

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

Detailed Title:	A Phase II, randomised, observer-blind, placebo-controlled, multi-centre study to evaluate the safety, reactogenicity and immunogenicity of GSK Biologicals' investigational vaccine GSK2838504A, when administered intramuscularly according to a 0, 2 month schedule to COPD patients aged 40 to 80 years
SAP version	Final
SAP date	09-FEB-2016
Scope:	All data pertaining to the above study.
Co-ordinating author:	PPD [REDACTED]
Other author(s):	
Adhoc reviewers:	PPD [REDACTED], Global Regulatory Lead (RA) PPD [REDACTED], PPD [REDACTED], Safety Representatives CTRS team
Approved by:	PPD [REDACTED], Clinical Research & Development Lead PPD [REDACTED], Lead Statistician PPD [REDACTED], Scientific Writer PPD [REDACTED], Statistician

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

TABLE OF CONTENTS

	PAGE
LIST OF ABBREVIATIONS	4
1. DOCUMENT HISTORY	7
2. STUDY DESIGN	7
3. OBJECTIVES	8
3.1. Primary objective	8
3.2. Secondary objective	8
3.3. Tertiary objectives	8
4. ENDPOINTS	9
4.1. Primary endpoints	9
4.2. Secondary endpoints	9
4.3. Tertiary endpoints	9
5. STUDY POPULATION	10
5.1. Total vaccinated cohort	10
5.2. ATP cohort for analysis of immunogenicity	11
5.3. ATP cohort for analysis of persistence of immunogenicity	11
5.4. ATP cohort for analysis of efficacy	11
6. STATISTICAL METHODS	12
6.1. Analysis up to Visit 6 (Day 90)	12
6.1.1. Analysis of demographics	12
6.1.2. Analysis of safety	13
6.1.3. Analysis of immunogenicity	14
6.1.4. Analysis of HRQOL, lung capacity and exercise capacity	15
6.1.5. Analysis of selected biomarkers	16
6.2. Final analysis (up to day 450)	16
6.2.1. Analysis of demographics	16
6.2.2. Analysis of safety	16
6.2.3. Analysis of immunogenicity	18
6.2.4. Analysis of impact of the investigational vaccine on AECOPD incidence	19
6.2.5. AECOPD visits and daily questionnaires	20
6.2.6. Analysis of HRQOL, lung capacity and exercise capacity	21
6.2.7. Analysis of selected biomarkers	21
6.2.8. Analysis of sputum	21
6.2.8.1. Sputum sample collection and quality	21
6.2.8.2. Sputum bacterial/viral results from culture or from PCR	22

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

6.2.8.3.	Impact of previous antibiotic administration on sputum bacterial results.....	22
6.2.8.4.	Impact of induction of sample on sputum bacterial results	22
6.2.8.5.	Impact of previous occurrence of Hi bacterial pathogen in sputum on AECOPDs	22
7.	STATISTICAL CALCULATIONS	23
7.1.	Derived and transformed data.....	23
7.1.1.	Demography	23
7.1.2.	Safety	23
7.1.3.	Immunogenicity.....	24
7.1.4.	NTHi associated AECOPD.....	25
7.1.5.	HRQOL.....	25
7.1.6.	EXACT-PRO derived AECOPD definitions and calculation rules.....	25
7.1.7.	“e-diary derived AECOPD” definitions and calculation rules.....	26
7.1.8.	End/Duration of an e-diary signal of an exacerbation:	27
7.1.9.	Time between 2 e-diary signals for exacerbations/evaluation of baseline symptom score:	27
7.1.10.	Missing data	28
7.1.11.	AECOPDs definitions and calculation rules.....	28
7.1.11.1.	AECOPD definition (from clinical evaluation)	28
7.1.11.2.	Characterization of inconsistencies between morning e-diary signals and exacerbation site visits	29
7.1.12.	Grading of severity of an exacerbation.....	30
7.2.	Data presentation description	30
8.	CONDUCT OF ANALYSES.....	31
8.1.	Sequence of analyses.....	31
8.2.	Statistical considerations for interim analyses	31
9.	CHANGES FROM PLANNED ANALYSES.....	32
10.	REFERENCES.....	32

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

LIST OF ABBREVIATIONS

AE	Adverse event
AECOPD	Acute exacerbation of COPD
ALT	Alanine aminotransferase
ATP	According-To-Protocol
ANCOVA	Analysis of covariance
CAT	COPD assessment test
CI	Confidence Interval
CMI	Cell-mediated immunity
COPD	Chronic obstructive pulmonary disease
CRF	Case Report Form
CTRS	Clinical Trial Registry
EF	Internal GSK database code for Efficacy analysis (elimination codes) link to FORM-BIO-CLIN-9004-05 Criteria for eliminating subjects from the analyses
EL.U/ml	ELISA unit per milliliter
ELISA	Enzyme-linked immunosorbent assay
Eli Type	Internal GSK database code for type of elimination code
EXACT-PRO	EXAcerbations of Chronic Pulmonary Disease Tool - Patient Reported Outcome
FEV ₁	Forced expiratory volume in 1 second
FVC	Forced vital capacity
FU	Internal GSK database code for Follow-up analysis (elimination codes) link to FORM-BIO-CLIN-9004-05 Criteria for eliminating subjects from the analyses
GMC	Geometric mean antibody concentration
GOLD	Global Initiative for Chronic Obstructive Lung Disease
GSK	GlaxoSmithKline
<i>H. influenzae</i>	<i>Haemophilus influenzae</i>
HRQOL	Health-Related Quality of Life
ICF	Informed consent form
ICS	Intracellular cytokine staining

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

IU/ml	International units per milliliter
LL	Lower Limit of the confidence interval
MA	Internal GSK database code for Main analysis (elimination codes) link to FORM-BIO-CLIN-9004-05 Criteria for eliminating subjects from the analyses
MATEX	MATerial EXcellence
<i>M. catarrhalis</i>	<i>Moraxella catarrhalis</i>
MedDRA	Medical Dictionary for Regulatory Activities
mMRC scale	mMRC scale
N.A.	Not Applicable
NTHi	Non-Typeable <i>Haemophilus influenzae</i>
PBMC	Peripheral blood mononuclear cell
PCR	Polymerase chain reaction
PD	Protein D
PE	Protein E
PI	Prescribing information
PilA	Type IV pili subunit of non-typeable <i>Haemophilus influenzae</i>
pIMD	Potential immune-mediated disease
PT	Preferred term
SAE	Serious adverse event
<i>S. aureus</i>	<i>Staphylococcus aureus</i>
SAP	Statistical Analysis Plan
SBIR	GSK Biological's Internet Randomization System
SD	Standard Deviation
SR	Study Report
STGG	Skim milk, tryptone, glucose, and glycerin transport medium
SYN	Synopsis
TFL	Tables Figures and Listing template annexed to SAP
TVC	Total vaccinated cohort
UL	Upper Limit of the confidence interval
VE	Vaccine efficacy

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

6MWT

6-Minute Walk Test

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

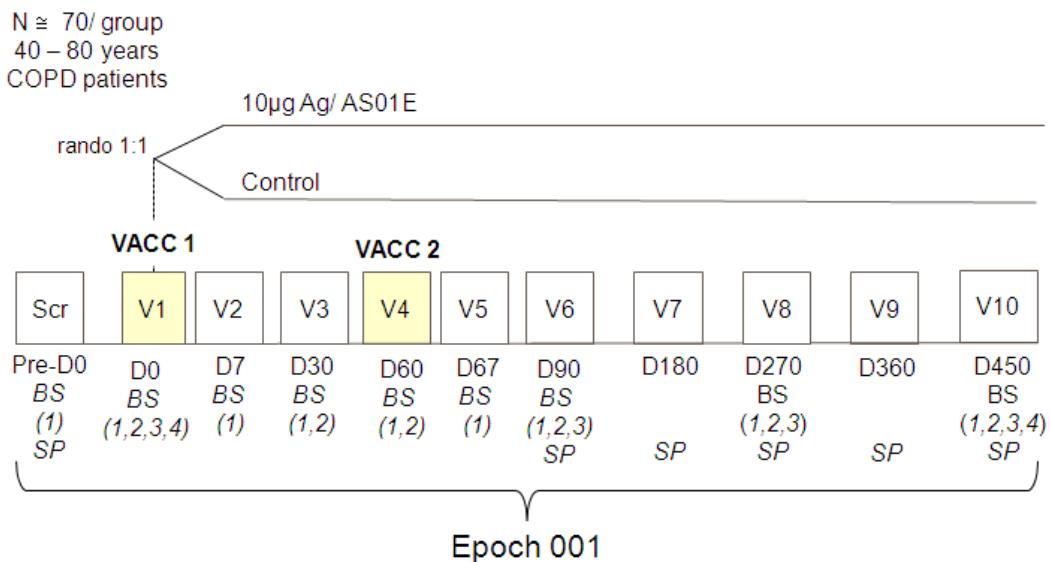
The complete statistical analysis plan and results presentation is divided into 2 parts: the first part detailing the analyses to be performed (known as SAP, current document) and a second part, annex (-es) (called TFL) describing the flow and format of tables, figures and listings to be annexed to the SR.

1. DOCUMENT HISTORY

Date	Description	Protocol Version
29-JAN-2016	Version 1	Amendment 2 - 15-DEC-2014

2. STUDY DESIGN

Figure 1 Study design overview



Scr = Screening; **V** = Visit; **D** = Day; **random** = randomisation; **VACC** = vaccination; **BS(1)** = blood sample for safety; **BS(2)** = blood sample for humoral immunogenicity; **BS(3)** = blood sample for cell-mediated immunity (CMI); **BS(4)** = blood sample for biomarkers; **SP** = sputum sample

In addition to the scheduled study visits depicted above, *ad hoc AECOPD-driven study visit(s) and/ or phone contact(s)* will take place for each AECOPD occurring from first vaccination up to study conclusion. An AECOPD visit will be scheduled as soon as possible after the onset of AECOPD symptoms (maximum 96 hours after and, if applicable, preferably before starting treatment with antibiotics). In addition, follow-up phone call(s) and/ or visit(s) will take place to determine the end of the AECOPD. These contacts will take place at least every 2 weeks, until the AECOPD has resolved.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

Experimental design: Phase II randomised, observer-blind, placebo-controlled, multi-centre study with 2 parallel groups.

Duration of the study: for each subject enrolled, the study will last approximately 16 months from Visit 1 up to study end

The following group names will be used for the statistical analyses:

Group order in tables	Group label in tables	Group definition for footnote
1	10-AS01E	2 doses of AS01e adjuvanted vaccine containing 10 mcg of each antigen
2	Control	2 doses of saline solution

3. OBJECTIVES

3.1. Primary objective

- To describe the safety and reactogenicity of the investigational vaccine.

3.2. Secondary objective

- To describe the humoral and cellular immunogenicity of the investigational vaccine.

3.3. Tertiary objectives

- To explore the impact of the investigational vaccine on AECOPD.
- To explore the impact of the investigational vaccine on NTHi presence in sputum.
- To explore the impact of the investigational vaccine on health-related quality of life (HRQOL).
- To explore the impact of the investigational vaccine on lung function.
- To explore the impact of the investigational NTHi vaccine on exercise capacity.
- To describe selected biomarkers in stable COPD and during AECOPD.
- To collect blood and sputum samples for assay development, for disease diagnostic purpose, for lung microbiome analysis and/ or for additional evaluation of the immune responses to the investigational vaccine and to other potential pathogens involved in AECOPD.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

4. ENDPOINTS

4.1. Primary endpoints

- Occurrence of each solicited local adverse event (AE), during the 7-day follow-up period (*i.e.* day 0 - 6) following each vaccination, in all subjects.
- Occurrence of each solicited general AE, during the 7-day follow-up period (*i.e.* day 0 - 6) following each vaccination, in all subjects.
- Occurrence of any unsolicited AE, during the 30-day follow-up period (*i.e.* day 0 - 29) following each vaccination, in all subjects.
- Occurrence of each haematological/ biochemical laboratory abnormality at Day 0, Day 7, Day 30, Day 60, Day 67, Day 90, Day 270 and Day 450, in all subjects.
- Occurrence of any potential immune-mediated disease (pIMD) from first vaccination up to study conclusion, in all subjects.
- Occurrence of any serious adverse event (SAE) from first vaccination up to study conclusion, in all subjects.

4.2. Secondary endpoints

- Anti-PD, anti-PE and anti-PilA total IgG antibody concentrations as measured by ELISA, at Day 0, Day 30, Day 60, Day 90, Day 270 and at Day 450, in all subjects.
- NTHi-specific cell-mediated immune responses as measured by flow cytometry intracellular cytokine staining (ICS) (frequency of specific CD4⁺/CD8⁺ T-cells expressing two or more markers, such as IL-2, IL-13, IL-17, IFN- γ , TNF- α and CD40L), at Day 0, Day 90, Day 270 and at Day 450, in a sub-cohort of subjects.

