PUFFIN trial

CLINICAL TRIAL STATISTICAL ANALYSIS PLAN

To be approved and reviewed by:

Trial name:	Pharmacogenetics Use For Further treatment Improvement in children (PUFFIN) trial
Trial registration number:	NL63849.018.17
Protocol title:	Pharmacogenetics Use For Further treatment Improvement in children (PUFFIN) trial, version 6
SAP version number:	1
SAP date:	15/08/2019

	Name	**Signature**	**Date**
Principal Investigator	Prof. dr. A.H. Maitland van der zee		
Coordinating investigator (author)	E.M.A. Slob, MSc PharmD		
Senior Clinical Trial Statistician	Prof. dr. M.J.C. Eijkemans		

REVISION HISTORY

Version number	Revision history	Author	Date
1			DD/MMM/YYYY

TABLE OF CONTENTS

SECTION 1: ADMINISTRATIVE INFORMATION

1. TITLE AND TRIAL REGISTRATION

Title: Statistical Analysis Plan of the Pharmacogenetics Use For Further treatment Improvement in children (PUFFIN) trial

ClinicalTrials.gov Identifier: NCT03654508
Nederlands Trial register: NTR6727

2. SAP VERSION

SAP version: 1
Date: 10-01-2019

3. PROTOCOL VERSION

Protocol version: 6

4. SAP REVISIONS

4A. SAP revision history

4B. Justification for each SAP revision

5. ROLES AND RESPONSIBILITY

Chief Investigator Trial:
Prof. dr. A.H. Maitland-van der Zee, Department of Respiratory Medicine, Amsterdam University Medical Centers, University of Amsterdam, P.O. Box 22700, NL-1100 DE Amsterdam, The Netherlands

Senior Clinical Trial statistician:
Prof. dr. M.J.C. Eijkemans, Biostatistics and Research Support, Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, Heidelberglaan 100, 3584CX Utrecht, The Netherlands

Coordinating Investigator:
E.M.A. Slob, MSc PharmD, Department of Respiratory Medicine, Ansterdam University Medical Centers, University of Amsterdam, P.O. Box 22700, NL-1100 DE Amsterdam, The Netherlands

6. SIGNATURES

6A. Coordinating Investigator:

E.M.A. Slob Date:

6B. Senior Clinical Trial statistician

M.J.C. Eijkemans Date:

6C. Chief investigator / clinical lead.

A.H. Maitland-van der Zee Date:

SECTION 2: INTRODUCTION

7. BACKGROUND AND RATIONALE

There is large variability between patients in the level of symptom control or lung function improvement upon asthma maintenance treatment. In patients with poorly controlled asthma, improving adherence and inhaler technique are the first steps to improve asthma. However, even in clinical trials in which adherence to treatment is closely monitored, subgroups of patients remain symptomatic despite maintenance treatment (1). Already in 2000, Drazen et al. suggested that up to 80% of the interindividual variance in lung function response upon treatment in asthmatic patients could be due to genetic variations. Since then, several candidate gene studies and a handful of genome-wide association studies (GWAS) have described genetic variants associated with response to asthma treatment.

One of these variants in a gene encoding for the beta-2 adrenergic receptor (ADRB2), has been positively associated poor response to long-acting beta-2 agonists (2-4). This variant (rs1042713) is known as Arg16Gly since the 16th amino acid of the receptor is changed from glycine into arginine and the homozygous Arg16 variant is present in approximately 1 in 6 children (4). A recent meta-analysis in the Pharmacogenomics in Childhood Asthma Consortium (PiCA) of 5 populations with 4,226 children and young adults of white Northern European and Latino origin showed that this variant was associated with an increased risk of asthma exacerbations when treated with LABA as add on treatment (4). Per copy of the risk allele patients exposed to LABA had an increased risk of 52% for a severe asthma exacerbation. There was no increased risk for severe asthma exacerbations if patients were not exposed to LABA. An important observation from this meta-analysis is that the adverse effects of LABA were also observed in heterozygote carriers of the Arg16 variant, thereby providing evidence that over 60 percent of the population (17% homozygotes and 50% heterozygotes) is potentially at risk. This observational study indicated that a large group of paediatric asthma patients might not benefit from LABA and even suffer from more exacerbations, which may be preventable if they would have been treated differently. Besides that this leads to the research question whether *ADRB2* genotyping is effective in children not under control at step 2 of asthma treatment, there is another question that remains unanswered. It is unclear how to treat the patients with the heterozygous variant. This is why it is very important to include heterozygous patients in our study.

Other clinical trials

One previous trial has been performed to address the effects of LABA in relation to *ADRB2* genotype, but only with children homozygous for the risk variant (5). Sixty-two asthmatic children with the Arg16Arg genotype were randomized to treatment with ICS plus LABA or ICS plus LTRA and followed for 1 year. The trial showed that children treated with LTRA had fewer school absences, used less rescue medication, had less symptoms and a better quality of life compared to the group treated with LABA, with no effect on lung function scores between both study arms. The difference between both treatment groups could already be observed after 3 months (figure 1).

Why should children be studied?

It is important to study children with asthma instead of adults since ample data indicate that this genetic risk effect is mainly observed in the paediatric population. The adverse effects of LABA might be more prominent in children than in adults, which was clearly shown in a meta-analysis on the risk of LABA of 110 controlled clinical trials with 60.954 patients performed by the Food and Drug Administration (FDA). It could be that adults with asthma are less vulnerable to the negative effects of LABA due to the influence of other modifying factors such as increased airway wall rigidity (caused by airway remodeling over time), long-term inflammation or a different affinity of the beta-2 adrenergic receptors to their agonists (6). However, a subsequent safety trial mandated by the FDA found no significant difference between the risk of serious asthma events in children receiving a combination of LABA and ICS compared to children who only received an ICS (7).

Therefore, it is important to study the effect of this genetic variant on the treatment outcome in children. This study will include children with asthma to test whether *ADRB2*-genotype guided treatment will lead to better and faster asthma control.

Assessing the costs and benefits of ADRB2-genotyping
Alongside this RCT, a cost-utility analysis will be conducted in order to quantify the incremental costs and benefits of introducing a genetic-based diagnostic tool into clinical asthma practice from a societal perspective. A key research question that we will consider is if *ADRB2*-prospective genotyping is found to be clinically effective, is whether this translates into economically important differences in patients' health, their quality of life, healthcare resource utilization and costs.

References:

1. Szefler SJ, Phillips BR, Martinez FD, Chinchilli VM, Lemanske RF, Strunk RC, et al. Characterization of within-subject responses to fluticasone and montelukast in childhood asthma. J Allergy Clin Immunol. 2005;115(2):233-42.
2. Basu K, Palmer CN, Tavendale R, Lipworth BJ, Mukhopadhyay S. Adrenergic beta(2)-receptor genotype predisposes to exacerbations in steroid-treated asthmatic patients taking frequent albuterol or salmeterol. J Allergy Clin Immunol. 2009;124(6):1188-94.e3.
3. Zuurhout MJ, Vijverberg SJ, Raaijmakers JA, Koenderman L, Postma DS, Koppelman GH, et al. Arg16 ADRB2 genotype increases the risk of asthma exacerbation in children with a reported use of long-acting β2-agonists: results of the PACMAN cohort. Pharmacogenomics. 2013;14(16):1965-71.
4. Turner S, Francis B, Vijverberg S, Pino-Yanes M, Maitland-van der Zee AH, Basu K, et al. Childhood asthma exacerbations and the Arg16 β2-receptor polymorphism: A meta-analysis stratified by treatment. J Allergy Clin Immunol. 2016;138(1):107-13.e5.
5. Lipworth BJ, Basu K, Donald HP, Tavendale R, Macgregor DF, Ogston SA, et al. Tailored second-line therapy in asthmatic children with the Arg(16) genotype. Clin Sci (Lond). 2013;124(8):521-8.
6. McMahon AW, Levenson MS, McEvoy BW, Mosholder AD, Murphy D. Age and Risks of FDA-Approved Long-Acting β2-Adrenergic Receptor Agonists. Pediatrics. 2011;128(5):e1147-e54.
7. Stempel DA, Szefler SJ, Pedersen S, Zeiger RS, Yeakey AM, Lee LA, et al. Safety of Adding Salmeterol to Fluticasone Propionate in Children with Asthma. N Engl J Med. 2016;375(9):840-9.

