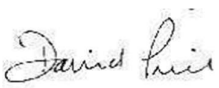
PREVAIL UK: Statistical Analysis Plan

Study title: A 3-year cluster randomised controlled trial of the impact of a quality improvement and clinical decision support package versus usual care for patients with modifiable high-risk chronic obstructive pulmonary disease with or without a current diagnosis

Document history

Version number	Change	Date
0.1	n/a	4/4/2024
0.2	Editing/clarifying following statistical review.	19/4/2024
0.3	Clarifying following AZ review & further team discussion	21/3/2025
0.4	Reinstating tables from global protocol, altering analysis sets and providing additional detail, adding a sensitivity analysis	01/08/2025
0.5	Editing following AZ review & further team discussion	16/09/2025
0.6	Including reference to data obtained from NHSE, modifying outcomes, modifying analysis set definitions and adding subgroup analyses	26/05/2026
1.0	Editing following AZ review & further team discussion	31/05/2026
1.1	Clarified i) linked primary/secondary care analyses will use a subgroup of patients, ii) ICD10 codes to be used for linked primary/secondary care analyses, iii) use of propensity score weighting for all analysis sets	01/06/2026

	Signature	Date
Prepared by:		
Dr Andy Dickens		01/06/2026
Prof William Henley		01/06/2026
Approved by:		

Prof David Price		01/06/2026

Sponsor	Observational & Pragmatic Research International Ltd.
Funder	AstraZeneca
Funding reference number	ESR-21-21637
Chief Investigator	David Price
Sponsor reference number	OPRIUK-2003
REC reference	295908
ISRCTN	ISRCTN15819828
Based on protocol version	1.5, 29/05/2026
Independent Steering Committee	Global: John Hurst, Dave Singh, Fernando Martinez, Marc Miravitlles, Luis Alves, Anita Sharma, Tonya Winders, Rongchang Chen, Shigeo Muro, Mohit Bhutani, Daiana Stolz UK: David Halpin, Katherine Hickman, James Chalmers, Graham Devereux, Martin Gibson, Tom Wilkinson, Mukesh Singh, Mona Bafadhel, Jenny Quint, Michael Crooks, Chris Gale

Abbreviations

Abbreviation	Term
AD	Already Diagnosed
A&E	Accident and Emergency
CAT	COPD Assessment Test
CI	Confidence interval
CONQUEST	Collaboration on a Quality Improvement Initiative to Achieve Excellence in Standards of COPD Care
CONSORT	Consolidated Standards of Reporting Trials
COPD	Chronic Obstructive Pulmonary Disease
EMR	Electronic Medical Record
FEV1	Forced Expiratory Volume in 1 second
FVC	Forced Vital Capacity
GERD	Gastro-oesophageal reflux disease
GP	General Practitioner
HES	Hospital Episode Statistics
ICD-10	International Statistical Classification of Diseases and Related Health Problems (10th revision)
ICS	Inhaled corticosteroid
IHD	Ischaemic Heart Disease
IMD	Index of Multiple Deprivation
ITT	Intention To Treat
IQR	Inter-Quartile Range
LABA	Long-Acting Beta-Agonist
LAMA	Long-Acting Muscarinic Antagonist
LMM	Linear Mixed Model
MACE	Major Adverse Cardiac Events
MACRE	Major Adverse Cardiac or Respiratory Events
MHR	Modifiable High-Risk
ND	Newly Diagnosed

Abbreviation	Term
NHS	National Health Service
OCS	Oral Corticosteroid
OPC	Optimum Patient Care
OPCRD	Optimum Patient Care Research Database
OPRI	Observational Pragmatic Research International
OSA	Obstructive Sleep Apnoea
PAS	Primary Analysis Set
PCS	Primary Care Site
PEF	Peak Expiratory Flow
PREVAIL	Pragmatic Evaluation of a Quality Improvement Programme for People Living with Modifiable High-risk COPD
PS	Propensity Score
QI	Quality Improvement
QIP	Quality Improvement Program
QOF	Quality Outcomes Framework
PEF	Peak Expiratory Flow
PI	Principal Investigator
RCT	Randomized Controlled Trial
REC	Research Ethics Committee
SAP	Statistical Analysis Plan
SCS	Systemic Corticosteroid
SD	Standard Deviation
T2D	Type 2 diabetes
UK	United Kingdom

Contents

Document history 1
Abbreviations 3
Contents 5
1 Introduction 7
1.1 Study background 7
1.2 Study aim 7
1.3 SAP objective 7
2 Study Objectives and Endpoints 7
2.1 Primary objective 7
2.2 Secondary objectives 8
2.3 Exploratory objectives 8
2.4 Detailed description of primary, secondary and exploratory endpoints 9
3 Study design 34
3.1 Target population for PREVAIL trial 34
3.2 Description of the treatment arms 34
3.2.1 Intervention arm 35
3.2.2 Control arm 35
3.3 Randomisation 35
3.4 Study time periods 36
3.4.1 Randomisation date 36
3.4.2 Practice initiation date 36
3.4.3 Risk identification period 36
3.4.4 Implementation period 36
3.4.5 Outcome evaluation period 37
3.4.6 Trial end date 37
4 Sample size and power 37
4.1 Annual Exacerbation Rate (combined primary analyses cohorts) 37
5 Populations and Analysis sets 38
5.1 Primary analysis set (PAS) 40
5.2 Secondary analysis set 1: early vs late CONQUEST intervention 40
5.3 Secondary analysis set 2: already diagnosed patients in PAS population 40
5.4 Secondary analysis set 3: newly diagnosed patients in PAS population 41
6 Statistical Analysis 43

6.1 General principles 43
6.2 Confounding adjustment 43
6.3 Accounting for clustering by GP practice 43
6.4 Use of primary care and secondary care data 44
6.5 Primary and secondary comparisons 44
6.6 Descriptive analysis 49
6.7 Analysis of the primary outcome 49
6.8 Analysis of the secondary outcomes 49
6.8.1 MACRE events (secondary objective 1) 49
6.8.2 MACE events (secondary objective 2) 50
6.8.3 Annual systemic corticosteroid dose used to treat exacerbations (secondary objective 3) 50
6.8.4 Annual rate of systemic corticosteroid courses used to treat exacerbations (secondary objective 4) 51
6.8.5 Incident type 2 diabetes (secondary objective 5) 51
6.8.6 COPD exacerbations (time to first event) (secondary objective 6) 51
6.9 Analysis of the exploratory outcomes 51
6.9.1 Explanatory diagnoses (exploratory objective 1) 51
6.9.2 Unscheduled respiratory events (exploratory objective 2) 52
6.9.3 Change in lung function (exploratory objective 3) 52
6.9.4 Change in disease-specific health-related quality of life (COPD Assessment Test [CAT]) (exploratory objective 4) 53
6.9.5 Incident pneumonia (exploratory objective 5) 53
6.10 Subgroup analysis 54
6.11 Sensitivity analysis 54
6.12 Missing data 54
6.13 Software 54
7 Exploratory analysis sets 54
8 Tables 57
9 Appendix 58
9.1 Corticosteroid conversion table 58
9.2 COPD Assessment Test[15] 59
9.3 Exacerbation algorithm: high-risk patient identification 60
9.4 Exacerbation algorithm: endpoint 62
10 References 64

1 Introduction

1.1 Study background

Chronic obstructive pulmonary disease (COPD) is a frequently underdiagnosed major health problem, responsible for over 250 million cases of disease. Despite resulting in more than £1-3 billion in healthcare costs in the UK alone, and causing significant morbidity and lost days of work, COPD remains a globally under-recognised condition, with an estimated 60% of cases undetected at any one time, and frequently misdiagnosed in smokers and people with asthma. Even once diagnosed, two thirds or more of patients will have already experienced significant lung function decline or previous serious exacerbations, indicating that opportunities for earlier diagnosis may have been lost. COPD patients that are symptomatic and have frequent exacerbations represent a high-risk cohort due to risk of future respiratory events, and yet are often undertreated and insufficiently monitored. Such 'modifiable high-risk' patients would benefit from further assessment, and initiation of appropriate medication and other therapies.

1.2 Study aim

The aim of the PREVAIL trial is to evaluate the impact of the CONQUEST quality improvement program (QIP) on COPD exacerbations and major adverse cardiac or respiratory events (MACREs) and other clinical outcomes, in both diagnosed and undiagnosed modifiable high-risk COPD patients. The QIP will identify these patient populations and support their management through a multicomponent package incorporating guideline-based Clinical Decision Support, delivered via clinical reviews.

1.3 SAP objective

The objective of this SAP is to describe the statistical analyses to be carried out for the PREVAIL trial. Future amendments to the protocol will be reviewed for their impact on this SAP, which will be updated only if necessary and any changes to the SAP documented under version control.

2 Study Objectives and Endpoints

The broad study objectives are listed below, and details are provided in Table 1 to Table 12. These tables include information about the specific endpoints and summarise the proposed analyses and analysis sets. Definitions are provided later in the SAP regarding time periods (Section 3.4), and full details of the analysis sets and proposed analyses are provided in Sections 5 and 6 respectively.

2.1 Primary objective

- To evaluate the impact of the CONQUEST programme on the annual rate of moderate or severe COPD exacerbations, compared to usual clinical care

2.2 Secondary objectives

- To evaluate the impact of the CONQUEST programme on the time to first major adverse cardiac or respiratory event (MACRE) and its disaggregated components, compared to usual clinical care
- To evaluate the impact of the CONQUEST programme on the time to first major adverse cardiac event (MACE), compared to usual clinical care
- To assess the impact of the CONQUEST programme on cumulative systemic corticosteroid (SCS) dose used to treat COPD exacerbations, compared to usual clinical care
- To assess the impact of the CONQUEST programme on the annual rate of systemic corticosteroid (SCS) courses to treat COPD exacerbations, compared to usual clinical care
- To assess the impact of the CONQUEST programme on incident type 2 diabetes, compared to usual clinical care
- To assess the impact of the CONQUEST programme on the time to first moderate/severe COPD exacerbation, compared to usual clinical care

2.3 Exploratory objectives

- To estimate rates of incident explanatory diagnoses other than COPD
- To evaluate the impact of the CONQUEST programme on the annual rate of unscheduled respiratory care events, compared to usual clinical care
- To evaluate the impact of the CONQUEST programme on lung function, compared to usual clinical care
- To evaluate the impact of the CONQUEST programme on COPD health status, as measured by the COPD Assessment Test (CAT), compared to usual clinical care
- To assess the impact of the CONQUEST programme on incident pneumonia, compared to usual clinical care

2.4 Detailed description of primary, secondary and exploratory endpoints

Table 1: Annual rate of moderate/severe COPD exacerbations (primary objective)

<table>
<tr><td rowspan="4">Primary objective:
To evaluate the impact of the CONQUEST programme on the annual rate of moderate or severe COPD exacerbations, compared to usual clinical care</td><td>Endpoint:
The term “exacerbation” will be used to describe both COPD exacerbations in diagnosed patients and respiratory events in undiagnosed patients, requiring oral corticosteroids and/or antibiotics, or resulting in hospital visits and admissions. Exacerbations* will be categorized as:
Moderate exacerbations defined as
● Record of prescription with ORAL prednisolone at an acute dose (prescription for a 3- to-14-day regimen (25-50 mg/d) or equivalent) and/or
● Record of prescription for antibiotic within ± 3 days of a lower respiratory consultation,
● Hospital/emergency room visit for a COPD-related respiratory event (no admittance), or
Severe exacerbations defined as
● Record of hospitalisation or death for a COPD-related respiratory event

Complicated exacerbations defined as
● Exacerbations requiring additional hospitalisation or treatment with acute doses of oral corticosteroids or antibiotics between 8 and 28 days after the start of the initial exacerbation will be defined as a complicated exacerbation.</td></tr>
<tr><td>Analysis:
In each group, the incidence rates will be estimated as the total number of COPD exacerbations during the outcome evaluation period divided by the total person-time of follow-up. These rates will then be reported as annual rates, that is the rate per patient per year.
The effectiveness of the intervention, as compared to usual care, will be measured as the rate ratio as described in section 6.7 .
Secondary analyses will be conducted assessing incidence rates of moderate exacerbations, severe exacerbations and complicated exacerbations separately, expressed as annual rates per patient per year.</td></tr>
<tr><td>Analysis timeframe:
Start date: 6 months after practice initiation date; End date: end of outcome evaluation period</td></tr>
<tr><td>Data source to be used:</td></tr>
</table>

	This outcome will be analysed exclusively using primary care EMR data. Blinded comparison of HES vs EMR data showed that using EMR data alone would miss 2% of exacerbations, whereas using linked EMR/HES data would lose 10% of practices due to consent/technical issues. Retaining trial power has been prioritised.
	Diagnostic codes and/or code list, and rule set (if appropriate): Exacerbations: 'cl_conquest_exacerbation_listv3' tab of 'PREVAIL code lists_2026.xlsx'. The specified code list relates to the algorithm outlined in section 9.4 Exacerbations of diagnosed COPD and of potential COPD will be identified using respiratory codes included in the endpoint algorithm (see Section 9.4). The assumed duration of an exacerbation will be 14 days. Events recorded 15 or more days after the start of the initial exacerbation, with a 14-day event-free window, will be considered as separate events.
	Analysis sets included in analysis (see section 5 and section 7 for description. Unless specified, analysis sets include already diagnosed and newly diagnosed patients): • Primary analysis set (PAS): intervened early MHR (intervention) vs all identified MHR (control) • Secondary analysis set 1: intervened early MHR (intervention) vs intervened late MHR (control) • Secondary analysis set 2: intervened early MHR AD patients (intervention) vs all identified MHR AD patients (control) • Secondary analysis set 3: intervened early MHR ND patients (intervention) vs all identified MHR ND patients (control) • Exploratory analysis set 1: intervened early MHR ND patients (intervention) vs intervened late MHR ND patients (control) • Exploratory analysis set 2: implemented any time MHR (intervention) vs all identified MHR (control) • Exploratory analysis set 3: implemented any time MHR (intervention) vs intervened late MHR (control) • Exploratory analysis set 4: questionnaire/QI report MHR (intervention) vs intervened late MHR (control) • Exploratory analysis set 5: all identified MHR (intervention vs control) • Exploratory analysis set 7: broad already/undiagnosed +/-MHR (intervention vs control)

