et al.* Causes of death in patients with asthma and asthma-chronic obstructive pulmonary disease overlap syndrome. *Int J Chron Obstruct Pulmon Dis* 2015;**10**:595–602. https://doi.org/10.2147/COPD.S77491

55. Singh D, D'Urzo AD, Donohue JF, Kerwin EM. Weighing the evidence for pharmacological treatment interventions in mild COPD; a narrative perspective. *Respir Res* 2019;**20**:141. https://doi.org/10.1186/s12931-019-1108-9

56. Gartlehner G, Hansen RA, Carson SS, Lohr KN. Efficacy and safety of inhaled corticosteroids in patients with COPD: a systematic review and meta-analysis of health outcomes. *Ann Fam Med* 2006;**4**:253–62. https://doi.org/10.1370/afm.517

57. Jones PW, Willits LR, Burge PS, Calverley PM, Inhaled Steroids in Obstructive Lung Disease in Europe study investigators. Disease severity and the effect of fluticasone propionate on chronic obstructive pulmonary disease exacerbations. *Eur Respir J* 2003;**21**:68–73. https://doi.org/10.1183/09031936.03.00013303

58. 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. https://doi.org/10.1093/ije/dyz034

Appendix 1 Additional tables

TABLE 17 Characteristics of the aged > 80 years cohort before and after propensity score matching for the mortality analysis (FP-SAL vs. SAL) (CPRD non-interventional population)

	Before propensity score matching			After propensity score matching		
Variable	Exposed to SAL (*N* = 763 time periods from 194 people)	Exposed to FP-SAL (*N* = 1645 exposed time periods from 670 people)	Standardised difference	Exposed to SAL (*N* = 94 people)	Exposed to FP-SAL (*N* = 94 people)	Standardised difference
Age (years), median (IQR)	85 (83–88)	85 (83–88)	0.038	84 (82–87)	85 (82–87)	0.013
Sex (male), *n* (%)	517 (68)	976 (59)	0.176	49 (66)	40 (54)	0.25
Lung function, median (IQR)						
FEV_1 per cent of predicted	52 (46–59)	52 (41–59)	0.085	51 (43–57)	46 (37–58)	0.153
FEV_1 : FVC per cent	55 (47–63)	55 (45–64)	0.092	52 (47–62)	50 (45–61)	0.294
BMI (kg/m^2), median (IQR)	25 (22–28)	25 (22–28)	0.016	25 (23–28)	24 (21–27)	0.233
Prior exacerbations, mean (SD)	0.60 (0.97)	0.73 (1.16)	0.124	0.70 (1.17)	0.99 (1.31)	0.229
Cardiovascular disease, *n* (%)						
Coronary heart disease	55 (7)	391 (24)	0.47	8 (11)	14 (19)	0.229
Peripheral vascular disease	21 (3)	158 (10)	0.288	3 (4)	13 (18)	0.446
Cerebrovascular disease	18 (2)	184 (11)	0.357	3 (4)	7 (9)	0.217
Other atherosclerosis	0 (0)	3 (0)	0.06	0 (0)	0 (0)	0 (0)
Statin prescription, *n* (%)	328 (43)	694 (42)	0.016	29 (39)	27 (36)	0.056
Aspirin prescription, *n* (%)	308 (40)	700 (43)	0.044	30 (41)	30 (41)	0.000
Other COPD medication prescriptions, *n* (%)						
LABA	675 (88)	14 (1)	3.728	10 (14)	10 (14)	0.000
ICS	297 (39)	108 (7)	0.837	32 (43)	8 (11)	0.784
LAMA	244 (32)	901 (55)	0.473	29 (39)	44 (59)	0.414
ICS plus LABA	19 (2)	71 (4)	0.101	5 (7)	0 (0)	0.381