8. OBJECTIVES

Primary Objective:

- To assess whether *ADRB2* genotype-guided treatment leads to better asthma control after 3 months compared to usual care in children who are uncontrolled despite adherent and adequate use of ICS

Secondary Objectives:

- To assess whether *ADRB2* genotype-guided treatment leads to better asthma control at 6 months.
- To assess whether *ADRB2* genotype-guided treatment leads to improved quality of life (QoL), fewer school absences, fewer exacerbations, and better lung function compared to usual care in children at 3 and 6 months
- To assess whether *ADRB2* genotype-guided treatment leads to fewer changes in asthma therapy at 3 months, compared to usual care.
- To assess whether *ADRB2* genotype-guided treatment leads to a shorter time to reach asthma control, compared to usual care
- To assess the cost-utility of *ADRB2*-genotype guided treatment
- To identify –omics-biomarkers for non-response to ICS treatment

SECTION 3: STUDY METHODS

9. TRIAL DESIGN

Study design
National, multi-centre, double-blind randomized controlled trial

Duration
6 months, with 3 visits in the hospitals (at t=0, t=3 months and t=6 months)

Setting
Patients are recruited at out-patient asthma clinics in secondary and tertiary care hospitals in the Netherlands.

Description
Three hundred ten children (6 to 17 years of age) with a doctor's diagnosis of asthma and uncontrolled asthma symptoms despite adherent and adequate use of ICS for at least three months (step 2 asthma treatment) will be recruited by secondary and tertiary care centers in the Netherlands. All participants are eligible for step-up asthma treatment (from step 2 to step 3) as assessed by the treating paediatrician/paediatric pulmonologist Participants will be randomized to a genotype-guided treatment arm (n=155) or to a usual care, non-genotype guided, arm (n=155) (Figure 2) and followed for 6 months.

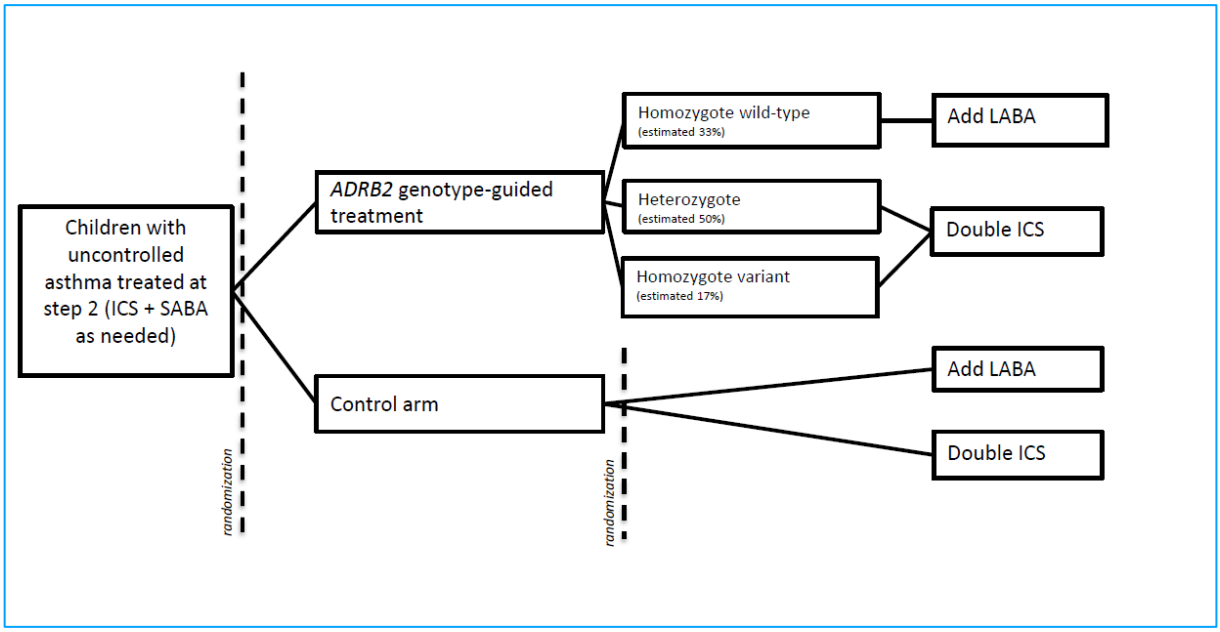

Figure 1. PUFFIN study design

Genotyping before start treatment
During the baseline visit in the hospital, clinical data and biological samples (including a DNA sample) will be collected. Upon this visit, the DNA sample will be send to the Clinical Chemistry department of the Erasmus MC (Head: Prof. R. van Schaik) to perform genotyping of the *ADRB2* gene within one week. The treating physician will adapt the treatment regime of the participant based on the treatment advice of the study coordinator (Table 1). For the children in the genotype-arm, this will be based on the genotype. The treating physician will not know (be blinded) whether the treatment advice was based on the genotype (intervention arm) or based on randomization (control arm). The participant will be followed for 6 months. If the participant is still uncontrolled at t=3 months, treatment will be adapted according to Table 1. All children will be genotyped, in order to assess the influence of the genotype on treatment outcome in the usual arm group retrospectively. The children should use the same inhalation device during the study to avoid confusion on how they should inhale their medication.

Table 1. Treatment regimes

	Arg16Arg or Arg16Gly	**Gly16Gly**	
Therapy: month 0- 3	Double ICS	Double ICS	ICS+LABA
Therapy for month 4-6 if still uncontrolled at 3 months	Normal dosage ICS and LTRA	Normal dosage of ICS and LABA	Double ICS

Intervention arm: ADRB2 genotype-guided treatment arm
In the genotype-stratified arm, children will be treated based on their *ADRB2* genotype. Children homozygous for the risk variant Arg16 and heterozygotes (Arg16Gly) will be treated with doubling dosages of their ICS. Children homozygous for the wild type allele (Gly16Gly) will receive LABA.

Control arm: Non genotype-guided treatment arm
In the control arm, genotyping will be performed for retrospective analysis, but the genotype information will not be used to guide treatment. Children in this study arm will be randomized again between doubling ICS dosage (n=75) or LABA treatment (n=75), the two most commonly preferred add-on options among paediatric pulmonologists in the Netherlands. We choose to randomize between both treatments options, since international guidelines do not agree on the preferred treatment option.
Furthermore, to test our hypothesis it is necessary to have enough children in the control group with Arg16Arg or Arg16Gly to be treated with LABA. The amount of children treated with LABA and ICS should be equal in the control group. Therefore we decided to randomise children in the control group over doubling ICS (n=77) and adding LABA (n=77). This will lead to an estimated number of children with Arg16Arg or Arg16Gly of 51 who will get LABA add on. In this way the power is high enough to determine the effectivity of both treatment options in the three genotypes. We find it important to define effectivity next to the question whether genotyping benefits children with asthma. In the control group DNA samples will be obtained for retrospective analysis.
It is safe to randomise the children again who are randomised within our control arm, because treatment with a double dose of ICS and adding a LABA are both standard of care. A Cochrane review from 2009 has shown that both treatments have proven to be equally effective in both children and adults
Randomisation in the control arm is important because it would be futile if the children in this arm would be treated with the same therapy by accident. Randomisation is necessary to make the trial as small and effective as possible. At this moment physicians do not have the tools to determine which therapy is the best for every child. This is why we think it is correct to randomise in the control arm.

Based on the previous studies of Lipworth et al. (5), and Turner et al. (4), we hypothesize that children with one or two Arg16 alleles in *ADRB2* will experience less asthma control, thus randomization in the control arm will ensure sufficient children with these *ADRB2* genotypes to be exposed to LABA in order to test our hypothesis. We will perform interim genetic analysis to verify that we include sufficient children with these risk genotypes in our study. Our trial is designed to reflect clinical practice as closely as possible, therefore the choice of the type of ICS or LABA will be done by the treating physician.

10. RANDOMIZATION

Randomisation between intervention and control arm
Participants will be randomized 1:1 to the intervention arm or the control arm. Block randomization with randomly chosen block sizes and stratified per center (academic/non-academic) will be applied. ALEA is used to generate randomization codes.

