Protocol B1851202

CT-45

A Low Intervention Study of the Effectiveness Of 13-Valent Pneumococcal Conjugate Vaccine (PCV13) Against Vaccine Type Pneumococcal Hospitalised Community Acquired Pneumonia (CAP) in Adults 60 Years and Older Using A Test Negative Design Study in A Well-Defined Area of the South of Madrid Region

Statistical Analysis Plan
(SAP)

Version: 3.0
Date: 07 Sep 2023

TABLE OF CONTENTS

LIST OF TABLES

1. VERSION HISTORY

Table 1. Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
1.0/ 27 September 2021	Version 2.0, 13 Apr 2021	N/A	N/A
2.0/ 20 Jan 2023	V3.0, 03 Feb 2022	Protocol Amendment	Updated study design section 2.2 Updated population section 2.2.1 Inclusion of new exploratory objective: RSV PCR testing in saliva samples collected for pneumococcal testing. Added sections 3.3.3, 6.3.9, and 6.3.9.1. Minor wording and grammatical changes
3.0/ 07 Sep 2023	V3.0, 03 Feb 2022	A correction in the analysis population	Updated few things in the Section 4. Analysis Sets (Populations for analysis) Added Section 5.4 Updated the vaccination status in section 6.1.1.1 to specify who should be excluded due to receiving the PCV13 and PPV23 at interval discordant as per the guidelines, and also those participants who received PCV20. Updated the analysis in the section 6.1.1.1 specifying that logistic regression will be used instead of Generalized Estimating Equation (GEE)

2. INTRODUCTION

This statistical analysis plan (SAP) provides the detailed methodology for summary and statistical analyses of the data collected in Study B1851202. This document may modify the plans outlined in the protocol; however, any major modifications of the primary endpoint definition or its analysis will also be reflected in a protocol amendment. Note that any text taken directly from the protocol is *italicized*, with exceptions note where it appears.

This study is designed to evaluate the effectiveness of 13-valent pneumococcal conjugate vaccine (PCV13) in preventing hospitalised vaccine-type (VT) community acquired pneumococcal (CAP) disease among adults aged ≥60 years in Madrid using a prospective study to collect data for test -negative design (TND) analysis, overall and among immunocompetent persons only. Further this study will also measure the proportion of CAP that is associated with Respiratory Syncitial Virus (RSV) infection.

Pneumococcal vaccination status will be based on whether receipt of ⩾1 dose of PCV13 or PPV23 could be confirmed by electronic registries. All available vaccination data will be collected from the registry on subjects and time from vaccination will be incorporated in the analysis phase. Patients will be included regardless of the timing of their vaccination relative to their qualifying pneumonia admission but will be excluded from vaccine effectiveness (VE) and serotype distribution analysis if they received pneumococcal vaccination ⩽30 days before hospitalisation for CAP because of risk of false positive urinary antigen detection (UAD) test.

2.1. Study Objectives

2.1.1. Primary Objectives

- *To determine the effectiveness of PCV13 to prevent hospitalised VT- pneumococcal CAP among adults aged ≥60 years in Madrid using a TND, overall and among immunocompetent persons only.*

2.1.2. Secondary Objectives

- *To describe the distribution of SP serotypes using blood, high-quality respiratory cultures, and a serotype-specific UAD assay among adults 60 years of age and older with CAP.*
- *To determine the proportion of persons with respiratory pneumococcal carriage among adults 60 years of age and older presenting with CAP by testing saliva specimens using both conventional culture and the sensitive molecular method of polymerase chain reaction (PCR) for the detection of two target genes (lytA and piaB).*
- *To describe the proportion of participants with CAP and with S. pneumoniae (SP+CAP) who present with underlying at-risk and high-risk medical conditions.*

- *To describe the frequency and type of antibiotic resistance among SP isolates, overall and by specimen type.*

2.1.3. Exploratory Objectives

- *To estimate incidence rates for hospitalised CAP and SP+CAP (both overall and radiologically confirmed) in the surveillance region within the limitations of the expected under ascertainment of hospital-based screening process and current diagnostic testing.*

- *To compare the effectiveness of PPV23 to that of PCV13 in preventing cases of hospitalised VT- pneumococcal CAP among adults aged ≥60 years using a test-negative study design, overall and among immunocompetent persons only.*

- *To compare pneumococcal serotype results from high-quality respiratory specimens (standard-of-care), UAD1/2 (study procedure), and saliva specimens taken for carriage testing (study procedure).*

- *To estimate incidence rates for hospitalised CAP among adults aged ≥60 years with underlying at-risk and high-risk medical conditions, in aggregate and for individual risk conditions within the limitations of the expected under ascertainment of hospital-based screening process and current diagnostic testing.*

2.1.4. Additional COVID-19 Exploratory Objectives

- *To determine if the frequency of pneumococcal infection differs between hospitalised CAP patients with COVID-19 infection compared to those without COVID-19 infection, overall and stratified by age, sex, and pneumococcal risk status.*