Table 2: Annual rate of major adverse cardiac or respiratory events (secondary objective 1)

<table>
<tr><td rowspan="5">Secondary objective 1:
To evaluate the impact of the CONQUEST programme on the time to first major adverse cardiac or respiratory event (MACRE), compared to usual clinical care</td><td>Endpoint:
MACRE outcomes will consist of any of the following events. The first of any of these pre-defined events will denote a MACRE occurrence.
New diagnosis for heart failure
Hospitalisation for heart failure
Non-fatal myocardial infarction
Non-fatal ischaemic stroke
All-cause mortality:
sudden death
cardiac death
respiratory death
none of these
Hospitalisation (admittance) for respiratory event
Complicated exacerbations</td></tr>
<tr><td>Analysis:
Cumulative incidence of the composite MACRE outcome will be reported over the outcome evaluation period. Incidence of MACRE will be analysed using a Cox regression model, which will be reported as a time to first event with a Kaplan Meier plot.
The effectiveness of the intervention, as compared to usual care, will be measured as a hazard ratio as specified in section 6.8.1.
Secondary analyses will be conducted assessing incidence of MACRE components separately, expressed as time to first event. Incidence of the individual components of MACRE will be estimated as above. MACRE components will be explored and reported, if recorded for a minimum of 10 patients in each treatment arm.</td></tr>
<tr><td>Analysis timeframe:
Start date: 6 months after practice initiation date; End date: end of outcome evaluation period</td></tr>
<tr><td>Data source to be used:
This outcome will be analysed exclusively using linked primary and secondary care data. Please refer to section 6.4 for discussion.</td></tr>
<tr><td>Diagnostic codes and/or code list, and rule set (if appropriate):
<u>Secondary care data (ICD-10 codes):</u></td></tr>
</table>

	New diagnosis for heart failure (Admitted Patient Care): • 'hf_dx' tab of 'PREVAIL ICD10 code lists_2026.xlsx' *Hospitalisation for heart failure (Admitted Patient Care):* • Hospitalisation relating to 'hf_dx' tab of 'PREVAIL ICD10 code lists_2026.xlsx' *Non-fatal myocardial infarction:* • 'mi_dx' tab of 'PREVAIL ICD10 code lists_2026.xlsx' *Non-fatal ischaemic stroke:* • 'stroke_dx' tab of 'PREVAIL ICD10 code lists_2026.xlsx' *Mortality (Civil Registration of Deaths):* • Sudden death: 'death_sudden' tab of 'PREVAIL ICD10 code lists_2026.xlsx' • Cardiac death: recorded death relating to 'death_cardiac' tab of 'PREVAIL ICD10 code lists_2026.xlsx' • Respiratory death: recorded death relating to 'hosp_death_resp_event' tab of 'PREVAIL ICD10 code lists_2026.xlsx' • Not sudden/cardiac/respiratory death: any recorded death excluding those with codes relating to sudden death, cardiac or respiratory *Hospital admission for respiratory event (Admitted Patient Care):* • 'hosp_death_resp_event' tab of 'PREVAIL ICD10 code lists_2026.xlsx' *Complicated exacerbation:* • Both primary care and secondary care data will be used to define complicated exacerbations. The 'comp_exa_hes' tab of 'PREVAIL ICD10 code lists_2026.xlsx' will identify hospitalisations not recorded in the primary care data, between 8 and 28 days after the start of the initial exacerbation <u>Primary care EMR data:</u> *New diagnosis for heart failure:* • 'cl_heart_failure' tab of 'PREVAIL code lists_2026.xlsx' *Complicated exacerbation:* • Both primary care and secondary care data will be used to define complicated exacerbations i.e. exacerbations requiring additional hospitalisation or treatment with acute doses of oral corticosteroids or antibiotics between 8 and 28 days after the start of the initial exacerbation. New heart failure diagnosis = no monitoring or diagnostic heart failure clinical codes recorded in available primary/secondary care records prior to the event

	Analysis sets included in analysis (see section 5 and section 7 for description. Unless specified, analysis sets include already diagnosed and newly diagnosed patients): • Primary analysis set (PAS): intervened early MHR (intervention) vs all identified MHR (control) • Secondary analysis set 2: intervened early MHR AD patients (intervention) vs all identified MHR AD patients (control) • Secondary analysis set 3: intervened early MHR ND patients (intervention) vs all identified MHR ND patients (control) • Exploratory analysis set 2: implemented any time MHR (intervention) vs all identified MHR (control) • Exploratory analysis set 5: all identified MHR (intervention vs control) • Exploratory analysis set 7: broad already/undiagnosed +/-MHR (intervention vs control)

Table 3: Annual rate of major adverse cardiac events (secondary objective 2)

Secondary objective 2:
To evaluate the impact of the CONQUEST programme on the time to first major adverse cardiac event (MACE), compared to usual clinical care
	Endpoint:
	MACE outcomes will consist of any of the following events. The first of any of these pre-defined events will denote a MACE occurrence.
	● New diagnosis for heart failure
	● Hospitalisation for heart failure
	● Non-fatal myocardial infarction
	● Non-fatal ischaemic stroke
	● All-cause mortality:
	  ○ sudden death
	  ○ cardiac death
	  ○ none of these

	Analysis: Cumulative incidence of the composite MACE outcome will be reported over the outcome evaluation period. Incidence of MACE will be analysed using a Cox regression model, which will be reported as a time to first event with a Kaplan Meier plot. The effectiveness of the intervention, as compared to usual care, will be measured as a hazard ratio as specified in section 6.8.2. Secondary analyses will be conducted assessing incidence of MACE components separately, expressed as time to first event. Incidence of the individual components of MACE will be estimated as above. MACE components will be explored and reported, if recorded for a minimum of 10 patients in the intervention arm. Small number suppression will be applied to counts in the control arm if appropriate.
	Analysis timeframe: Start date: 6 months after practice initiation date; End date: end of outcome evaluation period
	Data source to be used: This outcome will be analysed exclusively using linked primary and secondary care data. Please refer to section 6.4 for discussion.
	Diagnostic codes and/or code list, and rule set (if appropriate): <u>Secondary care data (ICD-10 codes):</u> *New diagnosis for heart failure (Admitted Patient Care):* • 'hf_dx' tab of 'PREVAIL ICD10 code lists_2026.xlsx' *Hospitalisation for heart failure (Admitted Patient Care):*

	• 'hf_dx' tab of 'PREVAIL ICD10 code lists_2026.xlsx' *Non-fatal myocardial infarction:* • 'mi_dx' tab of 'PREVAIL ICD10 code lists_2026.xlsx' *Non-fatal ischaemic stroke:* • 'stroke_dx' tab of 'PREVAIL ICD10 code lists_2026.xlsx' *Mortality (Civil Registration of Deaths):* • Sudden death: 'death_sudden' tab of 'PREVAIL ICD10 code lists_2026.xlsx' • Cardiac death: recorded death relating to 'death_cardiac' tab of 'PREVAIL ICD10 code lists_2026.xlsx' • Not sudden/cardiac death: any recorded death excluding those with codes relating to sudden death, or cardiac <u>Primary care EMR data:</u> *New diagnosis for heart failure:* • 'cl_heart_failure' tab of 'PREVAIL code lists_2026.xlsx' New heart failure diagnosis = no monitoring or diagnostic heart failure clinical codes recorded in available primary/secondary care records prior to the event.
	Analysis sets included in analysis (see section 5 and section 7 for description. Unless specified, analysis sets include already diagnosed and newly diagnosed patients): • Primary analysis set (PAS): intervened early MHR (intervention) vs all identified MHR (control) • Secondary analysis set 2: intervened early MHR AD patients (intervention) vs all identified MHR AD patients (control) • Secondary analysis set 3: intervened early MHR ND patients (intervention) vs all identified MHR ND patients (control) • Exploratory analysis set 2: implemented any time MHR (intervention) vs all identified MHR (control) • Exploratory analysis set 5: all identified MHR (intervention vs control) • Exploratory analysis set 7: broad already/undiagnosed +/-MHR (intervention vs control)

Table 4: Systemic corticosteroid use (secondary objective 3)

Secondary objective 3: To assess the impact of the CONQUEST programme on systemic corticosteroid (SCS) use to treat COPD exacerbations, compared to usual clinical care	**Endpoint:** SCS exposure will be measured as the average annual dose, in milligrams, of prednisolone or equivalent taken via oral route available in primary care electronic medical records (EMR). The annual dose is the cumulative steroid dose related to exacerbations i.e. regimens started within 3 days of a coded respiratory event or specific exacerbation code. Use of SCS other than prednisolone will be converted to prednisolone equivalent dosing using a conversion table (see Section 9.1). Use of injectable steroids for rheumatological or orthopaedic conditions will be specifically excluded.
	Analysis: In each group, the average annual cumulative dose will be estimated as the total annual SCS dose during the outcome evaluation period divided by the total person-time of follow-up. The average annual cumulative dose in each treatment group will then compared using an appropriate generalised linear mixed model (see section 6.8.3).
	Analysis timeframe: Start date: 6 months after practice initiation date; End date: end of outcome evaluation period
	Data source to be used: This outcome will be analysed exclusively using primary care EMR data. Please refer to section 6.4 for discussion.
	Diagnostic codes and/or code list, and rule set (if appropriate): 'dl_oral_steroids' tab of 'PREVAIL code lists_2026.xlsx'
	Analysis sets included in analysis (see Section 5 and section 7 for description. Unless specified, analysis sets include already diagnosed and newly diagnosed patients): • Primary analysis set (PAS): intervened early MHR (intervention) vs all identified MHR (control) • Secondary analysis set 1: intervened early MHR (intervention) vs intervened late MHR (control) • Secondary analysis set 2: intervened early MHR AD patients (intervention) vs all identified MHR AD patients (control) • Secondary analysis set 3: intervened early MHR ND patients (intervention) vs all identified MHR ND patients (control) • Exploratory analysis set 1: intervened early MHR ND patients (intervention) vs intervened late MHR ND patients (control) • Exploratory analysis set 2: implemented any time MHR (intervention) vs all identified MHR (control)

	• Exploratory analysis set 3: implemented any time MHR (intervention) vs intervened late MHR (control) • Exploratory analysis set 4: questionnaire/QI report MHR (intervention) vs intervened late MHR (control) • Exploratory analysis set 5: all identified MHR (intervention vs control) • Exploratory analysis set 7: broad already/undiagnosed +/-MHR (intervention vs control)

Table 5: Annual rate of SCS courses (secondary objective 4)

Secondary objective 4: To assess the impact of the CONQUEST programme on the annual rate of systemic corticosteroid (SCS) courses to treat COPD exacerbations, compared to usual clinical care	**Endpoint:** The number of SCS courses will be measured as the number of prescriptions of prednisolone or equivalent taken via oral/injection route available in primary care electronic medical records (EMR). The annual rate is the number of courses related to exacerbations i.e. regimens started within 3 days of a coded respiratory event or specific exacerbation code.
	Analysis: In each group, the incidence rates will be estimated as the total number of SCS courses during the outcome evaluation period divided by the total person-time of follow-up. These rates will then be reported as annual rates, that is the rate per patient per year. The effectiveness of the intervention, as compared to usual care, will be measured as the rate ratio as described in section 6.7 .
	Analysis timeframe: Start date: 6 months after practice initiation date; End date: end of outcome evaluation period
	Data source to be used: This outcome will be analysed exclusively using primary care EMR data. Please refer to section 6.4 for discussion.
	Diagnostic codes and/or code list, and rule set (if appropriate): 'dl_oral_steroids' tab of 'PREVAIL code lists_2026.xlsx'
	Analysis sets included in analysis (see section 5 and section 7 for description. Unless specified, analysis sets include already diagnosed and newly diagnosed patients): • Primary analysis set (PAS): intervened early MHR (intervention) vs all identified MHR (control) • Secondary analysis set 1: intervened early MHR (intervention) vs intervened late MHR (control) • Secondary analysis set 2: intervened early MHR AD patients (intervention) vs all identified MHR AD patients (control) • Secondary analysis set 3: intervened early MHR ND patients (intervention) vs all identified MHR ND patients (control) • Exploratory analysis set 1: intervened early MHR ND patients (intervention) vs intervened late MHR ND patients (control) • Exploratory analysis set 2: implemented any time MHR (intervention) vs all identified MHR (control) • Exploratory analysis set 3: implemented any time MHR (intervention) vs intervened late MHR (control)

	• Exploratory analysis set 4: questionnaire/QI report MHR (intervention) vs intervened late MHR (control) • Exploratory analysis set 5: all identified MHR (intervention vs control) • Exploratory analysis set 7: broad already/undiagnosed +/-MHR (intervention vs control)