Randomisation within control arm
Children in the control arm will be randomized 1:1 to a) doubling ICS dosage or b) adding LABA to the treatment regime. Block randomization with randomly chosen block sizes and stratified per center (academic/non-academic) will be applied. Randomization software will be used to generate randomization codes.

11. SAMPLE SIZE

A sample size analysis was performed for Arg16 homozygotes double dose ICS vs. LABA treatment (our main research question), with power 80%, α=0.05 2-sided, and based on the ACT scores and correlation between repeated measurements in a previous performed RCT in children with asthma (the BATMAN study) (SD=3.6, r=0.58). Taking into account a minimal clinically relevant significant difference of 3 points on the ACT scores, that 16% of the children in the PACMAN study were carrying Arg16Arg, and that ACT will be measured 4 times (baseline t=0, t=1, t=2 and t=3) before the primary endpoint (at t=3 months), a minimum of 153 children need to be included in each study arm. When the analysis will be adjusted for the baseline value, this will reduce the variance of the estimate of difference between the group by a factor 1-r^2, and thereby increase the power. Because not all children might be willing to complete the trial we will include a total of 310 children. The power is based on the inclusion of 16% of children carrying Arg16Arg. We will count the amount of children with this genotype when 310 children have been included. If this amount is not reached after 310 children we will continue the trial until we have included the expected 49 children with the Arg16Arg genotype.

12. FRAMEWORK

The PUFFIN trial protocol states that the primary objective is: “To assess whether ADRB2 genotype-guided treatment leads to better asthma control after 3 months compared to usual care in children who are uncontrolled despite adherent and adequate use of ICS.” Therefore, the primairy outcome is testing for superiority.

The PUFFIN trial protocol states that one of the secondary objectives is “To assess whether ADRB2 genotype-guided treatment leads to better asthma control at 6 months.” Therefore, this objective is testing for superiority.

The PUFFIN trial protocol states that one of the secondary objectives is “To assess whether ADRB2 genotype-guided treatment leads to improved quality of life (QoL), fewer school absences, fewer exacerbations, and better lung function compared to usual care in children at 3 and 6 months.” Therefore, this objective is testing for superiority.

The PUFFIN trial protocol states that one of the secondary objectives is “To assess whether ADRB2 genotype-guided treatment leads to fewer changes in asthma therapy at 3 months, compared to usual care.” Therefore, this objective is testing for superiority.

The PUFFIN trial protocol states that one of the secondary objectives is “To assess whether ADRB2 genotype-guided treatment leads to a shorter time to reach asthma control, compared to usual care.” Therefore, this objective is testing for superiority.

The PUFFIN trial protocol states that one of the secondary objectives is “To assess the cost-utility of ADRB2-genotype guided treatment.” The intention of this objective is to show that ADRB2 genotype guided treatment is cost-effective. Therefore, this objective is testing for non-inferiority rather than superiority for the other objectives.

13. STATISTICAL INTERIM ANALYSES AND STOPPING GUIDANCE

13A. Information on interim analyses

One formal statistical interim analysis is planned to be conducted after inclusion of 310 patients, to count whether 16% of the patients has the Arg16Arg genotype.

13B. Adjustment to the significance level due to interim analyses

No adjustment to the significance level is needed, as the data will only be analyzed until 49 children with the Arg16Arg genotype are included for the final study outcomes.

13C. Details for stopping the trial early

The trial will not be stopped early as it is a minimal invasive intervention and no harm will be done to the included children.

14. TIMING OF FINAL ANALYSIS

All analyses will be performed collectively after all the 310 included children including at least 49 Arg16Arg patients are included and followed for 6 months.

15. TIMING OF OUTCOME ASSESSMENTS

The study consists of 3 clinical visits (t=0, t=3 months (± 2 weeks) , t=6 months (± 2 weeks)) and monthly online questionnaires (Figure 3).

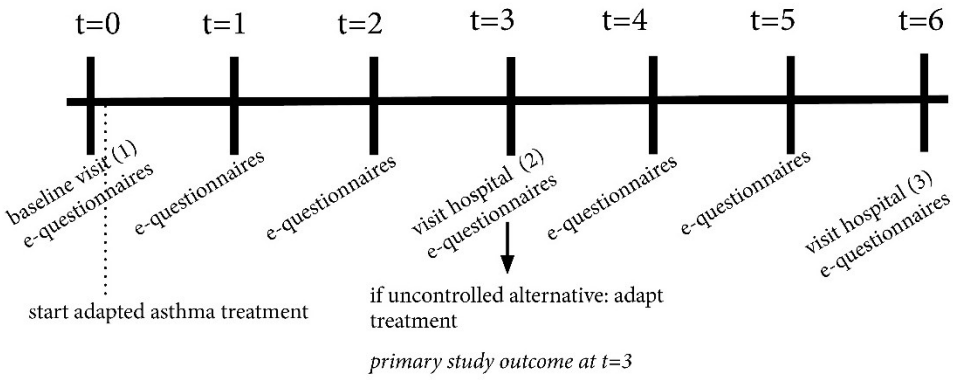

Figure 2: Timeline PUFFIN trial

Informing participants

During a regular care visit, the treating physician will inform the parents and the patient about the study. The physician will contact the PUFFIN investigators. Patients and parents will get at least 24 hours to consider participation. A research assistant of the hospital of the child will contact the patients and parents by phone to inform whether they would like to participate. In case they are willing to participate, a study visit will be planned within 2 weeks.

Measurements during clinical visit 1

Screening

Upon informed consent, the patient will be screened to check whether he/she fulfills the inclusion criteria (e.g. current asthma symptoms, adequate inhalation technique). In case he/she fulfills these criteria, the following study measurements will be performed:

Lung function testing (in case this has not been performed during the last clinical routine visit, < 2 weeks preceding the study visit):

- Spirometry: Forced expiratory Volume in 1 second (FEV_1), Forced Vital Capacity (FVC) and FEV_1/FVC ratio, Forced Expiratory Flow after 75% of the exhaled volume (FEF75) before and after inhalation of salbutamol will be measured.
- Fraction of exhaled nitric oxide (FeNO) will be measured before spirometry in all applicable sites.

Children will be instructed to stop their short acting bronchodilators at least 8 hours prior to lung function testing. Their long acting bronchodilators should be discontinued at least 24 hours prior to lung function testing.

NB: A lung function measurement could have been assessed during the screening visit. In this case, this measurement is eligible to count for visit 1.

Clinical review

Measure height/weight.

Questionnaires:

Children and parents are asked to complete several questionnaires (if possible online):

- *Asthma control test*

The validated Dutch version of the Asthma Control Test (ACT, for children of ≥ 12 years) or Childhood Asthma Control Test (C-ACT, for children < 12 years) inquires on asthma symptoms in the past 4 weeks.

- *Exacerbations*

Severe exacerbations (asthma-related unscheduled health care visits, use of OCS, admissions), as well as mild-to-moderate exacerbations (sudden increase of symptoms, asthma attack requiring additional rescue medication, unplanned visits to general practitioner for asthma) and school absences (in days) due to asthma symptoms will also be recorded.

- *Asthma medication use and adherence to maintenance treatment*

Questions on current asthma medication will be included. Parents and children are asked for consent to extract medication dispensing data of the child from the local pharmacies

- *Asthma-Quality of Life*

In children aged 12 years and older, asthma-related quality of life will be measured with the 13-item self-reported Dutch validated version of the Paediatric Asthma-Related Quality of Life Questionnaire (PAQLQ) for children and expressed as overall asthma-related quality of life. In children aged below 12 years, we use the Paediatric Asthma-Related Caregiver Quality of Life Questionnaire (PACQLQ).

- *Productivity loss parents*

Modules of the Productivity Cost Questionnaire (PCQ) to assess the loss of productivity of the caregivers will be included in the questionnaire.

- *Fatigue*

The PedsQL Multidimensional Fatigue Score will be included to assess symptoms of fatigue in the past month.

- Control of Allergic Rhinitis

In case, children suffer from allergic rhinitis in addition to asthma, The Control of Allergic Rhinitis and Asthma Test (CARAT) will be used to assess the level of control of allergic rhinitis.

- *Other questions*

Furthermore we will include questions on allergy and rhinitis complaints, environmental factors, pre- and postnatal factors, demographics and diet (related to the microbiome).

