Statistical and Health Economics Analysis Plan for

The PACE Study

Primary care use of a C-Reactive Protein (CRP) Point of Care Test (POCT) to help target antibiotic prescribing to patients with Acute Exacerbations of Chronic Obstructive Pulmonary Disease (AECOPD) who are most likely to benefit

EudraCT/ISRCTN No:	ISRCTN24346473	Version Number:	1.0

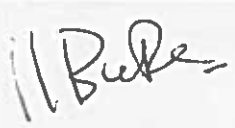
Draft Plan

Based on protocol version: 5.0 (15/03/2016)

SHEAP Revision History

Protocol version	Updated Sap version no.	Section number changed	Description and reason for change	Date changed

ROLES AND RESPONSIBILITIES

Trial Statistician: David Gillespie			
Role: Author of SHEAP			
Date:	08/03/2017	Signature:	
Senior Statistician: Kerenza Hood			
Role: Approve proposed analysis			
Date:	8/3/17	Signature:	
Health Economist: Bernadette Sewell			
Role: Author of SHEAP			
Date:		Signature:	
Chief Investigator: Christopher Butler			
Role: Review and approve proposed analysis			
Date:	9/3/17	Signature:	
Chief Investigator: Nick Francis			
Role: Review and approve proposed analysis			
Date:	5/4/17	Signature:	

ROLES AND RESPONSIBILITIES

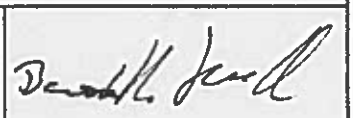
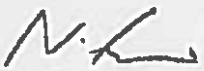
Trial Statistician: David Gillespie			
Role: Author of SHEAP			
Date:	08/03/2017	Signature:	
Senior Statistician: Kerenza Hood			
Role: Approve proposed analysis			
Date:	8/3/17	Signature:	
Health Economist: Bernadette Sewell			
Role: Author of SHEAP			
Date:	10/04/2017	Signature:	
Chief Investigator: Christopher Butler			
Role: Review and approve proposed analysis			
Date:	9/3/17	Signature:	
Chief Investigator: Nick Francis			
Role: Review and approve proposed analysis			
Date:	5/4/17	Signature:	

TABLE OF CONTENTS

1. INTRODUCTION

This statistical analysis plan provides guidelines for the final presentation and analysis for the PACE trial. This plan, along with all other documents relating to the analysis of this trial, will be stored in the Statistical Analysis Master File electronically and/or in hard signed copy formats.

2. BACKGROUND

2.1 RATIONALE AND RESEARCH QUESTION

Better targeting of antibiotics for acute exacerbations of chronic obstructive pulmonary disease (AECOPD) represents a major opportunity for antimicrobial stewardship. Over 70% of patients presenting with AECOPD in primary care are prescribed an antibiotic. Current antibiotic prescribing recommendations for GPs in the UK are based on symptoms alone. However, these have insufficient diagnostic accuracy to predict which patients can safely be managed without antibiotics. C-reactive protein (CRP) Point-Of-Care Tests (POCTs) are widely available and commonly used to help guide antibiotic prescribing decisions, including for AECOPD in primary care in a number of European countries. However, CRP testing in conjunction with clinical examination has not yet been evaluated. We have developed a CRP algorithm based on best available evidence and consensus statements by a panel of experts. Our trial will establish whether a CRP-POCT used within the remits of our proposed algorithm can be used to target antibiotic treatment for AECOPD in primary care.

2.2 OBJECTIVES

The primary objective of PACE is to determine whether the addition of a CRP POCT (with training on test use and interpretation), to usual care for managing an AECOPD, leads to a reduction in antibiotic consumption for AECOPD without negatively impacting on COPD health status, compared with current best practice alone. To meet this objective, we will assess:

1. Antibiotic consumption (any consumption of antibiotics for AECOPD vs. no consumption of antibiotics for AECOPD) over the first four weeks following randomisation. Actual consumption (rather than prescribing or dispensing) is the driver of AMR, and a four-week window will allow us to capture consumption of both antibiotics prescribed at the initial consultation and those that are related to the AECOPD episode in question, but are obtained or started at a later date;
2. Recovery in terms of COPD health status, which will be assessed two-weeks post-randomisation using the CCQ. (Van Der Molen et. al., 2003) Two-weeks is the time

when most patients will have recovered (in terms of improvement in health status and physiological parameters) and therefore the point at which a difference would be most indicative of a delayed recovery.

Secondary objectives are to assess the effect of using a CRP POCT for AECOPD in primary care on:

1. Prevalence of resistant bacteria in sputum (or throat swab) at 4 weeks;
2. COPD health status over time (weeks 1, 2 and 4);
3. Health utility, measured using the EQ-5D-5L at 1, 2 and 4 weeks and at 6 months; (Herdman et. al., 2011)
4. All-cause antibiotic consumption during the first four weeks;
5. Antibiotic prescribing at the index consultation;
6. Antibiotic prescribing during the first 4 weeks post randomisation
7. Use of other COPD treatments including oral steroids during the first four weeks;
8. Adverse effects from antibiotics and other medication prescribed for their AECOPD during the first four weeks;
9. Primary and secondary care consultations (including out of hours, A&E visits and hospitalisations) during the subsequent 6 months;
10. Costs and cost-effectiveness from a health service perspective;
11. Incidence of pneumonia during the first 4 weeks and from the 4-week follow up to 6 months;
12. Disease-specific Health-Related Quality of Life (HRQoL) (CRQ-SAS) at 6 months. (Schunemann et. al., 2003)

3. STUDY MATERIALS

3.1 TRIAL DESIGN

Two-arm individually (1:1 ratio) randomised controlled trial. Patients with AECOPD randomised to be managed by usual care alone or with the addition of a CRP POCT and protocol-based training in test use and interpretation to guide decisions about the use of antibiotic treatment for AECOPD.

Co-primary outcomes (antibiotic consumption for AECOPD within four-weeks post-randomisation and COPD health status two weeks-post randomisation) will be used to answer the primary research question of the study. Between-group differences in antibiotic consumption will be investigated for superiority, while differences in COPD health status will be investigated for non-inferiority.