4.3. Tertiary endpoints

- Number of cases of NTHi-associated moderate and severe AECOPD, over a period starting 1 month post-Dose 2 and lasting for 1 year, in all subjects.
- Number of cases of NTHi-associated AECOPD (any severity), over a period starting 1 month post-Dose 2 and lasting for 1 year, in all subjects.
- Number of cases of moderate and severe AECOPD (any cause), over a period starting 1 month post-Dose 2 and lasting for 1 year, in all subjects.
- Number of cases of AECOPD (any cause, any severity), over a period starting 1 month post-Dose 2 and lasting for 1 year, in all subjects.
- Time to first AECOPD, in all subjects.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

- Assessment of EXACT-PRO score, daily at bedtime throughout the study, in all subjects.
- NTHi presence in sputum, at Screening, Day 90, Day 180, Day 270, Day 360 and Day 450, and from first vaccination to study conclusion for each AECOPD, in all subjects.
- Assessment of *CAT* score, at Screening, Day 270 and Day 450, and from first vaccination to study conclusion for each AECOPD, in all subjects.
- Assessment of SGRQ-C score, at Screening, Day 270 and Day 450, in all subjects.
- Assessment of mMRC scale, at Screening, Day 270 and Day 450, and from first vaccination to study conclusion for each AECOPD, in all subjects.
- Occurrence of rescue medication use, over a period starting 1 month post-Dose 2 and lasting for 1 year, in all subjects.
- Occurrence of healthcare use for AECOPD, over a period starting 1 month post-Dose 2 and lasting for 1 year, in all subjects.
- Assessment of FEV1% of predicted normal value at Screening, Day 270 and Day 450, in all subjects.
- Assessment of 6MWT score, at Screening and Day 450, in all subjects.
- Occurrence of selected biomarkers in a subset of blood samples, at Day 0 and Day 450, and from first vaccination to study conclusion for each AECOPD, in all subjects.

5. STUDY POPULATION

The following study cohorts will be evaluated.

5.1. Total vaccinated cohort

The Total vaccinated cohort (TVC) will include all subjects with at least 1 study vaccine administration documented:

- A **safety** analysis based on the TVC will include all vaccinated subjects.
- An **immunogenicity/ efficacy** analysis based on the TVC will include all vaccinated subjects for whom immunogenicity/ efficacy data are available.

The TVC analysis will be performed per treatment actually administered at Dose 1.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

5.2. ATP cohort for analysis of immunogenicity

The ATP cohort for immunogenicity will include all subjects in the TVC:

- Who met all eligibility criteria.
- For whom the administration route and site of the vaccines was according to protocol.
- Who complied with the vaccination schedule.
- Who received the study vaccines according to protocol procedures.
- Who did not receive a concomitant medication/ product leading to elimination from the ATP analysis for immunogenicity up to the 1 month post-Dose 2 visit (Day 90).
- Who did not present with an intercurrent medical condition leading to elimination from the ATP analysis for immunogenicity up to the 1 month post-Dose 2 visit (Day 90).
- Who complied with the blood sample timings, at the 1 month post-Dose 2 visit (Day 90).
- For whom post-vaccination immunogenicity results are available for at least 1 assay.

5.3. ATP cohort for analysis of persistence of immunogenicity

The ATP cohort for persistence of immunogenicity (starting at Day 180) will include all evaluable subjects, *i.e.*, those who were included in the ATP cohort for immunogenicity, or were excluded from this cohort solely because they had no blood samples taken or because of incompliance with blood sample timings up to the 1 month post-Dose 2 visit, and:

- Who did not receive a concomitant medication/ product leading to elimination from the ATP analysis for immunogenicity.
- Who did not present with an intercurrent medical condition leading to elimination from the ATP analysis for immunogenicity.
- Who complied with the blood sample timings after the 1 month post-last vaccination visit.
- For whom persistence immunogenicity results are available for at least 1 assay.

5.4. ATP cohort for analysis of efficacy

The ATP cohort for analysis of efficacy will include all subjects in the TVC:

- Who met all eligibility criteria.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

- For whom the administration route and site of the vaccines was according to protocol.
- Who complied with the vaccination schedule
- Who received the study vaccine according to protocol procedures.
- Not having received a medication/ product/ vaccine that may lead to elimination from the ATP analysis for efficacy.

The list of applicable elimination codes for each cohort can be found in the study specific form FORM-BIO-CLIN-9004-05 Criteria for eliminating subjects from the analyses.

Cohort	Elimination codes	Eli Type
ATP cohort for analysis of immunogenicity	1010-2500	MA
ATP cohort for analysis of efficacy	1010-2500	EF
ATP cohort for analysis of persistence of immunogenicity	1010-2500	FU

6. STATISTICAL METHODS

6.1. Analysis up to Visit 6 (Day 90)

6.1.1. Analysis of demographics

Demographic characteristics (age at the first dose in years, gender and geographical ancestry), cohort description and other characteristics such as smoking/exposure history status, pulmonary function test baseline results and COPD history, history of exacerbation present/past medical history/intercurrent morbidity, and previous vaccination with a pneumococcal vaccine or an influenza vaccine will be summarized by group using descriptive statistics:

- Frequency tables will be generated for categorical variable such as geographical ancestry;
- Mean, median and standard error will be provided for continuous data such as age.

Other characteristics such as height, weight, BMI selected results and several risk factors from the six-minutes-walk test, and distribution of several severity stages using CAT or mMRC or SGRQ-C at enrolment visit will also be summarized.

The number of subjects enrolled into the study as well as the number of subjects excluded from ATP analyses will be tabulated.

The distribution of subjects enrolled among the study sites will be tabulated as a whole and per group.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

The numbers of withdrawn subjects will be tabulated according to the reason for withdrawal.

6.1.2. Analysis of safety

The analysis will be performed on the TVC. If the percentage of vaccinated subjects excluded from the ATP cohort for efficacy is at least 10%, a second analysis will be performed on the ATP cohort for efficacy to complement the first analysis.

The percentage of subjects with at least one **local AE** (solicited and unsolicited), with at least one **general AE** (solicited and unsolicited) and with any AE during the 7-day (day 0 – day 6) or the 30-day (day 0 – day 29) follow-up period will be tabulated after each vaccination and overall with exact 95% confidence interval (CI). The same computations will be done for Grade 3 AEs, for related AEs and for Grade 3 related AEs.

The percentage of subjects/doses reporting each individual solicited local (any grade, Grade 3) and general (any grade, Grade 3, related, Grade 3 related) AE during the 7-day (day 0 to day 6) follow-up period will be tabulated for each group as follows:

- Overall, the percentage of subjects with the symptom and its exact 95% CI.
- Overall, the percentage of doses with the symptom and its exact 95% CI.
- Per study vaccine dose, the percentage of subjects with the symptom and its exact 95% CI.

The exact 95% CIs will be calculated assuming independence between doses. For fever, additional analyses will be performed by 0.5°C increments.

As during the study, 22 subjects have completed wrong diary cards (paediatric ones instead of adults), a sensitivity analysis will be done on subjects that have completed the correct diary cards only, to be sure that they were no impact on results.

The verbatim reports of **unsolicited** symptoms will be reviewed by a physician and the signs and symptoms will be coded according to the MedDRA Dictionary for Adverse Reaction Terminology. Every verbatim term will be matched with the appropriate preferred term (PT). The percentage of subjects with unsolicited symptoms during the 30-day follow-up period after any study vaccine dose with its exact 95% CI will be tabulated by group and by MedDRA PT. Similar tabulation will be done for the percentage of doses, for Grade 3 unsolicited symptoms, for unsolicited symptoms that resulted in a medically attended visit, for Grade 3 and causally related unsolicited symptoms and for unsolicited symptoms causally related to vaccination.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

For each group and for each **haematology/ biochemistry parameter**:

- The percentage of subjects having results below or above the local laboratory normal ranges will be tabulated by time point.
- The maximum grading from Screening up to 30 days post-Dose 2 will be tabulated. Grades will be based on the FDA Guidance for Industry “Toxicity Grading Scale for Healthy Adults and Adolescent Volunteers Enrolled in Preventive Vaccine Clinical Trials” (See Protocol).

The number of subjects who experienced any **pIMD** or any **SAE** from first vaccination up to 30 days post-Dose 2 will be reported.

The number of subjects who experienced any **AE leading to study withdrawal**, from first vaccination up to study conclusion, or any **SAE related to study participation of concurrent GSK medication/ vaccination**, during the entire study period, will be reported.

Pregnancy reports from first vaccination up to 30 days post-Dose 2 and pregnancy outcomes will be described in detail.

The percentage of subjects/ dose using **concomitant medication/ product** (any medication/ product, any antipyretic and any antipyretic taken prophylactically, respectively) during the 7-day follow-up period (day 0 – day 6) and during the 30-day follow-up period (day 0 – day 29) will be summarised per group for each dose and overall per dose.

6.1.3. Analysis of immunogenicity

The analysis will be performed on the ATP cohort for immunogenicity. If the percentage of vaccinated subjects with serological results excluded from the ATP cohort for analysis of immunogenicity is more than 10%, a second analysis will be performed on the TVC.

Humoral immune response

Within group evaluation

For each group, at each time point during which blood samples are collected for humoral immune response and for each assay:

- Seropositivity rate and their exact 95% CI will be tabulated.
- GMCs and their 95% CI will be calculated.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

For each group, at each time point during which blood samples are collected for humoral immune response:

- Antibody concentrations distribution will be investigated using Reverse Cumulative Curves.

Between groups evaluation

Comparative analyses will be exploratory with the aim to characterise the difference between the 2 groups in humoral immune response.

The difference in terms of GMCs will be evaluated at 1 month post-Dose 2 by computing the 95% CIs of the GMC ratio between groups by using an ANCOVA model that consider heterogeneity of variances (via Proc Mixed) on the logarithm10 transformation of the concentrations. This model will include the country, the group, age category (40 - 59 years or 60 - 80 years), number of moderate and severe AECOPD in the year before vaccine Dose 1 (< 2 or \geq 2), GOLD grade (GOLD 2 or GOLD 3) and pre-Dose 1 concentration (as covariate) as fixed effects.

However, these differences should be interpreted with caution considering that there will be no adjustment for multiplicity of endpoints.

Cell-mediated immune response

The frequency of specific CD4 $^{+}$ / CD8 $^{+}$ T-cells will be summarised (mean, SD, minimum, Q1, median, Q3, and maximum) by group, at each time point during which blood samples are collected for CMI (descriptive statistics).

Each data will be obtained by subtracting the summary of the background value to each summary value. Values less than or equal to zero will be set to 1 for the purpose of geometric mean calculation and graphical representation.

6.1.4. Analysis of HRQOL, lung capacity and exercise capacity

The analysis will be performed on the TVC. If the percentage of vaccinated subjects excluded from the ATP cohort for efficacy is at least 10%, a second analysis will be performed on the ATP cohort for efficacy to complement the first analysis.

Descriptive statistics (median, mean, range, standard deviation, first and third quartiles) on the **EXACT-PRO**, **CAT**, **SGRQ-C** and **mMRC** scores, on **FEV₁% of predicted normal value** and on the **6MWT** scores will be tabulated by specified visit.

Descriptive summaries **rescue medication and healthcare use for AECOPD** will be provided.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

6.1.5. Analysis of selected biomarkers

The analysis will be performed on the ATP cohort for efficacy. If the percentage of vaccinated subjects excluded from the ATP cohort for efficacy is at least 10%, a second analysis will be performed on the TVC to complement the first analysis.

For each type of timepoint (Pre, Exacerbation visit) during which blood samples are collected for analysis of biomarkers, descriptive statistics (median, mean, range, standard deviation, first and third quartiles) will be tabulated for each selected biomarker.

6.2. Final analysis (up to day 450)

6.2.1. Analysis of demographics

Withdrawal status will be summarised by group using descriptive statistics:

The number of subjects enrolled into the study as well as the number of subjects excluded from ATP analyses will be tabulated.

The numbers of withdrawn subjects will be tabulated according to the reason for withdrawal.

6.2.2. Analysis of safety

The analysis will be performed on the TVC. If the percentage of vaccinated subjects excluded from the ATP cohort for efficacy is at least 10%, a second analysis will be performed on the ATP cohort for efficacy to complement the first analysis.

The percentage of subjects with at least one **local AE** (solicited and unsolicited), with at least one **general AE** (solicited and unsolicited) and with any AE during the 7-day (day 0 – day 6) or the 30-day (day 0 – day 29) follow-up period will be tabulated after each vaccination and overall with exact 95% confidence interval (CI). The same computations will be done for Grade 3 AEs, for related AEs and for Grade 3 related AEs.

