Statistical Analysis Plan

Study alias & e-track number(s): EPI-STREP-064 BOD ES (115813)

Detailed Title:	A multi-centre, hospital-based, cross-sectional epidemiology study to identify and characterise bacteria in the lower airways of children aged ≥ 6 months to < 6 years with suspected chronic lower respiratory tract infections (LRTIs) in Spain.
SAP version	Version 1
SAP date	27-NOV-2015
Scope:	All data pertaining to the above study.
Co-ordinating author:	PPD
Other author(s):	PPD
Approved by:	PPD, Lead Statistician PPD, Senior Manager, Epidemiology PPD, Scientific writer

TABLE OF CONTENTS

PAGE

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.

LIST OF ABBREVIATIONS

AE	Adverse event
ATP	According-To-Protocol
BAL	Bronchoalveolar Lavage
CI	Confidence Interval
CRF	Case Report Form
CRP	C-Reactive Protein
CTRS	Clinical Trial Registry
ESR	Erythrocyte Sedimentation Rate
GSK	GlaxoSmithKline
H. influenzae	*Haemophilus influenzae*
LL	Lower Limit of the confidence interval
LRTIs	Lower Respiratory Tract Infections
M. catarrhalis	*Moraxella catarrhalis*
MedDRA	Medical Dictionary for Regulatory Activities
MIC	Minimum Inhibitory Concentration
N.A.	Not Applicable
NP	Nasopharyngeal
SAE	Serious adverse event
SAP	Statistical Analysis Plan
S. pneumoniae	*Streptococcus pneumoniae*
SR	Study Report
SYN	Synopsis
TFL	Tables Figures and Listing template annexed to SAP
UL	Upper Limit of the confidence interval

WBC	White Blood Cell

1. DOCUMENT HISTORY

Date	Description	Protocol Version
27-NOV-2015	Version 1	Amendment 2 Final: 05 OCT2015

2. STUDY DESIGN

- Type of design: Epidemiological, interventional, multi-centre, hospital-based, cross-sectional study with prospective subject recruitment in Spain.
- Study population: Subjects aged ≥ 6 months to < 6 years visiting the study hospital with suspected chronic LRTI and for whom a BAL sample will be collected as per the recommendation of the clinician.

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

Identifier	Sub-group order in tables	Sub-group label in tables	Sub-group definition for footnote
Age	1	6-11M	6-11 months old subjects
	2	12-23M	12-23 months old subjects
	3	24-35M	24-35 months old subjects
	4	36-47M	36-47 months old subjects
	5	48-59M	48-59 months old subjects
	6	60-71M	60-71 months old subjects
Age2*	1	6-23M	6-23 months old subjects
	2	24-47M	24-47 months old subjects
	3	48-71M	48-71months old subjects
Gender	1	F	Female
	2	M	Male
Pneumococcal vaccination history	1	Yes	Subjects received at least two doses or only one dose after one year of age of pneumococcal vaccination
	2	No	Subjects did not receive any dose or received only one dose within the first year of life of pneumococcal vaccination
	3	Unknown	History of pneumococcal vaccination is unknown
S. pneumoniae	1	*S. pneumoniae* +	Episode with *S. pneumoniae* positive by culture

status	2	*S. pneumoniae* -	Episode with *S. pneumoniae* negative by culture
H. influenzae status	1	*H. influenzae* +	Episode with *H. influenzae* positive by culture
	2	*H. influenzae* -	Episode with *H. influenzae* negative by culture
M. catarrhalis status	1	*M. catarrhalis* +	Episode with M. *catarrhalis* positive by culture
	2	*M. catarrhalis* -	Episode with *M. catarrhalis* negative by culture

* The age will be grouped into the following classes 6-11, 12-23, 24-35, 36-47, 48-59 and 60-71 months; however, if there are too few subjects in any of the categories, the classification of age2 will be used instead for subgroup analysis by age.

3. OBJECTIVES

3.1. Primary objective

1) To characterise the bacterial aetiology of BAL fluid samples in subjects visiting the hospital with suspected chronic LRTIs.

Refer to Section 4.1 for the definition of the primary endpoint.