- *Among the subset of hospitalised CAP patients with COVID-19 identified, to determine if the proportion who experience severe clinical outcomes (eg, intensive care unit [ICU] stay, mechanical ventilation, prolonged hospitalisation, and sequelae) differs between those with pneumococcal co-infection compared to those without pneumococcal co-infection, overall and stratified by age, sex, and pneumococcal risk status.*

- *Using a TND to determine the VE of PCV13 against severe COVID-19 clinical outcomes among COVID-19 positive hospitalised CAP patients, overall and among immunocompetent persons only.*

- *To assess if frequency of pneumococcal carriage differs for those with current or past COVID-19 infection, overall and stratified by age, sex, and pneumococcal risk status.*

2.1.5. Additional 2022 RSV Exploratory Objectives

- *To determine the proportion of persons with RSV infection among adults 60 years of age and older presenting with CAP by testing saliva specimens using the PCR method.*

2.2. Study Design

This study is a low interventional, prospective, multicentre, hospital-based study involving adults 60 years of age and older hospitalised with CAP at participating sites. Hospital surveillance data will be used for incidence rates calculations. Vaccine effectiveness will be calculated using a TND. The approximate study duration will be 4 years.

Assuming 15% PCV13 vaccination coverage among controls, 60% VE, 9% VT serotype prevalence, a ratio of 1:10 case and controls and with alpha=0.025 and 80% power, enrolment of 170 cases and 1,692 controls would be required to allow appropriate evaluation of the study objective.

2.2.1 Population

This study will prospectively enroll adults, aged 60 years and over, who are hospitalised with CAP in one of the study hospitals.

2.2.2. Variables

The screening log will collect demographic data, including eligibility of the participants for the study, whether consented, and confirmation of diagnosis by chest rediology. If consent is given then further data would be collected including: medical and vaccination history, smoking and alcohol use as well as details of current illness including all symptoms, treatment, severity score, vital signs and test results. At follow up visits confirmation of CAP or alternative diagnosis, event disposition and vital status will be recorded at 30 and 180 days.

2.2.3. Data sources

In this study, the sources of data will include both CRF and non-CRF. The CRF data include hospital records, including radiology and laboratory results, as well as primary care medical records. The non-CRF sources will be the Screening Log, reference lab, andlocal sources. Local sources will provide the data on the count of population living in the surveillance area, overall, by age group, and with the risk condition(s). The database named SISPAL maintained by the Regional Health System will be used to obtain vaccination histories.

3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

3.1. Primary Endpoint

- *VE will be calculated as 1 minus the OR comparing the odds of having received PCV13 for cases and controls, multiplied by 100%, overall and restricted to immunocompetent subjects only.*

Cases and Controls for VE against PCV13

The cases are the patients with CAP (i.e., enrolled population who have a final diagnosis consistent with CAP) and in whom PCV13 serotypes are identified by any method including UADs, routine culture of blood, pleural fluid, or other high-quality respiratory tract specimen. PCV13 serotypes are as follows: 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A,

19F and 23F. Only sputum isolates from high-quality specimens (Appendix 4) will be considered in assigning case status. Participants with VT serotype who also have a Non-Vaccine Type (NVT) serotype will also be classified as cases.

The following criteria must be met to be a control:

- hospitalized patient with CAP
- met the inclusion/exclusion criteria, and
- PCV13 or related serotype (e.g., 6C) are not identified from any sources

A second model will be developed excluding immunocompromised persons (Appendix 2.4) to evaluate PCV13 VE without those who may not be able to respond to the vaccine.

Cases and Controls for VE against PPV23

The cases are the patients with CAP in whom any one of the 23 serotypes in PPV23: 1, 2, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F, 18C, 19A, 19F, 20, 22F, 23F, and 33F is identified (including those where a NVT serotype is also identified).

The following criteria must be met to be a control for PPV23:

- hospitalized patient with CAP,
- met the inclusion/exclusion criteria, and
- PPV23 are not identified from any source

3.2. Secondary Endpoint(s)

- *Proportion of CAP events for each pneumococcal serotype, PCV-13, and 20-valent pneumococcal conjugate vaccine (PCV20) serotypes among all CAP subjects with UAD1/2 testing (overall) and among all CAP subjects with a pneumococcus identified.*

  The denominators for this endpoint are:
  i) all CAP patients having their UAD 1 or 2 results available,
  ii) all CAP patients with a pneumococcus identified (*Sp*+) from UAD, Binax, blood culture, or high-quality respiratory cultures

  Note: If a participant, for example, is positive for a PCV20 serotype (e.g., 8) and a non-PCV20 serotype (e.g., 17F), the participant will be counted as PCV20 *Sp*+ and not non-PCV20 *Sp*+.

- *Proportion of CAP events where the pneumococcus was identified from saliva by culture or PCR divided by all CAP events where the subject had a valid saliva specimen test result.*

The denominator will be the CAP population with valid saliva specimen test result.

- *Proportion of CAP subjects with underlying at-risk and high-risk medical conditions, overall and restricted to those with pneumococcus isolated from blood, high-quality respiratory cultures, and a serotype-specific urinary antigen detection (UAD) assay.*

  For overall, the denominator will be all CAP population. For participants restricted to pneumococcus, the denominator will be the CAP population in whom a pneumococcus is identified from culture, UAD, or Binax.