3.2 RANDOMISATION

Participants will be randomised in a 1:1 ratio to receive either current best clinical management alone (control) or best clinical assessment with the addition of CRP POCT (intervention). Randomisation will take place remotely using minimisation, with a random element set at 80% to improve the integrity of the randomisation process. The Anthonisen criteria (categorised as type 1, 2 or 3) will be used as a minimisation variable, so that balance is achieved with respect to differing levels of COPD exacerbation severity. Remote allocation will maintain allocation concealment from both the participant and the treating clinician prior to allocation, as this is an unblinded study.

3.3 SAMPLE SIZE

We aim to have sufficient power to detect a 15% reduction from a current estimated antibiotic consumption rate (proportion that take any antibiotics) for AECOPD from 70% to 55% in the four weeks following randomisation. Current estimates suggest that approximately 80% of patients with AECOPD in primary care are prescribed antibiotics (Llor et. al., 2012) and the majority of patients with this illness are likely to initiate their treatment. To show a difference in proportions between 0.70 and 0.55 at the 5% significance level and with 90% power we would need a total of 434 participants, inflating to 544 to account for the loss to follow-up of approximately 20% of participants. We have also aimed to have sufficient power to demonstrate that participants managed with CRP POCT are no worse (non-inferior), compared to those managed without CRP POCT, in terms of their COPD health status two weeks following randomisation. Assuming an expected difference between groups of zero, a non-inferiority margin of 0.3 (lower than the minimal clinically important difference (Kocks et. al., 2006) and a common standard deviation of 1.1 (Kocks et. al., 2006) then based on a one-sided significance level of 0.05 and 90% power we would need 462 participants, inflating to 580 to account for the loss to follow-up of approximately 20% of participants.

Formulating our overall hypothesis using the Intersection-Union test, (Offen et. al., 2007) we will carry out our individual sub-hypothesis tests at the 5% level, and if both are significant conclude overall significance at the 5% level. However, power will be affected by the level of correlation between the two outcomes and their respective effect sizes. The impact on overall power is at its greatest when there is zero correlation between outcomes and effect sizes are identical (the overall power is the product of the powers for testing each individual sub-hypothesis), and is increasingly negligible the more correlated outcomes are and the more different effect sizes become. We do not expect our effect sizes to be similar, as our co-primary outcomes are two very different constructs. We also anticipate that the outcomes will not be entirely independent. We will therefore aim to recruit at least 650 participants to maintain an overall power between 81 and 90%.

Participants will be recruited from approximately 10 practices during the internal pilot, with each practice recruiting an approximate average of 7 participants over 6 months, and approximately 70 practices in the substantive trial. Our final sample size will provide adequate power to account for the clustering of antibiotic prescribing by practice,

assuming an ICC of 0.02, though as this is an individually randomised controlled trial and practices were not balanced on at randomisation, we did not set out to inflate our sample size to account for clustering. Nevertheless, we will investigate clustering and account for it in our analysis. (Kahan and Morris, 2013)

3.4 FRAMEWORK

For the primary outcomes, between-group differences in antibiotic consumption will be investigated for superiority, while differences in COPD health status will be investigated for non-inferiority. All subsequent outcomes will be investigated for superiority.

3.5 INTERIM ANALYSES

No interim analyses are planned.

3.5.1 PLANNED SAMPLE SIZE ADJUSTMENT

Not applicable.

3.5.2 STOPPING RULES

Not applicable.

3.6 TIMING OF FINAL ANALYSIS

Final analysis of the primary and majority of the secondary outcomes will take place once all randomised participants have completed their four-week follow-up assessment, all forms have been received, cleaned and the datasets have been locked.

Analysis of the health utility, primary / secondary care consultations, incidence of pneumonia, and disease-specific HRQoL will take place once all randomised participants have completed their six-month follow-up assessment, all forms have been received, cleaned and the datasets have been locked.

Reporting and publication of all primary and secondary analyses will occur together.

3.7 TIMING OF OUTCOME ASSESSMENT

Outcomes are assessed at the following time points:

- Immediately post-randomisation (during the index consultation)
- One-week post-randomisation (-1/+2 days)
- Two-weeks post-randomisation (-1/+7 days)

- Four-weeks post-randomisation (-3/+14 days)
- Six-months post-randomisation
    - Postal questionnaire: -7 days/+21 days
    - Notes review: a one-month lag is added to the notes review to allow for any delay in events occurring and events appearing in medical notes

Outcome data collected outside of the specified windows will be included in the intention-to-treat and modified intention-to-treat analyses (see Section 4.3). However, sensitivity analyses based on a per-protocol population will exclude participants who provide data outside the stated windows.

4. STATISTICAL PRINCIPLES

4.1 LEVELS OF CONFIDENCE AND P-VALUES

For the primary analysis of the between-group comparison of CCQ two-weeks post-randomisation, a one-sided 95% confidence interval will be calculated. For all other analyses, two-sided 95% confidence intervals and p-values will be calculated. P-values will be reported to three decimal places, and a 5% significance level will be used to categorise hypothesis tests as statistically significant.

4.1.1 ADJUSTMENT FOR MULTIPLICITY

There will be no adjustments for multiple testing.

4.2 ADHERENCE AND PROTOCOL DEVIATIONS

4.2.1 DEFINITION AND ASSESSSMENT OF ADHERENCE

The recruiting (and treating) clinician will report CRP use at the index consultation.

Participants will be asked during their Week 1 and 2 telephone follow-ups, as well as at their Week 4 consultation, whether or not they received a finger prick blood test for their COPD during their initial consultation.

Logs of CRP cartridge use are also kept by the practice staff. In addition, the usage of cartridges will be retained on the CSV files of the Afinion machine. Where participants state they received an intervention that is contrary to their allocation (i.e. they were allocated to usual care and received a finger prick blood test, or allocated to CRP and received no finger prick blood test), the answer could be verified with the logs.