Variable	Before propensity score matching			After propensity score matching		
	Exposed to SAL (*N* = 763 time periods from 194 people)	Exposed to FP-SAL (*N* = 1645 exposed time periods from 670 people)	Standardised difference	Exposed to SAL (*N* = 94 people)	Exposed to FP-SAL (*N* = 94 people)	Standardised difference
Type 2 diabetes, *n* (%)	103 (13)	195 (12)	0.049	7 (9)	10 (14)	0.127
History of cancer, *n* (%)	282 (37)	545 (33)	0.08	20 (27)	22 (30)	0.06
Chronic kidney disease, *n* (%)	223 (29)	675 (41)	0.249	18 (24)	19 (26)	0.031
Health-care utilisation, median (IQR)						
Number of GP consultations	49 (35–72)	19 (12–31)	1.161	42 (28–63)	32 (25–57)	0.368
Number of distinct medications	6 (3–9)	6 (3–10)	0.108	6 (3–9)	10 (6–14)	0.827
Number of hospitalisations	0 (0–1)	0 (0–1)	0.106	0 (0–1)	0 (0–1)	0.009
Number of hospital procedures	0 (0–1)	0 (0–2)	0.143	0 (0–1)	0 (0–0)	0.005

IQR, interquartile range; SD, standard deviation.

TABLE 18 Characteristics of the aged > 80 years cohort before and after propensity score matching for the pneumonia analysis (FP-SAL vs. SAL) (CPRD non-interventional population)

Variable	Before propensity score matching			After propensity score matching		
	Exposed to SAL (*N* = 763 time periods from 194 people)	Exposed to FP-SAL (*N* = 1645 exposed time periods from 670 people)	Standardised difference	Exposed to SAL (*N* = 92 people)	Exposed to FP-SAL (*N* = 92 people)	Standardised difference
Age (years), median (IQR)	85 (83–88)	85 (83–88)	0.038	85 (82–86)	84 (82–87)	0.014
Sex (male), *n* (%)	517 (68)	976 (59)	0.176	40 (61)	41 (62)	0.031
Lung function, median (IQR)						
FEV_1 per cent of predicted	52 (46–59)	52 (41–59)	0.085	52 (46–58)	50 (37–59)	0.06
FEV_1 : FVC per cent	55 (47–63)	55 (45–64)	0.092	53 (47–63)	52 (44–58)	0.231
BMI (kg/m^2), median (IQR)	25 (22–28)	25 (22–28)	0.016	26 (23–28)	24 (22–27)	0.224
Prior exacerbations, mean (SD)	0.60 (0.97)	0.73 (1.16)	0.124	0.86 (1.24)	0.76 (1.02)	0.093
Cardiovascular disease, *n* (%)						
Coronary heart disease	55 (7)	391 (24)	0.47	10 (15)	16 (24)	0.23
Peripheral vascular disease	21 (3)	158 (10)	0.288	3 (5)	10 (15)	0.362
Cerebrovascular disease	18 (2)	184 (11)	0.357	3 (5)	6 (9)	0.181
Other atherosclerosis	0 (0)	3 (0)	0.06	0 (0)	0 (0)	0 (0)
Statin prescription, *n* (%)	328 (43)	694 (42)	0.016	24 (36)	31 (47)	0.216
Aspirin prescription, *n* (%)	308 (40)	700 (43)	0.044	28 (42)	29 (44)	0.031
Other COPD medication prescriptions, *n* (%)						
LABA	675 (88)	14 (1)	3.728	10 (15)	10 (15)	0.000
ICS	297 (39)	108 (7)	0.837	28 (42)	4 (6)	0.937
LAMA	244 (32)	901 (55)	0.473	29 (44)	40 (61)	0.338
ICS plus LABA	19 (2)	71 (4)	0.101	3 (5)	3 (5)	0.000