- *Proportion of S. pneumoniae isolates with antibiotic resistance identified by standard of care testing, overall and by resistance type.*

The analysis will be conducted among participants who are *Sp*+ by culture among the CAP population for whom the data on antibiotic resistance are available. If the antibiotic resistance information is missing or not done for a participant, then the participant will be excluded from the analysis. For participants whose *Sp* isolates were shipped to the Reference Lab, antibiotic resistance results will be used in the analysis. Participants who are detected as *Sp*+ via culture method but without serotype information ('Not Serotyped') will also be included in the analysis.

Antibiotic resistance will be reported in the eCRF as one of the following categories:

- SUSCEPTIBLE
- RESISTANT
- NOT APPLICABLE
- UNKNOWN

The following antimicrobials will be considered: Penicillin, Amoxicillin, Cefotaxime, Erythromycin, Clindamycin, Tetracycline, and Levofloxacin.

Antibiotic susceptibility profile will be reported in the analysis as follows:

Report in the analysis	**Reported in the eCRF**
Susceptible	Susceptible
Resistant	Resistant
Multidrug resistant	Resistance to penicillin and resistance to two classes of non-ß-lactam antibiotics (i.e. two of the following: Erythromycin, Levofloxacin, Tetracycline)

3.3. Other Endpoints

3.3.1. Exploratory Endpoints

- *Number of CAP events identified by screening log divided by number of persons ≥60 years in surveillance region.*

The denominator will be the number of individuals aged ⩾60 years living in the hospitals' catchment area. The data on the denominator will be obtained from INE (National Statistical Institute).

- *Final adjusted VE point estimate for PPV23 subtracted from final adjusted VE subtracted for PCV13, overall and among immunocompetent persons only.*

Cases and controls for VE against PCV13

The cases are the patients with CAP in whom PCV13 serotypes are identified by any method including UADs, routine culture of blood, pleural fluid, or other high-quality respiratory tract specimen. PCV13 VTs are as follows: 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F and 23F. Only sputum isolates from high-quality specimens (Appendix 4) will be considered in assigning case status. Participants with VT serotype who have also an NVT serotype will also be classified as cases.

The controls are the patients with CAP without a VT serotype (including the CAP events for which the pneumococcus was not identified or pneumococcus was present but the serotype was unknown) that met the inclusion and exclusion criteria. These will serve as test-negative controls. **Participants with 6C serotype identified (VT serotype) will be excluded from the controls.**

The definition of the immunocompromised persons is given in Appendix 2.4.

Cases and controls for VE against PPV23

The cases are the patients with CAP in whom any one of the 23 serotypes in PPV23: 1, 2, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F, 18C, 19A, 19F, 20, 22F, 23F, and 33F is identified.

The controls are the patients with NVT episode of CAP. Participants having both VT and NVT serotypes will not be included in the controls.

- *Percent agreement in serotype results between specimen types for events that have both types available (Three comparisons: 1. high-quality respiratory culture [standard-of-care specimen] versus UAD1/2; 2. UAD1/2 versus saliva carriage sample; 3. High-quality respiratory culture versus saliva carriage sample).*

This analysis will include the participants who have at least two of the following with a valid test result: high quality respiratory culture, UAD1/2, or saliva samples.

- *Number of CAP events identified by screening log multiplied by the proportion of CAP subjects with specific risk condition(s) divided by the estimated number of persons ≥60 years residents in the hospitals' catchement area with the same specific risk condition(s), in aggregate for at-risk and high-risk conditions and for individual risk groups if the size of risk group is adequate for a stable estimate.*

The denominator will be the population living in the hospitals' catchment area with the risk condition(s), which will be obtained from local source(s) depending on availability.

3.3.2. COVID-19 Related Endpoints

- *Percentage of CAP events with pneumococcal infection identified by any means (e.g., UAD, BinaxNOW®, bacterial culture) among CAP events associated with a positive COVID-19 test compared to percentage of CAP events with pneumococcal infection identified by any means (e.g., UAD, BinaxNOW®, bacterial culture) among CAP events associated with a negative COVID-19 test, overall and stratified by age, sex, and pneumococcal risk status.*

The following calculations will be made for overall and stratified by age group, sex, and pneumococcal risk status among COVID-19 subpopulation, i.e. all enrolled population with a valid results from COVID-19 viral pathogen testing.

- *Among CAP subjects with COVID-19 infection, percentage of CAP subjects with severe clinical outcomes (overall and by individual outcome) among those with pneumococcal co-infection identified by any means (e.g., UAD, BinaxNOW®, bacterial culture) compared to percentage of CAP subjects with serious clinical outcomes among those without pneumococcal co-infection identified by any means (e.g., UAD, BinaxNOW®, bacterial culture), overall and stratified by age, sex, and pneumococcal risk status.*

  In this analysis, the patients with COVID-19 will be divided into two groups:
  - those with coinfection with *Sp* identified by any means (e.g., UAD, Binax, bacterial culture).
  - those without coinfection with *Sp* identified by any means (e.g., UAD, Binax, bacterial culture).