Participants will be considered as having adhered to their allocated treatment if they:

- Were allocated to the CRP arm and received a CRP test at their index consultation
- Were allocated to the control arm and received no CRP test between their index consultation and Week 4 consultation.

All data sources will be checked for agreement without knowledge of the participant's allocation. Where disagreement occurs, this will be queried with the corresponding site.

4.2.2 PRESENTATION OF ADHERENCE

Adherence will be presented, both overall and split by trial arm, as frequencies and percentages. Where non-adherence occurs in the control arm, this will be explored further and reported as whether it occurred at the index consultation or during a subsequent consultation.

4.2.3 DEFINITION OF PROTOCOL DEVIATION

Potential protocol deviations will include:

- Participants who are randomised but do not meet eligibility criteria
- Participants for whom data are collected outside of the windows specified in Section 3.7

4.2.4 PRESENTATION OF PROTOCOL DEVIATIONS

The frequency of protocol deviations will be presented both overall and by trial arm. Where deviations occur, further details will be provided (e.g. the reason why a participant was ineligible, how far outside the window their data were collected, etc.)

4.3 ANALYSIS POPULATION

Several analysis populations will be considered.

- Intention-to-treat population (ITT): All randomised participants, regardless of protocol deviations, intervention received, and outcome availability
- Modified intention-to-treat population (MITT): All randomised participants who provide outcome data, regardless of protocol deviations, intervention received.
- Per-protocol (PP): All randomised participants who provide outcome data within the specified time windows, with no other protocol deviations and received the intervention to which they were allocated.

The primary antibiotic consumption analysis will be based on the MITT population.

A Complier Average Causal Effect (CACE), which will extend our MITT analysis by taking into account departures from randomised treatment while maintaining a comparison of groups as randomised, will be estimated for the primary CCQ analysis. (Angrist et. al., 1996) The conclusions drawn on the primary CCQ analysis will be based on both the MITT and CACE analyses (i.e. the upper limit of the one-sided 95% confidence interval will have to exclude 0.3 in both analyses for non-inferiority to be concluded).

Missing outcome data will be imputed using a variety of assumptions about likely missing mechanisms (see Section 6.3 for more detail), in order to obtain a secondary analysis of the co-primary outcomes based on the full ITT population.

Further sensitivity analyses on the co-primary outcomes will also consider a PP population, as defined above.

Analysis of secondary outcomes will be based on the MITT population.

5. STUDY POPULATION

5.1 SCREENING DATA

The number of practices approached and enrolled will be reported. The number of participants approached and recruited / followed up will also be reported. Screening logs will be used to describe characteristics of participants approached but not recruited.

5.2 ELIGIBILITY

Summary statistics on eligibility, recruitment, withdrawal and dropout will be collated for both trials arms and will form the basis of the CONSORT flow diagram for clinical trial reporting. (Schulz, et. al., 2010)

5.3 RECRUITMENT

Recruitment will be reported both overall and by trial arm.

5.4 WITHDRAWAL/FOLLOW UP

5.4.1 LEVEL OF WITHDRAWAL

Several levels of withdrawal are possible:

- From trial intervention
- From further follow-up
- From entire study, but with consent to use data up until point of withdrawal
- From entire study, without consent to use data up until point of withdrawal

All withdrawals will be tabulated both overall and by trial arm, and will be classified according to the above levels.

5.4.2 TIMING OF WITHDRAWAL

The timing of withdrawal will be presented in the CONSORT diagram.

5.4.3 REASONS FOR WITHDRAWAL

Reasons for withdrawal will be presented in the CONSORT diagram.

5.4.4 PRESENTATION OF WITHDRAWAL/LOSS TO FOLLOW-UP

Withdrawals will be reported as frequencies and percentages, both overall and by trial arm. Level, timing, and reason for withdrawal will also be reported, as previously described.

5.5 BASELINE PARTICPANT CHARACTERISTICS

5.5.1 LIST OF BASELINE DATA

The following data collected at baseline (i.e. pre-randomisation) will be reported both overall and by trial arm:

- Age
- Gender
- Number of days that participant has had symptoms of their current acute exacerbation of COPD
- Co-morbid illness (heart failure, coronary heart disease, diabetes, chronic kidney disease, hyper tension, another chronic disease)
- Clinical examination findings:
    - Temperature
    - Pulse rate
    - Oxygen saturation on room air
    - Participant able to complete a full sentence without having to stop due to shortness of breath
- Detection of abnormal chest findings (crackles, wheeze, diminished vesicular sounds, evidence of consolidation)
- Collection of throat swab
- Collection of sputum sample
- Current sputum colour (according to Bronko test)
    - Clinician-assessed (if sputum sample obtained)
    - Participant-assessed (if sputum sample not obtained)
- Participant-assessed sputum colour when not exacerbating (according to Bronko test)
- CCQ total score
- EQ-5D-5L
- Exacerbation features (Anthonisen criteria used for randomisation):
    - Presence of worsened shortness of breath, increased sputum volume, increased sputum purulence
    - Number of exacerbation features present

Additional pre-randomisation variables that are collected via a notes review at six months post-randomisation include:

- Most recent spirometry assessment (FVC, FEV1, predicted FVC and FEV1, height, weight, and date of all measurements)
- Most recent eosinophil count
- Antibiotic prescriptions within the past 12 months
- Oral steroid prescriptions within the past 12 months

5.5.2 DESCRIPTIVE STATISTICS

Participant characteristics will be summarised using frequencies and percentages, means and standard deviations, or medians and interquartile ranges, and min / mix as appropriate.

In addition to presenting baseline characteristics for all randomised participants, they will also be presented for those participants included in the primary analyses in order to assess whether attrition has potentially induced selection bias.

6. ANALYSIS

6.1 OUTCOME DEFINITIONS

6.1.1 PRIMARY OUTCOMES

- Antibiotic consumption for AECOPD during the first four-weeks post-randomisation;
- CCQ total score two-weeks post-randomisation.