The percentage of subjects/doses reporting each individual solicited local (any grade, Grade 3) and general (any grade, Grade 3, related, Grade 3 related) AE during the 7-day (day 0 to day 6) follow-up period will be tabulated for each group as follows:

- Overall, the percentage of subjects with the symptom and its exact 95% CI.
- Overall, the percentage of doses with the symptom and its exact 95% CI.
- Per study vaccine dose, the percentage of subjects with the symptom and its exact 95% CI.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

The exact 95% CIs will be calculated assuming independence between doses. For fever, additional analyses will be performed by 0.5°C increments.

As during the study, some subjects have completed wrong diary cards (paediatric ones instead of adults), a sensitivity analysis will be done on subjects that have completed the correct diary cards to be sure that they were no impact on results.

The verbatim reports of **unsolicited** symptoms will be reviewed by a physician and the signs and symptoms will be coded according to the MedDRA Dictionary for Adverse Reaction Terminology. Every verbatim term will be matched with the appropriate preferred term (PT). The percentage of subjects with unsolicited symptoms during the 30-day follow-up period after any study vaccine dose with its exact 95% CI will be tabulated by group and by MedDRA PT. Similar tabulation will be done for the percentage of doses, for Grade 3 unsolicited symptoms, for unsolicited symptoms that resulted in a medically attended visit, for Grade 3 and causally related unsolicited symptoms and for unsolicited symptoms causally related to vaccination.

For each group and for each **haematology/ biochemistry parameter**:

- The percentage of subjects having results below or above the local laboratory normal ranges will be tabulated by time point.
- The maximum grading from Screening up to study conclusion will be tabulated. Grades will be based on the FDA Guidance for Industry “Toxicity Grading Scale for Healthy Adults and Adolescent Volunteers Enrolled in Preventive Vaccine Clinical Trials” (see Appendix C of the protocol).

The number of subjects who experienced any **pIMD** or any **SAE** from first vaccination up to 6 months post-Dose 2 will be reported will be reported.

The number of subjects who experienced any **AE leading to study withdrawal**, from first vaccination up to study conclusion, or any **SAE related to study participation of concurrent GSK medication/ vaccination**, during the entire study period, will be reported.

Pregnancy reports from first vaccination up to 6 months after last vaccination and pregnancy outcomes will be described in detail.

The percentage of subjects/ dose using **concomitant medication/ product** (any medication/ product, any antipyretic and any antipyretic taken prophylactically, respectively) during the 7-day follow-up period (day 0 – day 6) and during the 30-day follow-up period (day 0 – day 29) will be summarised per group for each dose and overall per dose.

Statistical Analysis Plan



GlaxoSmithKline

Study alias & e-track number(s): NTHi-004 (200157)

6.2.3. Analysis of immunogenicity

The analysis will be performed on the ATP cohort for immunogenicity. If the percentage of vaccinated subjects with serological results excluded from the ATP cohort for analysis of immunogenicity is more than 10%, a second analysis will be performed on the TVC.

Humoral immune response

Within group evaluation

For each group, at each time point during which blood samples are collected for humoral immune response and for each assay:

- Seropositivity rate and their exact 95% CI will be tabulated.
- GMCs and their 95% CI will be calculated.

For each group, at each time point during which blood samples are collected for humoral immune response:

- Antibody concentrations distribution will be investigated using Reverse Cumulative Curves.

Between groups evaluation

Comparative analyses will be exploratory with the aim to characterise the difference between the 2 groups in humoral immune response.

The difference in terms of GMCs will be evaluated at 1 month post-Dose 2 and at Day 450 by computing the 95% CIs of the GMC ratio between groups by using an ANCOVA model that consider heterogeneity of variances (via Proc Mixed) on the logarithm10 transformation of the concentrations. This model will include the country, the group, age category (40 - 59 years or 60 - 80 years), number of moderate and severe AECOPD in the year before vaccine Dose 1 (< 2 or \geq 2), GOLD grade (GOLD 2 or GOLD 3) and pre-Dose 1 concentration (as covariate) as fixed effects.

However, these differences should be interpreted with caution considering that there will be no adjustment for multiplicity of endpoints.

Cell-mediated immune response

The frequency of specific CD4 $^{+}$ / CD8 $^{+}$ T-cells will be summarised (mean, SD, minimum, Q1, median, Q3, and maximum) by group, at each time point during which blood samples are collected for CMI (descriptive statistics).

Each data will be obtained by subtracting the summary of the background value to each summary value. Values less than or equal to zero will be set to 1 for the purpose of geometric mean calculation and graphical representation.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

6.2.4. Analysis of impact of the investigational vaccine on AECOPD incidence

The analysis will be performed on the ATP cohort for efficacy. If the percentage of vaccinated subjects excluded from the ATP cohort for efficacy is more than 10%, a second analysis based on the TVC will be performed to complement the first analysis.

The following incidence rates will be computed over a period starting 1 month post-Dose 2 and lasting for 1 year, with 95%CI:

- NTHi-associated moderate and severe AECOPD. (See definition in section 7.1.4)
- NTHi-associated all-severity AECOPD.
- All-cause moderate and severe AECOPD.
- All-cause, all-severity AECOPD.

Vaccine efficacy (VE) for AECOPD events will be defined as:

$$VE(AECOPD) = 1 - R_{vaccine} / R_{control}$$

with:

- $R_{vaccine}$ = average yearly incidence rate of AECOPD events per subject in the group 10-AS01E.
- $R_{control}$ = average yearly incidence rate of AECOPD events per subject in the Control group.

In case the 95% CI of the VE cannot be estimated, *e.g.* when no event is observed in the group 10-AS01E (leading to a VE of 100%), then an exact procedure will be used to estimate a VE and its 95% CI. This exact procedure will be based on the number of subjects with event instead of on the average number of event per subject in each group.

VE in the prevention of NTHi-associated moderate and severe AECOPD and its 95% CI will be calculated.

VE in the prevention of NTHi-associated all-severity AECOPD, of all-cause moderate and severe AECOPD and of all-cause, all-severity AECOPD and their 95% CI will also be calculated.

The 95%CI of the incidence rate will be computed using a model which accounts for repeated events. The Generalised linear model assuming the Negative Binomial distribution for the response variable with logarithm as link function, and the logarithm of time for follow-up as an offset variable will be used. In addition, the following covariates (as per randomization) will be taken into account:

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

- country,
- number of moderate/severe exacerbations reported in the 12 months prior to enrolment,
- COPD status at enrolment (derived by GSK from the Visit 1 spirometry results)
- Age (below or equal to median/above median)

If the model does not converge, the Poisson distribution will be used instead of the Negative Binomial one.

The frequency of all these AECOPDs will be presented. The estimated incidence rate with 95%CI of AECOPDs in which Hi (resp. Mcat, Sp) is present in culture and was not present in the previous visit estimated by a Generalised linear model assuming the Negative Binomial distribution will be estimated.

Time to first AECOPD event by baseline status - Cox regression model

The Cox's regression model including the risk factors:

- country,
- number of moderate/severe exacerbations reported in the 12 months prior to enrolment,
- COPD status at enrolment (derived by GSK from the Visit 1 spirometry results)
- Age (below or equal to median/above median)

as covariates (as per randomization) will be performed to model the time until the first occurrence of AECOPD episodes. Additional stratification/risk factor may be defined during the analyses like sex factor as exploratory analysis.

6.2.5. AECOPD visits and daily questionnaires

The time until first AECOPD as of 1 month post-Dose 2 will be computed with exact 95% CI by group. Subjects who exacerbate before 1 month post-Dose 2 will not be taken into account in this analysis.

The number of morning e-dairy (PHT) alerts (resp. evening e-diary alerts (EXACT-PRO)) with and without corresponding site visits will be tabulated. The reason for inconsistency between morning e-diaries and site visits will also be presented.

The number of exacerbation site visits with confirmed AECOPD with and without morning e-diary alerts (resp. evening e-diary alerts (EXACT-PRO)) will be tabulated along with the reason for no e-diary alert.

Concordance/discordance of exacerbations periods as defined clinically or via the morning diary cards or the evening diary cards will be tabulated and presented graphically.

Statistical Analysis Plan



GlaxoSmithKline

Study alias & e-track number(s): NTHi-004 (200157)

The percentage of spontaneous exacerbation visits will be tabulated as well as the percentage of monthly visits with recovered/stable status.

6.2.6. Analysis of HRQOL, lung capacity and exercise capacity

The analysis will be performed on the TVC. If the percentage of vaccinated subjects excluded from the ATP cohort for efficacy is at least 10%, a second analysis will be performed on the ATP cohort for efficacy to complement the first analysis.

Descriptive statistics (median, mean, range, standard deviation, first and third quartiles) on the **EXACT-PRO, CAT, SGRQ-C** and **mMRC** scores, on **FEV₁% of predicted normal value** and on the **6MWT** scores will be tabulated by specified visit.

Descriptive summaries **rescue medication and healthcare use for AECOPD** will be provided.

6.2.7. Analysis of selected biomarkers

The analysis will be performed on the ATP cohort for efficacy. If the percentage of vaccinated subjects excluded from the ATP cohort for efficacy is at least 10%, a second analysis will be performed on the TVC to complement the first analysis.

For each type of timepoint (Pre, Exacerbation visit, Day 450) during which blood samples are collected for analysis of biomarkers, descriptive statistics (median, mean, range, standard deviation, first and third quartiles) will be tabulated for each selected biomarker.

6.2.8. Analysis of sputum

6.2.8.1. Sputum sample collection and quality

The proportion of subjects (at each stable and exacerbation visit) who provided a sputum sample will be computed; overall and by the sample method (spontaneous or induced).

The proportion of subjects with induced sputum will be tabulated by COPD status at enrolment and by severity of the AECOPD.

The distribution of the time (in days) between the start of exacerbation and the day of the sputum sample collection will be presented.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

6.2.8.2. Sputum bacterial/viral results from culture or from PCR

The distribution of the proportion of sputum samples positive given the number of days between start of exacerbation and sample collection will be presented.

The proportion of sputum samples obtained at each - stable and exacerbation - visit and positive for specific bacterial/viral pathogens (overall and by bacterial culture, bacterial PCR, or viral species) will be computed. The average infection rate by type of visit will also be calculated.

The bacterial presence in culture versus the bacterial presence in PCR detected in sputum will be tabulated.

The viral infections versus the bacterial presence in PCR detected in sputum will be tabulated.

The incidence of exacerbations and of exacerbations containing Hi, Mcat, Sp, and all bacterial and/or virus that are planned to be tested by the laboratory will be presented.

For specific bacterial pathogens, the quantitative assessment of PCR results will be provided overall, for sputum samples positive in culture, or negative in culture of the specific pathogen.

6.2.8.3. Impact of previous antibiotic administration on sputum bacterial results

The impact of antibiotic administration before sputum collection on the bactericidal culture or PCR results will be evaluated, for stable (scheduled) visits, exacerbation visits, or any visit.

6.2.8.4. Impact of induction of sample on sputum bacterial results

The impact of sputum induction on the bactericidal culture or PCR results will be evaluated, for stable (scheduled) visits, exacerbation visits, or any visit.

6.2.8.5. Impact of previous occurrence of Hi bacterial pathogen in sputum on AECOPDs

The impact of the occurrence or presence of respiratory pathogenic bacteria detected in sputum at the exacerbation visit given the presence of the bacteria at previous visits on the number of exacerbations will also be studied. In order to do this, 2 types of AECOPDs will be distinguished:

- AECOPDs in which Hi is present and was not present any previous visit
- AECOPDs in which Hi is present and was not present in the previous stable visit;

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

7. STATISTICAL CALCULATIONS

7.1. Derived and transformed data

The study groups will be defined by treatment actually administered at Dose 1.

7.1.1. Demography

- For a given subject and a given demographic variable, missing measurement will not be replaced.
- Age: Age at the reference activity, computed as the number of units between the date of birth and the reference activity In case of partial dates of any of these 2 dates:
 - 15th of month, If only the day is missing
 - 30th of June, if day and months are missing

7.1.2. Safety

- For solicited symptoms, the analysis will exclude subjects with missing or un-evaluable measurements (*i.e.* total analysis of solicited symptoms will include all vaccinated subjects with documented solicited symptom sheets).
- For the unsolicited symptoms and concomitant medications/ products/ vaccinations, subjects who miss reporting symptoms/concomitant medications/ products/ vaccinations will be treated as subjects without unsolicited symptoms or concomitant medications/ products/ vaccinations, respectively.
- Note that for all tables described in this section, the way the percentage of subjects will be derived will depend on the event analysed (see table below for details). As a result, the N value will differ from one table to another.