The definitions of the severe clinical outcomes are given below:

- Admitted at ICU: admission at ICU, irrespective of duration of ICU stay.
- PaO_2/FiO_2 ratio: PaO_2/FiO_2 ratio<250.
- Respiratory rate: respiratory rate>30 breaths/min.
- Outcome of illness: death.
- Use of mechanical ventilation: any type ventilation (intubation or non-invasive positive pressure).
- Diagnosed as Acute Respiratory Distress Syndrome (ARDS): diagnosis of ARDS.

- Any severe outcome: If the participant falls in any of the categories of severe clinical outcomes as described above, then the participant will fall in this category.

- *VE calculated as 1 minus the OR comparing the odds of having received PCV13 by cases and controls, multiplied by 100%, overall and among immunocompetent persons only, adjusted for confounders. (Cases=COVID-19-positive CAP subjects with severe clinical outcomes; Controls=COVID-19-positive CAP subjects without severe clinical outcomes.*

  This will be a nested case-control study, and the COVID-19 status will be determined from standard-of-care viral testing. The cases are CAP patients with COVID-19 with severe clinical outcomes, and the controls are the CAP patients with COVID-19 without severe clinical outcomes. Several clinical outcomes will be used in the definition of cases and controls as stated below:

  - Patients admitted to the ICU
  - Patients with PaO2/FiO2 ratio <250 mmHg
  - Patients with respiratory rate >30 breaths/min
  - Patients died
  - Patients received mechanical ventilation
  - Patients diagnosed as ARDS
  - Any severe clinical outcome mentioned above

- *Frequency of pneumococcal carriage among subjects with current/recent COVID-19 infection compared to frequencyof pneumococcal carriage among subjects without current/recent COVID-19 infection, overall and stratified by age, sex, and pneumococcal risk status.*

  In this analysis, the denominator will be the Saliva Carriage Subpopulation with COVID-19 status to be determined from standard-of-care viral testing.

3.3.3. Additional RSV Exploratory Endpoints

- *Proportion of CAP events where RSV was identified from saliva by PCR divided by all CAP events where the subject had a valid saliva specimen test result. If we have any other viruses detected within the scope of this study, then we will also describe those viruses.*

The denominator for this endpoint will be the CAP population for which a valid saliva specimen test results from PCR is available.

3.4. Baselines Variables

Demography

- Birth date
- Sex
- Nationality
- Enrolled previously in the current study.

Disposition – Screening

- Date of Completion/Discontinuation/Death
- Status (Completed, Physician decision, Screen Failure, Withdrawal by Subject, Death, Other)
- Will the subject continue to the next phase of this study?

Pre-Specified Significant Medical History

High-risk (immunocompromised)

- Asplenia
- Cancer/Malignancy, Hematologic
- Cancer/Malignancy, Solid Tumor
- Chronic Kidney Disease
- Human Immunodeficiency Virus (HIV) – AIDS
- Human Immunodeficiency Virus (HIV) – No AIDS
- Immunodeficiency
- Immunosuppressant Drug Therapy
- Organ Transplantation
- Multiple Myeloma

At Risk (immunocompetent)

- Asthma
- Alcoholism
- Celiac Disease
- Chronic Liver Disease without Hepatic Failure
- Chronic Liver Disease with Hepatic Failure
- Chronic Obstructive Pulmonary Disease
- Cochlear Implant
- Congestive Heart Failure
- Coronary Artery Disease (CAD)
- Chronic Neurologic Diseases
- Coagulation factor replacement therapy
- CSF Leak
- Diabetes Treated with Medication
- Down syndrome
- Institutionalized in nursing home or LTC facility (Nursing home or long-term care facility for those with disability or dependency on subject characteristics/risk determinants eCRF page)
- Occupational risk with exposure to metal fumes
- Other Chronic Heart Disease
- Other Chronic Lung Disease
- Other pneumococcal disease risk factors
- Previous Invasive Pneumococcal Disease
- Tobacco smoking (Tobacco/E-Cigarettes)

The details of the pneumococcal risk classifications have been provided in Appendix 2.4.

Subject Characteristics - Risk Determinants

- Where do you currently live?
    - Nursing Home [at risk- immunocompetent]
    - Private Residence
    - Long-term Care Facility for those with Disability or Dependency [at risk immunocompetent]
    - Other
- Healthcare Facility Exposure in the past 3 months
- Weekly exposure to children less than 5 years of age
- Antibiotics Used treatment within 14 days prior to admission
- Mental Status (confused/not confused)

Substance Use

At Risk (immunocompetent)

- Tobacco (current user)
- E-cigarettes (current user)

Not classified as risk

- Recreational Marijuana
- Illicit Drugs, not specified

Concomitant Medications

- Influenza Vaccine History (in last year only)
- Vaccination History PPV23
- Vaccination History PCV13
- COVID-19 Marketed Vaccine History (COVID MRKT)
- COVID-19 Investigation Vaccine History (COVID INV)

Vital Signs

- Weight (kg)
- Height (cm)

- Systolic
- Diastolic
- Heart Rate
- Respiratory Rate
- Temperature (Celsius)
- SpO2 pulse oximetry percentage
- CRB-65 Score will be derived using the following elements:

Clinical factor	Point
Mental status is confused	1
Respiratory rate >30 breaths per minute	1
Systolic blood pressure <90 mm Hg or Diastolic blood pressure ≤60 mm Hg	1
Age ≥65 years	1

- Charlson Co-morbidity index will be calculated using the table given in Appendix 6.