Noninvasive procedures

Two nasal epithelial swab (Copan Flocked Swabs), nr 56380CS01 per nostril will be taken for DNA and RNA isolation. This sample will be taken by gently rotating the flocked swab at the site of the lower inferior turbinate. Furthermore, a saliva sample will be taken for genomic DNA isolation. We will also ask the children to send a feces sample to the researchers after the study visit for microbiomics/metabolomics analysis. Within the AMC inclusion site, breath will be analyzed using the SpiroNose. For this measurement, children just have to breathe into the SpiroNose.

Laboratory testing:

Nasal epithelial swabs: DNA and RNA isolation will be performed. Whole genome gene-expression and epigenomics analyses will be performed.

Saliva: DNA will be isolated and genotyped for the *ADRB2* gene.

Feces: short-chain fatty acids will be measured, 16s RNA sequencing will be performed to analyze the microbiome.

Clinical visit 2 (t=3 months (± 2 weeks))

During this visit asthma control is assessed. In case children are still not controlled with current treatment, they will receive alternative treatment according to Table 1.

The following measurements will be performed during this visit:

- clinical review (similar to clinical visit 1)
- lung function measurements (similar to clinical visit 1), FeNO measurement
- online asthma questionnaire (less extensive as clinical visit 1, but including questions on current asthma medication use, symptoms, asthma-related quality of life, fatigue and exacerbations)
- Nose swabs for DNA and RNA extraction.

Clinical visit 3 (t=6 months (± 2 weeks))

Same measurements as clinical visit 2.

SECTION 4: STATISTICAL PRINCIPLES

16. LEVEL OF STATISTICAL SIGNIFICANCE

All applicable statistical tests will be 2-sided and will be performed using a 5% significance level (α=0.05).

17. DESCRIPTION AND RATIONALE FOR ANY ADJUSTMENT FOR MULTIPLICITY

Depending on the amount of tests that are conducted, corrections for multiplicity will be performed.

18. CONFIDENCE INTERVALS

All confidence intervals presented will be 95% and two-sided.

19. ADHERENCE AND PROTOCOL DEVIATIONS

19A. Definition of adherence

Adherence is one of the inclusion criteria and should be considered sufficient (>70%) according to the doctor's believes.

19B. Definition of protocol deviations

The following are pre-defined major protocol violations with a direct bearing on the primary outcome:

1) The included patient does not meet the in- and/or exclusion criteria
2) Other respiratory medication/montelukast being prescribed to the patient not according to the description of the study medication in the PUFFIN trial protocol.

19C. Summarised protocol deviations

Protocol deviations are classified prior to unblinding of treatment. The number (and percentage) of patients with major and minor protocol deviations will be summarized by treatment group with details of type of deviation provided. Patients that did not meet the in- or exclusion criteria will not be included in the analysis. Patients that are included in the intention to treat analysis data set will be used as the denominator to calculate the percentages. No formal statistical testing will be undertaken.

20. ANALYSIS POPULATIONS

The population will be analysed according to the intention to treat principle. Additionally, a per protocol analysis will be performed.

SECTION 5: TRIAL POPULATION

21. SCREENING DATA

The following enrolment summaries will be presented for all screened patients: the number of days recruiting, the number of patients screened, the number of patients recruited, the number of patients recruited per day, the number of screened patients not recruited and the reason for non-recruitment. This summary will be provided overall and by study center.

22. ELIGIBILITY

The number of ineligible patients randomized will be reported with reasons for ineligibility.

23. RECRUITMENT

A CONSORT flow diagram will be used to summarise the number of patients who were:

- assessed for eligibility at screening
    - eligible at screening
    - ineligible at screening*
- eligible and randomized
- eligible but not randomised*
- received the randomised allocation
- did not receive the randomised allocation*
- lost to follow-up*
- discontinued the intervention*
- randomised and included the primary analysis
- randomized and excluded from the primary analysis*

* reasons will be provided.

24. WITHDRAWAL / FOLLOW-UP

24A. Level of withdrawal

The level of consent withdrawal will be tabulated (classified as “consent to continue follow-up and data collection”, “consent to continue data collection only”, “complete – no further follow-up or data collection”).

24B. Timing of withdrawal

This will be presented in CONSORT diagram format rather than as a table, with numbers and reasons for withdrawal and/or exclusion from analysis given at each stage (visit 1, t=1, t=2, visit 2, t=4, t=5 or visit 3).

24C. Presentation of withdrawal/lost to follow-up data

The numbers (with reasons) of losses to follow-up (drop-outs and withdrawals) over the course of the trial will be summarised by treatment arm.

25. BASELINE PATIENT CHARACTERISTICS

25A. List of baseline characteristics to be summarised

Patients will be described with respect to age, gender, body mass index, ethnicity, baseline (c-)ACT-score, allergic vs. non-allergic asthma. These variables can be found in Castor at visit 1, ‘Demografie’.

25B. Details of how baseline characteristics will be summarised

Categorical data will be summarised by numbers and percentages. Continuous data will be summarised by mean, SD and range if data are normal and median, IQR and range if data are paranormally distributed. Minimum and maximum values will also be presented for continuous data. Baseline differences will be tested with Wilcoxon Ranks Sum Tests or Chi-squared tests.

SECTION 6: ANALYSIS

26. DATASETS

26.1 General data set
The General data set includes all patients who have had at least one measurement during the PUFFIN trial. This set will be used to summarise demographics, baseline characteristics, and adverse events.

26.2 Modified general data set
The modified general data set is defined as all patients who have been randomized for the PUFFIN trial. This analysis set will be used for efficacy analysis and cost-effectiveness analysis.

26.3 Per protocol set
The per protocol set is a subset of the modified general data set (26.2), excluding patients with major protocol deviations. This population may be used to perform a sensitivity analysis to summarize efficacy when patients are treated per protocol. Protocol deviations will be classified in the following categories:

- Did not fulfil eligibility criteria
- Met criteria to switch medication after three months but continued the medication of the initial randomisation
- Received incorrect medication despite the advice after randomisation/dose
- Received prohibited concomitant medication
- Protocol-required procedure not adhered to
- Others

The per protocol analysis set will be identified prior to database lock.

27. OUTCOME DEFINITIONS

All outcomes will be calculated for both treatment arms: the genotype guided treatment arm and the control arm, unless otherwise stated

Primary Outcome:

- Change in asthma control, based on (c-)ACT score difference, after 3 months compared to baseline

(c-)ACT score
The (c-)ACT contains of five questions with a 5-point scale (for symptoms and activities: 1=all the time to 5= not at all; for asthma control rating: 1=not controlled at all to 5=completely controlled). The total score ranges from 5 (poor control of asthma) to 25 (complete control of asthma) points, with higher scores reflecting greater asthma control. An (c-)ACT score >19 points indicates well-controlled asthma.

Required variables
For this outcome the (c-)ACT score at t=0, t=1, t=2 and t=3 (obtained during visit 1 and 2 and via e-questionaires) is required. If one of these questionnaires is not completed, the analysis will be based on the non-missing data. If the baseline data of this outcome is not completed, the data will be regarded as missing and the baseline data will be imputed with multiple imputation.

Included trial patients
To assess whether children with a variant genotype have a poorer improvement in asthma control we will assess:
1. The improvement in asthma control is different between patients with Arg16Arg treated with LABA and treated with double dose ICS
2. The improvement in asthma control is different between patients with Arg16Gly treated with LABA and treated with double dose ICS
3. The improvement in asthma control is different between patients with Gly16Gly treated with LABA and treated with double dose ICS

To assess the clinical impact of a genotype-guided strategy on improvement in asthma control we will assess whether:

4. The improvement in asthma control is different between patients in the genotype-guided treatment arm compared to the control arm.

Secondary Outcomes:

- Change in asthma control, based on (c-)ACT score difference, after 6 months compared to baseline

  Required variables
  For this outcome the (c-)ACT score at t=0 and t=6 (obtained during visit 1 and 3) is required. The mean/median difference of the outcome at t=6 and t=0 will be calculated. If one of these questionnaires is not completed, the analysis will be based on the non-missing data. If the baseline data of this outcome is not completed, the data will be regarded as missing and the baseline data will be imputed with multiple imputation.

  Included trial patients
  The same groups of patients will be used for the assessments as for the primary outcome.