Event	N used for deriving % per subject
Concomitant vaccination	All subjects with study vaccine administered
Solicited general symptom	All subjects with at least one solicited general symptom documented as either present or absent (<i>i.e.</i> symptom screen completed)
Solicited local symptom	All subjects with at least one solicited local symptom documented as either present or absent (<i>i.e.</i> symptom screen completed)
Unsolicited symptom	All subjects with study vaccine administered
Concomitant medication	All subjects with study vaccine administered

Study alias & e-track number(s): NTHi-004 (200157)

7.1.3. Immunogenicity

- For a given subject and the analysis of a given immunogenicity measurement, missing or un-evaluatable measurements will not be replaced.
- A seronegative subject is defined as a subject whose antibody concentration is below the assay cut-off value.
- A seropositive subject is defined as a subject whose antibody concentration is greater than or equal to the assay cut-off value.
- Antibody concentrations below the assay cut-off will be given an arbitrary value of half the assay cut-off for the purpose of geometric mean concentration (GMC) calculation.
- Calculation of the GMCs will be performed by taking the anti-logarithm in base 10 (anti-log10) of the mean of the log10 concentration transformations.
- The assay cut-off is the value under which there is no quantifiable result available. For an assay with a specific 'cut_off' , numerical immuno result is derived from a character field (rawres):
 - If rawres is 'NEG' or '-' or '(-)', numeric result= cutt_off/2,
 - if rawres is 'POS' or '+' or '(+)', numeric result = cut_off,
 - if rawres is '< value' and value<=cut_off, numeric result =cut_off/2,
 - if rawres is '< value' and value>cut_off, numeric result =value,
 - if rawres is '> value' and value<cut_off, numeric result =cut_off/2,
 - if rawres is '> value' and value>=cut_off, numeric result =value,
 - if rawres is '<= value' or '>= value' and value<cut_off, numeric result =cut_off/2,
 - if rawres is '<= value' or '>= value' and value>=cut_off, numeric result =value,
 - if rawres is a value < cut_off, numeric result = cut_off/2,
 - if rawres is a value >= cut_off, numeric result = rawres,
 - else numeric result is left blank.
- All CI computed will be two-sided 95% CI.

Impact of the investigational vaccine on AECOPD

- For a given subject and the analysis of a given laboratory assay, missing or un-evaluatable measurements will not be replaced.
- Subjects who did not have any sputum collected during at least 1 AECOPD (if at least 1 AECOPD occurred) will not be taken into account for the NTHI-associated AECOPD analyses.
- For summaries of the number of AECOPD occurring during the 1 year follow-up period starting 1 month post-Dose 2, an exacerbation rate will be calculated for each subject. The number of exacerbations during the 1-year follow-up period will be

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

imputed for subjects withdrawing from the study to provide an estimate of the number of exacerbations over the follow-up period. This calculation will only be performed for purposes of reporting summary statistics for the rate of exacerbations during the 1 year follow-up period since the modelling of exacerbations takes into account the number of exacerbations and the time of follow-up for each subject. The calculation of exacerbation rate will be based on follow-up period intervals to avoid obtaining high imputed rates if a subject withdrew very early in the follow-up period after experiencing an exacerbation. Since treatment courses for moderate/severe exacerbations are $\leq 2 - 4$ weeks when appropriate, calculated numbers of exacerbations for subjects withdrawing from the study will be based on 4-week intervals of the follow-up period. For subjects followed for less than 1 year, the number of exacerbations during the 1-year follow-up period will be calculated by multiplying the number of exacerbations experienced by the subject by 13 and dividing by the number of 4-week periods the subject was followed up:

Number of exacerbations per year = Number of exacerbations * 13 /Number of 4-week treatment period intervals

7.1.4. NTHi associated AECOPD

An NTHI associated to AECOPD is not yet well defined at the time of the SAP is written. Here is the first approach:

- An AECOPD for which NTHi was observed: Hi confirmed first by culture and then by PCR.

7.1.5. HRQOL

The CAT index will be derived as the sum of the ratings recorded for each of the eight individual items. Each of these items has 6 possible scores (0, 1, 2, 3, 4 or 5), leading to a range of 0 to 40 for CAT score.

7.1.6. EXACT-PRO derived AECOPD definitions and calculation rules

Start of an EXACT-PRO signal for exacerbation (as in EXACT-PRO user Manual and in Makay, et al., 2014):

1. *An EXACT-PRO exacerbation is defined as a 12 point increase above baseline for 2 consecutive days, or a nine point increase for 3 consecutive days.*
2. *The baseline score corresponds to the average EXACT score during a baseline period defined as days 14 to 8 preceding exacerbation onset*
3. *Exacerbation onset will be taken as the first day in which at least two symptoms (at least 1 major) are present and have not been discounted (see above).*

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

4. Only symptom data collected during the study period will be used to define the start of an exacerbation.

End/recovering of an EXACT-PRO signal of an exacerbation:

1. The end/recovering of an EXACT-PRO signal of an exacerbation corresponds to the day at the EXACT-PRO score return to its baseline value.
2. The length of a signal of an exacerbation will be at least of 2 days. So rule 1 will only apply from the third day onwards.

Missing data:

Days with missing data will not be ignored when defining start of a signal: e.g., if a patient has two qualifying symptoms followed by a missing day, if the next day with data has two qualifying symptoms this will not be defined as an exacerbation.

7.1.7. “e-diary derived AECOPD” definitions and calculation rules

The definition of an AECOPD case definition from section 5.3.1.2 of the Protocol ***is what we will call here an “e-diary derived (AECOPD) signal”***. In this paragraph, only the derivation of occurrence and length of ***e-diary*** signals is described. Derivation of occurrence and length of AECOPDs is described in Section 7.1.11.

Start of an e-diary signal for exacerbation:

1. An exacerbation is defined as a worsening of at least 2 symptoms for ≥ 2 consecutive days where at least one of the symptoms is major.
2. The same two symptoms do not have to be present on both days as long as at least one major symptom is present on both days.
3. Any symptom that has occurred consecutively for 5 days or more prior to a potential exacerbation should be discounted from the symptoms that define an exacerbation.
4. Exacerbation onset will be taken as the first day in which at least two symptoms (at least 1 major) are present and have not been discounted (see above).
5. Only symptom data collected during the study period will be used to define the start of an exacerbation.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

7.1.8. End/Duration of an *e-diary* signal of an exacerbation:

1. The end/recovering of an *e-diary* signal of an exacerbation corresponds to the day at which the 3-day moving average of the daily symptom count score (average between the day the day before and the day after) to return to baseline symptom count score (the mean daily symptom count score over days 14 to 8 preceding exacerbation onset). For a specific day, the daily symptom count score is defined as the sum of the symptom scores experienced in this day (a major symptom accounts for 2 points, a minor symptom accounts for 1 point, as in Trappenburg *et al.*, 2011, but contrary to Seemungal *et al.* 2000 and Mackay *et al.* 2014, in which all symptoms accounted to 1). Note that the symptoms that have occurred consecutively for 5 days or more prior to a potential exacerbation and that are discarded for these exacerbations are not discarded from the symptoms counts.
2. The length of an *e-diary* signal of an exacerbation will be at least of 2 days. So rule 1 will only apply from the third day onwards.
3. The baseline symptom count score is calculated as the average count score of the observed days in the period used to calculate the baseline symptom count score. For example, if an exacerbation occurs from day 10 onwards, the baseline period will be the average between day1 and day2 (day10 - 8 days= day2, day10 - 9 days = day1, Day1 being the first study day). For exacerbations before day 9 but after day 2, the baseline symptom score will be considered to be day1
4. If the first *e-diary* signal of an exacerbation is occurring before day 3: duration will be counted until the day in which there is at most 2 majors or 1 major and one minor symptom.
5. Recovery time is calculated as the number of days from exacerbation *e-diary* signal onset (included) to exacerbation *e-diary* signal end (not included).

7.1.9. Time between 2 *e-diary* signals for exacerbations/evaluation of baseline symptom score:

The minimum duration between 2 *e-diary* signals with distinct baseline value is 15 days, which is the minimal period for which a new full exacerbation-free baseline can be calculated for the next *e-diary* episode.

If a new worsening of symptoms occurs between 6 and 14 days of a previous exacerbation, the baseline period used will be the same baseline as the one of the original exacerbation. Therefore, it will last until the daily symptom count score has gone below the baseline symptom score of the previous event.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

7.1.10. Missing data

1. Days with missing data will not be ignored when defining start of an **e-diary** signal: e.g., if a patient has two qualifying symptoms followed by a missing day, if the next day with data has two qualifying symptoms this will not be defined as an exacerbation. Hence, the start of an **e-diary** signal of an exacerbation will require 2 consecutive non-missing days of worsening. If there is partially recorded symptom data at a visit, however, only the observed values will be considered but the day will not be discarded.
2. Days with missing data will be ignored when defining end of an **e-diary** signal of an exacerbation: e.g., if a patient has two days missing in the baseline period and one day missing in the 3- day moving average, both baseline and 3-day moving average values will be calculated using observed data only. Similar rules will apply when there is partially recorded symptom data at a visit (and daily average values will be calculated among the observed values).
3. All available data will be used for each patient.
4. As a general rule, the baseline symptom count score, and 3 day moving averages are calculated as the average count score of the observed days in the baseline period. For example, if an **e-diary** signal occurs from day 10 onwards, the baseline period will be the average between day1 and day2 (day1 being the first study day). If day 13 is missing, the 3 day moving average at day 14 will only take into account day 14 and day 15.
5. If there are greater than 2 consecutive days of missing data after an **e-diary** signal not resolved, the **e-diary** signal will not be considered resolved until the next day with observed data if the corresponding moving average is below the baseline value. In other words, if the 3-day average daily score is missing, the day will be counted in the duration of the episode.

7.1.11. AECOPDs definitions and calculation rules

7.1.11.1. AECOPD definition (from clinical evaluation)

Due to the sensitivity and specificity of the e-diary device, false alerts may occur, and exacerbations resulting in site visits may also occur without e-diary alert. Therefore, only unscheduled AECOPD visits with clinical confirmation of AECOPD (with or without associated signals) will be considered as AECOPD events in the primary analysis.

In addition, missed AECOPD visits may have occurred and can be:

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

- *a signal from the e-diary (PHT) without a site visit but that cannot be discounted as a possible exacerbation signal after reconciliation,*
and/or
- *an exacerbation reported retrospectively in the Safety database or medical records for which there has not been in the same period of retrospective reporting a corresponding exacerbation visit declared.*

7.1.11.2. Characterization of inconsistencies between morning e-diary signals and exacerbation site visits

The reconciliation between morning e-diary (PHT) signals and site visits for exacerbation will aim at characterizing the following 2 cases:

Case A : an e-diary alert with absent/delayed exacerbation visit (i.e. no exacerbation visit within the calculated length of the e-diary episode)

Case B : a exacerbation visit with absent/delayed e-diary exacerbation alert (i.e. RDE exacerbation not within e-diary exacerbation period)

The different reasons of inconsistencies of Case A will be categorized by the site after review according to the following explanation:

- 1) *Not an exacerbation after further evaluation* : eg. when a patient does not feel well but this is due to increased exercise, or patient does not feel well but this is not clinically significant.
- 2) *Ongoing exacerbation* : e.g., previous exacerbation is not terminated according to the clinician.
- 3) *RDE or other technical issues*: e.g., due to technical issues e-diary alerts were not sent and when the site was aware of it was too late to determine whether this was an exacerbation or not.
- 4) *Other reasons: specify* : e.g. subject was on holidays / subject unable to visit the site / delayed RDE visit/early RDE entry.

The different reasons of inconsistencies of Case B will be categorized by the site after review according to the following explanation:

- 5) *Spontaneous visit clinically deemed exacerbating* i.e. exacerbation visit took place before the e-diary alert (spontaneous visit, visit triggered by GP / Patient call asked to be seen/ felt too unwell to complete e-diary abrupt onset)
- 6) *Planned monthly visit clinically deemed exacerbating*

Statistical Analysis Plan



GlaxoSmithKline

Study alias & e-track number(s): NTHi-004 (200157)

7) *e-diary diary card recording issue or other technical issues* : e.g. late e-diary alerts, subject missed or did not correctly record diary card, broken log pad device.

8) *Other reasons: specify*

For a Case A, no laboratory data has been collected, but if the answer about the inconsistency is 3) or sometimes 4) (depending on the specification), it will qualify as AECOPD. All Cases B will be considered to be an AECOPD if the AECOPD is confirmed by the site.

7.1.12. Grading of severity of an exacerbation

Severity of exacerbations is defined as (Protocol Section 5.3.1.2):

- Mild: Worsening symptoms of COPD that are self-managed by the patient.
- Moderate: Worsening symptoms of COPD that require treatment with oral corticosteroids and/or antibiotics.
- Severe: Worsening symptoms of COPD that require treatment with in-patient hospitalisation or home care intervention.