Table 6: Incidence of type 2 diabetes (secondary objective 5)

Secondary objective 5: To assess the impact of the CONQUEST programme on incident type 2 diabetes, compared to usual clinical care	**Endpoint:** The rate of incident type 2 diabetes (T2D) recorded during the outcome evaluation period. Incident diagnoses are the first occurrence of a diagnostic code in a patient's available primary/secondary care records.
	Analysis: Cumulative incidence of T2D will be reported over the outcome evaluation period. Incidence of T2D will be analysed using a Cox regression model, which will be reported as a time to first event with a Kaplan Meier plot. The effectiveness of the intervention, as compared to usual care, will be measured as a hazard ratio as specified in section 6.8.5.
	Analysis timeframe: Start date: 6 months after practice initiation date; End date: end of outcome evaluation period
	Data source to be used: This outcome will be analysed exclusively using primary care EMR data. Please refer to section 6.4 for discussion.
	Diagnostic codes and/or code list, and rule set (if appropriate): • 'dl_diab_t2' tab of 'PREVAIL code lists_2026.xlsx'
	Analysis sets included in analysis (see section 5 and section 7 for description. Unless specified, analysis sets include already diagnosed and newly diagnosed patients): • Primary analysis set (PAS): intervened early MHR (intervention) vs all identified MHR (control) • Secondary analysis set 1: intervened early MHR (intervention) vs intervened late MHR (control) • Secondary analysis set 2: intervened early MHR AD patients (intervention) vs all identified MHR AD patients (control) • Secondary analysis set 3: intervened early MHR ND patients (intervention) vs all identified MHR ND patients (control) • Exploratory analysis set 1: intervened early MHR ND patients (intervention) vs intervened late MHR ND patients (control) • Exploratory analysis set 2: implemented any time MHR (intervention) vs all identified MHR (control) • Exploratory analysis set 3: implemented any time MHR (intervention) vs intervened late MHR (control) • Exploratory analysis set 4: questionnaire/QI report MHR (intervention) vs intervened late MHR (control) • Exploratory analysis set 5: all identified MHR (intervention vs control) • Exploratory analysis set 7: broad already/undiagnosed +/-MHR (intervention vs control)

Table 7: Time to first moderate or severe exacerbation (secondary objective 6)

Secondary objective 6: To evaluate the impact of the CONQUEST programme on the time to first moderate or severe COPD exacerbation, compared to usual clinical care	**Endpoint:** The term "exacerbation" will be used to describe both COPD exacerbations in diagnosed patients and respiratory events in undiagnosed patients, requiring oral corticosteroids and/or antibiotics, or resulting in hospital visits and admissions. Exacerbations* will be categorized as: Moderate exacerbations defined as ● Record of prescription with ORAL prednisolone at an acute dose (prescription for a 3- to-14-day regimen (25-50 mg/d) or equivalent) and/or ● Record of prescription for antibiotic within ± 3 days of a lower respiratory consultation, ● Hospital/emergency room visit for a COPD-related respiratory event (no admittance), or Severe exacerbations defined as ● Record of hospitalisation or death for a COPD-related respiratory event
	Analysis: Cumulative incidence of moderate/severe COPD exacerbations will be reported over the outcome evaluation period. Incidence of COPD exacerbations will be analysed using a Cox regression model, which will be reported as a time to first event with a Kaplan Meier plot. The effectiveness of the intervention, as compared to usual care, will be measured as a hazard ratio as specified in section 6.8.6. Secondary analyses will be conducted assessing incidence of moderate and severe COPD exacerbations separately, expressed as time to first event. Incidence of moderate and severe COPD exacerbations will be reported over the outcome evaluation period.
	Analysis timeframe: Start date: 6 months after practice initiation date; End date: end of outcome evaluation period
	Data source to be used: This outcome will be analysed exclusively using primary care EMR data. Please refer to section 6.4 for discussion.
	Diagnostic codes and/or code list, and rule set (if appropriate): Exacerbations: 'cl_conquest_exacerbation_listv2' tab of 'PREVAIL code lists_2026.xlsx' Exacerbations of diagnosed COPD and of potential COPD will be identified using respiratory codes included in the endpoint algorithm (see Section 9.4).

	Analysis sets included in analysis (see section 5 and section 7 for description. Unless specified, analysis sets include already diagnosed and newly diagnosed patients): • Primary analysis set (PAS): intervened early MHR (intervention) vs all identified MHR (control) • Secondary analysis set 1: intervened early MHR (intervention) vs intervened late MHR (control) • Secondary analysis set 2: intervened early MHR AD patients (intervention) vs all identified MHR AD patients (control) • Secondary analysis set 3: intervened early MHR ND patients (intervention) vs all identified MHR ND patients (control) • Exploratory analysis set 1: intervened early MHR ND patients (intervention) vs intervened late MHR ND patients (control) • Exploratory analysis set 2: implemented any time MHR (intervention) vs all identified MHR (control) • Exploratory analysis set 3: implemented any time MHR (intervention) vs intervened late MHR (control) • Exploratory analysis set 4: questionnaire/QI report MHR (intervention) vs intervened late MHR (control) • Exploratory analysis set 5: all identified MHR (intervention vs control) • Exploratory analysis set 7: broad already/undiagnosed +/-MHR (intervention vs control)

Table 8: Explanatory diagnoses amongst patients who were COPD negative after case-finding spirometry (exploratory objective 1)

Exploratory objective 1: To estimate rates of incident explanatory diagnoses other than COPD	**Endpoint:** The rate of individual incident explanatory diagnoses other than COPD, recorded during the 6-month period after the first spirometry clinic. This endpoint is restricted to initially undiagnosed potential MHR COPD patients, who attended a case-finding spirometry clinic and were COPD negative, and who remained COPD negative 6 months after the first spirometry clinic or at the time of censoring due to leaving practice or death. 'Explanatory' diagnoses are those that may explain why patients were identified as experiencing potential COPD exacerbations i.e. recorded episodes of respiratory symptoms treated with courses of antibiotics and/or steroids. Incident diagnoses are the first occurrence of a diagnostic code in a patient's available primary/secondary care records. Incident explanatory diagnoses will include: • Asthma • Lung cancer • Heart failure Other potential explanatory diagnoses will be explored and reported, if recorded for a minimum of 10 patients in the intervention arm. Small number suppression will be applied to counts in the control arm if appropriate.
	Analysis: Analyses will be restricted to the intervention arm, as case-finding spirometry clinics were not conducted within the control arm practices until the end of the outcome evaluation period. Incident rate of explanatory diagnoses will be calculated with 95% confidence intervals using the exact Poisson method. Incidence rates and 95% confidence intervals will be presented.
	Analysis timeframe: Start date: 6 months after practice initiation date; End date: end of outcome evaluation period
	Data source to be used: This outcome will be analysed exclusively using primary care EMR data. Please refer to section 6.4 for discussion.
	Diagnostic codes and/or code list, and rule set (if appropriate): For all below code lists, the first occurrence of a diagnostic code in a patient's available primary care record will be defined as an incident diagnosis.

	Asthma: 'cl_asthma' tab of 'PREVAIL code lists_2026.xlsx' Lung cancer: 'cl_lung_cancer' tab of 'PREVAIL code lists_2026.xlsx' Heart failure: 'cl_heart_failure' tab of 'PREVAIL code lists_2026.xlsx'
	Analysis sets included in analysis (see section 7 for description): • Exploratory analysis set 6: undiagnosed deemed COPD negative at spirometry clinic (intervention)

Table 9: Annual rate of unscheduled respiratory care events (exploratory objective 2)

Exploratory objective 2: To evaluate the impact of the CONQUEST programme on the annual rate of unscheduled respiratory care events, compared to usual clinical care	**Endpoint:** Unscheduled respiratory care events will be defined as any of the following: • Respiratory-related A&E attendance • Respiratory-related unscheduled hospitalisation, except outpatient • Respiratory-related primary care attendance with no supporting COPD review code or record of dyspnoea (mMRC)/disease-specific health status (COPD Assessment Tool) score
	Analysis: In each group, the incidence rates will be estimated as the total number of unscheduled respiratory care events during the outcome evaluation period divided by the total person-time of follow-up. These rates will then be reported as annual rates, that is the rate per patient per year. The effectiveness of the intervention, as compared to usual care, will be measured as the rate ratio as described in section 6.9.2.
	Analysis timeframe: Start date: 6 months after practice initiation date; End date: end of outcome evaluation period
	Data source to be used: This outcome will be analysed exclusively using primary care EMR data. Please refer to section 6.4 for discussion.
	Diagnostic codes and/or code list, and rule set (if appropriate): <u>Primary care EMR data:</u> *Respiratory-related primary care attendance:* • 'cl_resp_pc_attend' tab of 'PREVAIL code lists_2026.xlsx'
	Analysis sets included in analysis (see section 5 and section 7 for description. Unless specified, analysis sets include already diagnosed and newly diagnosed patients): • Primary analysis set (PAS): intervened early MHR (intervention) vs all identified MHR (control) • Secondary analysis set 1: intervened early MHR (intervention) vs intervened late MHR (control) • Secondary analysis set 2: intervened early MHR AD patients (intervention) vs all identified MHR AD patients (control) • Secondary analysis set 3: intervened early MHR ND patients (intervention) vs all identified MHR ND patients (control)

	• Exploratory analysis set 1: intervened early MHR ND patients (intervention) vs intervened late MHR ND patients (control) • Exploratory analysis set 2: implemented any time MHR (intervention) vs all identified MHR (control) • Exploratory analysis set 3: implemented any time MHR (intervention) vs intervened late MHR (control) • Exploratory analysis set 4: questionnaire/QI report MHR (intervention) vs intervened late MHR (control) • Exploratory analysis set 5: all identified MHR (intervention vs control)

Table 10: Lung function (exploratory objective 3)

Exploratory objective 3: To evaluate the impact of the CONQUEST programme on lung function, compared to usual clinical care	**Endpoint:** Lung function will be assessed using spirometry values. Where recorded, values with specific post-bronchodilator codes will be used for analysis. In the absence of specific pre-/post-bronchodilator coding, values from patients on inhaled therapy will be classified as post-bronchodilator. The following lung function spirometry endpoints will be analysed: • FEV1 % predicted (main value of interest) • FEV1 in millilitres • FVC in millilitres • FVC % predicted • FEV1/FVC ratio, % • PEF in litres • PEF % predicted
	Analysis: Regression models will be conducted for all parameters, and each model will use all available data in patients' EMR. Linear mixed models will be used to estimate the between-group difference (intervention vs control) at 6, 12, 24 months and end of outcome evaluation period and, within-group differences (practice initiation date to end of outcome evaluation period).
	Analysis timeframe: Between-group comparison – 6, 12, and 24 months after practice initiation date Within-group comparison – Start date: practice initiation date; End date: 24 months after practice initiation date
	Data source to be used: This outcome will be analysed exclusively using primary care EMR data. Please refer to section 6.4 for discussion.
	Diagnostic codes and/or code list, and rule set (if appropriate): 'cl_spiro' tab of 'PREVAIL code lists_2026.xlsx'
	Analysis sets included in analysis (see section 5 and section 7 for description. Unless specified, analysis sets include already diagnosed and newly diagnosed patients): • Primary analysis set (PAS): intervened early MHR (intervention) vs all identified MHR (control) • Secondary analysis set 1: intervened early MHR (intervention) vs intervened late MHR (control)

	• Secondary analysis set 2: intervened early MHR AD patients (intervention) vs all identified MHR AD patients (control) • Secondary analysis set 3: intervened early MHR ND patients (intervention) vs all identified MHR ND patients (control) • Exploratory analysis set 1: intervened early MHR ND patients (intervention) vs intervened late MHR ND patients (control) • Exploratory analysis set 2: implemented any time MHR (intervention) vs all identified MHR (control) • Exploratory analysis set 3: implemented any time MHR (intervention) vs intervened late MHR (control) • Exploratory analysis set 4: questionnaire/QI report MHR (intervention) vs intervened late MHR (control) • Exploratory analysis set 5: all identified MHR (intervention vs control)

Table 11: COPD health status (exploratory objective 4)

Exploratory objective 4: To evaluate the impact of the CONQUEST programme on COPD health status, compared to usual clinical care	**Endpoint:** COPD health status will be assessed using the COPD Assessment Test (CAT). The CAT is an 8-question instrument measured on a 0-5 point Likert scale with a total score of 0 – 40. Higher scores indicate worse health-related quality of life.
	Analysis: Regression models will be conducted for all parameters, and each model will use all available data in patients' EMR. Linear mixed models will be used to estimate the between-group difference (intervention vs control) at 6, 12, 24 months and end of outcome evaluation period and, within-group differences (practice initiation date to end of outcome evaluation period).
	Analysis timeframe: Between-group comparison – 6, 12, 24 months after practice initiation date and end of study outcome evaluation period Within-group comparison – Start date: practice initiation date; End date: study end date
	Data source to be used: This outcome will be analysed exclusively using primary care EMR data. Please refer to section 6.4 for discussion.
	Diagnostic codes and/or code list, and rule set (if appropriate): 'cl_CAT' tab of 'PREVAIL code lists_2026.xlsx'
	Analysis sets included in analysis (see section 5 and section 7 for description. Unless specified, analysis sets include already diagnosed and newly diagnosed patients): • Primary analysis set (PAS): intervened early MHR (intervention) vs all identified MHR (control) • Secondary analysis set 1: intervened early MHR (intervention) vs intervened late MHR (control) • Secondary analysis set 2: intervened early MHR AD patients (intervention) vs all identified MHR AD patients (control) • Secondary analysis set 3: intervened early MHR ND patients (intervention) vs all identified MHR ND patients (control) • Exploratory analysis set 1: intervened early MHR ND patients (intervention) vs intervened late MHR ND patients (control) • Exploratory analysis set 2: implemented any time MHR (intervention) vs all identified MHR (control) • Exploratory analysis set 3: implemented any time MHR (intervention) vs intervened late MHR (control) • Exploratory analysis set 4: questionnaire/QI report MHR (intervention) vs intervened late MHR (control) • Exploratory analysis set 5: all identified MHR (intervention vs control)