3.2. Secondary objectives

In a hospital setting and among subjects with suspected chronic LRTIs:

1) To describe the bacterial load of *H. influenzae*, *S. pneumoniae* and *M. catarrhalis* detected by quantitative culture and by molecular techniques (PCR) in the BAL fluid.
2) To describe the presence of other bacterial pathogens detected by qualitative culture in the BAL fluid.
3) To describe the bacterial load of *H. influenzae*, *S. pneumoniae* and *M. catarrhalis* detected by quantitative culture and by molecular techniques (PCR) in the nasopharyngeal swab samples.
4) To describe the presence of other bacterial pathogens detected by qualitative culture in the nasopharyngeal swab samples.
5) To describe colonisation of the upper airways by characterising *H. influenzae*, *S. pneumoniae, M. catarrhalis* and other pathogens in nasopharyngeal swab samples.
6) To determine the serotypes of *H. influenzae* and *S. pneumoniae* identified from BAL fluid samples.
7) To determine the serotypes of *H. influenzae* and *S. pneumoniae* identified from nasopharyngeal swab samples.
8) To determine the antibiotic susceptibility profile for *H. influenzae, S. pneumoniae* and *M. catarrhalis* identified from BAL fluid and nasopharyngeal swab samples.
9) To describe, according to microbiological results observed:
   - Age and gender
   - Clinical symptoms and radiological evidence
   - Pneumococcal conjugate vaccine, *H. influenzae* type b vaccine and influenza vaccine status
   - Medical history and co-morbidities
   - Information on feeding and day care practice, environmental exposure and smoking environment

- History of antibiotic use in the past six months as well as other treatments for chronic lower respiratory disease
- Laboratory results [including white blood cell counts, C-Reactive Protein (CRP) level and erythrocyte sedimentation rate, if available].

Refer to Section 4.2 for the definition of the secondary endpoints.

4. ENDPOINTS

4.1. Primary endpoint

1) Occurrence of *H. influenzae, S. pneumoniae, M. catarrhalis* and other bacteria in BAL fluid samples of subjects aged ≥ 6 months to < 6 years visiting the hospital with suspected chronic LRTIs and an indication for BAL
   - *H. influenzae, S. pneumoniae* and *M. catarrhalis* confirmed by bacterial load >10^4 cfu/mL if present alone or 10^5 cfu/mL if present as co-infection.

4.2. Secondary endpoints

In a hospital setting and among subjects aged ≥ 6 months to < 6 years with suspected chronic LRTIs and an indication for BAL:

1) Bacterial load of *H. influenzae*, *S. pneumoniae* and *M. catarrhalis* in the BAL fluid as determined by quantitative culture and PCR.
2) Description of other bacterial pathogens in the BAL fluid as determined by qualitative culture.
3) Bacterial load of *H. influenzae*, *S. pneumoniae* and *M. catarrhalis* in nasopharyngeal swab samples as determined by quantitative culture and PCR.
4) Description of other bacterial pathogens in nasopharyngeal swab samples as determined by qualitative culture.
5) Occurrence of *H. influenzae*, *S. pneumoniae, M. catarrhalis* and other pathogens detected by culture in the nasopharyngeal swab samples.
6) Occurrence of *H. influenzae* and *S. pneumoniae* serotypes identified from BAL fluid samples and nasopharyngeal swab samples.
7) Occurrence of *H. influenzae, S. pneumoniae* and *M. catarrhalis* minimum inhibitory concentration (MIC) values identified from BAL fluid and nasopharyngeal swab samples.
8) Demographic characteristics (age, gender, etc), clinical characteristics (including medical history, LRTI symptoms, treatment history, vaccination status, etc),

radiological evidence and laboratory results (including white blood cell counts, CRP level and erythrocyte sedimentation rate, if available).

5. STUDY POPULATION

5.1. Screened cohort

The screened cohort will include all subjects who visit the hospital with suspected chronic LRTIs and have an indication for BAL. All the information (anonymised) for these subjects will be collected in the aggregated logbook.

5.2. Total cohort

The total cohort will include all enrolled subjects.