Oxygenation Parameters

- Date of assessment
- FiO2 (Fraction of Inhaled Oxygen)
- PaO2 (Arterial Blood Gases) will be derived from SpO2 as shown in Appendix 5. **Note that if the timing of measurement for FiO2 and SpO2 are not same then PaO2 will not be calculated and will be treated as missing.**

Oxygen Support

- Oxygen Source (room air/oxygen therapy)
- Date & time of assessment
- Name of Procedure (Low flow of oxygen delivery/high flow of delivery)

Imaging

- Type of imaging (CT Scan/X-Ray/MRI/Other)

- Assessment (Normal/Not evaluable/abnormal)
- If abnormal (Pneumonia without Pleural Effusion/ Pneumonia with Pleural Effusion /Other)

Hospitalization Details

- Hospitalization Category
- Hospital Duration (Discharge Date Time-Admission Date Time)
- Hospital Ongoing?
- Did the patient have evidence of Acute Respiratory Distress Syndrome (ARDS)?

Hospitalization Details – Ventilation Details

- Mechanical Ventilation use
- Type of Mechanical Ventilation
- Ventilation duration(End Date Time- Start Date Time)
- Mechanical Ventilation Ongoing?

Hospitalization Details – Intensive Care Unit (ICU)

- ICU duration (Discharge Date Time-Admission Date Time)
- Ongoing?

Microbiology Culture – Detection RSV SARS

- Specimen Type
- If Nasal Cavity, please provide Directionality
- Method of Test
- Test of Pathogen
- Test Name
- Test Result
- Not Tested

Microbiology Culture - Other Respiratory Viruses

- Specimen Type
- If Nasal Cavity, please provide directionality

- Method of test
- Was pathogen isolated?
- Other pathogen specify

Microbiology Culture Susceptibility - Pathogen Detection

- Specimen Type
- Method of Test
- Were Isolates Identified?
- Isolate Category
- Isolate Name
- Other Pathogen
- Serotype

Microbiology Culture Susceptibility - Pathogen Detection

- Antimicrobial Susceptibility Test Drug
- Susceptibility for Drug
- Method to Determine Drug Susceptibility

Pneumonia Severity Index Scale

- PSI Score
- PSI Grade will be derived from the PSI Score as follows:
    - Grade I: PSI Score=0
    - Grade II: PSI Score=1-70
    - Grade III: PSI Score=71-90
    - Grade IV: PSI Score=91-130
    - Grade V: PSI Score>130

Survival Follow-up Vital Status

- Current subject status (dead/alive)

Final Diagnosis

- Final Diagnosis
- If Final Diagnosis is CAP, specify if radiologically confirmed

Event Outcome Status

- Event outcome status

Adverse Event Report

- AE ID
- Adverse Event Term
- Is the adverse event still ongoing?
- Severity (Mild/Moderate/Severe)
- Is the adverse event serious?
- Do research related injury criteria apply?
- What was the outcome of this adverse event?
- This event is due to:
    - Other (specify)
- Did the adverse event cause the subject to be discontinued from the study?

Disposition – Follow-up

- Phase of disposition
- Status of disposition

Results of Culture and Urine Assays for *S. pneumoniae*

Sp+ is defined as the identification of *Streptococcus pneumoniae* via any of the following three diagnostic methods:

1. Microbiology culture: As part of standard of care, all blood and respiratory specimens (e.g., blood, pleural fluid, transtracheal aspirate, bronchoalveolar lavage or other respiratory tract specimen) will undergo bacterial culture at the local site laboratory for the identification of *S. pneumoniae* according to their standard methodology. Results of blood cultures (including antibiotic susceptibility) from this standard of care bacterial culture testing will be recorded in the CRF.

For sputum specimens collected as part of SOC testing, specimen quality will be assessed using the specimen's standard-of-care Gram stain results. **Only bacterial culture results from high-quality sputum specimens will be used in case of categorization for the VE analysis**. However, all pneumococcal isolates will be serotyped regardless of specimen quality.

Pneumococcal bacterial isolates from standard of care specimens will be shipped and analyzed in the Pneumococcal Reference Laboratory from the Centro Nacional de Microbiología of the Instituto de Salud Carlos III (ISCIII). The following antimicrobials will be evaluated: penicillin, amoxicillin, cefotaxime, erythromycin, clindamycin, tetracycline, and levofloxacin.

2. Binax assay: If a positive result is issued from the urine testing, when analysed by the Pfizer VRD laboratory, then it is considered *Sp*+. This assay does not provide *Sp* serotype information.