- Time to (c-)ACT ≥ 20

  Required variables
  For this outcome the (c-)ACT score at visit 1, t=1, t=2, visit 2, t=4, t=5 and visit 3 is required. The first month number (e.g. for month 2 as the first month it is 2, and for month 3 as the first month it is 3) from baseline that the ACT ≥ 20 will be counted as the outcome. If one of these questionnaires is not completed, the analysis will be based on the non-missing data. If the baseline data of this outcome is not completed, the data will be regarded as missing and the baseline data will be imputed with multiple imputation.

  Included trial patients
  The same groups of patients will be used for the assessments as for the primary outcome.

- Change in asthma-related quality of life (PA(C)QLQ) scores after 3 and after 6 months

  PA(C)QLQ
  The PAQLQ has 23 questions in 3 domains (symptoms, activity limitation and emotional function). The activity domain contains 3 'patient-specific' questions and is designed for children of twelve years or older. The PACQLQ (designed for the parents of the children of eleven years or younger) has 13 questions and is designed to be filled in by the parents of children of eleven years or younger. Children and caregivers are asked to think about how they have been during the previous week and to respond to each of the 23 or 13 questions respectively on a 7-point scale (7 = not bothered at all - 1 = extremely bothered). The overall PA(C)QLQ score is the mean of all 23 responses and the individual domain scores are the means of the items in those domains. The minimally clinical important difference is reflected by a change in score of approximately 0.5 on a 7 point scale.

  Required variables
  For this outcome the mean/median outcome of the 23 PAQLQ or 13 PACQLQ questions per patient will be calculated for t=0, t=3 and t=6. The difference will be calculated between the PA(C)QLQ scores at t=3 and t=0 for after 3 months and between t=6 and t=0 for after 6 months. If one of these questionnaires is not completed, the analysis will be based on the non-missing data. If the baseline data of this outcome is not completed, the data will be regarded as missing and the baseline data will be imputed with multiple imputation.

  Included trial patients
  The same groups of patients will be used for the assessments as for the primary outcome.

- Change in fatigue score after 3 and after 6 months

PedsQL
Measured by the PedsQL questionnaire, a 23 item generic score scale consisting of physiological functioning, emotional functioning, social functioning and school functioning questions with multidimensional scales. Subscales are summed. This modular instrument uses a 5-point scale: from 0 (never) to 4 (almost always). Items are reversed scored and linearly transformed to a 0-100 scale as follows: 0=100, 1=75, 2=50, 3=25, 4=0. Four dimensions (physical, emotional, social, & school functioning) are scored. The minimally clinical important difference is reflected by a change in score of approximately 5 on the total score.

Required variables
For this outcome the mean/median per dimension should be calculated as a sum of the item over the number of items answered for t=0, t=3 and t=6. The mean/median difference will be calculated between the PedsQL scores at t=3 and t=0 for after 3 months and between t=6 and t=0 for after 6 months. The total score is the sum of all the items over the number of items answered on all the scales. If one of these questionnaires is not completed, the analysis will be based on the non-missing data. If the baseline data of this outcome is not completed, the data will be regarded as missing and the baseline data will be imputed with multiple imputation.

Included trial patients
The same groups of patients will be used for the assessments as for the primary outcome.

- Change in school absences after 3 and after 6 months

Questionnaire about school absences
This secondary outcome is measured by the question: "*How many school days (or part of school days) did your child miss due to his/her asthma complaints during the last 3 months?*"

Required variables
For this outcome the mean/median difference in amount of days between t=0 and t=3 and between t=3 and t=6 will be calculated and the proportion of patients with school absences will be counted per group. If one of these questionnaires is not completed, the analysis will be based on the non-missing data. If the baseline data of this outcome is not completed, the data will be regarded as missing and the baseline data will be imputed with multiple imputation.

Included trial patients
The same groups of patients will be used for the assessments as for the primary outcome.

- Change in exacerbations (oral corticosteroids use, ER visits, hospital admissions)

<u>Oral corticosteroids use</u>

Questionnaire about oral corticosteroids use
The oral corticosteroids use was recorded during study visit 2 and visit 3 by: *"Did your child use prednisone during the last three months? Yes/No"* If the answer to this question is yes, it is followed by: *"How many days did the usage last?"*

Required variables
For this outcome the mean/median difference in amount of oral corticosteroids users will be calculated between t=3 and t=0 and t=6 and t=0. The mean amount of days of oral corticosteroid usage will be calculated between t=3 and t=0 and between t=6 and t=3 will be calculated. If one of these questionnaires is not completed, the analysis will be based on the non-missing data. If the baseline data of this outcome is not completed, the data will be regarded as missing and the baseline data will be imputed with multiple imputation.

Included trial patients
The same groups of patients will be used for the assessments as for the primary outcome.

ER visits

Questionnaire about ER visits
The ER visits are recorded by the online e-question: *"Has your child visited the ER during the last year due to acute airway complaints? Yes/No"* If the answer to this question is yes, it is followed by: "*How many times did you visit the ER?"*

Required variables

For this outcome the mean/median difference in amount of patients with ER visits will be calculated between t=3 and t=0 and between t=6 and t=0. The difference in mean amount of ER visit per patient will be calculated between t=3 and t=6 and between t=6 and t=0. If one of these questionnaires is not completed, the analysis will be based on the non-missing data. If the baseline data of this outcome is not completed, the data will be regarded as missing and the baseline data will be imputed with multiple imputation.

Included trial patients
The same groups of patients will be used for the assessments as for the primary outcome.

Hospital admissions

Questionnaire about hospital admissions
The hospital admissions are recorded by the online e-question: "*Was your child admitted to the hospital during the last year due to acute airway complaints? Yes/No?"* If the answer to this question is yes, it is followed by: "*How many times did you visit the hospital?"*

Required variables
For this outcome the mean/median difference in amount of patients with hospitalization will be calculated between t=3 and t=0 and between t=6 and t=0. The difference in mean amount of hospitalizations per patient will be calculated between t=3 and t=6 and between t=6 and t=0. If one of these questionnaires is not completed, the analysis will be based on the non-missing data. If the baseline data of this outcome is not completed, the data will be regarded as missing and the baseline data will be imputed with multiple imputation.

Included trial patients
The same groups of patients will be used for the assessments as for the primary outcome.

- Time to first exacerbation

Exacerbations outcome
The exacerbations are recorded by the online e-questions described above. This will be a composite outcome of oral corticosteroids use, ER visits and hospital admissions. For the three outcomes, when the first question is answered with yes, the patient is asked with the following question when the starting date was. The date that is the first date of the three of these outcomes, will be used to describe the first time to exacerbation.

For this outcome the mean/median difference in days between the date of the first exacerbation and date of informed consent for the PUFFIN trial will be calculated. If one of these questionnaires is not completed, the analysis will be based on the non-missing data. If the baseline data of this outcome is not completed, the data will be regarded as missing and the baseline data will be imputed with multiple imputation.

Included trial patients
The same groups of patients will be used for the assessments as for the primary outcome.

- Amount of changes in therapy at t=3 months

Required variable

During visit 2 the variable '*veranderingstudiemedicatie*' will be answered. For this outcome the means/medians of changes in therapy will be calculated at visit 2. If one of these questionnaires is not completed, the analysis will be based on the non-missing data. If the baseline data of this outcome is not completed, the data will be regarded as missing and the baseline data will be imputed with multiple imputation.

Included trial patients
We would like to test whether the amount of changes in therapy is different between patients in the genotype-guided treatment arm compared to the control arm.

- Change in lung function (FEV1 pre- and postbronchodilator) at t=3 and t=6 months
  Means/medians of differences in change in FEV1 pre- and postbronchodilator reversibility (%) at t=3 and t=6 will be assessed between patients with different genotypes treated with LABA or double dose ICS, and between patients in the intervention arm and the control arm, using the following formula:

  $$\% \text{ FEV1 Reversibility} = \frac{\text{Post-bronchodilator FEV1} - \text{Pre bronchodilator FEV1}}{\text{Pre-bronchodilator FEV1}} \times 100$$

  If one of these questionnaires is not completed, the analysis will be based on the non-missing data. If the baseline data of this outcome is not completed, the data will be regarded as missing and the baseline data will be imputed with multiple imputation.