Table 12: Incident rate of pneumonia (exploratory objective 5)

Exploratory objective 5: To assess the impact of the CONQUEST programme on incident pneumonia, compared to usual clinical care	**Endpoint:** The rate of incident pneumonia recorded during the outcome evaluation period. Incident diagnoses are the first occurrence of a diagnostic code in a patient's available primary/secondary record.
	Analysis: Cumulative incidence of pneumonia will be reported over the outcome evaluation period. Incidence of pneumonia will be analysed using a Cox regression model, which will be reported as a time to first event with a Kaplan Meier plot. The effectiveness of the intervention, as compared to usual care, will be measured as a hazard ratio as specified in section 6.9.5.
	Analysis timeframe: Start date: 6 months after practice initiation date; End date: end of outcome evaluation period
	Data source to be used: This outcome will be analysed exclusively using linked primary and secondary care data. Please refer to section 6.4 for discussion.
	Diagnostic codes and/or code list, and rule set (if appropriate): <u>Secondary care data (ICD-10 codes):</u> J100, J110, J12-18, J851 <u>Primary care EMR data:</u> 'cl_pneumonia' tab of 'PREVAIL code lists_2026.xlsx'
	Analysis sets included in analysis (see section 5 and section 7 for description. Unless specified, analysis sets include already diagnosed and newly diagnosed patients): • Primary analysis set (PAS): intervened early MHR (intervention) vs all identified MHR (control) • Secondary analysis set 1: intervened early MHR (intervention) vs intervened late MHR (control) • Secondary analysis set 2: intervened early MHR AD patients (intervention) vs all identified MHR AD patients (control) • Secondary analysis set 3: intervened early MHR ND patients (intervention) vs all identified MHR ND patients (control) • Exploratory analysis set 1: intervened early MHR ND patients (intervention) vs intervened late MHR ND patients (control)

	• Exploratory analysis set 2: implemented any time MHR (intervention) vs all identified MHR (control) • Exploratory analysis set 3: implemented any time MHR (intervention) vs intervened late MHR (control) • Exploratory analysis set 4: questionnaire/QI report MHR (intervention) vs intervened late MHR (control) • Exploratory analysis set 5: all identified MHR (intervention vs control)

3 Study design

PREVAIL is a multicentre, cluster randomised trial. GP practices (the unit of randomisation) recruited from England will be randomly allocated to the intervention arm or the control arm (routine care) of the trial.

3.1 Target population for PREVAIL trial

The target population for the PREVAIL trial includes:

- Patients aged ≥40 and ≤80 years, with a recorded COPD diagnosis prior to the start of the trial i.e. already diagnosed COPD
- Patients aged ≥40 and ≤80 years, with their first COPD diagnosis recorded since the start of the trial i.e. newly diagnosed COPD

For patients to be included in the trial as newly diagnosed COPD, they needed to have had a positive smoking history (≥10 years smoking duration or ≥10 packs years) and not be actively treated and diagnosed with asthma in the 12 months prior to COPD diagnosis.

Patients are eligible for the trial if they meet the below modifiable high-risk (MHR) criteria prior to the start of the trial:

- are prescribed dual COPD maintenance therapy or less, and
- have evidence of ≥2 moderate or ≥1 severe exacerbations (or potential exacerbations) in the previous 24 months, including at least one exacerbation in the last 12 months of the 24 month period.

Operational criteria are also applied to prevent contacting inappropriate patients, such as those with recent cancer diagnoses or cognitive impairment.

Analysis sets will be derived from identified populations; definitions including patient selection criteria for control and intervention arm practices are defined in Section 5.

3.2 Description of the treatment arms

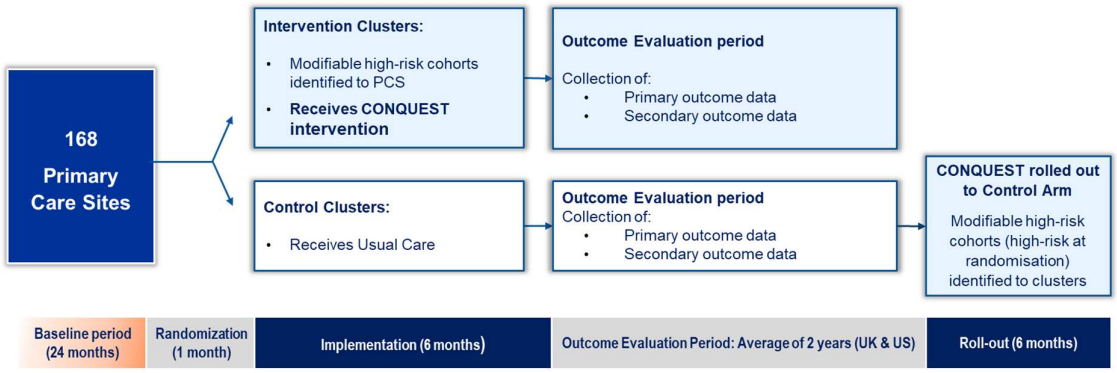


Figure 1: PREVAIL trial design overview

3.2.1 Intervention arm

CONQUEST is a quality improvement programme designed to support primary healthcare professionals to diagnose, assess, and optimise the management of their patients with modifiable high-risk COPD, through the promotion of guideline-based care.

CONQUEST will promote management according to four "Quality Standards" for the care of modifiable high-risk COPD patients, developed by the CONQUEST study team in collaboration with local and international respiratory experts on the programme Steering Committees. A detailed description of the quality standards have been previously published[1], and can be summarised as the timely and appropriate:

- Identification of modifiable high-risk target population
- Assessment of disease and quantification of future risk
- Non-pharmacological and pharmacological intervention
- Clinical follow-up

The CONQUEST programme will promote adherence to these standards through several intervention elements:

- use of electronic medical record (EMR)-based algorithms for the identification of high-risk patients,
- supported respiratory and cardiovascular assessments for these patients

Both underpinned by a clinical decision support that provides practices with tools such as assessment checklists, assimilation of guideline recommendations for COPD care, and considerations for cases where specialist referral may be appropriate.

As the intervention is a quality improvement programme, the Study Management Group or Sponsor will not be responsible for the execution of any recommendations, all of which will occur within the routine clinical care processes at the PCS. All decisions about patient management are made entirely at the discretion of the clinicians and patients who together negotiate the preferred course of action during consultations in the implementation and outcomes evaluation phases of the trial.

3.2.2 Control arm

Practices in the control arm will continue to deliver usual care to all of their patients, throughout the implementation and outcome evaluation periods, in accordance with local and national guidelines[2].

As control sites will be linked to the OPC network, they may participate in alternative quality improvement packages during the implementation and outcome evaluation periods, on condition that such packages will not affect the outcomes of interest to this trial. The CONQUEST programme, including spirometry clinics for patients with potential undiagnosed COPD, will be rolled-out to control practices near the end of the outcome evaluation period.

3.3 Randomisation

Stratified randomisation using random permuted blocks with random block size will be used to randomise the practices. Firstly, consenting sites (PCSs) will be stratified into "low" and "high"

groups based on two practice-level characteristics: deprivation (Index of Multiple Deprivation [IMD] score 2019) and opportunity for triple therapy (QOF COPD list size – number of patients currently prescribed triple therapy). For the latter, October 2020 data will be used due to the increase in triple therapy prescriptions for asthma in subsequent months. For both characteristics, the median value from nationally available data will be used as a cut-off to categorise PCSs into high and low strata.

Within strata, sites will be randomised at a ratio of 1:1 using permuted blocks of different sizes to receive either the CONQUEST programme (intervention arm; n=84 practices), or to continue to provide usual care during the outcome evaluation period (control arm; n=84 practices).

Randomisation will be overseen by a PREVAIL trial statistician not involved with the conduct of the trial.

3.4 Study time periods

An overview of the trial design is provided in Figure 1, and each phase is described below.

3.4.1 Randomisation date

The date that practices will be randomly allocated to the intervention or control arm of the trial.

3.4.2 Practice initiation date

The practice initiation date is applied at the practice-level, meaning all patients within a practice share the same initiation date.

For practices in the intervention arm the initiation date is the date that OPC advised the practice of their treatment arm allocation.

For practices in the control arm, the practice initiation date will be the same as that for the intervention practice randomised closest to them in time.

3.4.3 Risk identification period

EMR data in intervention and control arm practices will be retrospectively searched over a 24-month period prior to the practice initiation date to identify patients meeting trial eligibility criteria.

3.4.4 Implementation period

An implementation period, beginning on the practice initiation date, will be allowed for the QIP to be implemented in intervention arm practices. QIP implementation includes case-finding, clinical consultations for disease assessment, and initiation of appropriate treatment. The implementation period will be six months for already diagnosed patients and 12 months for newly diagnosed patients; allowing comparable time for QIP elements to occur after diagnosis.

3.4.5 Outcome evaluation period

The outcome evaluation period will start six months after the practice initiation date in both intervention and control practices i.e. at the end of the implementation period. Trial outcomes will be assessed during the outcome evaluation period, which is intended to last an average of 2 years across all intervention and control arm practices. Patients' outcome data will be captured between practice initiation date and the first occurrence of: patient death, end of patient registration, withdrawal of patient consent for data sharing or, trial end date.

3.4.6 Trial end date

The trial end date will be 28th February 2026, signifying the end of the outcome evaluation period.

4 Sample size and power

Sample size estimates for the primary analyses cohort are based on the primary objective – the impact of CONQUEST on annual rate of moderate or severe COPD exacerbations, over the outcome evaluation period. The below estimates are indicative of the potential power to detect impact of the intervention using existing primary care medical record data for modifiable high-risk COPD patients held in the Optimum Patient Care Research databases (OPCRD).

All sample size calculations account for the clustered nature of the trial design using appropriate formulas for event rates. Conservative parameter estimates are based on a database snapshot of modifiable high-risk patients in 2015 in 514 GP practices with electronic medical record (EMR) data in OPCRD. The anticipated yield of new COPD cases was informed by a previously developed case-finding tool[3] and subsequent modelling analysis of OPCRD data[4] conducted prior to the COVID-19 pandemic, with a more conservative yield being adopted for the PREVAIL trial in line with existing literature. The anticipated reduction in exacerbation rate was informed by previous trials[5–8] regarding the impact of different therapy combinations in high-risk COPD patients.

Calculations have assumed a maximum allowed number of clusters of 84 per therapeutic arm, and a maximum of 1,008 high-risk COPD-diagnosed patients combined arising from the initially identified already-diagnosed and undiagnosed with potential modifiable high-risk COPD cohorts.

By enrolling 168 clusters with a mean of 12 patients meeting primary analyses cohort criteria per cluster we estimate:

4.1 Annual Exacerbation Rate (combined primary analyses cohorts)

90% power to detect a 30% reduction, equivalently a rate ratio of 0.7, (and 80% power to detect a 26% reduction (rate ratio of 0.74)) in the annual rate of exacerbations in patients of the primary analyses cohort. This is based on an estimated coefficient of variation between PCSs in this cohort of 0.56 and a background (control group annual exacerbation rate) over the subsequent 2-year period of 1.36[9].

Feasibility work has suggested that more than 12 patients per site meeting the primary analyses cohort criteria at trial baseline may be identified and assessed during the intervention and these patients will be included in subsequent analyses in that situation.

5 Populations and Analysis sets

Figure 2 presents the intervention and control arm populations that are included in the primary analysis set and secondary analysis sets. Below this is a description of these analysis sets, the patient selection criteria applied to the target population specified in Section 3.1 , and the research question addressed by these comparisons. Table 13 provides a matrix showing which trial outcomes will be investigated in each analysis set. Additional exploratory analysis sets will be developed (see Section 7) to explore the impact of the CONQUEST QIP on different patient populations, and will not be used for independent claims of effectiveness.