6.1.2 TIMING, UNITS, AND DERIVATION OF PRIMARY OUTCOMES

- Antibiotic consumption for AECOPD during the first four-weeks post-randomisation: At one-week, two-weeks, and four-weeks post-randomisation, participants are asked whether they have taken any antibiotics since their previous assessment (at one-week, this is since their index consultation). If it has not been possible to contact participants at one and/or two weeks, a minimum dataset is also collected at four weeks that covers these periods period (i.e. a reduced dataset that only captures medication use, CCQ, and EQ-5D). When answering these questions, they are also asked whether this antibiotic was taken for their COPD exacerbation. If a participant records a **YES** response to both of these questions at any of the above mentioned time points, they will be counted as having consumed an antibiotic for their AECOPD during the four-week follow-up period. If they record a **NO** response at every time point, they will be counted as not having consumed an antibiotic for their AECOPD during the four-week follow-up period;
- CCQ total score two-weeks post-randomisation: CCQ total score will be calculated according to its manual. See Appendix I for more detail. Data collected at two-

weeks post-randomisation (or via the minimum dataset at four-weeks) will be used to derive this outcome.

6.1.3 LIST OF SECONDARY OUTCOMES

There are several secondary outcomes:

- Prevalence of potentially pathogenic bacteria (incl *S.pneumoniae, H.spp* and *Enterobacteriacae*) cultured from sputum at 4 weeks
- Prevalence of resistant potentially pathogenic bacteria in sputum at four weeks;
- Prevalence of commensal organisms cultured from throat swabs at 4 weeks and proportion of bacteria that are resistant.
- COPD health status over time measured using the CCQ (total score measured at weeks 1, 2 and 4);
- CCQ symptoms domain (measured at weeks 1, 2, and 4);
- CCQ function state domain (measured at weeks 1, 2, and 4);
- CCQ mental state domain (measured at weeks 1, 2, and 4);
- Health utility measured using the EuroQol-5D (EQ-5D-5L) (measured at weeks 1, 2 and 4 and at month 6);
- All cause antibiotic consumption during the first four weeks (yes/no);
- Total antibiotic consumption during the first four weeks (number of days antibiotics consumed for AECOPD / any reason);
- Antibiotic prescribing at the index consultation;
- Antibiotic prescribing during the first 4 weeks post randomisation
- Any use of other COPD treatments including oral steroids (measured at weeks 1, 2 and 4);
- Adverse effects potentially attributable to antibiotics prescribed for their exacerbation (nausea, vomiting, diarrhoea, thrush, and rash) (measured at weeks 1, 2 and 4);
- Primary and secondary care consultations, including hospitalisations (measured at week 4 and month 6);
- Costs (total NHS cost) and cost-effectiveness (measured at month 6);
- Incidence of pneumonia (measured by patient and GP report at week 4 and month 6).
- Disease-specific health-related quality of life at six months, measured using CRQ-SAS (dyspnoea, fatigue, emotion function, mastery and total scores).

6.1.4 ORDER OF TESTING

Outcomes measured during the first four-weeks post-randomisation will be analysed first. Subsequent secondary outcomes collected at six-months post-randomisation will be analysed when they are available. This is to ensure reporting timelines are kept to, rather than any indication of a hierarchy of secondary outcomes.

6.1.5 TIMING, UNITS, AND DERIVATION OF SECONDARY OUTCOMES

Timing of secondary outcome measures are described in Section 6.1.3. Validated scales (CCQ, CRQ-SAS, and EQ-5D-5L) will be scored according to their manuals. See Appendix I for more details.

- Microbiology outcome operationalisation are given in Table 2
- All-cause antibiotic consumption during the first four-weeks post-randomisation: Similar definition as the primary outcome, but regardless of whether the participant stated that their antibiotic was to treat their AECOPD or not.
- Use of other COPD treatments including oral steroids: At one-week, two-weeks, and four-weeks post-randomisation, participants are asked whether they have taken any oral steroids, inhaled medications, or other medications to treat their COPD exacerbation. If a participant indicates that they have taken oral steroids, other medications, or have increased the use of inhaled medications by a bit or a lot, at any of the above mentioned time points, they will be counted as having used other treatments for their COPD exacerbation during the four-week follow-up period.
- Adverse effects: If a participant indicates that they have experienced at least one adverse effect potentially attributable to antibiotic treatment at either one, two, or four-weeks post-randomisation, they will be classed as having experienced an adverse effect.
- Primary and secondary care consultations: Participants will be classed as having had a primary or secondary care consultation if it is reported that they have had any consultations in either primary or secondary care at any point during the six-months post-randomisation.
- Incidence of pneumonia: Participants will be classed as having had pneumonia during the first six-months post-randomisation if it is reported that they have been diagnosed with pneumonia when followed up at either four-weeks or six-months.

6.2 ANALYSIS METHODS

6.2.1 LIST OF METHODS AND PRESENTATION

Primary outcomes

Between-group differences in the proportion of participants using antibiotics for their COPD exacerbation during the four-week follow-up period will be estimated using logistic regression. The analysis will also adjust for the Anthonisen criteria.

Between-group differences in the mean CCQ score at the two-week follow-up period will be estimated using linear regression, adjusting for the Anthonisen criteria and baseline CCQ score. A two-sided 90% confidence interval of the between-group difference in means will be estimated (CRP minus Usual Care), and non-inferiority concluded provided that the upper limit of the confidence interval (i.e. the one-sided 95% CI) does not contain 0.3 in both the MITT population and the CACE analysis.

The intervention will be deemed a success if antibiotics are used discernibly less in the CRP group, and the CCQ at two-weeks is no worse in the CRP group (that is, the 90% confidence interval for the mean difference in the CCQ at two-weeks does not contain the possibility that those allocated to the CRP group could have a mean score 0.3 points higher than those allocated to usual care) (Table 1).

For the primary analysis involving the CCQ, should the upper limit of the CI be between 0.3 and 0.4 (i.e. greater than the non-inferiority margin, but smaller than the established minimal clinically important difference), we will reflect on the magnitude of the between-group difference found in the primary analysis involving antibiotic use and the secondary outcomes.

Clustering of participants within practices will also be considered. If there is evidence of clustering by practice (ICC > 0), two-level models will be used in place of single-level models.