5.3. According-To-Protocol cohort

The ATP cohort will include all evaluable subjects (i.e. those meeting all eligibility criteria and complying with the procedures defined in the protocol during the study) with BAL fluid sampling and/ or nasopharyngeal swab sampling results available.

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
Total cohort	NA	MA
ATP cohort	900, 2010, 2040, 2050, 2060, 2100, 2120, 2500	MA

6. STATISTICAL METHODS

All statistical analyses will be done using SAS version 9.2 or later.

Continuous variables will be described using the mean, standard deviation, median, minimum and maximum. Categorical variables will be described using frequency tables (absolute numbers and percentages for levels of each variable, excluding missing values). For continuous and categorical variables, the number of missing observations will be recorded.

All the analyses will be performed on the ATP cohort. If the percentage of subjects from the Total effective cohort with serological results excluded from the ATP cohort is more

than 5%, a second analysis of the demography and the primary objective based on the Total effective cohort will be performed to support the ATP analysis.

6.1. Analysis of demographics/ baseline characteristics

Demographic, clinical and microbiological characteristics of subjects will be described using percentages for categorical variables and mean (standard deviation) or median (min-max) for continuous variables. The same variables will be presented descriptively, overall and by microbiological results.

6.1.1. Demographic characteristics

The number of screened and enrolled subjects as well as the number excluded from ATP analyses will be presented.

Demographic characteristics (age in months, gender, weight in kg, height in cm and BMI categories, such as extreme obesity, obesity, overweight normal, thinness and marked thinness (details see section 7.1))will be summarised using descriptive statistics (for all non-missing observations) overall and by center. Height and weight will also be tabulated by age groups.

6.1.2. Clinical characteristics

The following tables will be generated for the following endpoints:

- Chronic LRTI symptoms
    - The bronchoscopy details will be presented: the percentage of subjects for each indication (cough, wheezing, pathologic auscultation, infiltrates/atelectasis) as well as for each fibro-bronchoscopy indication will be tabulated overall and by center.
    - The percentage of subjects presenting respiratory or otitis media pre-existing conditions will be tabulated.
    - Number of previous episodes reported by subject within the last 12 months will be tabulated as well as the description of current cough.
- Medical history:
    - The percentage of subjects reporting any disease/conditions in the general medical history will be tabulated overall and by center. Summary of listed conditions will be tabulated overall and by center.

- Vaccination history (vaccination against Pneumococcal, *Haemophilus influenzae* B and Influenza) will be tabulated and percentage of pneumococcal vaccine administered (prevenar7, prevenar13, pneumovax or synflorix) presented.

- Feeding practices and day care practice:
  - If the child has ever been breastfed or not, the number of other children living in the same household for more than 3 days a week, and if the child is at a day care center and/or school at least 2 days a week or not, will be presented in a frequency table;

- Environmental exposure and smoking environment
  - The percentage of children living within 1 km of major pollution source and the source of the pollution (Major roadway, Highway, Area where trucks/vehicles idle, major industry area with smokestacks and other category) will be tabulated.
  - The percentage of children exposed more than 3 days a week to indoor cigarette smoke will also be tabulated.

6.1.3. Antibiotic therapy and other treatments

The percentage of subjects that took antibiotics within the 6 months prior to the visit will be tabulated. Trade/Generic name of Antibiotics will be tabulated. The same tables will be presented for other medications taken for this condition within the past 6 months.

6.1.4. Physical examination

Vital signs such as temperature, heart rate, blood pressure and respiratory rate will be presented.

6.1.5. Radiological results

The percentage of subjects that had a Chest X-Ray (CXR) done will be presented.

The findings from the CXR will be presented as frequency table stratified overall and by center. The percentage of subjects with an abnormal result from the CXR which was relevant to the decision to collect BAL sample will be presented.

6.1.6. Laboratory results: WBC, CRP, ESR, microbiological culture, antibiotic resistance test

The following results of laboratory tests in subjects for which they are available will be presented:

- Mean, Median and range of White Cell count level along with standard deviation will be presented overall and by center for subjects for whom blood sample is available and white cell count analyzed. The same will also be presented by age group.