Severity of the exacerbation will not be derived but taken directly from the information entered in the CRF.

7.2. Data presentation description

The following decimal description will be used for the demography and safety/reactogenicity.

Display Table	Parameters	Number of decimal digits
All summaries	% of count, including LL & UL of CI	1
All summaries	% of difference, including LL & UL of CI	2
Demographic characteristics	Mean, median age	1
Demographic characteristics	SD (age)	2
Immunogenicity	GMT ratio	2

The following data presentation description will be used for the immunogenicity analyses.

Component	GMC/T	Assay		Assay cut-off	Number of decimal digits		Tick marks on RCC's		
		method	Unit		GMC/T	GMC/T ratio			
					Low value	High value			
Anti-PD	GMC	ELISA	EL.U/mL	100	1	2	1	10000	
Anti-PE	GMC	ELISA	EL.U/mL	8	1	2	1	10000	
Anti-PiLA	GMC	ELISA	EL.U/mL	7	1	2	1	10000	

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

8. CONDUCT OF ANALYSES

8.1. Sequence of analyses

The analyses will be performed stepwise:

- The analysis of data up to Visit 6 (Day 90) will be performed in a first step. This analysis will include:
 - The final analysis of immunogenicity and solicited AEs post-Dose 1 and post-Dose 2.
 - The assessment of unsolicited AEs, SAEs and pIMDs up to 30 days post-Dose 1 and post-Dose 2 on as cleaned as possible data.

This analysis will be documented in a statistical report. At this point, the GSK statistician of the project (or delegates) will be unblinded (*i.e.* will have access to the individual subject treatment assignments).

- The final analysis of data up to study conclusion will be performed in a second step, once those data will be available and cleaned. This analysis will include:
 - The final analysis of immunogenicity at Day 270 and Day 450.
 - SAEs and pIMDs up to Day 450 on cleaned data.
 - Impact of the investigational vaccine on AECOPD, HRQOL, lung function, exercise capacity and biomarkers.

In addition, all previous analyses will be re-produced based on cleaned data at this point.

Individual listings will only be provided at this stage. A final study report containing data from the entire study will be written and will be made available to the investigators.

Description	Analysis ID (SDD & CARS sub-folder)
Analysis up to Visit 6 (Day 90)	E1_06
Final analysis (up to day 450)	E1_01

8.2. Statistical considerations for interim analyses

Not applicable.

Statistical Analysis Plan



GlaxoSmithKline

Study alias & e-track number(s): NTHi-004 (200157)

9. CHANGES FROM PLANNED ANALYSES

As during the study, some subjects have completed wrong diary cards (paediatric ones instead of adults), a sensitivity analysis will be done on subjects that have completed the correct diary cards to be sure that they were no impact on results.

10. REFERENCES

The exact 95% CIs for a proportion within a group will be calculated from Proc StatXact [Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of binomial. *Biometrika*. 1934;26:404-413].

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

Detailed Title:	A Phase II, randomised, observer-blind, placebo-controlled, multi-centre study to evaluate the safety, reactogenicity and immunogenicity of GSK Biologicals' investigational vaccine GSK2838504A, when administered intramuscularly according to a 0, 2 month schedule to COPD patients aged 40 to 80 years
SAP version	Amendment 1
SAP date	27-JUN-2016
Scope:	All data pertaining to the above study.
Co-ordinating author:	PPD [REDACTED]
Other author(s):	
Adhoc reviewers:	PPD [REDACTED], Global Regulatory Lead (RA) PPD [REDACTED], Project Leader PPD [REDACTED], CLS Clinical Read-Out Team Leader PPD [REDACTED], PPD [REDACTED], Safety Representatives
CTRS team	
Approved by:	PPD [REDACTED], Clinical and Epidemiology Project Lead PPD [REDACTED], Clinical Research & Development Lead PPD [REDACTED], Lead Statistician PPD [REDACTED], Scientific Writer PPD [REDACTED], Statistician

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

TABLE OF CONTENTS

	PAGE
LIST OF ABBREVIATIONS	4
1. DOCUMENT HISTORY	7
2. STUDY DESIGN	8
3. OBJECTIVES	9
3.1. Primary objective	9
3.2. Secondary objective	9
3.3. Tertiary objectives	9
4. ENDPOINTS	9
4.1. Primary endpoints	9
4.2. Secondary endpoints	10
4.3. Tertiary endpoints	10
5. STUDY POPULATION	11
5.1. Total vaccinated cohort	11
5.2. ATP cohort for analysis of immunogenicity	11
5.3. ATP cohort for analysis of persistence of immunogenicity	12
5.4. ATP cohort for analysis of efficacy	12
6. STATISTICAL METHODS	13
6.1. Analysis of demographics	13
6.1.1. Analysis of safety	13
6.1.2. Analysis of immunogenicity	15
6.1.3. Analysis of impact of the investigational vaccine on AECOPD incidence	16
6.1.4. AECOPD visits and daily questionnaires	18
6.1.5. Analysis of HRQOL, lung capacity and exercise capacity	18
6.1.6. Analysis of selected biomarkers	18
6.1.7. Analysis of sputum	19
6.1.7.1. Sputum sample collection and quality	19
6.1.7.2. Sputum bacterial/viral results from culture or from PCR	19
6.1.7.3. Impact of previous antibiotic administration on sputum bacterial results	20
6.1.7.4. Impact of induction of sample on sputum bacterial results	20
6.1.7.5. Impact of previous occurrence of Hi bacterial pathogen in sputum on AECOPDs	20

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

7. STATISTICAL CALCULATIONS	20
7.1. Derived and transformed data.....	20
7.1.1. Demography	20
7.1.2. Safety	21
7.1.3. Immunogenicity.....	21
7.1.4. NTHi associated AECOPD.....	23
7.1.5. HRQOL.....	23
7.1.6. EXACT-PRO derived AECOPD definitions and calculation rules.....	23
7.1.7. “e-diary derived AECOPD” definitions and calculation rules.....	24
7.1.8. End/Duration of an <i>e-diary</i> signal of an exacerbation:	24
7.1.9. Time between 2 <i>e-diary</i> signals for exacerbations/evaluation of baseline symptom score:	25
7.1.10. Missing data	25
7.1.11. AECOPDs definitions and calculation rules.....	26
7.1.11.1. AECOPD definition (from clinical evaluation)	26
7.1.11.2. Characterization of inconsistencies between morning <i>e-diary</i> signals and exacerbation site visits	26
7.1.12. Grading of severity of an exacerbation.....	27
7.1.13. Sputum Quality	28
7.2. Data presentation description	28
8. CONDUCT OF ANALYSES.....	28
8.1. Sequence of analyses.....	28
8.2. Statistical considerations for interim analyses	29
9. CHANGES FROM PLANNED ANALYSES.....	30
10. REFERENCES.....	30

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

LIST OF ABBREVIATIONS

AE	Adverse event
AECOPD	Acute exacerbation of COPD
ALT	Alanine aminotransferase
ATP	According-To-Protocol
ANCOVA	Analysis of covariance
CAT	COPD assessment test
CI	Confidence Interval
CMI	Cell-mediated immunity
COPD	Chronic obstructive pulmonary disease
CRF	Case Report Form
CTRS	Clinical Trial Registry
EF	Internal GSK database code for Efficacy analysis (elimination codes) link to FORM-BIO-CLIN-9004-05 Criteria for eliminating subjects from the analyses
EL.U/ml	ELISA unit per milliliter
ELISA	Enzyme-linked immunosorbent assay
Eli Type	Internal GSK database code for type of elimination code
EXACT-PRO	EXAcerbations of Chronic Pulmonary Disease Tool - Patient Reported Outcome
FEV ₁	Forced expiratory volume in 1 second
FVC	Forced vital capacity
FU	Internal GSK database code for Follow-up analysis (elimination codes) link to FORM-BIO-CLIN-9004-05 Criteria for eliminating subjects from the analyses
GMC	Geometric mean antibody concentration
GOLD	Global Initiative for Chronic Obstructive Lung Disease
GSK	GlaxoSmithKline
<i>H. influenzae</i>	<i>Haemophilus influenzae</i>
HRQOL	Health-Related Quality of Life
ICF	Informed consent form
ICS	Intracellular cytokine staining

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

IU/ml	International units per milliliter
LL	Lower Limit of the confidence interval
MA	Internal GSK database code for Main analysis (elimination codes) link to FORM-BIO-CLIN-9004-05 Criteria for eliminating subjects from the analyses
MATEX	MATerial EXcellence
<i>M. catarrhalis</i>	<i>Moraxella catarrhalis</i>
MedDRA	Medical Dictionary for Regulatory Activities
mMRC scale	mMRC scale
N.A.	Not Applicable
NTHi	Non-Typeable <i>Haemophilus influenzae</i>
PBMC	Peripheral blood mononuclear cell
PCR	Polymerase chain reaction
PD	Protein D
PE	Protein E
PI	Prescribing information
PilA	Type IV pili subunit of non-typeable <i>Haemophilus influenzae</i>
pIMD	Potential immune-mediated disease
PT	Preferred term
SAE	Serious adverse event
<i>S. aureus</i>	<i>Staphylococcus aureus</i>
SAP	Statistical Analysis Plan
SBIR	GSK Biological's Internet Randomization System
SD	Standard Deviation
SR	Study Report
STGG	Skim milk, tryptone, glucose, and glycerin transport medium
SYN	Synopsis
TFL	Tables Figures and Listing template annexed to SAP
TVC	Total vaccinated cohort
UL	Upper Limit of the confidence interval
VE	Vaccine efficacy

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

6MWT

6-Minute Walk Test

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

The complete statistical analysis plan and results presentation is divided into 2 parts: the first part detailing the analyses to be performed (known as SAP, current document) and a second part, annex (-es) (called TFL) describing the flow and format of tables, figures and listings to be annexed to the SR.

1. DOCUMENT HISTORY

Date	Description	Protocol Version
29-JAN-2016	Version 1	Amendment 2 - 15-DEC-2014
21-JUN-2016	<p>Amendment 1</p> <p>In order to see early effects of the vaccine on the microbiome, analysis on fresh sputum samples (culture results) will be done on all available data up to the data lock point of the interim analysis.</p> <p>In order to have a first look whether or not the investigational vaccine has an impact on AECOPD, AECOPD analyses will be done up to the data lock point of the interim analysis</p>	Amendment 3 - 15-APR-2016

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

2. STUDY DESIGN

Figure 1 Study design overview

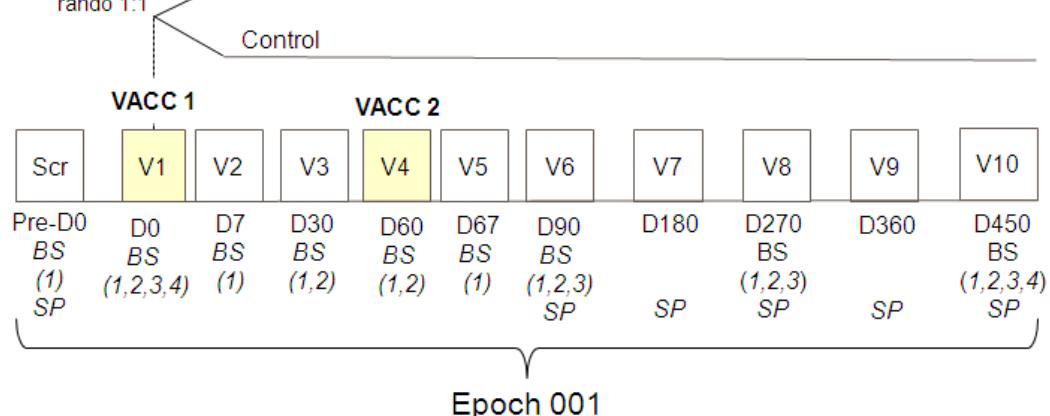
$N \geq 70/\text{group}$

40 – 80 years

10-30 years
COPD patients

10 μ g Ag/ AS01 E

108



Scr = Screening; **V** = Visit; **D** = Day; **rando** = randomisation; **VACC** = vaccination; **BS(1)** = blood sample for safety; **BS(2)** = blood sample for humoral immunogenicity; **BS(3)** = blood sample for cell-mediated immunity (CMI); **BS(4)** = blood sample for biomarkers; **SP** = sputum sample

In addition to the scheduled study visits depicted above, **ad hoc AECOPD-driven study visit(s) and/or phone contact(s)** will take place for each AECOPD occurring from first vaccination up to study conclusion. An AECOPD visit will be scheduled as soon as possible after the onset of AECOPD symptoms (maximum 96 hours after and, if applicable, preferably before starting treatment with antibiotics). In addition, follow-up phone call(s) and/or visit(s) will take place to determine the end of the AECOPD. These contacts will take place at least every 2 weeks, until the AECOPD has resolved.