Table 1: Criteria for judging the intervention to be successful or not

	CCQ worse in CRP group	**CCQ no worse in CRP group**
Antibiotic consumption higher in CRP group	Intervention unsuccessful	Intervention unsuccessful
No evidence that antibiotic consumption is different in either group	Intervention unsuccessful	Intervention unsuccessful
Antibiotic consumption lower in the CRP group	Intervention unsuccessful	Intervention successful

Secondary outcomes

Analysis of secondary outcomes will be based on the MITT population. All analysis will adjust for the Anthonisen criteria and will investigate clustering of participants within practices, fitting a two-level model if required. Table 2 describes the proposed analysis methods for the secondary outcomes.

Table 2: Analysis methods for secondary outcomes

Outcome	Analysis
Presence of potentially pathogenic bacteria from sputum at 4-weeks post-randomisation	Logistic regression (any versus none)
The percentage of tested antibiotics to which at least one cultured potentially pathogenic bacteria (from sputum) was resistant	Binomial regression with an identity link function (for risk difference)
The percentage of total bacteria load that grew on each antibiotic plate	Binomial regression with an identity link function (for risk difference) for each antibiotic plate separately
CCQ total score over time (1, 2, 4-weeks post-randomisation)	Linear mixed model, controlling for baseline CCQ total score
CCQ symptom domain over time (1, 2, 4-weeks post-randomisation)	Linear mixed model, controlling for baseline CCQ symptom domain
CCQ function state over time (1, 2, 4-weeks post-randomisation)	Linear mixed model, controlling for baseline CCQ functional state domain
CCQ mental state over time (1, 2, 4-weeks post-randomisation)	Linear mixed model, controlling for baseline CCQ mental state domain
EQ-5D at 1, 2, 4 weeks and 6-months post-randomisation	Mixed model, controlling for baseline EQ-5D
Antibiotic consumption (all cause) during first four-weeks post-randomisation	Logistic regression (any versus none)
Total number of days antibiotics were consumed during first four-weeks post-randomisation	Poisson regression

Total number of days that antibiotics were consumed for AECOPD during first four-weeks post-randomisation	Poisson regression
Antibiotic prescribing at index consultation	Logistic regression (any versus none)
Antibiotic prescribing during the first four-weeks post-randomisation	Logistic regression (any versus none)
Use of other COPD treatments during first four-weeks post-randomisation	Logistic regression (any versus none)
Adverse effects from antibiotics / other treatments during first four-weeks post-randomisation	Logistic regression (any versus none)
Primary and secondary care consultations during the six-months post-randomisation	Logistic regression (any versus none)
Incidence of pneumonia during the first four-weeks post-randomisation	Logistic regression
Incidence of pneumonia during the first six-months post-randomisation	Logistic regression
CRQ-SAS at six-months post-randomisation	Linear regression.

6.2.2 COVARIATE ADJUSTMENT

The Anthonisen criteria was the only pre-randomisation variable minimised on. This will be included in all analyses as a covariate. Clustering of participants within practices will also be investigated, with multilevel analysis that accounts for this clustering reported where it is present. Where other covariates are adjusted for, this is indicated this is indicated in Section 6.2.1.

6.2.3 ASSUMPTION CHECKING

Modelling assumptions and distributions of outcome variables and residuals will be checked prior to reporting. Transformations and model choice may vary from what is written, depending on the outcome of these checks. Any changes will be fully documented.

6.2.4 ALTERNATIVE METHODS IF DISTRIBTIONAL ASSUMPTIONS NOT MET

It is anticipated that the model proposed for the EQ-5D will not fulfil its assumptions. Should this be the case, transformations will be attempted. Should it not be possible to transform the outcome measure to fulfil the assumptions of the linear mixed model, a categorisation will be derived. This categorisation will be agreed between TMG and IDMC members and will be based on maximising the use of the data (i.e. favouring an ordinal rather than dichotomous outcome) while retaining interpretability (i.e. the categories will need to be meaningful to patients and clinicians). In the presence of over-dispersed count data, a negative binomial regression model will be fitted to the duration of antibiotic consumption outcome.

6.2.5 SENSITIVITY ANALYSES

Several secondary analyses of the primary outcomes are planned:

- A full ITT analysis for each primary outcome, with missing outcome data accounted for using multiple imputation (see Section 6.3 for more detail)
- To determine whether the change in inclusion/exclusion criteria (e.g. the relaxation of the need for spirometry-confirmed COPD) had any impact on trial findings, the primary analysis will be repeated including an explanatory variable that indicates whether a participant was recruited before or after the change. Its interaction with trial arm will also be considered to explore the impact of this change on the effect of the intervention.
- The primary antibiotic use analysis will be repeated, but instead of creating a composite outcome combining antibiotic use at each/any of the follow-up time points, the proportion who used antibiotics to treat their COPD exacerbation at one, two, and four-weeks post-randomisation will be estimated by fitting a generalised linear mixed model, with responses nested within individuals. The model will be fitted with and without an intervention group by time interaction, to investigate intervention effects averaged across all follow-up time points, and to determine whether there are any differential intervention effects over time.

In addition, the following sensitivity analyses will also be conducted:

- Where outcomes combine observations across multiple time points (e.g. all-cause antibiotic consumption and use of other COPD treatments during weeks 1, 2, and 4), mixed models will be used, with responses nested within individuals and with/without an intervention group by time interaction term, to estimate both

intervention effects averaged across all follow-up time points and to determine whether there are any differential intervention effects over time.

- Primary analyses will be repeated on the per-protocol population.
- The analysis of the primary outcomes based on the full ITT population will use multiple imputation in the presence of missing outcome data. This approach assumes data are MAR. The impact of departures from this assumption will be explored using a variety of different approaches, depending on the primary outcome:
    - For the CCQ primary outcome, pattern mixture models (Little, 1993) will be fitted, with various assumptions about the distribution of non-responders
    - For the antibiotic consumption outcome, various scenarios will be modelled. For example:
        - All non-responders used antibiotics
        - All non-responders did not use antibiotics
        - Non-responders used antibiotics only if they were prescribed at the index consultation
        - Non-responders in the CRP group used antibiotics and non-responders in the usual care group did not use antibiotics. This is the most conservative assumption we can make about non-responders, as it is the assumption which is most likely to move our estimate of effect away from demonstrating that the intervention is effective in lowering antibiotic use
- The number of consultations in primary/secondary care during the six-months post-randomisation will be modelled by fitting a Poisson or negative binomial regression model. The distribution will be examined, and zero-inflated models will be fitted, if necessary.