- Mean, Median and range of CRP level along with standard deviation will be presented for subjects for whom blood sample is available and CRP test was done. The same will also be presented by age group.

- CRP levels (see Table 1 below) will also be presented in groups overall and stratified by age group.

Table 1 CRP categories for the analysis

CRP	1	<40 mg/L	CRP <40 mg/L
	2	>=40 - <80 mg/L	40 mg/L <=CRP <80 mg/L
	3	>=80 - <120 mg/L	80 mg/L <=CRP <120 mg/L
	4	>=120 mg/L	CRP >=120 mg/L

- Mean, Median and range of Erythrocyte sedimentation rate (ESR) level along with standard deviation will be presented overall and by center for subjects for whom blood sample is available and ESR analyzed. The same will also be presented by age group

6.1.7. BAL fluid samples and nasopharyngeal swab

The percentage of subjects for whom BAL fluid sample was taken will be presented overall and by center. The same will be tabulated for nasopharyngeal swab samples taken.

6.2. Analysis of primary objective

The primary analysis will be based on the ATP cohort. If the percentage of subjects from the Total effective cohort with serological results excluded from the ATP cohort is more than 5%, a second analysis based on the Total effective cohort will be performed to support the ATP analysis.

The proportions and associated exact 2-sided 95% confidence intervals (CI) of *H. influenzae, S. pneumoniae, M. catarrhalis* and other bacteria confirmed by culture in BAL samples (*H. influenzae, S. pneumoniae* and *M. catarrhalis* confirmed by bacterial load >10^4 cfu/mL if present alone or 10^5 cfu/mL if present as co-infection) will be computed and presented for all study subjects.

The denominators will be the number of tested subjects with BAL sample results available.

6.3. Analysis of secondary objectives

All tables related to secondary objectives 1 to 8 will be presented overall as well as by age class, gender, pneumococcal vaccination status. The age will be grouped into the following classes 6-11, 12-23, 24-35, 36-47, 48-59 and 60-71 months; however, if there are too few subjects in any of the categories, the following classification will be used of 6-23, 24-47, 48-71 months instead.

In a hospital setting and among subjects with suspected chronic LRTIs:

1) To describe the bacterial load of *H. influenzae*, *S. pneumoniae* and *M. catarrhalis* detected by quantitative culture and by molecular techniques (PCR) in the BAL fluid.

The distribution (with number of non-missing observations, geometric mean, standard deviation, median, minimum and maximum and number of missing observations) of the bacterial load of *H. influenzae*, *S. pneumoniae, M. catarrhalis* in the BAL fluid will be described according to the testing method (culture and PCR).

2) To describe the presence of other bacterial pathogens detected by qualitative culture in the BAL fluid.

The presence (number and percentage) of other bacterial isolates detected in BAL fluid will be presented in frequency tables.

3) To describe the bacterial load of *H. influenzae*, *S. pneumoniae* and *M. catarrhalis* detected by quantitative culture and by molecular techniques (PCR) in the nasopharyngeal swab samples.

The distribution (with number of non-missing observations, geometric mean, standard deviation, median, minimum and maximum and number of missing observations) of the bacterial load of *H. influenzae*, *S. pneumoniae, M. catarrhalis* in the nasopharyngeal swab samples will be described according to the testing method (culture and PCR).

4) To describe the presence of other bacterial pathogens detected by qualitative culture in the nasopharyngeal swab samples.

The presence (number and percentage) of other bacterial isolates detected in the nasopharyngeal swab samples will be presented in frequency tables.

5) To describe colonisation of the upper airways by characterising *H. influenzae*, *S. pneumoniae, M. catarrhalis* and other pathogens in nasopharyngeal swab samples.

The occurrence of *H. influenzae*, *S. pneumoniae, M. catarrhalis* and other pathogens detected by culture in the nasopharyngeal swab samples will be presented:

The proportions and associated exact 2-sided 95% confidence intervals (CI) of *H. influenzae, S. pneumoniae, M. catarrhalis* and other pathogens confirmed by culture in the nasopharyngeal swab samples (*H. influenzae, S. pneumoniae* and *M. catarrhalis* confirmed by bacterial load >10^4 cfu/mL if present alone or 10^5 cfu/mL if present as co-infection) will be computed and presented for all study subjects.