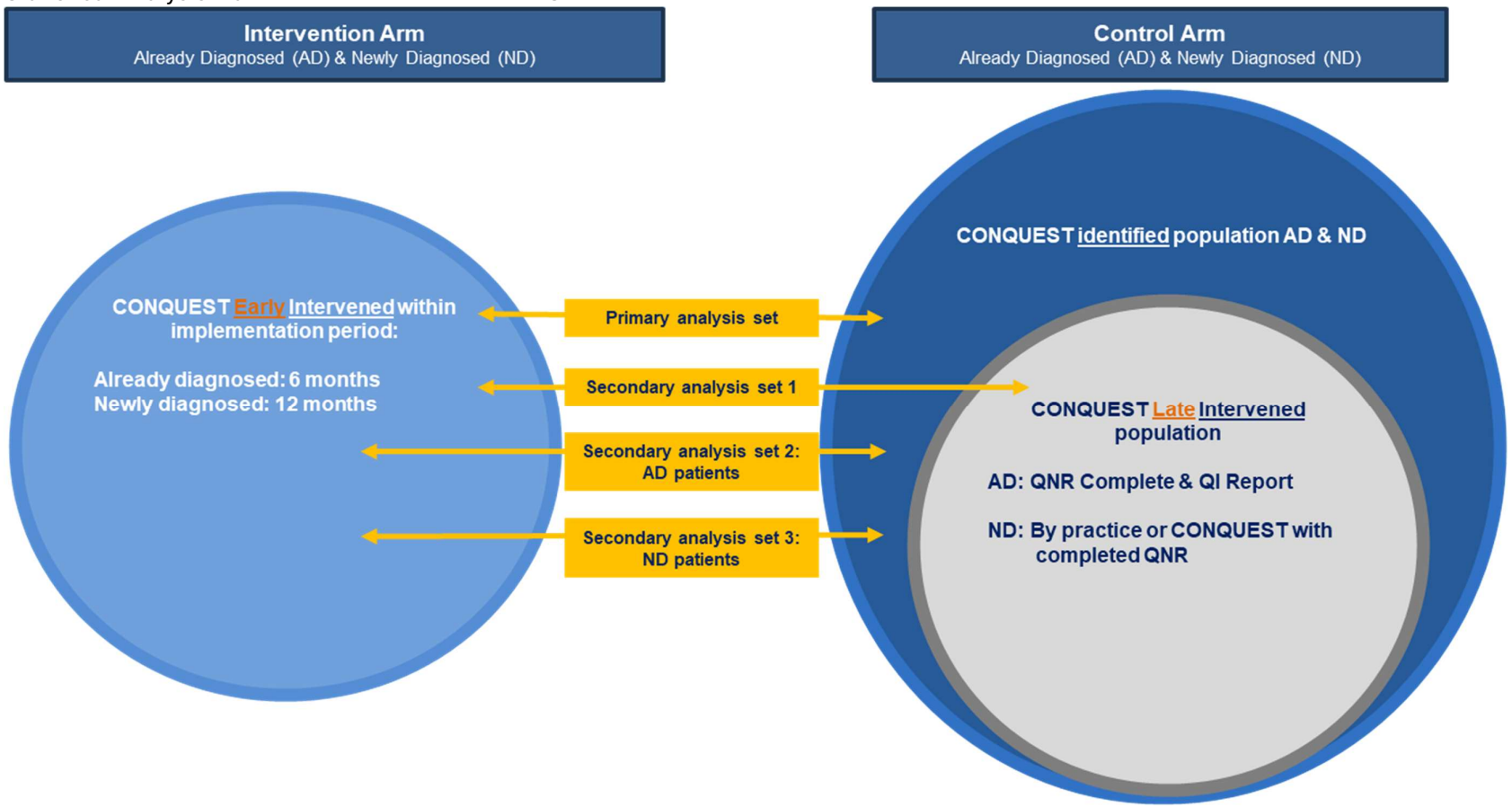


Figure 2: Populations used in primary analysis set and secondary analysis sets

5.1 Primary analysis set (PAS)

The primary analysis set (PAS) will include MHR already diagnosed and newly diagnosed COPD patients, comparing those in whom CONQUEST intervened within the implementation period in the intervention arm, against all identified patients in the control arm (see below).

Patient selection for primary analysis set

- Intervention arm
    - Already diagnosed MHR in whom CONQUEST intervened within 6 months of initiation date
    - Newly diagnosed MHR in whom CONQUEST intervened within 12 months of initiation date
- Control arm
    - Identified MHR patients with already diagnosed or newly diagnosed COPD (see Section 3.1)

Research question addressed by the primary analysis set

- What is the impact on MHR patient health outcomes of receiving the CONQUEST QIP within the intended implementation period, compared to MHR patients receiving routine clinical care?

5.2 Secondary analysis set 1: early vs late CONQUEST intervention

Secondary analysis set 1 will include MHR already diagnosed and newly diagnosed COPD patients, comparing those in whom CONQUEST intervened within the implementation period in the intervention arm, against those in whom CONQUEST intervened during the roll-out period in the control arm (see below).

- Intervention arm
    - Already diagnosed MHR in whom CONQUEST intervened within 6 months of initiation date
    - Newly diagnosed MHR in whom CONQUEST intervened within 12 months of initiation date
- Control arm
    - Already diagnosed MHR who received the CONQUEST intervention at study end
    - Newly diagnosed MHR who received the CONQUEST intervention at study end

Research question addressed by secondary analysis set 1

- What is the impact on MHR patient health outcomes of receiving the CONQUEST QIP within the intended implementation period, compared to MHR patients receiving delayed QIP intervention?

5.3 Secondary analysis set 2: already diagnosed patients in PAS population

Secondary analysis set 2 will include MHR already diagnosed COPD patients, comparing those in whom CONQUEST intervened within the implementation period in the intervention arm, against all identified already diagnosed COPD patients in the control arm (see below).

- Intervention arm
    - Already diagnosed MHR in whom CONQUEST intervened within 6 months of initiation date
- Control arm
    - Identified MHR patients with already diagnosed COPD (see Section 3.1)

Research question addressed by secondary analysis set 2

- What is the impact on MHR already diagnosed patient health outcomes of receiving the CONQUEST QIP within the intended implementation period, compared to MHR already diagnosed patients receiving routine clinical care?

5.4 Secondary analysis set 3: newly diagnosed patients in PAS population

Secondary analysis set 3 will include MHR newly diagnosed COPD patients, comparing those in whom CONQUEST intervened within the implementation period in the intervention arm, against all identified newly diagnosed COPD patients in the control arm (see below).

- Intervention arm
    - Newly diagnosed MHR in whom CONQUEST intervened within 6 months of initiation date
- Control arm
    - Identified MHR patients with newly diagnosed COPD (see Section 3.1)

Research question addressed by secondary analysis set 3

- What is the impact on MHR newly diagnosed patient health outcomes of receiving the CONQUEST QIP within the intended implementation period, compared to MHR newly diagnosed patients receiving routine clinical care?

Table 13: Primary and secondary analysis set by trial outcome matrix

	Primary Outcome	Secondary outcomes						Exploratory Outcomes				
Analysis Sets	Moderate or severe exacerbations	MACRE	MACE	SCS use	SCS courses	Type 2 diabetes	Time to exacerbation	Explanatory diagnoses	Unscheduled respiratory care events	Lung function†	CAT score	Pneumonia
Primary analysis set (PAS)	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓
Secondary analysis set 1	✓			✓	✓	✓	✓		✓	✓	✓	✓
Secondary analysis set 2	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓
Secondary analysis set 3	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓

MACRE: Major adverse cardiac / respiratory events; MACE: Major adverse cardiac events; SCS: Systemic corticosteroid; CAT: COPD assessment test

6 Statistical Analysis

6.1 General principles

All analyses will be conducted by two independent statisticians and governed by this comprehensive SAP, which will be agreed by the sponsor, independent steering committee and funder prior to any analyses being undertaken. There will be no formal interim analyses undertaken. Unless pre-specified, a 5% two-sided significance level will be used to denote statistical significance throughout. Any point estimates calculated will be presented alongside their 95% confidence intervals, for already diagnosed and newly diagnosed cohorts separately. Analyses will be conducted using patient outcome data collected during the outcome evaluation period (see Section 3.4.5) collected from GP practices according to their randomised trial arm. The scope of inference for each analysis is determined by the populations defined by the relevant analysis sets defined for each analysis set (section 5) and is detailed in the relevant sections below.

6.2 Confounding adjustment

Although practices will be broadly balanced on practice characteristics the recruited patients are not individually randomised and may show potential imbalances. To account for this source of potential confounding, each outcome will be analysed using two models; an unweighted and a propensity score weighted analysis[10] (see Table 14).

The unweighted (crude) model will include:

- A binary indicator identifying if the patient was in a CONQUEST practice or control practice
- Practice proportion of COPD patients not on triple therapy (≥median/<median) (stratification factor)
- Practice level of deprivation (IMD decile) (≥median/<median) (stratification factor)
- The practice identifier included as a clustering effect

The propensity score model will include the above variables, and a comprehensive set of candidate variables identified from baseline analysis of demographic/clinical characteristics and expert clinical opinion. Covariate selection will adopt the approach of Leyrat et al[11], to develop a parsimonious model to avoid overfitting problems.

Additional variables required for some analytical models, e.g., offset variables for the follow up time, will be described in the relevant analysis sections.

6.3 Accounting for clustering by GP practice

The primary analysis will use a random effects model with practice as a random effect (or frailty term in the time-to-event analyses) to account for clustering by practice. In the event of non-convergence, the analysis will instead use a marginal model with cluster-robust standard errors.

6.4 Use of primary care and secondary care data

Secondary care data will not be available for all practices (clusters) due to consent and technical constraints, and similarly not for all patients due to local/national data opt outs from data sharing. The estimated availability of secondary care data (<90% practices) will result in an excessive reduction in the number of practices. The preferred data source to analyse outcomes within the Primary Analysis Set will be the primary care EMR data, to maximise the number of practices included.

However, blinded comparison of key data points indicated that analysis of specific trial outcomes (MACRE, MACE, pneumonia) should be conducted exclusively on the linked primary and secondary care data, due to potential under-reporting of events within EMR data alone. The linked data will relate to a subgroup of practices and patients; hence this will not be used for independent claims of effectiveness.

6.5 Primary and secondary comparisons

The primary comparison will explore the treatment effect on all trial outcomes using the Primary Analysis Set, and the secondary comparison will repeat the same analyses using the secondary analysis sets (see section 5). Exploratory comparisons will repeat the analyses on all exploratory analysis sets (see section 7), to explore the impact on different patient populations and will not be used for independent claims of effectiveness. The matrix in Table 14 details the trial outcomes, data sources and analysis sets used within each comparison.

Table 14: Matrix of comparison hierarchy, trial outcomes, data sources and analysis sets

Comparison hierarchy	Endpoint scope	Trial outcome	Data source	Analysis set	Adjustment model
Primary	Primary	Rate of exacerbations	EMR	PAS: early intervened vs identified population	Propensity score model
Primary	Secondary 1	Time to MACRE	Linked EMR/HES data	PAS: early intervened vs identified population	Propensity score model
Primary	Secondary 2	Time to MACE	Linked EMR/HES data	PAS: early intervened vs identified population	Propensity score model
Primary	Secondary 3	SCS use	EMR	PAS: early intervened vs identified population	Propensity score model
Primary	Secondary 4	SCS courses	EMR	PAS: early intervened vs identified population	Propensity score model
Primary	Secondary 5	Type 2 diabetes	EMR	PAS: early intervened vs identified population	Propensity score model
Primary	Secondary 6	Time to exacerbation	EMR	PAS: early intervened vs identified population	Propensity score model
Primary	Exploratory 1	Explanatory diagnoses	n/a	*Outcome not analysed in this set*	n/a
Primary	Exploratory 2	Unscheduled respiratory events	EMR	PAS: early intervened vs identified population	Propensity score model
Primary	Exploratory 3	Lung function	EMR	PAS: early intervened vs identified population	Propensity score model
Primary	Exploratory 4	COPD health status	EMR	PAS: early intervened vs identified population	Propensity score model
Primary	Exploratory 5	Pneumonia	Linked EMR/HES data	PAS: early intervened vs identified population	Propensity score model
Secondary	Primary	Rate of exacerbations	EMR	Secondary analysis set 1: early vs late QIP intervened	Propensity score model
Secondary	Secondary 1	Time to MACRE	n/a	*Immortal time bias – outcome not analysed in this set*	n/a
Secondary	Secondary 2	Time to MACE	n/a	*Immortal time bias – outcome not analysed in this set*	n/a
Secondary	Secondary 3	SCS use	EMR	Secondary analysis set 1: early vs late QIP intervened	Propensity score model
Secondary	Secondary 4	SCS courses	EMR	Secondary analysis set 1: early vs late QIP intervened	Propensity score model
Secondary	Secondary 5	Type 2 diabetes	EMR	Secondary analysis set 1: early vs late QIP intervened	Propensity score model
Secondary	Secondary 6	Time to exacerbation	EMR	Secondary analysis set 1: early vs late QIP intervened	Propensity score model
Secondary	Exploratory 1	Explanatory diagnoses	n/a	*Outcome not analysed in this set*	n/a
Secondary	Exploratory 2	Unscheduled respiratory events	EMR	Secondary analysis set 1: early vs late QIP intervened	Propensity score model
Secondary	Exploratory 3	Lung function	EMR	Secondary analysis set 1: early vs late QIP intervened	Propensity score model
Secondary	Exploratory 4	COPD health status	EMR	Secondary analysis set 1: early vs late QIP intervened	Propensity score model
Secondary	Exploratory 5	Pneumonia	Linked EMR/HES data	Secondary analysis set 1: early vs late QIP intervened	Propensity score model
Secondary	Primary	Rate of exacerbations	EMR	Secondary analysis set 2: early intervened vs identified (AD)	Propensity score model
Secondary	Secondary 1	Time to MACRE	Linked EMR/HES data	Secondary analysis set 2: early intervened vs identified (AD)	Propensity score model
Secondary	Secondary 2	Time to MACE	Linked EMR/HES data	Secondary analysis set 2: early intervened vs identified (AD)	Propensity score model
Secondary	Secondary 3	SCS use	EMR	Secondary analysis set 2: early intervened vs identified (AD)	Propensity score model
Secondary	Secondary 4	SCS courses	EMR	Secondary analysis set 2: early intervened vs identified (AD)	Propensity score model
Secondary	Secondary 5	Type 2 diabetes	EMR	Secondary analysis set 2: early intervened vs identified (AD)	Propensity score model
Secondary	Secondary 6	Time to exacerbation	EMR	Secondary analysis set 2: early intervened vs identified (AD)	Propensity score model
Secondary	Exploratory 1	Explanatory diagnoses	n/a	*Outcome not analysed in this set*	n/a
Secondary	Exploratory 2	Unscheduled respiratory events	EMR	Secondary analysis set 2: early intervened vs identified (AD)	Propensity score model
Secondary	Exploratory 3	Lung function	EMR	Secondary analysis set 2: early intervened vs identified (AD)	Propensity score model