Variable	Before propensity score matching			After propensity score matching		
	Exposed to SAL (*N* = 763 time periods from 194 people)	Exposed to FP-SAL (*N* = 1645 exposed time periods from 670 people)	Standardised difference	Exposed to SAL (*N* = 92 people)	Exposed to FP-SAL (*N* = 92 people)	Standardised difference
Type 2 diabetes, *n* (%)	103 (13)	195 (12)	0.049	7 (11)	13 (20)	0.256
History of cancer, *n* (%)	282 (37)	545 (33)	0.08	17 (26)	33 (50)	0.516
Chronic kidney disease, *n* (%)	223 (29)	675 (41)	0.249	20 (30)	13 (20)	0.247
Health-care utilisation, median (IQR)						
Number of GP consultations	49 (35–72)	19 (12–31)	1.161	43 (28–81)	34 (25–53)	0.568
Number of distinct medications	6 (3–9)	6 (3–10)	0.108	7 (4–10)	7 (5–13)	0.271
Number of hospitalisations	0 (0–1)	0 (0–1)	0.106	0 (0–1)	0 (0–1)	0.178
Number of hospital procedures	0 (0–1)	0 (0–2)	0.143	0 (0–1)	0 (0–2)	0.066

IQR, interquartile range; SD, standard deviation.

TABLE 19 Characteristics of the cohort with an asthma diagnosis during the previous 5 years before and after propensity score matching for the mortality analysis (FP-SAL vs. SAL) (CPRD non-interventional population)

Variable	Before propensity score matching			After propensity score matching		
	Exposed to SAL (*N* = 5585 time periods from 1175 people)	Exposed to FP-SAL (*N* = 9719 exposed time periods from 3577 people)	Standardised difference	Exposed to SAL (*N* = 910 people)	Exposed to FP-SAL (*N* = 910 people)	Standardised difference
Age (years), median (IQR)	68 (61–73)	66 (59–72)	0.192	68 (61–73)	68 (61–74)	0.028
Sex (male), *n* (%)	2923 (52)	5409 (56)	0.067	1021 (52)	1050 (53)	0.029
Lung function, median (IQR)						
FEV_1 per cent of predicted	53 (42–62)	52 (41–61)	0.031	54 (44–63)	50 (40–59)	0.198
FEV_1 : FVC per cent	55 (45–63)	55 (46–63)	0.022	54 (45–63)	54 (45–63)	0.032
BMI (kg/m^2), median (IQR)	27 (24–31)	27 (24–32)	0.106	27 (24–31)	27 (23–31)	0.035
Prior exacerbations, mean (SD)	0.62 (1.02)	0.48 (1.01)	0.14	0.49 (0.93)	0.56 (1.07)	0.078
Cardiovascular disease, *n* (%)						
Coronary heart disease	190 (3)	268 (3)	0.037	65 (3)	51 (3)	0.042
Peripheral vascular disease	60 (1)	60 (1)	0.05	19 (1)	16 (1)	0.016
Cerebrovascular disease	73 (1)	163 (2)	0.031	27 (1)	28 (1)	0.004
Other atherosclerosis	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Statin prescription, *n* (%)	2141 (38)	3742 (39)	0.003	785 (40)	774 (39)	0.011
Aspirin prescription, *n* (%)	1281 (23)	2401 (25)	0.042	506 (26)	493 (25)	0.015
Other COPD medication prescriptions, *n* (%)						
LABA	5123 (92)	172 (2)	4.166	1787 (91)	89 (5)	3.389
ICS	4177 (75)	1071 (11)	1.685	631 (32)	662 (34)	0.033
LAMA	1630 (29)	5193 (53)	0.508	865 (44)	865 (44)	0.000
ICS plus LABA	145 (3)	507 (5)	0.136	74 (4)	131 (7)	0.130

Variable	Before propensity score matching			After propensity score matching		
	Exposed to SAL ($N = 5585$ time periods from 1175 people)	Exposed to FP-SAL ($N = 9719$ exposed time periods from 3577 people)	Standardised difference	Exposed to SAL ($N = 910$ people)	Exposed to FP-SAL ($N = 910$ people)	Standardised difference
Type 2 diabetes, *n* (%)	718 (13)	1429 (15)	0.054	283 (14)	263 (13)	0.029
History of cancer, *n* (%)	980 (18)	1629 (17)	0.021	362 (18)	353 (18)	0.012
Chronic kidney disease, *n* (%)	579 (10)	1142 (12)	0.044	221 (11)	245 (12)	0.038
Health-care utilisation, median (IQR)						
Number of GP consultations	13 (8–21)	16 (9–26)	0.259	15 (9–24)	15 (10–24)	0.036
Number of distinct medications	4 (2–7)	5 (3–8)	0.201	5 (3–7)	5 (3–8)	0.076
Number of hospitalisations	0 (0–1)	0 (0–1)	0.069	0 (0–1)	0 (0–1)	0.008
Number of hospital procedures	0 (0–0)	0 (0–1)	0.093	0 (0–0)	0 (0–0)	0.003

IQR, interquartile range; SD, standard deviation.