6.2.6 SUBGROUP ANALYSES

The analysis of the primary outcomes will be repeated on the subgroups of participants defined below in order to determine whether there are any differential intervention effects according to these. The subgroup analysis will extend the original analysis of the primary outcomes by including an additional subgroup variable as well as a subgroup by intervention interaction term. The subgroups of interest are:

- COPD severity (Gold I/II/III/IV), from most recent spirometry assessment
- Severity of COPD exacerbation (Anthonisen criteria type 1/2/3)
- Presence of a potentially pathogenic bacteria cultured from sputum sample at baseline

6.3 MISSING DATA

Missing items from validated scales will be handled according to the guidance in their corresponding scoring manuals. Where no guidance is available, a participant missing at

most 20% of their items from a scale will have their missing items substituted by the mean of their valid items.

Multiple imputation will be used to adjust findings for missing primary outcome data. This will assume that data are missing at random, given observed covariates (MAR). An imputation model will be built, based on variables contained in the corresponding analysis models and other variables found to be associated with the primary outcome being missing. The number of imputations to be run will depend on the percentage of incomplete cases (for example, if 10% of participants have missing outcome data, 10 imputations will be run). (White et. al., 2011) Sensitivity analyses will test the robustness of the findings to departures from a MAR assumption. See Section 6.2.5 for more details of this.

The analysis that involves modelling repeated observations over time using mixed models will also be valid under the MAR assumption.

6.4 ADDITIONAL ANALYSES

6.4.1 HEALTH ECONOMIC EVALUATION

A within-trial health economic analysis will be undertaken from a health service perspective (UK NHS). Costs due to patient absences from work will also be considered but reported separately. The health economic evaluation will include cost-effectiveness, cost-utility and cost-consequences analyses. A trial based budget impact analysis will be undertaken to estimate the likely financial impact of the use of CRP POCT in the management of antibiotic prescribing for COPD on NHS budgets.

CRP POCT implementation costs

Implementation costs will include costing of all resources used in the introduction and delivery of CRP POCT in general practices. This includes:

- General practice staff training

We will capture both trainer's and trainee's time spent on CRP POCT training together with the grade of health care professionals involved (e.g. GP or nurse). We will include relevant travel costs for any training provided outside of practices.

- CRP POCT kits

This will include cost of equipment, cost of kits and maintenance as well as consumables required. We will only include manufacturer representatives' time/cost involved in training GPs/nurses if the manufacturer charges separately for training.

- General practice staff opportunity cost

We will collect data on the staff time spent to prepare, run, interpret and report CRP POCT in general practice to estimate the additional cost to practices.

Resource use through CRP POCT implementation will be estimated through interviews with general practice staff, the manufacturer and the trial team. Costs will be assigned

according to published unit costs obtained from the Personal Social Service Research Unit (PSSRU) and manufacturer list prices.

Healthcare resource use

Data collection will include resource use of primary and secondary care services. These are likely to include prescribed and consumed antibiotics, use of other medications such as COPD treatments including oral steroids, GP visits, nurse visits (including location of visit; e.g. home or clinic visit), emergency department visit, use of outpatient clinics and hospitalisations. These resources will be collected using data from the in-trial CRFs, the 4 week follow up questionnaire and an adapted client service receipt inventory (CSRI) integrated in the 6 month note review. Costs will be applied in £ Sterling using appropriate unit costs published by the PSSRU, the British National Formulary (BNF) and the UK Department of Health.

The health care costs in both intervention and control arms will be summated to assess the change in profile of health care use as a result of the intervention compared to usual care, with mean difference per patient in costs (including 95% confidence intervals). As the trial is < 12 months, discounting will not be applied to costs or outcomes.

Missing Data

Missing data will be handled in line with the statistical analysis plan (section 6.3).

Cost-effectiveness analysis

The co-primary outcomes (assessed at 4 weeks) will be used in the cost-effectiveness analysis. Given the aim of the trial is whether CRP POCT-informed management of patients with AECOPD can reduce antibiotic consumption without negatively impacting on recovery, we will consider a range of scenarios.

A base case analysis will report an incremental cost-effectiveness ratio (ICER) presenting the additional cost of producing an extra unit (%) reduction in antibiotic prescribing and consumption (expressed as cost per unit % antibiotic prescription avoided and cost per unit % antibiotic consumption avoided).

If the main trial fails to demonstrate non-inferiority in terms of the CCQ (as defined in section 6.2.1) then the intervention would (if usual conditions apply) likely be regarded as not cost-effective. However, since the potential of a reduction in antimicrobial resistance could reduce costs of COPD for the health care provider, the intervention may still be considered to be worthwhile from an economic perspective (within limits) even if it does not meet conventional cost-effectiveness decision rules.

Thus, we will test a range of scenarios as part of the sensitivity analysis e.g. best case/worse case scenarios based on the results of the co-primary outcomes to explore the impact on the ICER.

- A threshold analysis will be undertaken to assess the willingness to accept the costs of obtaining a reduction in antibiotic prescribing and consumption, should the CCQ score fall between the values which will warrant further exploration within the main trial analysis.

- We will also explore possible scenarios to reflect subsequent adoption in routine general practice e.g. exclude the purchase and running costs of the CRP POCT equipment to reflect that the equipment may be used with general practices for a variety of POCT interventions.

Cost-utility analysis

We will also undertake a within-trial cost-utility analysis (CUA) to assess the incremental costs per quality adjusted life year (QALY) gained as a result of the use of CRP-POCT compared to usual care at 6 months using the EQ-5D-5L to generate QALYs. QALYs incorporate quantity of life (additional life years) and quality of life in one measure. Thus, by dividing the difference in costs by the difference in QALYs, cost per QALY can be calculated for each comparison.