3. UAD assay: The UAD-1 detects 13 serotypes (1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F, 23F) used in Pfizer's 13vPnC vaccine. The UAD-2 assay, using the same Luminex xMAP bead technology, detects 11 additional serotypes (2, 8, 9N, 10A, 11A, 12F, 15B, 17F, 20, 22F, 33F). A maximum of two serotypes can be reported from each one of the UADs.

   Any participant who is not determined as *Sp*+ by any of the three diagnostic methods will be considered as *Sp*-, unless they have missing results from all three diagnostic methods. If any culture identified *Sp*+, antibiotic susceptibility will be assessed at the local laboratory and results are to be entered in the CRF.

Antibiotic susceptibility and the minimum inhibitory concentrations (MICs) of each isolate will be obtained from local and the National Reference Lab. The data of the local lab will come through the CRF, and those from the National Reference Lab will be maintained in a separate electronic file (non-CRF). We will primarily use the data of the National Reference Lab, and if missing then we will use the data of the local lab.

Pneumococcal Carriage and RSV Infection Testing

The saliva samples are to be tested for the carriage study. The samples will be collected from all subjects and referred to the Instituto de Salud Carlos III (ISCIII) for culture and molecular biology study (for culture-negative samples). The results of the test will be used to understand the serotype distribution and prevalence in carriage among the study population. Additional PCR testing for RSV and other respiratory viruses will be performed on the same saliva sample to estimate the distribution of CAP events with RSV and also with other respiratory viruses; results will be maintained in a separate electronic file (non-CRF).

Results of the Respiratory Specimens for Viral Testing

Results of SOC testing for viral pathogens in respiratory specimens collected upon admission will be documented in the CRF. This includes documentation of COVID-19 results when available. SOC viral testing results via PCR and/or viral culture will also be documented. For RSV and COVID-19, both negative and positive test results will be documented.

Final Diagnosis and Event Outcome

The data on the final diagnosis will include:

- Community acquired pneumonia (CAP)
- Acute Bronchitis.
- Exacerbation of COPD.
- Empyema/lung abscess.
- Other Acute Lower Respiratory Infection.
- Acute pulmonary exacerbation of congestive heart failure.
- Non-infectious process

If final Diagnosis is CAP, please specify if radiologically confirmed:

- Confirmed
- Not confirmed
- Not done

Event outcome status at 30 days after admission will include:

- Fatal/deceased.
- Not recovered/not resolved.
- Recovered/resolved.
- Recovered/resolved with sequelae.
- Recovering/resolving.
- Unknown.

The following table will be produced with the Final Diagnosis and Event outcome

Final Diagnosis	Event Outcome

	Fatal/deceased N (%)	Not recovered/ not resolved N (%)	Recovered/ resolved N (%)	Recovered/ resolved with sequelae N (%)	Recovering/ resolving N (%)	Unknown N (%)
Community acquired pneumonia (CAP)						
Acute Bronchitis.						
Exacerbation of COPD						
Empyema/lung abscess						
Other Acute Lower Respiratory Infection						
Acute pulmonary exacerbation of congestive heart failure						
Non-infectious process						

3.5. Safety Endpoints

An AE is defined as any untoward medical occurrence and can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease, whether or not related to the patient's participation in the study. This study does not include interventions that pose more than a minimal risk or burden to the study participants. Adverse events and important research-related injuries (RRI) will be reported in CRF.

AEs and RRIs data will be coded using MedDRA dictionary. The SOC and preferred term will be reported.

4. ANALYSIS SETS (POPULATIONS FOR ANALYSIS)

4.1. Study Size

This is a low interventional, multicentre, prospective, hospital-based surveillance study in adults aged 60 years of age or older hospitalised with CAP at participating hospital sites. Based on a sample size calculation, it is estimated to enroll 2,000 to 5,500 participants, which is expected to take 1.7 - 4 years. Since the required number of participants is dependent on the proportion of CAP cases that are VT-CAP and the proportion of the population vaccinated, the required sample size may be adjusted based on ongoing information from laboratory test results and vaccine exposure data, if needed.

4.2. All Enrolled Population

The all enrolled population is the full analysis set consisting of all participants enrolled in the study. Any participants who were consented and recorded in the database with a subject ID are considered as 'enrolled'. Participants' disposition and demography will be presented for this population, as well as other population, if needed.

4.3. Per-Protocol Population

The per-protocol population consists of all participants in the all enrolled population who met all inclusion/exclusion criteria. Any participant who is included in the 'all enrolled population' will be excluded from per-protocol population if the participant did not meet the inclusion/exclusion criteria. A participant with an important protocol deviation as ascertained by the clinical team will also be excluded from the per-protocol population.

4.4. All Hospitalized CAP Population for Incidence Calculations

The hospitalized CAP population for incidence calculation will include:

1. Enrolled patients residents in the catchment area of the hospitals meeting inclusion and exclusion criteria and with final diagnosis of CAP from eCRF; and

2. Unenrolled patients residents in the catchment area of the hospitals meeting inclusion criteria 1, 2, 4 and no exclusion criteria from the Screening Log and with final diagnosis of CAP.