Secondary	Exploratory 4	COPD health status	EMR	Secondary analysis set 2: early intervened vs identified (AD)	Propensity score model
Secondary	Exploratory 5	Pneumonia	Linked EMR/HES data	Secondary analysis set 2: early intervened vs identified (AD)	Propensity score model
Secondary	Primary	Rate of exacerbations	EMR	Secondary analysis set 3: early intervened vs identified (ND)	Propensity score model
Secondary	Secondary 1	Time to MACRE	Linked EMR/HES data	Secondary analysis set 3: early intervened vs identified (ND)	Propensity score model
Secondary	Secondary 2	Time to MACE	Linked EMR/HES data	Secondary analysis set 3: early intervened vs identified (ND)	Propensity score model
Secondary	Secondary 3	SCS use	EMR	Secondary analysis set 3: early intervened vs identified (ND)	Propensity score model
Secondary	Secondary 4	SCS courses	EMR	Secondary analysis set 3: early intervened vs identified (ND)	Propensity score model
Secondary	Secondary 5	Type 2 diabetes	EMR	Secondary analysis set 3: early intervened vs identified (ND)	Propensity score model
Secondary	Secondary 6	Time to exacerbation	EMR	Secondary analysis set 3: early intervened vs identified (ND)	Propensity score model
Secondary	Exploratory 1	Explanatory diagnoses	n/a	*Outcome not analysed in this set*	n/a
Secondary	Exploratory 2	Unscheduled respiratory events	EMR	Secondary analysis set 3: early intervened vs identified (ND)	Propensity score model
Secondary	Exploratory 3	Lung function	EMR	Secondary analysis set 3: early intervened vs identified (ND)	Propensity score model
Secondary	Exploratory 4	COPD health status	EMR	Secondary analysis set 3: early intervened vs identified (ND)	Propensity score model
Secondary	Exploratory 5	Pneumonia	Linked EMR/HES data	Secondary analysis set 3: early intervened vs identified (ND)	Propensity score model
Exploratory	Primary	Rate of exacerbations	EMR	Exploratory analysis set 1	Propensity score model
Exploratory	Secondary 1	Time to MACRE	Linked EMR/HES data	*Immortal time bias – outcome not analysed in this set*	n/a
Exploratory	Secondary 2	Time to MACE	Linked EMR/HES data	*Immortal time bias – outcome not analysed in this set*	n/a
Exploratory	Secondary 3	SCS use	EMR	Exploratory analysis set 1	Propensity score model
Exploratory	Secondary 4	SCS courses	EMR	Exploratory analysis set 1	Propensity score model
Exploratory	Secondary 5	Type 2 diabetes	EMR	Exploratory analysis set 1	Propensity score model
Exploratory	Secondary 6	Time to exacerbation	EMR	Exploratory analysis set 1	Propensity score model
Exploratory	Exploratory 1	Explanatory diagnoses	n/a	*Outcome not analysed in this set*	n/a
Exploratory	Exploratory 2	Unscheduled respiratory events	EMR	Exploratory analysis set 1	Propensity score model
Exploratory	Exploratory 3	Lung function	EMR	Exploratory analysis set 1	Propensity score model
Exploratory	Exploratory 4	COPD health status	EMR	Exploratory analysis set 1	Propensity score model
Exploratory	Exploratory 5	Pneumonia	Linked EMR/HES data	Exploratory analysis set 1	Propensity score model
Exploratory	Primary	Rate of exacerbations	EMR	Exploratory analysis set 2	Propensity score model
Exploratory	Secondary 1	Time to MACRE	n/a	Exploratory analysis set 2	Propensity score model
Exploratory	Secondary 2	Time to MACE	n/a	Exploratory analysis set 2	Propensity score model
Exploratory	Secondary 3	SCS use	EMR	Exploratory analysis set 2	Propensity score model
Exploratory	Secondary 4	SCS courses	EMR	Exploratory analysis set 2	Propensity score model
Exploratory	Secondary 5	Type 2 diabetes	EMR	Exploratory analysis set 2	Propensity score model
Exploratory	Secondary 6	Time to exacerbation	EMR	Exploratory analysis set 2	Propensity score model
Exploratory	Exploratory 1	Explanatory diagnoses	n/a	*Outcome not analysed in this set*	n/a
Exploratory	Exploratory 2	Unscheduled respiratory events	EMR	Exploratory analysis set 2	Propensity score model
Exploratory	Exploratory 3	Lung function	EMR	Exploratory analysis set 2	Propensity score model
Exploratory	Exploratory 4	COPD health status	EMR	Exploratory analysis set 2	Propensity score model
Exploratory	Exploratory 5	Pneumonia	Linked EMR/HES data	Exploratory analysis set 2	Propensity score model

Exploratory	Primary	Rate of exacerbations	EMR	Exploratory analysis set 3	Propensity score model
Exploratory	Secondary 1	Time to MACRE	n/a	*Immortal time bias – outcome not analysed in this set*	n/a
Exploratory	Secondary 2	Time to MACE	n/a	*Immortal time bias – outcome not analysed in this set*	n/a
Exploratory	Secondary 3	SCS use	EMR	Exploratory analysis set 3	Propensity score model
Exploratory	Secondary 4	SCS courses	EMR	Exploratory analysis set 3	Propensity score model
Exploratory	Secondary 5	Type 2 diabetes	EMR	Exploratory analysis set 3	Propensity score model
Exploratory	Secondary 6	Time to exacerbation	EMR	Exploratory analysis set 3	Propensity score model
Exploratory	Exploratory 1	Explanatory diagnoses	n/a	*Outcome not analysed in this set*	n/a
Exploratory	Exploratory 2	Unscheduled respiratory events	EMR	Exploratory analysis set 3	Propensity score model
Exploratory	Exploratory 3	Lung function	EMR	Exploratory analysis set 3	Propensity score model
Exploratory	Exploratory 4	COPD health status	EMR	Exploratory analysis set 3	Propensity score model
Exploratory	Exploratory 5	Pneumonia	Linked EMR/HES data	Exploratory analysis set 3	Propensity score model
Exploratory	Primary	Rate of exacerbations	EMR	Exploratory analysis set 4	Propensity score model
Exploratory	Secondary 1	Time to MACRE	Linked EMR/HES data	*Immortal time bias – outcome not analysed in this set*	n/a
Exploratory	Secondary 2	Time to MACE	Linked EMR/HES data	*Immortal time bias – outcome not analysed in this set*	n/a
Exploratory	Secondary 3	SCS use	EMR	Exploratory analysis set 4	Propensity score model
Exploratory	Secondary 4	SCS courses	EMR	Exploratory analysis set 4	Propensity score model
Exploratory	Secondary 5	Type 2 diabetes	EMR	Exploratory analysis set 4	Propensity score model
Exploratory	Secondary 6	Time to exacerbation	EMR	Exploratory analysis set 4	Propensity score model
Exploratory	Exploratory 1	Explanatory diagnoses	n/a	*Outcome not analysed in this set*	n/a
Exploratory	Exploratory 2	Unscheduled respiratory events	EMR	Exploratory analysis set 4	Propensity score model
Exploratory	Exploratory 3	Lung function	EMR	Exploratory analysis set 4	Propensity score model
Exploratory	Exploratory 4	COPD health status	EMR	Exploratory analysis set 4	Propensity score model
Exploratory	Exploratory 5	Pneumonia	Linked EMR/HES data	Exploratory analysis set 4	Propensity score model
Exploratory	Primary	Rate of exacerbations	EMR	Exploratory analysis set 5	Propensity score model
Exploratory	Secondary 1	Time to MACRE	Linked EMR/HES data	Exploratory analysis set 5	Propensity score model
Exploratory	Secondary 2	Time to MACE	Linked EMR/HES data	Exploratory analysis set 5	Propensity score model
Exploratory	Secondary 3	SCS use	EMR	Exploratory analysis set 5	Propensity score model
Exploratory	Secondary 4	SCS courses	EMR	Exploratory analysis set 5	Propensity score model
Exploratory	Secondary 5	Type 2 diabetes	EMR	Exploratory analysis set 5	Propensity score model
Exploratory	Secondary 6	Time to exacerbation	EMR	Exploratory analysis set 5	Propensity score model
Exploratory	Exploratory 1	Explanatory diagnoses	n/a	*Outcome not analysed in this set*	n/a
Exploratory	Exploratory 2	Unscheduled respiratory events	EMR	Exploratory analysis set 5	Propensity score model
Exploratory	Exploratory 3	Lung function	EMR	Exploratory analysis set 5	Propensity score model
Exploratory	Exploratory 4	COPD health status	EMR	Exploratory analysis set 5	Propensity score model
Exploratory	Exploratory 5	Pneumonia	Linked EMR/HES data	Exploratory analysis set 5	Propensity score model
Exploratory	Primary	Rate of exacerbations	n/a	*Outcome not analysed in this set*	n/a
Exploratory	Secondary 1	Time to MACRE	n/a	*Outcome not analysed in this set*	n/a

Exploratory	Secondary 2	Time to MACE	n/a	*Outcome not analysed in this set*	n/a
Exploratory	Secondary 3	SCS use	n/a	*Outcome not analysed in this set*	n/a
Exploratory	Secondary 4	SCS courses	n/a	*Outcome not analysed in this set*	n/a
Exploratory	Secondary 5	Type 2 diabetes	n/a	*Outcome not analysed in this set*	n/a
Exploratory	Secondary 6	Time to exacerbation	n/a	*Outcome not analysed in this set*	n/a
Exploratory	Exploratory 1	Explanatory diagnoses	EMR	Exploratory analysis set 6	Propensity score model
Exploratory	Exploratory 2	Unscheduled respiratory events	n/a	*Outcome not analysed in this set*	n/a
Exploratory	Exploratory 3	Lung function	n/a	*Outcome not analysed in this set*	n/a
Exploratory	Exploratory 4	COPD health status	n/a	*Outcome not analysed in this set*	n/a
Exploratory	Exploratory 5	Pneumonia	n/a	*Outcome not analysed in this set*	n/a
Exploratory	Primary	Rate of exacerbations	EMR	Exploratory analysis set 7	Propensity score model
Exploratory	Secondary 1	Time to MACRE	Linked EMR/HES data	Exploratory analysis set 7	Propensity score model
Exploratory	Secondary 2	Time to MACE	Linked EMR/HES data	Exploratory analysis set 7	Propensity score model
Exploratory	Secondary 3	SCS use	EMR	Exploratory analysis set 7	Propensity score model
Exploratory	Secondary 4	SCS courses	EMR	Exploratory analysis set 7	Propensity score model
Exploratory	Secondary 5	Type 2 diabetes	EMR	Exploratory analysis set 7	Propensity score model
Exploratory	Secondary 6	Time to exacerbation	EMR	Exploratory analysis set 7	Propensity score model
Exploratory	Exploratory 1	Explanatory diagnoses	n/a	*Outcome not analysed in this set*	n/a
Exploratory	Exploratory 2	Unscheduled respiratory events	EMR	Exploratory analysis set 7	Propensity score model
Exploratory	Exploratory 3	Lung function	EMR	Exploratory analysis set 7	Propensity score model
Exploratory	Exploratory 4	COPD health status	EMR	Exploratory analysis set 7	Propensity score model
Exploratory	Exploratory 5	Pneumonia	Linked EMR/HES data	Exploratory analysis set 7	Propensity score model

6.6 Descriptive analysis

A CONSORT flow diagram will describe the flow of trial participants, withdrawals and post-randomisation exclusions.

Categorical variables will be described with number and % in each category. Continuous variables will be described with median, inter-quartile range (IQR) and minimum/maximum values The amount of missing data will be provided for each variable. Baseline data will be reported at the cluster level and individual patient level

In each treatment group, the incidence rates will be estimated as the total number of outcome events during the outcome evaluation period divided by the total person-time of follow-up (the sum of time that each patient is in the study). These rates will then be reported as annual rates, that is the rate per patient per year, and 95% confidence intervals will be calculated using the exact Poisson method. For the practice-level analysis, each practice will have the incidence rate and 95% confidence intervals calculated using data only from that practice. The mean incidence rate and 95% confidence interval for each treatment group will be calculated from the individual practice means using the methods described by Hayes and Moulton (2017)[12]. Both the individual- and practice-level results will be reported

6.7 Analysis of the primary outcome

The analysis of moderate and severe COPD exacerbations will be performed using a random effects Poisson model with the number of exacerbations as the outcome, a random effect for GP practice to account for the cluster trial design, and an offset variable containing the logarithm of the follow up time for each patient. Should significant overdispersion be detected, the analysis will use a negative binomial model[13].

An unweighted and propensity score (PS) weighted model will be analysed with the included covariates described in Section 6.2 . Unweighted and weighted rate ratios, 95% confidence intervals and p-values will be reported.