Generally, the UK National Institute for Health and Care Excellence (NICE) considers an intervention cost-effective if one of the following applies.

- The intervention is less costly and more clinically effective compared with all other relevant alternatives. In this case, no ICER is calculated as the strategy in question dominates the alternatives.
- The intervention has an ICER of less than £ 20,000 per QALY compared to the next best alternative. This means that an investment of up to £ 20,000 in order to achieve an additional QALY is considered cost-effective.

The ICER resulting from the CUA will be compared to the willingness to pay threshold of £20,000 per QALY gained as standardised by NICE. No conditions for non-inferiority will be applied in this analysis. Results will be reported as ICERs showing the extra cost of producing one extra QALY or the extra savings achieved by sacrificing one additional QALY.

Sensitivity analyses

For both analyses, deterministic sensitivity analysis will be undertaken to assess the extent to which parameter uncertainty affects the ICERs. Probabilistic sensitivity analyses will be run to estimate the probability of the ICERs falling below a range of willingness to pay (or accept) thresholds as standardised by NICE. For the cost-effectiveness analysis, no such threshold exists, thus we will examine the literature and opinion from the clinical team on what would constitute a suitable willingness to pay. Cost-effectiveness acceptability curves (CEACs) will be produced to visually present the probability of the ICER falling below a range of willingness to pay (accept) thresholds, including the current UK NICE willingness to pay threshold of £20,000 to £30,000 per QALY.

Cost-consequences analysis

A cost-consequence analysis will present all relevant primary and secondary outcomes alongside the costs in tabular form (without combining them into ICERs) to leave decision makers the option to form their own view of relative importance.

Budget impact analysis

A trial based budget impact analysis (BIA) will be undertaken to estimate the likely impact of the use of CRP POCT on NHS budgets through implementation costs, changes in

healthcare usage and potential reduction of antimicrobial resistance. The BIA will be based on the size and composition of the trial population and informed by trial data supplemented by the best available published evidence where required. Sensitivity analyses will be undertaken to estimate the range of a potential budget impact considering parameter uncertainty.

6.4.2 ANALYSIS OF PROCESS EVALUATION DATA

We will describe the use of the CRP machine and interpretation of test results. The distribution of test results will be summarised both numerically and graphically. The scores will also be categorised as <20mg/L, 20-40mg/L, >40mg/L, and antibiotic prescribing (and other management) decisions will be tabulated within these categories.

6.4.3 EXPLORATORY ANALYSIS

Analyses will be conducted that do not focus on a between-group comparison of trial outcomes. While separate, more detailed analysis plans will be written for these, they will be based on the following ideas:

In all participants:

- The association between antibiotic treatment (prescribing and/or use) and subsequent antibiotic resistance
    - Resistance at baseline and prescriptions in previous 12 months
    - Resistance at four weeks and prescription at baseline (and/or use at follow-up)
- Relationship between COPD severity (Gold criteria) and AMR at baseline
- Description of the natural history of acute exacerbations of COPD in primary care
- The association between anxiety at the index consultation (as measured by the anxiety items contained in the EQ-5D and/or the CCQ) and post-consultation behaviour (antibiotic use, use of other treatments, and subsequent healthcare resource use)
- Relationship between AMR at baseline and follow-up
    - Link to antibiotic use
    - The relationship between throat swabs and sputum samples
        - Overall (presence/absence and quantity of organisms grown, detection of likely pathogenic organism, detection of antibiotic resistance)
        - Stratified according to:
            - Order in which the tests were collected (sputum first, or sputum second)
            - Time between sample collection and receipt at laboratory
    - The association between pre-randomisation blood eosinophil count and COPD health status post-randomisation
    - Exploration of variation in presentation, management, and outcomes by practice

- Relationship between prescription and use of oral steroids and outcomes (COPD health status), microbiology (commensals and pathogens)
- Exploration of antibiotic use post-randomisation (initiation and persistence):
    - By COPD grade and exacerbation severity
- Relationship between sputum IL-6 levels at baseline and depression score (domains from CCQ & EQ-5D)
- Relationship between sputum IL-6 levels at week 4 and depression score (domains from CCQ & EQ-5D

In participants allocated to the CRP-arm:

- Predictors of deviations from the guidance around CRP readings and the impact on clinical outcomes (use of medication, healthcare resource use, recovery, antibiotic resistance)
- Relationship between CRP value, sputum colour, and microbiology
- Relationship between congruent antibiotic prescribing / use (given CRP value) and COPD health status
- Exploration of antibiotic use post-randomisation (initiation and persistence)
    - By CRP group and value
    - For high risk group (>40 mg/L), association between antibiotic use (initiation and persistence) and
        - CCQ at 1, 2, and 4 weeks
        - Presence of potentially pathogenic organisms at 4 weeks
        - Presence of antibiotic resistant organism at 4 weeks
        - Adverse effects at 1, 2, and 4 weeks
- Description of the management and outcomes of participants who receive a CRP reading that is in the intermediate range (20-40 mg/L)
    - Presentation characteristics that are associated with this reading
    - Management decisions
    - Outcomes (recovery, treatment use, antibiotic resistance, healthcare resource use, and anxiety)
    - Relationship between blood CRP levels at baseline and depression score (domains from CCQ & EQ-5D)
    - Relationship between blood CRP levels at week 4 and depression score (domains from CCQ & EQ-5D

These exploratory analyses will be completed following the completion of all trial analysis.

6.5 HARMS

Adverse events will be categorised according to seriousness and the extent to which the event was related to the intervention, and summarised according to intervention groups.

6.6 STATISTICAL SOFTWARE

IBM SPSS Statistics version 20 will be used for data management purposes. All analyses will be conducted using Stata version 13.0. Economic analysis will be carried out on SPSS, Stata and MS Excel platforms.

7. REFERENCES

7.1 NON-STANDARD STATISTICAL METHODS

Angrist, J.D., Imbens, G.W. and Rubin, D.B., 1996. Identification of causal effects using instrumental variables. Journal of the American statistical Association, 91(434), pp.444-455.