The denominators will be the number of tested subjects with nasopharyngeal swab samples results available.

6) To determine the serotypes of *H. influenzae* and *S. pneumoniae* identified from BAL fluid and nasopharyngeal swab samples

The serotype distribution (number and percentage) of *H. influenzae* and *S. pneumoniae* isolates detected in BAL fluid will be presented in frequency tables.

The serotype distribution (number and percentage) of *H. influenzae* and *S. pneumoniae* isolates detected in nasopharyngeal samples will be presented in frequency tables.

7) To determine the antibiotic susceptibility profile for *H. influenzae, S. pneumoniae* and *M. catarrhalis* identified from BAL fluid samples.

For the antibiotic susceptibility, *H. influenzae, S. pneumoniae and M. catarrhalis* isolates (detected by culture of BAL fluid and nasopharyngeal swab)will be classified as susceptible, intermediate or resistant (following Clinical and Laboratory Standards Institute guidelines ***[CLSI, 2015])*** see Table 2 and Table 3 below) and will be tabulated by bacterial pathogen and antimicrobial agent.

The percentage of subjects for whom antibiotic resistance test was performed and the results of the resistance test will be tabulated (Antibiotic sensitivity on BAL samples include penicillin, erythromycin, azithromycin, tetracycline, levofloxacin, trimethoprim/ sulfamethoxazole, amoxicillin/ clavulanate; and additionally a beta-lactamase test for *H. influenzae* and *M. catarrhalis*.).

The ranges for antimicrobial susceptibility for *S. pneumoniae* are the following:

Table 2 MIC ranges for antimicrobial susceptibility for *S. pneumoniae*

	Susceptible (S) (µg/mL)	Intermediate (I) (µg/mL)	Resistant (R) (µg/mL)
Penicilin	≤ 0.06	0.12 – 1.00	≥ 2.0

Amoxicillin / clavulanate	≤ 2/1	4/2	≥ 8/4
Erythromycin	≤ 0.25	0.5	≥ 1
Azithromycin	≤ 0.5	1	≥ 2
Tetracycline	≤ 1	2	≥ 4
Levofloxacin	≤ 2	4	≥ 8
Trimethoprim / sulfamethoxazole	≤ 0.5/9.5	1/19 – 2/38	≥ 4/76

The ranges for antimicrobial susceptibility for *H. influenzae* and *M. catarrhalis* are the following:

Table 3 MIC ranges for antimicrobial susceptibility for *H. influenzae* and *M. catarrhalis*

	Susceptible (S) (µg/mL)	Intermediate (I) (µg/mL)	Resistant (R) (µg/mL)
Penicilin	No CLSI standard – MIC values will be presented		
Amoxicillin / clavulanate	≤ 4/2	n.a.	≥ 8/4
Beta-lactamase (Nitrocefin)	Positive/Negative		
Erythromycin	No CLSI standard – MIC values will be presented		
Azithromycin	≤ 4	Non-susceptible (NS)	
Tetracycline	≤ 2	4	≥ 8
Levofloxacin	≤ 2	Non-susceptible (NS)	
Trimethoprim / sulfamethoxazole	≤ 0.5/9.5	1/19 – 2/38	≥ 4/76

8) To describe, according to microbiological results observed:

- Age, gender, weight, height and BMI categories
- Clinical symptoms and radiological evidence
- Pneumococcal conjugate vaccine, *H. influenzae* type b vaccine and influenza vaccine status
- Medical history and co-morbidities

- Information on feeding and day care practice, environmental exposure and smoking environment
- History of antibiotic use in the past six months as well as other treatments for chronic lower respiratory disease
- Laboratory results [including white blood cell counts, C-Reactive Protein (CRP) level and erythrocyte sedimentation rate, if available].