Experimental design: Phase II randomised, observer-blind, placebo-controlled, multi-centre study with 2 parallel groups.

Duration of the study: for each subject enrolled, the study will last approximately 16 months from Visit 1 up to study end

The following group names will be used for the statistical analyses:

Group order in tables	Group label in tables	Group definition for footnote
1	10-AS01E	2 doses of AS01e adjuvanted vaccine containing 10 mcg of each antigen
2	Control	2 doses of saline solution

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

3. OBJECTIVES

3.1. Primary objective

- To describe the safety and reactogenicity of the investigational vaccine.

3.2. Secondary objective

- To describe the humoral and cellular immunogenicity of the investigational vaccine.

3.3. Tertiary objectives

- To explore the impact of the investigational vaccine on AECOPD.
- To explore the impact of the investigational vaccine on NTHi presence in sputum.
- To explore the impact of the investigational vaccine on health-related quality of life (HRQOL).
- To explore the impact of the investigational vaccine on lung function.
- To explore the impact of the investigational NTHi vaccine on exercise capacity.
- To describe selected biomarkers in stable COPD and during AECOPD.
- To collect blood and sputum samples for assay development, for disease diagnostic purpose, for lung microbiome analysis and/ or for additional evaluation of the immune responses to the investigational vaccine and to other potential pathogens involved in AECOPD.

4. ENDPOINTS

4.1. Primary endpoints

- Occurrence of each solicited local adverse event (AE), during the 7-day follow-up period (*i.e.* day 0 - 6) following each vaccination, in all subjects.
- Occurrence of each solicited general AE, during the 7-day follow-up period (*i.e.* day 0 - 6) following each vaccination, in all subjects.
- Occurrence of any unsolicited AE, during the 30-day follow-up period (*i.e.* day 0 - 29) following each vaccination, in all subjects.
- Occurrence of each haematological/ biochemical laboratory abnormality at Day 0, Day 7, Day 30, Day 60, Day 67, Day 90, Day 270 and Day 450, in all subjects.
- Occurrence of any potential immune-mediated disease (pIMD) from first vaccination up to study conclusion, in all subjects.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

- Occurrence of any serious adverse event (SAE) from first vaccination up to study conclusion, in all subjects.

4.2. Secondary endpoints

- Anti-PD, anti-PE and anti-PiLA total IgG antibody concentrations as measured by ELISA, at Day 0, Day 30, Day 60, Day 90, Day 270 and at Day 450, in all subjects.
- NTHi-specific cell-mediated immune responses as measured by flow cytometry intracellular cytokine staining (ICS) (frequency of specific CD4⁺/CD8⁺ T-cells expressing two or more markers, such as IL-2, IL-13, IL-17, IFN- γ , TNF- α and CD40L), at Day 0, Day 90, Day 270 and at Day 450, in a sub-cohort of subjects.

4.3. Tertiary endpoints

- Number of cases of NTHi-associated moderate and severe AECOPD, over a period starting 1 month post-Dose 2 and lasting for 1 year, in all subjects.
- Number of cases of NTHi-associated AECOPD (any severity), over a period starting 1 month post-Dose 2 and lasting for 1 year, in all subjects.
- Number of cases of moderate and severe AECOPD (any cause), over a period starting 1 month post-Dose 2 and lasting for 1 year, in all subjects.
- Number of cases of AECOPD (any cause, any severity), over a period starting 1 month post-Dose 2 and lasting for 1 year, in all subjects.
- Time to first AECOPD, in all subjects.
- Assessment of EXACT-PRO score, daily at bedtime throughout the study, in all subjects.
- NTHi presence in sputum, at Screening, Day 90, Day 180, Day 270, Day 360 and Day 450, and from first vaccination to study conclusion for each AECOPD, in all subjects.
- Assessment of CAT score, at Screening, Day 270 and Day 450, and from first vaccination to study conclusion for each AECOPD, in all subjects.
- Assessment of SGRQ-C score, at Screening, Day 270 and Day 450, in all subjects.
- Assessment of mMRC scale, at Screening, Day 270 and Day 450, and from first vaccination to study conclusion for each AECOPD, in all subjects.
- Occurrence of rescue medication use, over a period starting 1 month post-Dose 2 and lasting for 1 year, in all subjects.
- Occurrence of healthcare use for AECOPD, over a period starting 1 month post-Dose 2 and lasting for 1 year, in all subjects.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

- Assessment of FEV1% of predicted normal value at Screening, Day 270 and Day 450, in all subjects.
- Assessment of 6MWT score, at Screening and Day 450, in all subjects.
- Occurrence of selected biomarkers in a subset of blood samples, at Day 0 and Day 450, and from first vaccination to study conclusion for each AECOPD, in all subjects.
- Presence of respiratory viral pathogens in sputum (including respiratory syncytial virus, parainfluenza virus, enterovirus/ rhinovirus, metapneumovirus, influenza virus, adenovirus, bocavirus and coronavirus) at Screening, Day 90, Day 180, Day 270, Visit Day 360 and Day 450 and at each AECOPD visit from first vaccination to study conclusion, in all subjects.

5. STUDY POPULATION

The following study cohorts will be evaluated.

5.1. Total vaccinated cohort

The Total vaccinated cohort (TVC) will include all subjects with at least 1 study vaccine administration documented:

- A **safety** analysis based on the TVC will include all vaccinated subjects.
- An **immunogenicity/ efficacy** analysis based on the TVC will include all vaccinated subjects for whom immunogenicity/ efficacy data are available.

The TVC analysis will be performed per treatment actually administered at Dose 1.

5.2. ATP cohort for analysis of immunogenicity

The ATP cohort for immunogenicity will include all subjects in the TVC:

- Who met all eligibility criteria.
- For whom the administration route and site of the vaccines was according to protocol.
- Who complied with the vaccination schedule.
- Who received the study vaccines according to protocol procedures.
- Who did not receive a concomitant medication/ product leading to elimination from the ATP analysis for immunogenicity up to the 1 month post-Dose 2 visit (Day 90).

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

- Who did not present with an intercurrent medical condition leading to elimination from the ATP analysis for immunogenicity up to the 1 month post-Dose 2 visit (Day 90).
- Who complied with the blood sample timings, at the 1 month post-Dose 2 visit (Day 90).
- For whom post-vaccination immunogenicity results are available for at least 1 assay.

5.3. ATP cohort for analysis of persistence of immunogenicity

The ATP cohort for persistence of immunogenicity (starting at Day 180) will include all evaluable subjects, *i.e.*, those who were included in the ATP cohort for immunogenicity, or were excluded from this cohort solely because they had no blood samples taken or because of incompliance with blood sample timings up to the 1 month post-Dose 2 visit, and:

- Who did not receive a concomitant medication/ product leading to elimination from the ATP analysis for immunogenicity.
- Who did not present with an intercurrent medical condition leading to elimination from the ATP analysis for immunogenicity.
- Who complied with the blood sample timings after the 1 month post-last vaccination visit.
- For whom persistence immunogenicity results are available for at least 1 assay.

5.4. ATP cohort for analysis of efficacy

The ATP cohort for analysis of efficacy will include all subjects in the TVC:

- Who met all eligibility criteria.
- For whom the administration route and site of the vaccines was according to protocol.
- Who complied with the vaccination schedule
- Who received the study vaccine according to protocol procedures.
- Not having received a medication/ product/ vaccine that may lead to elimination from the ATP analysis for efficacy.

The list of applicable elimination codes for each cohort can be found in the study specific form FORM-BIO-CLIN-9004-05 Criteria for eliminating subjects from the analyses.

Cohort	Elimination codes	Eli Type
ATP cohort for analysis of immunogenicity	1010-2500	MA

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

ATP cohort for analysis of efficacy	1010-2500	EF
ATP cohort for analysis of persistence of immunogenicity	1010-2500	FU

6. STATISTICAL METHODS

6.1. Analysis of demographics

Demographic characteristics (age at the first dose in years, gender and geographical ancestry), cohort description and other characteristics such as smoking/exposure history status, pulmonary function test baseline results and COPD history, history of exacerbation present/past medical history/intercurrent morbidity, and previous vaccination with a pneumococcal vaccine or an influenza vaccine will be summarized by group using descriptive statistics:

- Frequency tables will be generated for categorical variable such as geographical ancestry;
- Mean, median and standard error will be provided for continuous data such as age.

Other characteristics such as height, weight, BMI selected results and several risk factors from the six-minutes-walk test, and distribution of several severity stages using CAT or mMRC or SGRQ-C at enrolment visit will also be summarized.

The number of subjects enrolled into the study as well as the number of subjects excluded from ATP analyses will be tabulated.

The distribution of subjects enrolled among the study sites will be tabulated as a whole and per group.

The numbers of withdrawn subjects will be tabulated according to the reason for withdrawal.

Withdrawal status will be summarised by group using descriptive statistics:

6.1.1. Analysis of safety

The analysis will be performed on the TVC. If the percentage of vaccinated subjects excluded from the ATP cohort for efficacy is at least 10%, a second analysis will be performed on the ATP cohort for efficacy to complement the first analysis.

The percentage of subjects with at least one **local AE** (solicited and unsolicited), with at least one **general AE** (solicited and unsolicited) and with any AE during the 7-day (day 0 – day 6) or the 30-day (day 0 – day 29) follow-up period will be tabulated after each

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

vaccination and overall with exact 95% confidence interval (CI). The same computations will be done for Grade 3 AEs, for related AEs and for Grade 3 related AEs.

The percentage of subjects/doses reporting each individual solicited local (any grade, Grade 3) and general (any grade, Grade 3, related, Grade 3 related) AE during the 7-day (day 0 to day 6) follow-up period will be tabulated for each group as follows:

- Overall, the percentage of subjects with the symptom and its exact 95% CI.
- Overall, the percentage of doses with the symptom and its exact 95% CI.
- Per study vaccine dose, the percentage of subjects with the symptom and its exact 95% CI.

The exact 95% CIs will be calculated assuming independence between doses. For fever, additional analyses will be performed by 0.5°C increments.

As during the study, some subjects have completed wrong diary cards (paediatric ones instead of adults), a sensitivity analysis will be done on subjects that have completed the correct diary cards to be sure that they were no impact on results.

The verbatim reports of **unsolicited** symptoms will be reviewed by a physician and the signs and symptoms will be coded according to the MedDRA Dictionary for Adverse Reaction Terminology. Every verbatim term will be matched with the appropriate preferred term (PT). The percentage of subjects with unsolicited symptoms during the 30-day follow-up period after any study vaccine dose with its exact 95% CI will be tabulated by group and by MedDRA PT. Similar tabulation will be done for the percentage of doses, for Grade 3 unsolicited symptoms, for unsolicited symptoms that resulted in a medically attended visit, for Grade 3 and causally related unsolicited symptoms and for unsolicited symptoms causally related to vaccination.

For each group and for each **haematology/ biochemistry parameter**:

- The percentage of subjects having results below or above the local laboratory normal ranges will be tabulated by time point.
- The maximum grading from Screening up to study conclusion will be tabulated. Grades will be based on the FDA Guidance for Industry “Toxicity Grading Scale for Healthy Adults and Adolescent Volunteers Enrolled in Preventive Vaccine Clinical Trials” (see Appendix C of the protocol).

The number of subjects who experienced any **pIMD** or any **SAE** from from first vaccination up to 6 months post-Dose 2 will be reported will be reported.

The number of subjects who experienced any **AE leading to study withdrawal**, from first vaccination up to study conclusion, or any **SAE related to study participation of**

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

concurrent GSK medication/ vaccination, during the entire study period, will be reported.

Pregnancy reports from first vaccination up to 6 months after last vaccination and pregnancy outcomes will be described in detail.

The percentage of subjects/ dose using **concomitant medication/ product** (any medication/ product, any antipyretic and any antipyretic taken prophylactically, respectively) during the 7-day follow-up period (day 0 – day 6) and during the 30-day follow-up period (day 0 – day 29) will be summarised per group for each dose and overall per dose.

6.1.2. Analysis of immunogenicity

The analysis will be performed on the ATP cohort for immunogenicity. If the percentage of vaccinated subjects with serological results excluded from the ATP cohort for analysis of immunogenicity is more than 10%, a second analysis will be performed on the TVC.

Humoral immune response

Within group evaluation

For each group, at each time point during which blood samples are collected for humoral immune response and for each assay:

- Seropositivity rate and their exact 95% CI will be tabulated.
- GMCs and their 95% CI will be calculated.

For each group, at each time point during which blood samples are collected for humoral immune response:

- Antibody concentrations distribution will be investigated using Reverse Cumulative Curves.

Between groups evaluation

Comparative analyses will be exploratory with the aim to characterise the difference between the 2 groups in humoral immune response.