In addition to the primary analysis, the occurrence of moderate and severe exacerbations will be analysed separately using the same analytical model as the primary analysis.

6.8 Analysis of the secondary outcomes

6.8.1 MACRE events (secondary objective 1)

Cumulative incidence of the composite MACRE outcome will be reported over the outcome evaluation period. The distribution of time to events between the two arms will be compared using Kaplan Meier curves.

Incidence of MACRE will be analysed using a Cox regression model, with a frailty term to account for clustering by GP practice. The effectiveness of the intervention, as compared to usual care, will be measured as a hazard ratio. Unweighted and PS weighted hazard ratios, 95% confidence intervals and p-values will be reported for all analyses.

The primary analysis will be conducted for overall MACRE (combined events) incidence. In addition, a further analysis will be conducted for the incidence of individual components, namely:

- New heart failure diagnosis
- Hospitalisation for heart failure
- Non-fatal myocardial infarction
- Non-fatal ischaemic stroke
- All-cause mortality and further disaggregated by:
    - Cardiac death
    - Sudden death
    - Respiratory death
    - Other causes
- Hospitalization for respiratory events and complicated exacerbations

We will assess the proportional hazards assumption in survival analyses by performing the Grambsch-Therneau test and examining plots of the scaled Schoenfeld residuals against time. If the assumption of proportional hazards is not met an alternative flexible parametric model will be used.

6.8.2 MACE events (secondary objective 2)

Cumulative incidence of the composite MACE outcome will be reported over the outcome evaluation period. The distribution of time to events between the two arms will be compared using Kaplan Meier curves.

Incidence of MACE will be analysed using a Cox regression model, with a frailty term to account for clustering by GP practice. The effectiveness of the intervention, as compared to usual care, will be measured as a hazard ratio. Unweighted and PS weighted hazard ratios, 95% confidence intervals and p-values will be reported for all analyses.

Please refer to section 6.8.1 for how the proportional hazards assumption will be tested.

6.8.3 Annual systemic corticosteroid dose used to treat exacerbations (secondary objective 3)

SCS exposure will be measured as the average annual cumulative dose, in milligrams, of prednisolone taken via any systemic route. The annual dose is the cumulative steroid dose related to exacerbations i.e. regimens started within 3 days of a coded respiratory event or specific exacerbation code. Use of SCS other than prednisolone will be converted to prednisolone equivalent dosing using a conversion table (see Section 9.1). Data regarding SCS dose is available in primary care EMR but not in secondary care records; therefore, estimated doses may be an underestimate of the true annual cumulative dose.

To analyse SCS use, we will use a linear mixed model to compare the total annual SCS exposure between treatment arms. Random effects (normally distributed) for the practice and fixed effects for the stratification factors used in the randomisation algorithm will be included in the model. Unweighted and PS weighted mean differences in annual cumulative dose, 95% confidence intervals and p-values will be reported.

In the event of non-normality, the analysis will use either transformations or a generalised linear mixed model with gamma distribution and log link. Unweighted and PS weighted rate ratios, 95% confidence intervals and p-values will be reported.

6.8.4 Annual rate of systemic corticosteroid courses used to treat exacerbations (secondary objective 4)

The analysis of SCS course will be performed using a random effects Poisson model with the number of courses as the outcome, a random effect for GP practice to account for the cluster trial design, and an offset variable containing the logarithm of the follow up time for each patient. Should significant overdispersion be detected, the analysis will use a negative binomial model[13].

An unweighted and propensity score (PS) weighted model will be analysed with the included covariates described in Section 6.2 . Unweighted and weighted rate ratios, 95% confidence intervals and p-values will be reported.

6.8.5 Incident type 2 diabetes (secondary objective 5)

Cumulative incidence of type 2 diabetes (T2D) will be reported over the outcome evaluation period. The distribution of time to events between the two arms will be compared using Kaplan Meier curves.

Incidence of T2D will be analysed using a Cox regression model, with a frailty term to account for clustering by GP practice. The effectiveness of the intervention, as compared to usual care, will be measured as a hazard ratio. Unweighted and PS weighted hazard ratios, 95% confidence intervals and p-values will be reported for all analyses.

Please refer to section 6.8.1 for how the proportional hazards assumption will be tested.

6.8.6 COPD exacerbations (time to first event) (secondary objective 6)

In addition to the primary outcome, the time to first COPD exacerbation will be analysed as an exploratory outcome. Cumulative incidence COPD exacerbation will be reported over the outcome evaluation period. The distribution of time to events between the two arms will be compared using Kaplan Meier curves.

Incidence of COPD exacerbation will be analysed using a Cox regression model, with a frailty term to account for clustering by GP practice. The effectiveness of the intervention, as compared to usual care, will be measured as a hazard ratio. Unweighted and PS weighted hazard ratios, 95% confidence intervals and p-values will be reported for all analyses.

Please refer to section 6.8.1 for how the proportional hazards assumption will be tested.

6.9 Analysis of the exploratory outcomes

This section summarises the analysis methods to be applied to the exploratory endpoints described in Section 2.3 .

6.9.1 Explanatory diagnoses (exploratory objective 1)

The rate of incident explanatory diagnoses other than COPD, diagnosed during the implementation or outcome evaluation periods.

Explanatory diagnoses will include asthma, lung cancer and heart failure, and any others deemed relevant based on the observed data.

Incident rate of explanatory diagnoses will be calculated with 95% confidence intervals using the exact Poisson method. Incidence rates and 95% confidence intervals will be presented.

6.9.2 Unscheduled respiratory events (exploratory objective 2)

Unscheduled respiratory care events will be defined as any of the following:

- Respiratory-related A&E attendance
- Respiratory-related hospitalisation, except outpatient
- Respiratory-related primary care attendance with no supporting COPD review code or record of dyspnoea (mMRC)/disease-specific health-related quality of life (COPD Assessment Tool) score

The occurrence of unscheduled respiratory events will be analysed using the same random effects Poisson model as that described for the analysis of COPD exacerbations (see Section 6.7). Unweighted and PS weighted rate ratios and 95% confidence intervals will be presented.

6.9.3 Change in lung function (exploratory objective 3)

Change from practice initiation date will be assessed by fitting a linear mixed model (LMM) to the recorded lung function measurements.

The endpoints assessed by this method will be:

- FEV1 % predicted will be the main value of interest
- FEV
- FVC
- FVC % predicted
- FEV1/FVC ratio
- Peak expiratory flow (PEF)
- PEF % predicted

For each spirometry measure a LMM will be fit including the following variables as fixed effects:

- The spirometry measure (e.g., FEV1 % predicted) as the dependent variable
- A binary variable indicating the treatment group a patient has been randomised
- The time of the spirometry measure included as a continuous linear trend measured in months since the patient's study entry
- A time by treatment group interaction term

The following variables will be included as normally distributed random effects with an unstructured nested covariance matrix:

- The practice identifier
- Patient identifier as a random intercept

Should the data allow, a model including a random effect for time (random slopes model) will be explored.

The overall difference between treatment groups will be summarised as the estimated difference in the slopes of the linear trend between the treatment groups together with a 95% confidence interval and p-value. Differences from baseline within and between treatment groups will be estimated from the LMM and unweighted and PS weighted mean differences and 95% confidence intervals will be reported at 6, 12, 24 months.

6.9.4 Change in disease-specific health-related quality of life (COPD Assessment Test [CAT]) (exploratory objective 4)

The CAT instrument is a set of 8 questions with each question scores on a Likert scale of 0 to 5. The range of CAT scores lies between 0 and 40. An example of the CAT instrument is shown in Appendix Section 9.2 .

For this analysis the CAT scores are treated as continuous variables and modelled using a LMM using a similar approach to that of the spirometry measures.

The LMM will be fit including the following variables as fixed effects:

- The CAT score as the dependent variable
- A binary variable indicating the treatment group a patient has been randomised
- The time of the CAT score included as a continuous linear trend measured in months since the patient's study entry
- A time by treatment group interaction term

The following variables will be included as normally distributed random effects with an unstructured nested covariance matrix:

- The practice identifier
- Patient identifier as a random intercept

Should the data allow, a model including a random effect for time (random slopes model) will be explored.

The overall difference between treatment groups will be summarised as the estimated difference in the slopes of the linear trend between the treatment groups together with a 95% confidence interval and p-value. Differences from baseline within and between treatment groups will be estimated from the LMM and unweighted and PS weighted mean differences and 95% confidence intervals will be reported at 6, 12, 24 months.

6.9.5 Incident pneumonia (exploratory objective 5)

The rate of incident pneumonia, diagnosed during the implementation or outcome evaluation periods. The distribution of time to events between the two arms will be compared using Kaplan Meier curves.

Incidence of pneumonia will be analysed using a Cox regression model, with a frailty term to account for clustering by GP practice. The effectiveness of the intervention, as compared to usual care, will be measured as a hazard ratio. Unweighted and PS weighted hazard ratios, 95% confidence intervals and p-values will be reported for all analyses.

Please refer to section 6.8.1 for how the proportional hazards assumption will be tested.

6.10 Subgroup analysis

To assess the impact of the CONQUEST programme on patients on minimal therapy within the Primary Analysis Set (PAS), trial outcomes will be analysed in:

- Already diagnosed and newly diagnosed COPD patients in the PAS who were stepped up from no maintenance or single agent only maintenance therapy at baseline

The effectiveness of the intervention, as compared to usual care, will be measured using rate ratios or between-group differences, as described in previous sections. The subgroup analysis will not include the exploratory outcome of incident explanatory diagnoses other than COPD, as this outcome relates solely to intervention arm patients deemed COPD negative after spirometry.

6.11 Sensitivity analysis

To assess the impact of analysing the composite primary outcome of COPD exacerbations within EMR data alone, the component of severe exacerbations (recorded hospitalisation or death for a COPD-related respiratory event) will be analysed within the linked primary and secondary care data.

To assess the impact of including patients with less than 2 years' baseline data, analysis of the primary outcome within the PAS will be repeated after restricting to patients with 2 years' baseline data.

6.12 Missing data

Missing baseline data will be reported in the form of the count and percentage (n [%]) by variable and by randomised group.

6.13 Software

All analyses will be conducted using STATA version 14 or later, or the current version of R (at the time of writing, version 2022.02.0). Anonymised data will be stored and analysed within the NHS England Secure Data Environment, with access limited to those authorised to use it.

7 Exploratory analysis sets

Figure 3**Error! Reference source not found.** presents the additional populations in the intervention and control trial arms, and the comparisons that form the exploratory analysis sets. The exploratory analysis sets explore the impact on different patient populations and will not be used for independent claims of effectiveness. Table 15 provides a matrix showing which analysis sets will be used for each trial outcome.