- Change in FeNO at t=3 and t=6 months
  FeNO is measured as a continuous variable in parts per billion (ppb). We will calculate the median/mean difference level of FeNO between patients with different *ADRB2* genotypes treated with LABA or double dose ICS, and between patients in the intervention arm and in the control arm for this outcome measure.
  The means/median differences will be calculated for the difference between t=3 and t=0 and t=6 and t=0. If one of these questionnaires is not completed, the data will be imputed with multiple imputation to improve the balance of sample sizes in both arms. If one of these questionnaires is not completed, the analysis will be based on the non-missing data. If the baseline data of this outcome is not completed, the data will be regarded as missing and the baseline data will be imputed with multiple imputation.

  Included trial patients
  The same groups of patients will be used for the assessments as for the primary outcome.

- Incremental cost per quality adjusted life year (QALY)

  iPCQ and required variables
  The iPCQ will be used to calculate the productivity costs of the patients. The IPCQ measures:
  1. A*bsenteeism* (number of days missed from work in the 4 weeks due to illness) -- minimum is 0, maximum (in theory) is 28 days. Questionnaire 4, 5 and 6 are used for absenteeism. Question 4 asks for the working days that the participant was absent (short absenteeism). If the absenteeism started before the recall-period (question 5), in question 6 the date is asked when the absenteeism has started (long absenteeism).

     Short absenteeism = Number of working days absent * amount of hours in a working day * productivity value per hour

     Long absenteeism = ((Calendar day start absenteeism – calendar day of filling in iPCQ)/7) * amount of working days per week * amount of working hours per day * productivity value per hour

  2. *Presentism* (productivity lost while at work due to illness) -- minimum is 0, maximum is 224 hours. Questionnaire 7, 8 and 9 will be used to calculate presentism. Question 7 asks whether there was inconvenience due to the health problems during the work. If this question is answered with 'no', there was no productivity loss. If this question is answered with 'yes', question 8 will ask for the amount of working days affected by illness and question 9 will ask for the efficiency score (from 0 to 10).

presentism = number of working days affected by illness * (1-efficiency score/10)*number of working hours per working day * productivity value per hour

3. *Productivity loss due to unpaid work.* Question 10, 11 and 12 will be used for this outcome measure. Question 10 asks whether there was inconvenience due to health problems during the unpaid work. If this question is answered with 'no', there was no productivity loss in the unpaid work. If this question is answered with 'yes', question 11 will ask for the amount of working days affected by illness and question 12 will ask for the amount of hours lost.

   Productivity loss due to unpaid work = number of unpaid working days affected by illness * (1-efficiency score/10)*number of working hours per unpaid working day * standard hour loan for household care (in 2019: €11,07)

After measuring the productivity loss in hours and days, this is valued sing the friction cost method to obtain a monetary value for the productivity loss.

Included trial patients
The same groups of patients will be used for the assessments as for the primary outcome.

- Costs
  Volumes in resource use in terms of hospital admissions, ER visits and drug use are described elsewhere. The found volume of resource use is then valued
  according to guidelines for economic evaluation in healthcare in the Netherlands (ZIN). Costs will be presented in 2018 euros. Prices from previous years will be updated according to the Dutch consumer price index. The costs of the intervention were calculated by [material costs, time of specialists, capacity} .

  Included trial patients
  The same groups of patients will be used for the assessments as for the primary outcome.

  iPCQ and required variables
  The iPCQ will be used to calculate the productivity costs of the patients. The IPCQ measures:
  1. Absenteeism (number of days missed from work in the 4 weeks due to ilness) -- minimum is 0, maximum (in theory) is 28 days. Questionnaire 4, 5 and 6 are used for absenteeism. Question 4 asks for the working days that the participant was absent (short absenteeism). If the absenteeism started before the recall-period (question 5), in question 6 the date is asked when the absenteeism has started (long absenteeism).

  Short absenteeism = Number of working days absent * amount of hours in a working day * productivity value per hour

  Long absenteeism = ((Calenderday start absenteeism – calenderday of filling in iPCQ)/7) * amount of working days per week * amount of working hours per day * productivity value per hour

  2. Presenteeism (productivity lost while at work due to ilness) -- minimum is 0, maximum is 224 hours. Questionnaire 7, 8 and 9 will be used to calculate presenteeism. Question 7 asks whether there was inconvenience due to the health problems during the work. If this question is answered with 'no', there was no productivity loss. If this question is answered with 'yes', question 8 will ask for the amount of working days affected by illness and question 9 will ask for the efficiency score (from 0 to 10).

  presenteeism = number of working days affected by illness * (1-efficiency score/10)*number of working hours per working day * productivity value per hour

  3. Productivity loss due to unpaid work. Question 10, 11 and 12 will be used for this outcome measure. Question 10 asks whether there was inconvenience due to health problems during the unpaid work. If this question is answered with 'no', there was no productivity loss in the unpaid work. If this question is answered

with ‘yes’, question 11 will ask for the amount of working days affected by illness and question 12 will ask for the amount of hours lost.

Productivity loss due to unpaid work = number of unpaid working days affected by illness * (1-efficiency score/10)*number of working hours per unpaid working day * standard hour loan for household care (in 2019: €11,07)

After measuring the productivity loss in hours and days, this is valued sing the friction cost method in order to obtain a monetary value for the productivity loss.

Included trial patients
The same groups of patients will be used for the assessments as for the primary outcome.

28. ANALYSIS METHODS

In paragraph 28A, the general analysis methods will be discussed. The covariates used for these analyses will be discussed in paragraph 28B. In paragraph 27C the methods used for assumptions will be described, followed by paragraph 27D for the sensitivity analyses and 27E for the subgroup analyses.

28A. Analysis methods

Primary Outcome:

- Change in asthma control, based on (c-)ACT score difference, after 3 months compared to baseline

Change in asthma control after 3 months will be used to test for four different comparisons as described in paragraph 26. Repeated measures mixed models will be used to study whether there are significant differences between the groups using ACT scores at t=0, t=1, t=2 and t=3, adjusting for the baseline value of the (c-)ACT score.

If the repeated measures mixed model produces significant differences between the groups, then an analysis of covariance (ANCOVA), adjusting for the baseline (c-)ACT score, can be performed at each time point, to identify at which time points these differences become apparent. The mean/median difference in asthma control will be reported, including a 95% confidence interval and p-value.

Secondary Outcomes:

- Change in asthma control, based on (c-)ACT score difference, after 6 months compared to baseline
  Change in asthma control after 6 months will be reported and analysis in 4 comparisons as described in paragraph 27. Repeated measures mixed models will be used to study whether there are significant differences between the groups using ACT scores at t=0, t=1, t=2, t=3, t=4, t=5 and t=6.

  If the repeated measures mixed model produces significant differences between the groups, then an analysis of covariance (ANCOVA), adjusting for the baseline (c-)ACT score, can be performed at each time point, to identify at which time points these differences become apparent. The mean/median difference in asthma control will be reported, including a 95% confidence interval and p-value.

- Time to (c-)ACT ≥ 20
  For this outcome the variable described in paragraph 27 will be used to test for four different comparisons. A log-rank test or Cox semiparametric proportional hazards model will be used to study whether there are differences in time to reach a (c-)ACT score of 20 points or more. The Chi-square with one degree of freedom or estimated relative hazard will be reported, including 95% confidence intervals and p-values.

- Change in asthma-related quality of life (PA(C)QLQ) scores after 3 and after 6 months
  An analysis of covariance (ANCOVA), adjusting for baseline PA(C)QLQ will be carried out to identify whether the mean/median differences in change in asthma-related quality of life between the groups described in paragraph 27 are significantly different, including 95% confidence intervals and p-values.

- Change in fatigue score after 3 and after 6 months

An analysis of covariance (ANCOVA), adjusting for baseline PedsQL will be carried out to identify whether the mean/median differences in change in fatigue score between the groups described in paragraph 27 are significantly different, including 95% confidence intervals and p-values.

- Change in school absences after 3 and after 6 months
  An analysis of covariance (ANCOVA), adjusting for baseline school absences, will be carried out to identify whether the mean/median differences in change school absences between the groups described in paragraph 27 are significantly different, including 95% confidence intervals and p-values.

  A Chi-squared test with one degree of freedom will be carried out to test whether equal population proportions have school absences, including 95% confidence intervals and p-values.

- Change in exacerbations (oral corticosteroids use, ER visits, hospital admissions)

  Oral corticosteroids use
  An analysis of covariance (ANCOVA), adjusting for baseline oral corticosteroids use, will be carried out to identify whether the mean/median differences in use of oral corticosteroids between the groups described in paragraph 27 are significantly different, including 95% confidence intervals and p-values.