TABLE 20 Characteristics of the cohort with an asthma diagnosis during the previous 5 years before and after propensity score matching for the pneumonia analysis (FP-SAL vs. SAL) (CPRD non-interventional population)

Variable	Before propensity score matching			After propensity score matching		
	Exposed to SAL (*N* = 5585 time periods from 1175 people)	Exposed to FP-SAL (*N* = 9719 exposed time periods from 3577 people)	Standardised difference	Exposed to SAL (*N* = 573 people)	Exposed to FP-SAL (*N* = 573 people)	Standardised difference
Age (years), median (IQR)	68 (61–73)	66 (59–72)	0.192	68 (62–73)	68 (61–74)	0.022
Sex (male), *n* (%)	2923 (52)	5409 (56)	0.067	294 (52)	316 (56)	0.078
Lung function, median (IQR)						
FEV_1 per cent of predicted	53 (42–62)	52 (41–61)	0.031	52 (41–61)	51 (39–60)	0.055
FEV_1 : FVC per cent	55 (45–63)	55 (46–63)	0.022	54 (46–63)	54 (45–63)	0.019
BMI (kg/m^2), median (IQR)	27 (24–31)	27 (24–32)	0.106	27 (24–31)	27 (24–30)	0.106
Prior exacerbations, mean (SD)	0.62 (1.02)	0.48 (1.01)	0.14	0.64 (1.05)	0.73 (1.11)	0.085
Cardiovascular disease, *n* (%)						
Coronary heart disease	190 (3)	268 (3)	0.037	30 (5)	34 (6)	0.031
Peripheral vascular disease	60 (1)	60 (1)	0.05	10 (2)	10 (2)	0.000
Cerebrovascular disease	73 (1)	163 (2)	0.031	14 (2)	9 (2)	0.063
Other atherosclerosis	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Statin prescription, *n* (%)	2141 (38)	3742 (39)	0.003	195 (35)	195 (35)	0.000
Aspirin prescription, *n* (%)	1281 (23)	2401 (25)	0.042	136 (24)	126 (22)	0.042
Other COPD medication prescriptions, *n* (%)						
LABA	5123 (92)	172 (2)	4.166	124 (22)	127 (23)	0.013
ICS	4177 (75)	1071 (11)	1.685	366 (65)	166 (29)	0.76
LAMA	1630 (29)	5193 (53)	0.508	215 (38)	193 (34)	0.081
ICS plus LABA	145 (3)	507 (5)	0.136	49 (9)	33 (6)	0.109

Variable	Before propensity score matching			After propensity score matching		
	Exposed to SAL (*N* = 5585 time periods from 1175 people)	Exposed to FP-SAL (*N* = 9719 exposed time periods from 3577 people)	Standardised difference	Exposed to SAL (*N* = 573 people)	Exposed to FP-SAL (*N* = 573 people)	Standardised difference
Type 2 diabetes, *n* (%)	718 (13)	1429 (15)	0.054	69 (12)	61 (11)	0.044
History of cancer, *n* (%)	980 (18)	1629 (17)	0.021	102 (18)	97 (17)	0.023
Chronic kidney disease, *n* (%)	579 (10)	1142 (12)	0.044	52 (9)	49 (9)	0.019
Health-care utilisation, median (IQR)						
Number of GP consultations	13 (8–21)	16 (9–26)	0.259	15 (9–23)	14 (9–21)	0.066
Number of distinct medications	4 (2–7)	5 (3–8)	0.201	5 (3–8)	5 (3–8)	0.015
Number of hospitalisations	0 (0–1)	0 (0–1)	0.069	0 (0–1)	0 (0–0)	0.077
Number of hospital procedures	0 (0–0)	0 (0–1)	0.093	0 (0–0)	0 (0–0)	0.118

IQR, interquartile range; SD, standard deviation.