The difference in terms of GMCs will be evaluated at 1 month post-Dose 2 and at Day 450 by computing the 95% CIs of the GMC ratio between groups by using an ANCOVA model that considers heterogeneity of variances (via Proc Mixed) on the logarithm10 transformation of the concentrations. This model will include the country, the group, age category (40 - 59 years or 60 - 80 years), number of moderate and severe AECOPD in the year before vaccine Dose 1 (< 2 or \geq 2), GOLD grade (GOLD 2 or GOLD 3) and pre-Dose 1 concentration (as covariate) as fixed effects.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

However, these differences should be interpreted with caution considering that there will be no adjustment for multiplicity of endpoints.

Cell-mediated immune response

The frequency of specific CD4⁺/ CD8⁺ T-cells will be summarised (mean, SD, minimum, Q1, median, Q3, and maximum) by group, at each time point during which blood samples are collected for CMI (descriptive statistics).

Each data will be obtained by subtracting the summary of the background value to each summary value. Values less than or equal to zero will be set to 1 for the purpose of geometric mean calculation and graphical representation.

6.1.3. Analysis of impact of the investigational vaccine on AECOPD incidence

The analysis will be performed on the ATP cohort for efficacy. If the percentage of vaccinated subjects excluded from the ATP cohort for efficacy is more than 10%, a second analysis based on the TVC will be performed to complement the first analysis.

The following incidence rates will be computed over a period starting 1 month post-Dose 2 and lasting for 1 year, with 95%CI:

- NTHi-associated moderate and severe AECOPD. (See definition in section 7.1.4)
- NTHi-associated all-severity AECOPD.
- All-cause moderate and severe AECOPD.
- All-cause, all-severity AECOPD.

Vaccine efficacy (VE) for AECOPD events will be defined as:

$$VE (AECOPD) = 1 - R_{vaccine} / R_{control}$$

with:

- $R_{vaccine}$ = average yearly incidence rate of AECOPD events per subject in the group 10-AS01E.
- $R_{control}$ = average yearly incidence rate of AECOPD events per subject in the Control group.

In case the 95% CI of the VE cannot be estimated, *e.g.* when no event is observed in the group 10-AS01E (leading to a VE of 100%), then an exact procedure will be used to estimate a VE and its 95% CI. This exact procedure will be based on the number of subjects with event instead of on the average number of event per subject in each group.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

VE in the prevention of NTHi-associated moderate and severe AECOPD and its 95% CI will be calculated.

VE in the prevention of NTHi-associated all-severity AECOPD, of all-cause moderate and severe AECOPD and of all-cause, all-severity AECOPD and their 95% CI will also be calculated.

The 95%CI of the incidence rate will be computed using a model which accounts for repeated events. The Generalised linear model assuming the Negative Binomial distribution for the response variable with logarithm as link function, and the logarithm of time for follow-up as an offset variable will be used. In addition, the following covariates (as per randomization) will be taken into account:

- country,
- number of moderate/severe exacerbations reported in the 12 months prior to enrolment,
- COPD status at enrolment (derived by GSK from the Visit 1 spirometry results)
- Age (below or equal to median/above median)

If the model does not converge, the Poisson distribution will be used instead of the Negative Binomial one.

The frequency of all these AECOPDs will be presented. The estimated incidence rate with 95%CI of AECOPDs in which Hi (resp. Mcat, Sp) is present in culture and was not present in the previous visit estimated by a Generalised linear model assuming the Negative Binomial distribution will be estimated.

Time to first AECOPD event by baseline status - Cox regression model

The Cox's regression model including the risk factors:

- country,
- number of moderate/severe exacerbations reported in the 12 months prior to enrolment,
- COPD status at enrolment (derived by GSK from the Visit 1 spirometry results)
- Age (below or equal to median/above median)

as covariates (as per randomization) will be performed to model the time until the first occurrence of AECOPD episodes. Additional stratification/risk factor may be defined during the analyses like sex factor as exploratory analysis.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

6.1.4. AECOPD visits and daily questionnaires

The time until first AECOPD as of 1 month post-Dose 2 will be computed with exact 95% CI by group. Subjects who exacerbate before 1 month post-Dose 2 will not be taken into account in this analysis.

The number of morning e-dairy (PHT) alerts (resp. evening e-diary alerts (EXACT-PRO)) with and without corresponding site visits will be tabulated. The reason for inconsistency between morning e-diaries and site visits will also be presented.

The number of exacerbation site visits with confirmed AECOPD with and without morning e-diary alerts (resp. evening e-diary alerts (EXACT-PRO)) will be tabulated along with the reason for no e-diary alert.

Concordance/discordance of exacerbations periods as defined clinically or via the morning diary cards or the evening diary cards will be tabulated and presented graphically.

The percentage of spontaneous exacerbation visits will be tabulated as well as the percentage of monthly visits with recovered/stable status.

6.1.5. Analysis of HRQOL, lung capacity and exercise capacity

The analysis will be performed on the TVC. If the percentage of vaccinated subjects excluded from the ATP cohort for efficacy is at least 10%, a second analysis will be performed on the ATP cohort for efficacy to complement the first analysis.

Descriptive statistics (median, mean, range, standard deviation, first and third quartiles) on the **EXACT-PRO**, **CAT**, **SGRQ-C** and **mMRC** scores, on **FEV₁% of predicted normal value** and on the **6MWT** scores will be tabulated by specified visit.

Descriptive summaries **rescue medication and healthcare use for AECOPD** will be provided.

6.1.6. Analysis of selected biomarkers

The analysis will be performed on the ATP cohort for efficacy. If the percentage of vaccinated subjects excluded from the ATP cohort for efficacy is at least 10%, a second analysis will be performed on the TVC to complement the first analysis.

For each type of timepoint (Pre, Exacerbation visit, Day 450) during which blood samples are collected for analysis of biomarkers, descriptive statistics will be tabulated for each selected biomarker.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

6.1.7. Analysis of sputum

The analysis will be performed:

- By aggregating all laboratory sites where the sputum was tested
- By laboratory site in order to evaluate if differences exist in the way of testing sputum
- By country (aggregating all laboratory sites by country).

6.1.7.1. Sputum sample collection and quality

The proportion of subjects (at each stable and exacerbation visit) who provided a sputum sample will be computed; overall and by the sample method (spontaneous or induced).

The proportion of sputum of good quality will be computed.

The proportion of subjects with induced sputum will be tabulated by COPD status at enrolment and by severity of the AECOPD.

The distribution of the time (in days) between the start of exacerbation and the day of the sputum sample collection will be presented.

6.1.7.2. Sputum bacterial/viral results from culture or from PCR

The distribution of the proportion of sputum samples positive given the number of days between start of exacerbation and sample collection will be presented.

The proportion of sputum samples obtained at each - stable and exacerbation - visit and positive for specific bacterial/viral pathogens (overall and by bacterial culture, by bacterial PCR, or by viral species) will be computed. The average infection rate by type of visit will also be calculated.

The bacterial presence in culture versus the bacterial presence in PCR detected in sputum will be tabulated.

The viral infections versus the bacterial presence in PCR detected in sputum will be tabulated.

The incidence of exacerbations and of exacerbations containing Hi, Mcat, Sp, and all bacterial and/or virus that are planned to be tested by the laboratory will be presented.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

For specific bacterial pathogens, the quantitative assessment of PCR results will be provided overall, for sputum samples positive in culture, or negative in culture of the specific pathogen.

The same analyses should be performed by taking into account only the sputum that has a good quality.

6.1.7.3. Impact of previous antibiotic administration on sputum bacterial results

The impact of antibiotic administration before sputum collection on the bactericidal culture or PCR results will be evaluated, for stable (scheduled) visits, exacerbation visits, or any visit.

6.1.7.4. Impact of induction of sample on sputum bacterial results

The impact of sputum induction on the bactericidal culture or PCR results will be evaluated, for stable (scheduled) visits, exacerbation visits, or any visit.

6.1.7.5. Impact of previous occurrence of Hi bacterial pathogen in sputum on AECOPDs

The impact of the occurrence or presence of respiratory pathogenic bacteria detected in sputum at the exacerbation visit given the presence of the bacteria at previous visits on the number of exacerbations will also be studied. In order to do this, 2 types of AECOPDs will be distinguished:

- AECOPDs in which Hi is present and was not present any previous visit
- AECOPDs in which Hi is present and was not present in the previous stable visit;

7. STATISTICAL CALCULATIONS

7.1. Derived and transformed data

The study groups will be defined by treatment actually administered at Dose 1.

7.1.1. Demography

- For a given subject and a given demographic variable, missing measurement will not be replaced.
- Age: Age at the reference activity, computed as the number of units between the date of birth and the reference activity In case of partial dates of any of these 2 dates:
 - 15th of month, If only the day is missing

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

- 30th of June, if day and months are missing

7.1.2. Safety

- For solicited symptoms, the analysis will exclude subjects with missing or un-evaluable measurements (*i.e.* total analysis of solicited symptoms will include all vaccinated subjects with documented solicited symptom sheets).
- For the unsolicited symptoms and concomitant medications/ products/ vaccinations, subjects who miss reporting symptoms/concomitant medications/ products/ vaccinations will be treated as subjects without unsolicited symptoms or concomitant medications/ products/ vaccinations, respectively.
- Note that for all tables described in this section, the way the percentage of subjects will be derived will depend on the event analysed (see table below for details). As a result, the N value will differ from one table to another.

Event	N used for deriving % per subject
Concomitant vaccination	All subjects with study vaccine administered
Solicited general symptom	All subjects with at least one solicited general symptom documented as either present or absent (<i>i.e.</i> symptom screen completed)
Solicited local symptom	All subjects with at least one solicited local symptom documented as either present or absent (<i>i.e.</i> symptom screen completed)
Unsolicited symptom	All subjects with study vaccine administered
Concomitant medication	All subjects with study vaccine administered

7.1.3. Immunogenicity

- For a given subject and the analysis of a given immunogenicity measurement, missing or un-evaluable measurements will not be replaced.
- A seronegative subject is defined as a subject whose antibody concentration is below the assay cut-off value.
- A seropositive subject is defined as a subject whose antibody concentration is greater than or equal to the assay cut-off value.
- Antibody concentrations below the assay cut-off will be given an arbitrary value of half the assay cut-off for the purpose of geometric mean concentration (GMC) calculation.
- Calculation of the GMCs will be performed by taking the anti-logarithm in base 10 (anti-log10) of the mean of the log10 concentration transformations.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

- The assay cut-off is the value under which there is no quantifiable result for which no accuracy and precision can be reported. For an assay with a specific 'cut_off' , numerical immuno result is derived from a character field (rawres):
 - If rawres is 'NEG' or '-' or '(-)', numeric result= cutt_off/2,
 - if rawres is 'POS' or '+' or '(+)', numeric result = cut_off,
 - if rawres is '< value' and value<=cut_off, numeric result =cut_off/2,
 - if rawres is '< value' and value>cut_off, numeric result =value,
 - if rawres is '> value' and value<cut_off, numeric result =cut_off/2,
 - if rawres is '> value' and value>=cut_off, numeric result =value,
 - if rawres is '<= value' or '>= value' and value<cut_off, numeric result =cut_off/2,
 - if rawres is '<= value' or '>= value' and value>=cut_off, numeric result =value,
 - if rawres is a value < cut_off, numeric result = cut_off/2,
 - if rawres is a value >= cut_off, numeric result = rawres,
 - else numeric result is left blank.
- All CI computed will be two-sided 95% CI.

Impact of the investigational vaccine on AECOPD

- For a given subject and the analysis of a given laboratory assay, missing or un-evaluatable measurements will not be replaced.
- Subjects who did not have any sputum collected during at least 1 AECOPD (if at least 1 AECOPD occurred) will not be taken into account for the NTHI-associated AECOPD analyses.
- For summaries of the number of AECOPD occurring during the 1 year follow-up period starting 1 month post-Dose 2, an exacerbation rate will be calculated for each subject. The number of exacerbations during the 1-year follow-up period will be imputed for subjects withdrawing from the study to provide an estimate of the number of exacerbations over the follow-up period. This calculation will only be performed for purposes of reporting summary statistics for the rate of exacerbations during the 1 year follow-up period since the modelling of exacerbations takes into account the number of exacerbations and the time of follow-up for each subject. The calculation of exacerbation rate will be based on follow-up period intervals to avoid obtaining high imputed rates if a subject withdrew very early in the follow-up period after experiencing an exacerbation. Since treatment courses for moderate/severe exacerbations are $\leq 2 - 4$ weeks when appropriate, calculated numbers of exacerbations for subjects withdrawing from the study will be based on 4-week intervals of the follow-up period. For subjects followed for less than 1 year, the number of exacerbations during the 1-year follow-up period will be calculated by multiplying the number of exacerbations experienced by the subject by 13 and dividing by the number of 4-week periods the subject was followed up:

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

Number of exacerbations per year = Number of exacerbations * 13 /Number of 4-week treatment period intervals

7.1.4. NTHi associated AECOPD

An NTHi associated to AECOPD is not yet well defined at the time of the SAP is written. Here is the first approach:

- An AECOPD for which NTHi was observed: Hi confirmed first by culture and then by PCR.