  A Chi-squared test with one degree of freedom will be carried out to test whether equal population proportions have school absences, including 95% confidence intervals and p-values.

  ER visits
  An analysis of covariance (ANCOVA), adjusting for baseline ER visits, will be carried out to identify whether the mean/median differences in ER visits between the groups described in paragraph 27 are significantly different, including 95% confidence intervals and p-values.

  A Chi-squared test with one degree of freedom will be carried out to test whether the equal population proportions have ER visits, including 95% confidence intervals and p-values.

  Hospital admissions
  An analysis of covariance (ANCOVA), adjusting for baseline hospital admissions, will be carried out to identify whether the mean/median differences in hospital admissions between the groups described in paragraph 27 are significantly different, including 95% confidence intervals and p-values.

  A Chi-squared test with one degree of freedom will be carried out to test whether equal population proportions have hospital admissions, including 95% confidence intervals and p-values.

- Time to first exacerbation
  For this outcome the variable described in paragraph 27 will be used to test for four different comparisons. A log-rank test or Cox semiparametric proportional hazards model will be used to study whether there are differences in time to reach an exacerbation. The Chi-square with one degree of freedom or estimated relative hazard will be reported, including 95% confidence intervals and p-values.

- Amount of changes in therapy at t=3 months
  An analysis of covariance (ANCOVA) will be carried out to identify whether the mean/median differences in hospital admissions between the groups described in paragraph 27 are significantly different, including 95% confidence intervals and p-values.

- Change in lung function (FEV1 pre- and postbronchodilator) at t=3 and t=6 months
  An analysis of covariance (ANCOVA), adjusting for baseline FEV1 reversibility, will be carried out to identify whether the mean/median differences in FEV1 reversibility between the groups described in paragraph 27 are significantly different, including 95% confidence intervals and p-values.

- Change in FeNO at t=3 and t=6 months

An analysis of covariance (ANCOVA), adjusting for the baseline FeNO, will be carried out to identify whether the mean/median differences in FeNO changes between the groups described in paragraph 27 are significantly different, including 95% confidence intervals and p-values.

- Cost-effectiveness analyses.
  Results of the cost-effectiveness analysis will be expressed in terms of the incremental cost effectiveness ratio (ICER). These ICERs were calculated for two outcomes: 1) additional costs/avoided exacerbation as measured in the trial, 2) additional costs/QALY calculated with health utility values from literature
  The total costs included drug costs, GP, ER , hospitalisation costs and intervention costs, as well as productivity loss of the parents calculated over six months after the start date of the study. In order to analyse the uncertainty of the ICER results, a probabilistic sensitivity analysis (PSA) with 1000 replications with gamma distributions for all costs and exacerbations, a normal distribution for health utility values The resulting 1000 replicates will be plotted on the cost-effectiveness plane and used to construct a cost-effectiveness acceptability curve. The graphical presentation of the cost-effectiveness, is presented as the difference in costs on the vertical axis and the difference in effects on the horizontal axis. Deterministic probability analyses (DSA) were conducted for all different cost parameters to test the robustness of the analyses. Estimates for all different types of costs in both groups were varied between their 95% confidence intervals to assess the confidence. The resulting ranges of costs are presented in a tornado plot.

Extrapolation will be used to adjust for the period between the 3 months because the iPCQ is designed for 4 weeks.

28B. Adjustment for covariates

All analysis are by default adjusted from the baseline value of the respective outcome. Next to these adjusted analyses, adjusting for covariates will be performed.

<u>Primary Outcome:</u>

- Change in asthma control, based on (c-)ACT score difference, after 3 months compared to baseline
  In this analysis, there will be a correction for the following covariates: age, ethnicity, academic/non-academic hospital, baseline (c-)ACT, season in which the measurement was performed, and gender. The seasonality of having an asthma exacerbation episode will be explored by fitting season as a time-dependent covariate in the model.

<u>Secondary Outcomes:</u>

- Change in asthma control, based on (c-)ACT score difference, after 6 months compared to baseline
  In this analysis, there will be a correction for the following covariates: age, ethnicity, academic/non-academic hospital, baseline (c-)ACT, season in which the measurement was performed, and gender. The seasonality of having an asthma exacerbation episode will be explored by fitting season as a time-dependent covariate in the model.

- Time to (c-)ACT ≥ 20
  In this analysis, there will be a correction for the following covariates: age, ethnicity, academic/non-academic hospital, baseline (c-)ACT, season in which the measurement was performed, and gender. The seasonality of having an asthma exacerbation episode will be explored by fitting season as a time-dependent covariate in the model.

- Change in asthma-related quality of life (PA(C)QLQ) scores after 3 and after 6 months
  In this analysis, there will be a correction for the following covariates: age, ethnicity, academic/non-academic hospital, baseline PA(C)QLQ, season in which the measurement was performed, and gender. The seasonality of having an asthma exacerbation episode will be explored by fitting season as a time-dependent covariate in the model.

- Change in fatigue score after 3 and after 6 months
  In this analysis, there will be a correction for the following covariates: age, ethnicity, academic/non-academic hospital, baseline fatigue score, season in which the measurement was performed, and gender.

The seasonality of having an asthma exacerbation episode will be explored by fitting season as a time-dependent covariate in the model.

- Change in school absences after 3 and after 6 months
  In this analysis, there will be a correction for the following covariates: age, ethnicity, academic/non-academic hospital, baseline school absences level, season in which the measurement was performed, and gender. The seasonality of having an asthma exacerbation episode will be explored by fitting season as a time-dependent covariate in the model.

- Change in exacerbations (oral corticosteroids use, ER visits, hospital admissions)
  In this analysis, there will be a correction for the following covariates: age, ethnicity, academic/non-academic hospital, baseline exacerbation score, season in which the measurement was performed, and gender. The seasonality of having an asthma exacerbation episode will be explored by fitting season as a time-dependent covariate in the model.

- Time to first exacerbation
  In this analysis, there will be a correction for the following covariates: age, ethnicity, academic/non-academic hospital, baseline exacerbation score, season in which the measurement was performed, and gender. The seasonality of having an asthma exacerbation episode will be explored by fitting season as a time-dependent covariate in the model for time to first exacerbation.

- Amount of changes in therapy at t=3 months
  In this analysis, there will be a correction for the following covariates: ethnicity, academic/non-academic hospital, season in which the measurement was performed, and gender. The seasonality of having an asthma exacerbation episode will be explored by fitting season as a time-dependent covariate in the model.

- Change in lung function (FEV1 pre- and postbronchodilator) at t=3 and t=6 months
  In this analysis, there will be a correction for the following covariates: age, ethnicity, academic/non-academic hospital, baseline lung function, season in which the measurement was performed, and gender. The seasonality of having an asthma exacerbation episode will be explored by fitting season as a time-dependent covariate.

- Change in FeNO at t=3 and t=6 months
  In this analysis, there will be a correction for the following covariates: age, ethnicity, academic/non-academic hospital, baseline FeNO, season in which the measurement was performed, and gender. The seasonality of having an asthma exacerbation episode will be explored by fitting season as a time-dependent covariate in the model.

- Incremental cost per quality adjusted life year (QALY)
  No covariates will be taken into account.

- Incremental costs per avoided exacerbation
  No covariates will be taken into account.

28C. Methods used for assumptions

<u>Primary Outcome:</u>

- Change in asthma control, based on (c-)ACT score difference, after 3 months compared to baseline
  The following assumptions should be tested:
  - Tests for normality: a Shapiro Wilk's test will be carried out and QQ plots will be evaluated. When a clear pattern of deviation from normality is visible in the QQ plot, a suitable transformation of the outcome variable will be applied (e.g. squareroot or log).

<u>Secondary Outcomes:</u>

- Change in asthma control, based on (c-)ACT score difference, after 6 months compared to baseline
  The following assumptions should be tested:

- Tests for normality: a Shapiro Wilk's test will be carried out and QQ plots will be evaluated. When a clear pattern of deviation from normality is visible in the QQ plot, a suitable transformation of the outcome variable will be applied (e.g. squareroot or log).