Kahan, B.C. and Morris, T.P., 2013. Assessing potential sources of clustering in individually randomised trials. BMC medical research methodology, 13(1), p.1.

Little, R.J., 1993. Pattern-mixture models for multivariate incomplete data. Journal of the American Statistical Association, 88(421), pp.125-134.

Offen, W., Chuang-Stein, C., Dmitrienko, A., Littman, G., Maca, J., Meyerson, L., Muirhead, R., Stryszak, P., Baddy, A., Chen, K. and Copley-Merriman, K., 2007. Multiple co-primary endpoints: medical and statistical solutions: a report from the multiple endpoints expert team of the Pharmaceutical Research and Manufacturers of America. Drug Information Journal, 41(1), pp.31-46.

White, I.R., Royston, P. and Wood, A.M., 2011. Multiple imputation using chained equations: issues and guidance for practice. Statistics in medicine, 30(4), pp.377-399.

7.2 DATA MANAGEMENT PLAN

Data management plan location: S:\PCAPH\PCAPH\SEWTU Studies\PACE\15. Data Management and IS\DMP

7.3 TRIAL MASTER FILE AND STATISTICAL MASTER FILE

Trial Master File location: S:\PCAPH\PCAPH\SEWTU Studies\PACE

Statistical Master File location: S:\PCAPH\PCAPH\SEWTU Studies\PACE\16. Statistics

7.4 OTHER SOPS OR GUIDANCE DOCUMENTS

Trial protocol location: S:\PCAPH\PCAPH\SEWTU Studies\PACE\1. Study Protocol\Current protocol

Randomisation protocol location: S:\PCAPH\PCAPH\SEWTU Studies\PACE\16. Statistics\Randomisation

Drummond, M.F., Sculpher, M.J., Claxton, K., Stoddart, G.L. and Torrance, G.W., 2015. Methods for the economic evaluation of health care programmes. Oxford university press.

Herdman, M., Gudex, C., Lloyd, A., Janssen, M.F., Kind, P., Parkin, D., Bonsel, G. and Badia, X., 2011. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Quality of life research, 20(10), pp.1727-1736.

Kocks, J.W., Tuinenga, M.G., Uil, S.M., Van den Berg, J.W.K., Ståhl, E. and Van der Molen, T., 2006. Health status measurement in COPD: the minimal clinically important difference of the clinical COPD questionnaire. Respiratory research, 7(1), p.1.

Llor, C., Moragas, A., Hernández, S., Bayona, C. and Miravitlles, M., 2012. Efficacy of antibiotic therapy for acute exacerbations of mild to moderate chronic obstructive pulmonary disease. American journal of respiratory and critical care medicine, 186(8), pp.716-723.

Schulz, K.F., Altman, D.G. and Moher, D., 2010. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMC medicine, 8(1), p.1.

Schunemann, H.J., Griffith, L., Jaeschke, R., Goldstein, R., Stubbing, D., Austin, P. and Guyatt, G.H., 2003. A comparison of the original chronic respiratory questionnaire with a standardized version. Chest, 124(4), pp.1421-1429.

Van der Molen, T., Willemse, B.W., Schokker, S., Ten Hacken, N.H., Postma, D.S. and Juniper, E.F., 2003. Development, validity and responsiveness of the Clinical COPD Questionnaire. Health and quality of life outcomes, 1(1), p.1.)

SAP/ISAP DEVIATION LOG

Document number:		**Document version:**	
Reason for deviation:			

8. APPENDICES

Appendix I: Scoring and handling of missing items for validated scales

Scale	Scoring	Missing data
CCQ	There are 10 items and 3 domains. Responses range from 0 (very good health status) to 6 (extremely poor health status). Items 1, 2, 5, and 6 correspond to the symptoms domain. Items 7, 8, 9, and 10 correspond to the functional state domain. Items 3 and 4 correspond to the mental state domain. The total score includes all 10 items. Scores are calculated by taking the average (and therefore also range from 0 to 6).	Symptoms domain requires 3 items to be calculated. Functional state domain also requires 3 items. Mental state requires 2 items. Total score requires all domains to be calculable. Where items are missing (but a total score can still be calculated), the domain scores should be multiplied by the total number of items in that domain. These are then summed across all three domains and divided by 10.
CRQ-SAS	There are 20 items, each on a 7-point Likert scale. There are four domains (dyspnoea – Qs 1, 2, 3, 4, 5; fatigue – Qs 8, 11, 15, 17; emotion function – Qs 6, 9, 12, 14, 16, 18, 20; and mastery – Qs 7, 10, 13, 19). Each domain is scored by summing the items and dividing by the total number of questions. The domain scores therefore range from 1 (worst possible score) to 7 (best possible score).	Domains will be scored as long as at least 80% of applicable items are present.

	Items scored as "not done" will be interpreted as "not applicable" and will not be scored.	
EQ-5D-5L	There are five questions / dimensions (mobility, self-care, usual activities, pain / discomfort, anxiety / depression), with each dimension providing five response options, with 1 corresponding to the best health state and 5 corresponding to the worst. These will be dichotomised into no problems/problems and presented descriptively. Profiles of scores across dimensions are converted into index values. The average index value will be compared between trial arms. In addition, there is a Visual Analogue Scale that provides a rating of overall health on a scale from 0 (the worst health you can imagine) to 100 (the best health you can imagine). The average VAS score will also be compared between trial arms.	If any items are missing, the ED-5Q index value cannot be calculated.
Spirometry-confirmed COPD diagnosis and COPD severity (GOLD status)	Spirometry-confirmed COPD diagnosis will be defined as a participant having an FEV1/FVC ratio, dated prior to randomisation, which is less than 0.7. COPD GOLD classification will be defined as follows: Stage 1: FEV1 ≥ 80% predicted	Missing predicted FEV1 values will be derived from age and height (if available).

	Stage 2: FEV1 > 50 and < 80% predicted Stage 3: FEV1 > 30 and ≤ 50% predicted Stage 4: FEV1 ≤ 30% predicted	