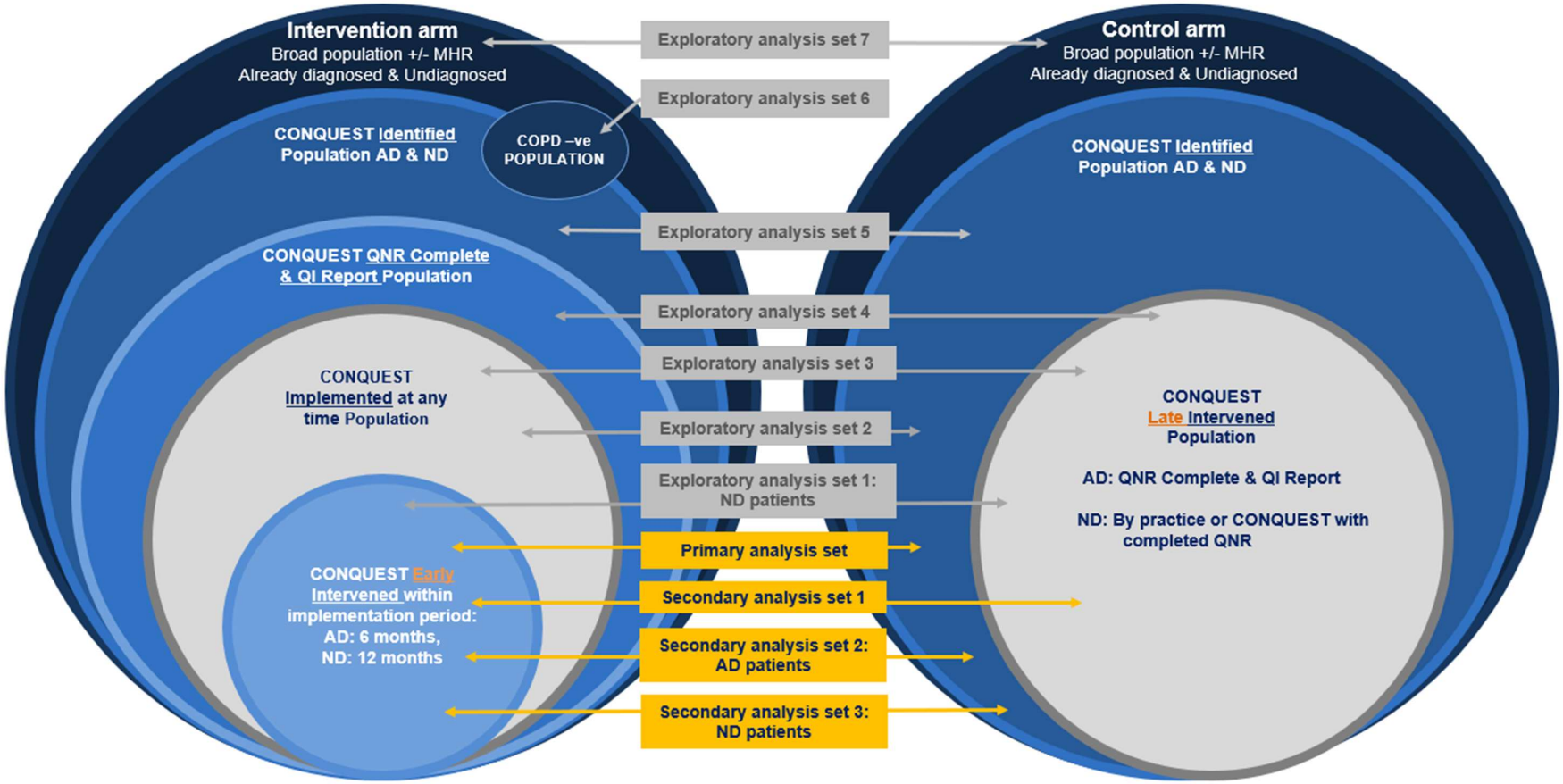


Figure 3: Population used in all analysis sets

Table 15: All analysis sets by trial outcome matrix

	Primary Outcome	Secondary outcomes						Exploratory Outcomes				
Analysis Sets	Moderate or severe exacerbations	MACRE	MACE	SCS use	SCS courses	Type 2 diabetes	Time to exacerbation	Explanatory diagnoses	Unscheduled respiratory care events	Lung function†	CAT score	Pneumonia
Primary analysis set (PAS)	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓
Secondary analysis set 1	✓			✓	✓	✓	✓		✓	✓	✓	✓
Secondary analysis set 2	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓
Secondary analysis set 3	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓
Exploratory analysis set 1	✓			✓	✓	✓	✓		✓	✓	✓	✓
Exploratory analysis set 2	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓
Exploratory analysis set 3	✓			✓	✓	✓	✓		✓	✓	✓	✓
Exploratory analysis set 4	✓			✓	✓	✓	✓		✓	✓	✓	✓
Exploratory analysis set 5	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓
Exploratory analysis set 6‡								✓§				
Exploratory analysis set 7	✓	✓	✓	✓	✓	✓	✓					

MACRE: Major adverse cardiac / respiratory events; MACE: Major adverse cardiac events; SCS: Systemic corticosteroid; CAT: COPD assessment test

† Repeat spirometry at end of the outcome evaluation period is not being performed within current CONQUEST/PREVAIL funding

8 Tables

The tables, listings and figures (TLFs) are shown in the attached Word file:
Tables_figures_v0.3_May2026.docx.

9 Appendix

9.1 Corticosteroid conversion table

Drug	Prednisolone equivalence factor
Betamethasone	6.67
Cortisone acetate	0.20
Dexamethasone	6.67
Hydrocortisone	0.25
Methylprednisolone	1.25
Prednisone	1.00

Table adapted from Voorham et al (2019)[14]

Conversion to Prednisolone-equivalent dosage = multiply drug dose by Prednisolone equivalence factor

Example 1 (Hydrocortisone 20mg dose): 20mg * 0.25 = 5mg Prednisolone

Example 2 (Betamethasone 0.75mg dose): 0.75mg * 6.67 = 5mg Prednisolone

9.2 COPD Assessment Test[15]

How is your COPD? Take the COPD Assessment Test™ (CAT)

This questionnaire will help you and your healthcare professional measure the impact COPD (Chronic Obstructive Pulmonary Disease) is having on your wellbeing and daily life. Your answers, and test score, can be used by you and your healthcare professional to help improve the management of your COPD and get the greatest benefit from treatment.

If you wish to complete the questionnaire by hand on paper, please click here and then print the questionnaire.

For each item below, place a mark (X) in the box that best describes you currently. Be sure to only select one response for each question.

Example: I am very happy 0 X 2 3 4 5 I am very sad

								SCORE
I never cough	0	1	2	3	4	5	I cough all the time	
I have no phlegm (mucus) in my chest at all	0	1	2	3	4	5	My chest is completely full of phlegm (mucus)	
My chest does not feel tight at all	0	1	2	3	4	5	My chest feels very tight	
When I walk up a hill or one flight of stairs I am not breathless	0	1	2	3	4	5	When I walk up a hill or one flight of stairs I am very breathless	
I am not limited doing any activities at home	0	1	2	3	4	5	I am very limited doing activities at home	
I am confident leaving my home despite my lung condition	0	1	2	3	4	5	I am not at all confident leaving my home because of my lung condition	
I sleep soundly	0	1	2	3	4	5	I don't sleep soundly because of my lung condition	
I have lots of energy	0	1	2	3	4	5	I have no energy at all	
							TOTAL SCORE	

Make sure you print your CAT before visiting your healthcare professional!

A COPD assessment test was developed by an interdisciplinary group of international COPD experts with support from GSK. GSK's activities in connection with the COPD assessment test are monitored by a supervisory council that includes external, independent experts, one of which is chair of the council.

9.3 Exacerbation algorithm: high-risk patient identification

Algorithm (hierarchy)	Explanation	Comments
1	**Number of COPD Exacerbations code**	QOF code with value attached specifying number of exacerbations
2	**COPD exacerbation code**	
3	**COPD hospital admission code**	Where specific COPD exacerbation hospital admission codes are available
4	**COPD code (COPD, chronic bronchitis, emphysema)**	code within 3 days of any of the below events
4.1	*With prescribed respiratory specific Antibiotic and OCS*	
4.2	*With prescribed OCS*	
4.3	*With prescribed respiratory specific Antibiotic*	
4.4	*With hospital admissions code*	
5	**LRTI code (including acute bronchitis)**	code within 3 days of any of the below events
5.1	*With prescribed respiratory specific Antibiotic and OCS*	
5.2	*With prescribed OCS*	
5.3	*With prescribed respiratory specific Antibiotic*	
5.4	*With hospital admissions code*	
6	**Other lower respiratory code (e.g., acute tests, symptoms)**	code within 3 days of any of the below events
6.1	*With prescribed respiratory specific Antibiotic and OCS*	
6.2	*With prescribed OCS*	
6.3	*With prescribed respiratory specific Antibiotic*	

6.4	*With hospital admissions code*	
7	**Influenza or COVID-19 infection**	code within 3 days of any of the below events
7.1	*With prescribed respiratory specific Antibiotic and OCS*	
7.2	*With prescribed OCS*	
7.3	*With prescribed respiratory specific Antibiotic*	
7.4	*With hospital admissions code*	
8	*No associated LR code step 1-7*	
8.1	Prescribed Antibiotic and OCS*	
8.2	Prescribed OCS*	
9	*No associated LR code step 1-7*	
	Prescribed Antibiotics**	

* Exclusion criteria for prescribed OCS:
- Dosage referencing “rheum” or “endocrin” or "cancer"
- OCS within 3 days of a code for OCS exclusion indications: Asthma, Gout, Iritis, Uveitis, Myocarditis, Pericarditis, Bell’s palsy, Myasthenia gravis, Scleritis, Lupus, Ankylosis spondylitis, Autoimmune hepatitis, Psoriatic arthropathy, Rheumatoid arthritis, Multiple sclerosis, IBD (Ulcerative Colitis and Crohn’s), Sarcoidosis, Amyloidosis, Vasculitis, Temporal arteritis, Sjogren’s syndrome, Autoimmune bullous disease, Nephrotic syndrome, Polymyalgia Rheumatica, Bullous Pemphigoid, Pemphigus vulgaris
- Chronic OCS prescription relating to temporal arteritis or polymyalgia

** Exclusion: within 3 days of a code indicating other infectious indications for AB, such as urinary tract infections and cellulitis. These codes will be identified using a manual frequency check.

9.4 Exacerbation algorithm: endpoint

Algorithm (hierarchy)	Explanation	Comments
1	**COPD exacerbation code**	
2	**COPD hospital admission code**	Where specific COPD exacerbation hospital admission codes are available
3	**COPD code (COPD, chronic bronchitis, emphysema)**	code within 3 days of any of the below events
3.1	*With prescribed respiratory specific Antibiotic and OCS*	
3.2	*With prescribed OCS*	
3.3	*With prescribed respiratory specific Antibiotic*	
3.4	*With hospital admissions code*	
4	**LRTI code (including acute bronchitis)**	code within 3 days of any of the below events
4.1	*With prescribed respiratory specific Antibiotic and OCS*	
4.2	*With prescribed OCS*	
4.3	*With prescribed respiratory specific Antibiotic*	
4.4	*With hospital admissions code*	
5	**Other lower respiratory code (e.g., acute tests, symptoms)**	code within 3 days of any of the below events
5.1	*With prescribed respiratory specific Antibiotic and OCS*	
5.2	*With prescribed OCS*	
5.3	*With prescribed respiratory specific Antibiotic*	
5.4	*With hospital admissions code*	

6	**Influenza or COVID-19 infection**	code within 3 days of any of the below events
6.1	*With prescribed respiratory specific Antibiotic and OCS*	
6.2	*With prescribed OCS*	
6.3	*With prescribed respiratory specific Antibiotic*	
6.4	*With hospital admissions code*	
7	*No associated LR code step 1-6*	
7.1	Prescribed Antibiotic and OCS*	
7.2	Prescribed OCS*	
8	*No associated LR code step 1-6*	
	Prescribed Antibiotics**	

* Exclusion criteria for prescribed OCS
- Dosage referencing “rheum” or “endocrin” or "cancer"
- Within 3 days of a code for OCS exclusion indications: Gout, Iritis, Uveitis, Myocarditis, Pericarditis, Bell’s palsy, Myasthenia gravis, Scleritis, Lupus, Ankylosis spondylitis, Autoimmune hepatitis, Psoriatic arthropathy, Rheumatoid arthritis, Multiple sclerosis, IBD (Ulcerative Colitis and Crohn’s), Sarcoidosis, Amyloidosis, Vasculitis, Temporal arteritis, Sjogren’s syndrome, Autoimmune bullous disease, Nephrotic syndrome, Polymyalgia Rheumatica, Bullous Pemphigoid, Pemphigus vulgaris
- Chronic OCS prescription relating to temporal arteritis or polymyalgia

**Exclusion: within 3 days of a code indicating other infectious indications for antibiotics, such as but not limited to urinary tract infections and cellulitis/skin infections.

10 References

1. Pullen, R. *et al*. CONQUEST Quality Standards: For the Collaboration on Quality Improvement Initiative for Achieving Excellence in Standards of COPD Care. *Int. J. Chron. Obstruct. Pulmon. Dis*. **Volume 16**, 2301–2322 (2021).
2. Overview | Chronic obstructive pulmonary disease in over 16s: diagnosis and management | Guidance | NICE. https://www.nice.org.uk/guidance/ng115 (2018).
3. Haroon, S., Adab, P., Riley, R. D., Fitzmaurice, D. & Jordan, R. E. Predicting risk of undiagnosed COPD: development and validation of the TargetCOPD score. *Eur. Respir. J*. **49**, (2017).
4. Ray, E. *et al*. *The Use of a Novel Case-Finding Algorithm in the Identification of Chronic Obstructive Pulmonary Disease (COPD) Patients in Primary Care – Baseline Results of the ASSIST Study*. (2018). doi:10.13140/RG.2.2.19494.16967.
5. Zheng, Y. *et al*. Triple therapy in the management of chronic obstructive pulmonary disease: systematic review and meta-analysis. *BMJ* **363**, k4388 (2018).
6. Mammen, M. J. *et al*. Triple Therapy versus Dual or Monotherapy with Long-Acting Bronchodilators for Chronic Obstructive Pulmonary Disease. A Systematic Review and Meta-analysis. *Ann. Am. Thorac. Soc*. **17**, 1308–1318 (2020).
7. Cazzola, M., Rogliani, P., Calzetta, L. & Matera, M. G. Triple therapy versus single and dual long-acting bronchodilator therapy in COPD: a systematic review and meta-analysis. *Eur. Respir. J*. **52**, (2018).
8. Koarai, A. *et al*. Triple versus LAMA/LABA combination therapy for patients with COPD: a systematic review and meta-analysis. *Respir. Res*. **22**, 183 (2021).
9. Rutterford, C., Copas, A. & Eldridge, S. Methods for sample size determination in cluster randomized trials. *Int. J. Epidemiol*. **44**, 1051–1067 (2015).

10. Eldridge, S. & Kerry, S. Analysis. in *A Practical Guide to Cluster Randomised Trials in Health Services Research* 99–136 (John Wiley & Sons, Ltd, 2012). doi:10.1002/9781119966241.ch6.

11. Leyrat, C., Caille, A., Foucher, Y. & Giraudeau, B. Propensity score to detect baseline imbalance in cluster randomized trials: the role of the c-statistic. *BMC Med. Res. Methodol.* **16**, 9 (2016).

12. Hayes, R. J. & Moulton, L. H. *Cluster Randomised Trials*. (Chapman and Hall/CRC, New York, 2017). doi:10.4324/9781315370286.

13. Keene, O. N., Calverley, P. M. A., Jones, P. W., Vestbo, J. & Anderson, J. A. Statistical analysis of exacerbation rates in COPD: TRISTAN and ISOLDE revisited. *Eur. Respir. J.* **32**, 17–24 (2008).

14. Voorham, J. *et al*. Healthcare resource utilization and costs associated with incremental systemic corticosteroid exposure in asthma. *Allergy* **74**, 273–283 (2019).

15. GlaxoSmithKline Services Unlimited. COPD Assessment Test (CAT). https://www.catestonline.org/patient-site-test-page-english.html.