7.1.5. HRQOL

The CAT index will be derived as the sum of the ratings recorded for each of the eight individual items. Each of these items has 6 possible scores (0, 1, 2, 3, 4 or 5), leading to a range of 0 to 40 for CAT score.

7.1.6. EXACT-PRO derived AECOPD definitions and calculation rules

Start of an EXACT-PRO signal for exacerbation (as in EXACT-PRO user Manual and in Makay, et al., 2014):

1. *An EXACT-PRO exacerbation is defined as a 12 point increase above baseline for 2 consecutive days, or a nine point increase for 3 consecutive days.*
2. *The baseline score corresponds to the average EXACT score during a baseline period defined as days 14 to 8 preceding exacerbation onset*
3. *Exacerbation onset will be taken as the first day in which at least two symptoms (at least 1 major) are present and have not been discounted (see above).*
4. *Only symptom data collected during the study period will be used to define the start of an exacerbation.*

End/recovering of an EXACT-PRO signal of an exacerbation:

1. *The end/recovering of an EXACT-PRO signal of an exacerbation corresponds to the day at the EXACT-PRO score return to its baseline value.*
2. *The length of a signal of an exacerbation will be at least of 2 days. So rule 1 will only apply from the third day onwards.*

Missing data:

Days with missing data will not be ignored when defining start of a signal: e.g., if a patient has two qualifying symptoms followed by a missing day, if the next day with data has two qualifying symptoms this will not be defined as an exacerbation.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

7.1.7. “e-diary derived AECOPD” definitions and calculation rules

The definition of an AECOPD case definition from section 5.3.1.2 of the Protocol is what we will call here an “e-diary derived (AECOPD) signal”. In this paragraph, only the derivation of occurrence and length of e-diary signals is described. Derivation of occurrence and length of AECOPDs is described in Section 7.1.11.

Start of an e-diary signal for exacerbation:

1. An exacerbation is defined as a worsening of at least 2 symptoms for ≥ 2 consecutive days where at least one of the symptoms is major.
2. The same two symptoms do not have to be present on both days as long as at least one major symptom is present on both days.
3. Any symptom that has occurred consecutively for 5 days or more prior to a potential exacerbation should be discounted from the symptoms that define an exacerbation.
4. Exacerbation onset will be taken as the first day in which at least two symptoms (at least 1 major) are present and have not been discounted (see above).
5. Only symptom data collected during the study period will be used to define the start of an exacerbation.

7.1.8. End/Duration of an e-diary signal of an exacerbation:

1. The end/recovering of an e-diary signal of an exacerbation corresponds to the day at which the 3-day moving average of the daily symptom count score (average between the day the day before and the day after) to return to baseline symptom count score (the mean daily symptom count score over days 14 to 8 preceding exacerbation onset). For a specific day, the daily symptom count score is defined as the sum of the symptom scores experienced in this day (a major symptom accounts for 2 points, a minor symptom accounts for 1 point, as in Trappenburg *et al.*, 2011, but contrary to Seemungal *et al.* 2000 and Mackay *et al.* 2014, in which all symptoms accounted to 1). Note that the symptoms that have occurred consecutively for 5 days or more prior to a potential exacerbation and that are discarded for these exacerbations are not discarded from the symptoms counts.
2. The length of an e-diary signal of an exacerbation will be at least of 2 days. So rule 1 will only apply from the third day onwards.
3. The baseline symptom count score is calculated as the average count score of the observed days in the period used to calculate the baseline symptom count score. For example, if an exacerbation occurs from day 10 onwards, the

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

baseline period will be the average between day1 and day2 (day10 - 8 days= day2, day10 - 9 days = day1, Day1 being the first study day). For exacerbations before day 9 but after day 2, the baseline symptom score will be considered to be day1

4. If the first e-diary signal of an exacerbation is occurring before day 3: duration will be counted until the day in which there is at most 2 majors or 1 major and one minor symptom.
5. Recovery time is calculated as the number of days from exacerbation e-diary signal onset (included) to exacerbation e-diary signal end (not included).

7.1.9. Time between 2 e-diary signals for exacerbations/evaluation of baseline symptom score:

The minimum duration between 2 e-diary signals with distinct baseline value is 15 days, which is the minimal period for which a new full exacerbation-free baseline can be calculated for the next e-diary episode.

If a new worsening of symptoms occurs between 6 and 14 days of a previous exacerbation, the baseline period used will be the same baseline as the one of the original exacerbation. Therefore, it will last until the daily symptom count score has gone below the baseline symptom score of the previous event.

7.1.10. Missing data

1. Days with missing data will not be ignored when defining start of an e-diary signal: e.g., if a patient has two qualifying symptoms followed by a missing day, if the next day with data has two qualifying symptoms this will not be defined as an exacerbation. Hence, the start of an e-diary signal of an exacerbation will require 2 consecutive non-missing days of worsening. If there is partially recorded symptom data at a visit, however, only the observed values will be considered but the day will not be discarded.
2. Days with missing data will be ignored when defining end of an e-diary signal of an exacerbation: e.g., if a patient has two days missing in the baseline period and one day missing in the 3- day moving average, both baseline and 3-day moving average values will be calculated using observed data only. Similar rules will apply when there is partially recorded symptom data at a visit (and daily average values will be calculated among the observed values).
3. All available data will be used for each patient.
4. As a general rule, the baseline symptom count score, and 3 day moving averages are calculated as the average count score of the observed days in the baseline

Study alias & e-track number(s): NTHi-004 (200157)

period. For example, if an **e-diary** signal occurs from day 10 onwards, the baseline period will be the average between day1 and day2 (day1 being the first study day). If day 13 is missing, the 3 day moving average at day 14 will only take into account day 14 and day 15.

5. If there are greater than 2 consecutive days of missing data after an **e-diary** signal not resolved, the **e-diary** signal will not be considered resolved until the next day with observed data if the corresponding moving average is below the baseline value. In other words, if the 3-day average daily score is missing, the day will be counted in the duration of the episode.

7.1.11. AECOPDs definitions and calculation rules

7.1.11.1. AECOPD definition (from clinical evaluation)

Due to the sensitivity and specificity of the e-diary device, false alerts may occur, and exacerbations resulting in site visits may also occur without e-diary alert. Therefore, only unscheduled AECOPD visits with clinical confirmation of AECOPD (with or without associated signals) will be considered as AECOPD events in the primary analysis.

In addition, missed AECOPD visits may have occurred and can be:

- a signal from the e-diary (PHT) without a site visit but that cannot be discounted as a possible exacerbation signal after reconciliation,

and/or

- an exacerbation reported retrospectively in the Safety database or medical records for which there has not been in the same period of retrospective reporting a corresponding exacerbation visit declared.

7.1.11.2. Characterization of inconsistencies between morning e-diary signals and exacerbation site visits

The reconciliation between morning e-diary (PHT) signals and site visits for exacerbation will aim at characterizing the following 2 cases:

Case A : an e-diary alert with absent/delayed exacerbation visit (i.e. no exacerbation visit within the calculated length of the e-diary episode)

Case B : a exacerbation visit with absent/delayed e-diary exacerbation alert (i.e. RDE exacerbation not within e-diary exacerbation period)

The different reasons of inconsistencies of Case A will be categorized by the site after review according to the following explanation:

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

- 1) *Not an exacerbation after further evaluation* : eg. when a patient does not feel well but this is due to increased exercise, or patient does not feel well but this is not clinically significant.
- 2) *Ongoing exacerbation* : e.g., previous exacerbation is not terminated according to the clinician.
- 3) *RDE or other technical issues*: e.g., due to technical issues e-diary alerts were not sent and when the site was aware of it was too late to determine whether this was an exacerbation or not.
- 4) *Other reasons: specify* : e.g. subject was on holidays / subject unable to visit the site / delayed RDE visit/early RDE entry.

The different reasons of inconsistencies of Case B will be categorized by the site after review according to the following explanation:

- 5) *Spontaneous visit clinically deemed exacerbating* i.e. exacerbation visit took place before the e-diary alert (spontaneous visit, visit triggered by GP / Patient call asked to be seen/ felt too unwell to complete e-diary abrupt onset)
- 6) *Planned monthly visit clinically deemed exacerbating*
- 7) *e-diary diary card recording issue or other technical issues* : e.g. late e-diary alerts, subject missed or did not correctly record diary card, broken log pad device.
- 8) *Other reasons: specify*

For a Case A, no laboratory data has been collected, but if the answer about the inconsistency is 3) or sometimes 4) (depending on the specification), it will qualify as AECOPD. All Cases B will be considered to be an AECOPD if the AECOPD is confirmed by the site.

7.1.12. Grading of severity of an exacerbation

Severity of exacerbations is defined as (Protocol Section 5.3.1.2):

- Mild: Worsening symptoms of COPD that are self-managed by the patient.
- Moderate: Worsening symptoms of COPD that require treatment with oral corticosteroids and/or antibiotics.
- Severe: Worsening symptoms of COPD that require treatment with in-patient hospitalisation or home care intervention.

Severity of the exacerbation will not be derived but taken directly from the information entered in the CRF.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

7.1.13. Sputum Quality

The quality of sputum is defined as follow:

- < 10 squamous epithelial cells/field: sample of good quality
- 10-25 squamous epithelial cells/field with significant numbers of cells derived from the lower respiratory tractus (ciliary epithelial cells, bronchial cells, round cells): sample of moderate quality
- > 25 squamous epithelial cells/field: sample of bad quality

7.2. Data presentation description

The following decimal description will be used for the demography and safety/reactogenicity.

Display Table	Parameters	Number of decimal digits
All summaries	% of count, including LL & UL of CI	1
All summaries	% of difference, including LL & UL of CI	2
Demographic characteristics	Mean, median age	1
Demographic characteristics	SD (age)	2
Immunogenicity	GMT ratio	2

The following data presentation description will be used for the immunogenicity analyses.

Component	GMC/T	Assay		Assay cut-off	Number of decimal digits		Tick marks on RCC's	
		method	Unit		GMC/T	GMC/T ratio		
Anti-PD	GMC	ELISA	EL.U/mL	153	1	2	1	10000
Anti-PE	GMC	ELISA	EL.U/mL	8	1	2	1	10000
Anti-PilA	GMC	ELISA	EL.U/mL	7	1	2	1	10000

8. CONDUCT OF ANALYSES

8.1. Sequence of analyses

The analyses will be performed stepwise:

- The analysis of data up to Visit 6 (Day 90) will be performed in a first step. This analysis will include:

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

- The final analysis of immunogenicity and solicited AEs post-Dose 1 and post-Dose 2.
- The assessment of unsolicited AEs, SAEs and pIMDs up to 30 days post-Dose 1 and post-Dose 2 on as cleaned as possible data.

This analysis will be documented in a statistical report. At this point, the GSK statistician of the project (or delegates) will be unblinded (*i.e.* will have access to the individual subject treatment assignments).

- The analyses of the impact of the investigational vaccine on AECOPD will be performed in parallel with the first step. This will include:
 - the analyses of fresh sputum samples (culture) obtained up to the data lock point of the interim analysis.
 - The analyses of AECOPD obtained up to the data lock point of the interim analysis.
- The final analysis of data up to study conclusion will be performed in a second step, once those data will be available and cleaned. This analysis will include:
 - The final analysis of immunogenicity at Day 270 and Day 450.
 - SAEs and pIMDs up to Day 450 on cleaned data.
 - Impact of the investigational vaccine on AECOPD, HRQOL, lung function, exercise capacity and biomarkers.

In addition, all previous analyses will be re-produced based on cleaned data at this point.

Individual listings will only be provided at this stage. A final study report containing data from the entire study will be written and will be made available to the investigators.

Description	Analysis ID (SDD & CARS sub-folder)
Analysis up to Visit 6 (Day 90)	E1_06
Final analysis (up to day 450)	E1_01

8.2. Statistical considerations for interim analyses

Not applicable.

Statistical Analysis Plan



Study alias & e-track number(s): NTHi-004 (200157)

9. CHANGES FROM PLANNED ANALYSES

As during the study, some subjects have completed wrong diary cards (paediatric ones instead of adults), a sensitivity analysis will be done on subjects that have completed the correct diary cards to be sure that they were no impact on results.

10. REFERENCES

The exact 95% CIs for a proportion within a group will be calculated from Proc StatXact [Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of binomial. *Biometrika*. 1934;26:404-413].