TABLE 21 Characteristics of the mild COPD cohort before and after propensity score matching for the mortality analysis (FP-SAL vs. SAL) (CPRD non-interventional population)

Variable	Before propensity score matching			After propensity score matching		
	Exposed to SAL (*N* = 3876 exposed time periods from 877 people)	Exposed to FP-SAL (*N* = 4783 exposed time periods from 2013 people)	Standardised difference	Exposed to SAL (*N* = 2652 people)	Exposed to FP-SAL (*N* = 2652 people)	Standardised difference
Age (years), median (IQR)	66 (60–72)	66 (59–73)	0.003	65 (58–71)	66 (59–72)	0.099
Sex (male), *n* (%)	2104 (54)	2483 (52)	0.047	656 (48)	714 (52)	0.085
Lung function, median (IQR)						
FEV_1 per cent of predicted	76 (69–87)	76 (67–87)	0.041	76 (69–86)	75 (66–87)	0.091
FEV_1 : FVC per cent	68 (61–75)	70 (63–77)	0.143	68 (61–74)	69 (62–77)	0.124
BMI (kg/m^2), median (IQR)	28 (25–32)	28 (25–32)	0.055	28 (24–32)	28 (24–32)	0.019
Prior exacerbations, mean (SD)	0.50 (0.87)	0.29 (0.72)	0.263	0.26 (0.69)	0.30 (0.71)	0.061
Cardiovascular disease, *n* (%)						
Coronary heart disease	173 (4)	925 (19)	0.472	157 (12)	143 (10)	0.033
Peripheral vascular disease	52 (1)	256 (5)	0.224	48 (4)	43 (3)	0.02
Cerebrovascular disease	52 (1)	271 (6)	0.237	49 (4)	38 (3)	0.046
Other atherosclerosis	0 (0)	1 (0)	0.02	0 (0)	0 (0)	0.038
Statin prescription, *n* (%)	1589 (41)	2202 (46)	0.102	476 (35)	530 (39)	0.082
Aspirin prescription, *n* (%)	1172 (30)	1391 (29)	0.025	279 (20)	325 (24)	0.081
Other COPD medication prescriptions, *n* (%)						
LABA	3455 (89)	36 (1)	3.871	1197 (88)	16 (1)	3.569
ICS	1177 (30)	544 (11)	0.481	158 (12)	233 (17)	0.158
LAMA	965 (25)	1795 (38)	0.275	456 (33)	482 (35)	0.04
ICS plus LABA	37 (1)	152 (3)	0.157	30 (2)	36 (3)	0.029

Variable	Before propensity score matching			After propensity score matching		
	Exposed to SAL (*N* = 3876 exposed time periods from 877 people)	Exposed to FP-SAL (*N* = 4783 exposed time periods from 2013 people)	Standardised difference	Exposed to SAL (*N* = 2652 people)	Exposed to FP-SAL (*N* = 2652 people)	Standardised difference
Type 2 diabetes, *n* (%)	401 (10)	598 (13)	0.068	132 (10)	152 (11)	0.048
History of cancer, *n* (%)	634 (16)	768 (16)	0.008	180 (13)	230 (17)	0.103
Chronic kidney disease, *n* (%)	426 (11)	602 (13)	0.049	119 (9)	147 (11)	0.069
Health-care utilisation, median (IQR)						
Number of GP consultations	14 (9–21)	16 (10–25)	0.230	15 (10–23)	16 (10–25)	0.082
Number of distinct medications	4 (2–7)	5 (3–8)	0.21	5 (3–8)	5 (3–8)	0.172
Number of hospitalisations	0 (0–1)	0 (0–1)	0.109	0 (0–1)	0 (0–1)	0.081
Number of hospital procedures	0 (0–0)	0 (0–1)	0.093	0 (0–0)	0 (0–0)	0.062

IQR, interquartile range; SD, standard deviation.