The demography characteristics (age, gender, weight in kg, height in cm and BMI categories), clinical symptoms and radiological evidence, Pneumococcal conjugate vaccine history status, *H. influenzae* type b vaccine history status and influenza vaccine history status, Medical history, Information on feeding and day care practice, environmental exposure and smoking environment, History of antibiotic use in the past six months as well as other treatments for chronic lower respiratory disease and Laboratory results will be presented by microbiological results of (*S. pneumoniae* positive/negative, *H. influenzae* positive/negative and *M. catarrhalis* positive/negative, co-infection or not co-infection), and by type of sample used (BAL or nasopharyngeal sample) (Details see section 6.1).

7. STATISTICAL CALCULATIONS

7.1. Derived and transformed data

Age at time of enrolment in the study will be computed as the difference between the date of enrolment [date when the ICF was signed by the parent(s)/ LAR(s)] and the date of birth. The age will be expressed in months.

The age will be grouped into the following classes 6-11, 12-23, 24-35, 36-47, 48-59 and 60-71 months; and also another category of 6-23, 24-47, 48-71 months just in case there will be too few subjects in the younger groups of 6-11 or 12-23 months.

BMI will be classified as below:

- Marked thinness: <z-score (-3)
- Thinness: ≥z-score (-3) and <z-score (-2)
- Normal: ≥z-score (-2) and ≤z-score (1)
- Overweight: >z-score (1) and ≤z-score (2)
- Obesity: >z-score (2) and ≤z-score (3)
- Extreme obesity: >z-score (3)

The source for the z-scores for age from 0 to 60 months is available on the WHO website BMI-for-age tables Girls z-scroes
http://www.who.int/childgrowth/standards/b_f_a_tables_z_girls/en/
BMI-for-age tables Boys z-scroes
http://www.who.int/childgrowth/standards/b_f_a_tables_z_boys/en/

The source for the z-scores for age from 61 months to < 6 years old is available on the WHO website of http://www.who.int/growthref/who2007_bmi_for_age/en/, enter the website, go to Tables, then z-scroes: girls; and z-scores:boys (see for more details)

The pneumococcal conjugate vaccination status will defined as:

- Vaccinated, if the subject either received at least two doses or received only one dose after one year of age;
- Unvaccinated, if the subject did not receive any dose or received only one dose within the first year of life;
- Unknown otherwise.

Temperature will be converted to axillary route temperature (if route is rectal or tympanic rectal, then value will be 0.5°C). If this derived value of temperature is ≥ 37.5°C, then the subject is considered to have fever.

7.2. Methods for CIs calculation

- The exact 95% CIs for a proportion within a group will be calculated from Proc StatXact [Clopper, 1934].
- Proc StatXact will be used to derive the standardised asymptotic 95% CI for thegroup difference in proportions [Robert G, 1998, method six]. The standardised asymptotic method used within GSK Biologicals is the method six.

8. CONDUCT OF ANALYSES

8.1. Sequence of analyses

The analyses will be performed in one step, when all data, will be available and cleaned. These analyses and associated individual data will be presented in a final study report.

Description	Analysis ID (SDD & CARS sub-folder)
Final	ANALYSIS_E1_01

8.2. Statistical considerations for interim analyses

No interim analysis is planned for this study.

9. CHANGES FROM PLANNED ANALYSES

N.A

10. REFERENCES

Clinical and Laboratory Standards Institute (CLSI) M100-S25 Performance Standards for Antimicrobial Susceptibility Testing. January 2015.

Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of binomial. Biometrika. 1934;26:404-413

Robert G. Newcombe, interval estimation for the difference between independent proportions: comparison of eleven methods, Statist Med. 1998; 17, 873-890

11. APPENDIX

These pdf files for BMI categories for age by gender will be stored in the same location as SAP in eTMF

- BMI_boys_0_13_zscores.pdf
- BMI_boys_0_2_zscores.pdf
- BMI_boys_2_5_zscores.pdf
- BMI_girls_0_13_zscores.pdf
- BMI_girls_0_2_zscores.pdf
- BMI_girls_2_5_zscores.pdf
- BMIfa_boys_5_19years_z.pdf
- BMIfa_girls_5_19years_z.pdf