4.5. The CAP Population

The CAP population will include the per-protocol population and who have a final diagnosis consistent with CAP (CAP-radiologically confirmed; CAP not radiologically confirmed, or CAP-no chest radiology done).

4.6. Radiologically Confirmed CAP (RAD+CAP) Population

The RAD+CAP population will include the per-protocol population and who have a final diagnosis of radiologically-confirmed CAP.

4.7. Serotype Distribution CAP Population

This population will include the per-protocol population and those who:

- have a final diagnosis consistent with CAP (CAP-radiologically confirmed; CAP not

radiologically confirmed or CAP-no chest radiology done).

- did not receive any pneumococcal vaccination ≤30 days prior to urine sample collection, and

- pneumococcal serotype identified from UAD1/2 test results, Binax, blood culture or high-quality respiratory culture (e.g., from bacterial isolate).

Events with the same serotype result involving rehospitalization ≤30 days from the date of discharge of the earlier hospitalization will be considered same event in the analysis.

4.8. Sub-Populations

4.8.1. Urinary Antigen Detection (UAD) Subpopulation

The UAD subpopulation will include all participants in the per-protocol population that have at least one valid UAD result, i.e., the participants should have UAD results as 'POS' or 'NEG' for any of the 24 serotypes. For each serotype, there are 5 possible values:

1. "POS" – Positive
2. "NEG" – Negative
3. "INDETERMINATE"
4. "QNS" – quantity not sufficient
5. "NOT DONE" – urine was received by VRED laboratory but the test was not done on the sample

The UAD subpopulation will exclude the participants that have

1. "INDETERMINATE" results for all 24 serotypes, or
2. QNS results for all 24 serotypes, or
3. NOT DONE results for all 24 serotypes, or
4. had any combination of QNS/ indeterminate/NOT DONE results for all 24 serotypes

4.8.2. Binax Subpopulation

The Binax subpopulation will include all participants from the per-protocol population that have valid Binax results ('POS' or 'NEG'). There will only be one result per urine sample test. There are four possible values:

1. "POS" – Positive
2. "NEG" – Negative
3. "QNS" – quantity not sufficient
4. "NOT DONE" – urine was received by VRED laboratory but the test was not run on the sample.

The Binax subpopulation will exclude the participants that have QNS results, or NOT DONE

4.8.3. Microbiology Culture Subpopulation

The microbiology culture subpopulation will include all participants that are included in the per-protocol population and have non-missing culture results (i.e., 'No Growth', '*Streptococcus Pneumoniae*' or growth attributable to another pathogen).

4.8.4. Viral Pathogen Testing Subpopulation

The viral testing subpopulation will include all participants that are included in the per-protocol population and have non-missing viral pathogen testing results (i.e., 'Not done').

4.8.5. *Sp*+ Subpopulation

This subpopulation is composed of the per-protocol population that are also *Sp*+ by any of the following three diagnostic methods:

- microbiology culture results of *S. pneumoniae* as a pathogen from local lab entered in the eCRF; or
- 'POS' results from Binax assay; or
- 'POS' results from UAD assay for any of the 24 serotypes.

4.8.6. *Sp*+ RAD+CAP Subpopulation

This subpopulation is composed of participants in the RAD+CAP population that are also *Sp*+ by any of the following three diagnostic methods:

- microbiology culture results of *S. pneumoniae* as a pathogen from local lab entered in the eCRF; or
- 'POS' results from BinaxNOW assay; or
- 'POS' results from UAD assay for any of the 24 serotypes.

4.8.7. *Sp*- RAD+CAP Subpopulation

This subpopulation is composed of the participants in the RAD+CAP population who are not included in the *Sp*+ RAD+CAP subpopulation.

4.8.8. COVID-19 Subpopulation

This subpopulation will include participants in the per-protocol population with a valid result from COVID-19 viral pathogen testing.

4.8.9. COVID-19 Postive Subpopulation

This subpopulation will include participants in the per-protocol population and who are positive for COVID-19 from a valid result from COVID-19 viral pathogen testing.

4.8.10. COVID-19 Negative Subpopulation

This subpopulation will include participants in the per-protocol population and who are negative for COVID-19 from a valid result from COVID-19 viral pathogen testing.

4.8.11. Pneumococcal Carriage Subpopulation

This subpopulation will include the per-protocol population for whom a valid penumoccocal carriage result is available from saliva samples.

4.8.12. RSV Subpopulation

This subpopulation will include the per-protocol population for whom a valid RSV result is available either from the study testing or SOC.

4.8.13. RSV Positive Subpopulation

This subpopulation will include the per-protocol population with a positive RSV result either from the study testing or SOC.

4.8.14. RSV Negative Subpopulation

This subpopulation will include the per-protocol population with a negative RSV result either from the study testing or SOC.

5. GENERAL METHODOLOGY AND CONVENTIONS

5.1. Hypotheses and Decision Rules

The decision rule determines under what circumstance to reject the null hypotheses. There is no hypothesis test and no decision rule applied in this study. The objective of this study is to evaluate the PCV13 VE against hospitalised VT-CAP among adults aged ≥60 years.