TABLE 22 Characteristics of the mild COPD cohort before and after propensity score matching for the pneumonia analysis (FP-SAL vs. SAL) (CPRD non-interventional population)

Variable	Before propensity score matching			After propensity score matching		
	Exposed to SAL (*N* = 3876 exposed time periods from 877 people)	Exposed to FP-SAL (*N* = 4783 exposed time periods from 2013 people)	Standardised difference	Exposed to SAL (*N* = 2652 people)	Exposed to FP-SAL (*N* = 2652 people)	Standardised difference
Age (years), median (IQR)	66 (60–72)	66 (59–73)	0.003	66 (59–73)	66 (59–72)	0.038
Sex (male), *n* (%)	2104 (54)	2483 (52)	0.047	217 (51)	222 (52)	0.024
Lung function, median (IQR)						
FEV_1 per cent of predicted	76 (69–87)	76 (67–87)	0.041	74 (67–84)	75 (66–86)	0.017
FEV_1 : FVC per cent	68 (61–75)	70 (63–77)	0.143	67 (61–75)	71 (63–78)	0.171
BMI (kg/m^2), median (IQR)	28 (25–32)	28 (25–32)	0.055	28 (24–32)	27 (24–31)	0.022
Prior exacerbations, mean (SD)	0.50 (0.87)	0.29 (0.72)	0.263	0.59 (0.97)	0.56 (0.90)	0.033
Cardiovascular disease, *n* (%)						
Coronary heart disease	173 (4)	925 (19)	0.472	56 (13)	41 (10)	0.111
Peripheral vascular disease	52 (1)	256 (5)	0.224	12 (3)	8 (2)	0.062
Cerebrovascular disease	52 (1)	271 (6)	0.237	15 (4)	11 (3)	0.055
Other atherosclerosis	0 (0)	1 (0)	0.02	0 (0)	0 (0)	0.069
Statin prescription, *n* (%)	1589 (41)	2202 (46)	0.102	173 (41)	158 (37)	0.072
Aspirin prescription, *n* (%)	1172 (30)	1391 (29)	0.025	122 (29)	121 (28)	0.005
Other COPD medication prescriptions, *n* (%)						
LABA	3455 (89)	36 (1)	3.871	31 (7)	31 (7)	0.000
ICS	1177 (30)	544 (11)	0.481	187 (44)	92 (22)	0.490
LAMA	965 (25)	1795 (38)	0.275	137 (32)	142 (33)	0.025
ICS plus LABA	37 (1)	152 (3)	0.157	16 (4)	19 (4)	0.036

Variable	Before propensity score matching			After propensity score matching		
	Exposed to SAL (*N* = 3876 exposed time periods from 877 people)	Exposed to FP-SAL (*N* = 4783 exposed time periods from 2013 people)	Standardised difference	Exposed to SAL (*N* = 2652 people)	Exposed to FP-SAL (*N* = 2652 people)	Standardised difference
Type 2 diabetes, *n* (%)	401 (10)	598 (13)	0.068	49 (12)	40 (9)	0.069
History of cancer, *n* (%)	634 (16)	768 (16)	0.008	68 (16)	63 (15)	0.033
Chronic kidney disease, *n* (%)	426 (11)	602 (13)	0.049	49 (12)	46 (11)	0.022
Healthcare utilisation, median (IQR)						
Number of GP consultations	14 (9–21)	16 (10–25)	0.230	15 (10–24)	16 (10–24)	0.011
Number of distinct medications	4 (2–7)	5 (3–8)	0.21	5 (3–8)	5 (3–8)	0.037
Number of hospitalisations	0 (0–1)	0 (0–1)	0.109	0 (0–1)	0 (0–1)	0.073
Number of hospital procedures	0 (0–0)	0 (0–1)	0.093	0 (0–0)	0 (0–0)	0.087

IQR, interquartile range; SD, standard deviation.

EME
HS&DR
HTA
PGfAR
PHR

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

Published by the NIHR Journals Library