- Time to (c-)ACT ≥ 20
  The following assumptions should be tested:
  - Tests for normality: a Shapiro Wilk's test will be carried out and QQ plots will be evaluated. When a clear pattern of deviation from normality is visible in the QQ plot, a suitable transformation of the outcome variable will be applied (e.g. squareroot or log).

- Change in asthma-related quality of life (PA(C)QLQ) scores after 3 and after 6 months
  The following assumptions should be tested:
  - Tests for normality: a Shapiro Wilk's test will be carried out and QQ plots will be evaluated. When a clear pattern of deviation from normality is visible in the QQ plot, a suitable transformation of the outcome variable will be applied (e.g. squareroot or log).

- Change in fatigue score after 3 and after 6 months
  The following assumptions should be tested:
  - Tests for normality: a Shapiro Wilk's test will be carried out and QQ plots will be evaluated. When a clear pattern of deviation from normality is visible in the QQ plot, a suitable transformation of the outcome variable will be applied (e.g. squareroot or log).

- Change in school absences after 3 and after 6 months
  The following assumptions should be tested:
  - Tests for normality: a Shapiro Wilk's test will be carried out and QQ plots will be evaluated. When a clear pattern of deviation from normality is visible in the QQ plot, a suitable transformation of the outcome variable will be applied (e.g. squareroot or log).

- Change in exacerbations (oral corticosteroids use, ER visits, hospital admissions)
  The following assumptions should be tested for each exacerbation outcome:
  - Tests for normality: a Shapiro Wilk's test will be carried out and QQ plots will be evaluated. When a clear pattern of deviation from normality is visible in the QQ plot, a suitable transformation of the outcome variable will be applied (e.g. squareroot or log).

- Time to first exacerbation
  The following assumptions should be tested:
  - Tests for normality: a Shapiro Wilk's test will be carried out and QQ plots will be evaluated. When a clear pattern of deviation from normality is visible in the QQ plot, a suitable transformation of the outcome variable will be applied (e.g. squareroot or log).

- Amount of changes in therapy at t=3 months
  The following assumptions should be tested:
  - Tests for normality: a Shapiro Wilk's test will be carried out and QQ plots will be evaluated. When a clear pattern of deviation from normality is visible in the QQ plot, a suitable transformation of the outcome variable will be applied (e.g. squareroot or log).

- Change in lung function (FEV1 pre- and postbronchodilator) at t=3 and t=6 months
  The following assumptions should be tested:
  - Tests for normality: a Shapiro Wilk's test will be carried out and QQ plots will be evaluated. When a clear pattern of deviation from normality is visible in the QQ plot, a suitable transformation of the outcome variable will be applied (e.g. squareroot or log).

- Change in FeNO at t=3 and t=6 months
  The following assumptions should be tested:

- Tests for normality: a Shapiro Wilk's test will be carried out and QQ plots will be evaluated. When a clear pattern of deviation from normality is visible in the QQ plot, a suitable transformation of the outcome variable will be applied (e.g. squareroot or log).

- Incremental cost per quality adjusted life year (QALY)
  No methods for assumption are needed in this analysis.

- Incremental costs per avoided exacerbation
  No methods for assumption are needed in this analysis.

28D. Sensitivity analysis

- Difference in treatment type in the control arm
A sensitivity analysis will be performed for each outcome to study whether there is a difference between the treatment groups in the control arm (e.g. the group treated with ICS + LABA and the group treated with a double dose of ICS). To this end, an extra term is added to the mixed model, indicating the treatment group in the control arm.

- Pairwise comparisons of genotype treatment interactions
By matching patients with a genotype in the genotype guided treatment arm with patients with a similar genotype in the control arm, we can make a pairwise comparison per outcome variable as described in 28A and 28B.

28E. Subgroup analysis
A subgroup analysis will be performed for all outcomes including only the heterozygous patients. We would like to investigate how to treat patients with the heterozygous variant (Arg16Gly). We will compare the heterozygous patients treated with LABA and ICS with the patients with a double dose ICS using the primary and secondary outcomes. These will be presented as described in 28A and 28B.

Another subgroup analysis will be performed to study differences in young (age 6-12 years of age) versus older children (13-17 years of age). We will study whether the outcomes change in one of these subgroups. These outcome of these analysis will be presented as described in 28A and 28B.

29. MISSING DATA

Patients who are not genotyped will be excluded for further analysis. In paragraph 27 it is stated per outcome how missing data will be handled.

30. ADDITIONAL ANALYSES

The following additional analyses will be performed:

1. Change in nasal gene expression and nasal gene methylation in relation to the treatment effect at t=3
For each sample, DNA methylation level at each CpG site will be calculated in percentage by B= (M/M+U) * 100%. Where M is the signal strength of methylated CpG given by Illumina HumanMethylation450 BeadChip array, and U is the signal strength of unmethylated CpG. For each marker a T-test will be used to assess methylation variation between responders and non-responders to treatment at t=3 months. Multiple comparison adjustment will be performed adopting a Bonferroni correction for the number of probes that will be successfully analysed after quality control.

2. change in microbiome profile and treatment effect at t=3
Using 16S RNA sequencing, each sample sequence set will be sub-sampled to 8,700 sequences. Differences in abundance will be detected using a Kruskal-Wallis generating a Bonferroni corrected p-value, whereby responders and non-responders to treatment at t=3 months are compared.

3. Breathome analysis for AMC site
A breathomics analysis will be performed for patients from the Academic Medical Center (AMC) using offline pattern recognition software.

31. HARMS

The number and percentage of patients experiencing each AE/SAE will be presented for each treatment arm categorized by severity. For each patient, only the maximum severity experienced of each type of AE will be displayed. The number and percentage of occurrences of each AE/SAE will also be presented for each treatment arm. No formal statistical testing will be undertaken.

32. STATISTICAL SOFTWARE

The analysis will be carried out using SPSS version 25 (IBM, North Castle). Other packages such as R or Stata will be used if necessary.

33. REFERENCES

32A. References to be provided for nonstandard statistical methods
This is not applicable for the PUFFIN trial.

32B. Reference to Data Management Plan
Data management Plan version 2.0 was used to write this statistical analysis plan. It is located at our communal G-drive within the Academic Medical Center: G:\diva\Longziekten_Research\PUFFIN\Data management.

32C. Reference to the Trial Master File and Statistical Master File
The Statistical Master File can be found within the Trial Master File, located at the AmsterdamUMC, location Academic Medical Center, room F5-259, Meibergdreef 9, 1105AZ Amsterdam-Zuidoost.

32D. Reference to other standard operating procedures or documents to be adhered to.

C. Bouwmans et al. Handleiding iMTA Productivity Cost Questionnaire (iPCQ). Rotterdam: iMTA, Erasmus Universiteit, 2013.

European Medicines Agency. Guideline on Missing Data in Clinical trials. EMA/CPMP/EWP/1776/99 Rev. 1. Committee for Meidicinal Products for Human Use, London, 2 July 2010.

European Medicines Agency. Guideline on adjustment for baseline covariates in clinical trials. EMA/CHMP/295050/2013. Committee for Medicinal Products for Human Use (CHMP), London, 26 February 2015.

European Medicines Agency. Points to consider on multiplicity issues in clinical trials. CPMP/EWP/908/99. Committee for proprietary medicinal products, London, 19 September 2002.

European Medicines Agency. Note for guidance on statistical principles for clinical trials. CPMP/ICH/363/96. London, September 1998.

Gamble C, Krishan A, Stocken D, Lewis S, Jusczak E, Doré C. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. JAMA 2017; 318(23): 2337-2343.

Groenwold RHH, Moons KGM, Vandenbroucke JP. Randomized trials with missing outcome data: how to analyze and what to report. CMAJ 2014; 186(15): 1153-1157.

Universiteit van Amsterdam (UvA). UvA Wiki methodologiewinkel – Normaliteit. Access via: https://wiki.uva.nl/methodologiewinkel/index.php/Normaliteit. Accessed 04-01-2019.

Universiteit van Amsterdam (UvA). UvA Wiki methodologiewinkel – Onafhankelijkheid. Access via: https://wiki.uva.nl/methodologiewinkel/index.php/Onafhankelijkheid. Accessed 04-01-2019.

LIST OF ABBREVIATIONS

DSMB	Data and Safety Monitoring Board
IQR	Interquartile range
MREC	Medical Research and Ethics Committee
PI	Principal investigator
SD	Standard deviation
SAP	Statistical analysis plan