5.2. General Methods

5.2.1. Analyses for Binary Endpoints

Binary data will be summarized with counts and percentages. For the purpose of comparing two binary variables, 2×2 contingency tables will be used with cell frequencies and percentages. Percentages will be based on of the total number of participants used in the analysis (i.e., participants with both binary variables recorded). The number and percent of concordant and discordant results will also be calculated. Some of the endpoints may be summarized with exact 95% confidence intervals (CIs) for the proportion of interest. The exact 95% CI will be calculated based on the Clopper-Pearson method (Appendix 3.1)

5.2.2. Analyses for Categorical Endpoints

Categorical data will be summarized as the number of participants, as well as number of isolates, included in the analysis, frequencies and percentages of the individual categories. Pneumonia

severity will be categorized as grade I (PSI Score=0), II (PSI Score=1-70), III (PSI Score=71-90), IV (PSI Score=91-130), and V (PSI Score>130) and analyzed also as the categorical variable. All percentages will be based on the total number of participants or isolates included in the analysis.

5.2.3. Analyses for Continuous Endpoints

Continuous variables will be summarized as the number of participants included in the analysis, mean, standard deviation, median and range. The applicable variables will include demographics, e.g., age at enrollment, weight, height, BMI, duration of ICU stay, duration on ventilator, duration of hospital stay, and PSI score. Participants with a missing value will be excluded from the analysis for that variable.

5.2.4. Analysis of Incidence Rates

An incidence rate is defined as the proportion of new disease cases out of the total number of individuals at risk of developing that disease over a specified period of time. The numerator will be the number of disease cases (CAP events captured from eCRF and Screening Log) residing in the hospitals' catchment area and the denominator will be the total number of individuals living in the hospitals' catchment area. No adjustment will be made, such as market share, to adjust the population because the hospitals serve a defined population based on the structure of health service delivery in this region. This data on the denominator will come from government data sources (e.g., census) or other population data sources. Note that the start of the incidence surveillance period will be when surveillance is considered complete at both study sites, i.e, when it is confirmed that all pneumonia events are being captured on the screening log.

The incidence rate will be calculated per 100,000 population per year in each specific age groups (60-74 years, 75-84 years and ≥85 years) as well as all ages of the study. The formula for calculating the incidence rate is given and the analysis section, and 95% CIs for the incidence rate will be based on the assumption that the actual count of cases arises from a Poisson distribution. Further details of the approach to CIs calculations are given in Appendix 3.2.

5.3. Methods to Manage Missing Data

Analyses will be based on all available data. Participants who have data missing for a certain analysis will be excluded only from that analysis. For example, if a participants has missing UAD assay results but has results available from Binax and culture, then the participants will be excluded from any analyses requiring the UAD results, but will be included in any analyses requiring results from Binax or culture.

5.4. Serotype Determination

A subject may have 2 different serotypes identified from UAD-1. Similarly, the participants may have 2 different serotypes identified from UAD-2, totaling a maximum of 4 different serotypes from UADs. Participants can also have multiple Sp isolates from microbiology cultures. All Sp isolates are sent to the Reference Lab for serotyping. If multiple Sp isolates are serotyped and found to be the same serotype they will be counted one in the serotype distribution. If multiple isolates are sent to the Reference Lab and have different serotypes, they will be counted once for

each unique serotype. However, if no Sp isolates were serotyped, then 'not serotyped' will be considered in the serotype distribution.

Subjects may also be identified with 4 or more different isolates/serotypes from the two diagnostic methods (culture and UADs). For subjects who are Sp+ via culture method but without serotype information ('Not Serotyped'), the following approaches are used:

a. If subjects had multiple isolates identified as Sp+ from culture which are 'Not Serotyped', those will be counted as only one in the numerator when summarizing 'Not Serotyped' in the analysis of 'percent of isolates'
b. If subjects are identified as Sp+ from UAD, serotypes from UAD will be included in the numerator of the numbered serotype(s).
c. If subjects were NOT identified as Sp+ from UAD, the subjects with isolates not serotyped will be included in the numerator when summarizing 'Not Serotyped' and 'Any Sp+' in the analysis of 'percent of subjects'. The isolates will also be included in the numerator of 'Not Serotyped' for the analysis of ''Percent of Isolates'. However, only one "Not serotyped" isolate will be counted in the numerator.

For subjects who are detected as Sp+ via culture and UAD, the following approaches are used:

a. If the same serotype is identified, the serotype will only be counted once in the analysis of 'percent of serotypes'.
b. If different serotypes are identified, all unique serotypes from both methods will be in analysis of 'percent of serotypes

Serotype at group level

Vaccine type (VT): If any of the serotypes covered in the 20-valent pneumococcal conjugate vccine (PCV20), 13-valent pneumococcal conjugate vaccine (PCV13), the 7-valent pneumococcal conjugate vaccine (PCV7) or the 23-valent pneumococcal polysaccharide vaccine (PPV23) are detected either by UAD or culture.

PCV20: If any of the following serotypes 1, 3, 4, 5, 6A, 6B, 7F, 8, 9V, 10A, 11A, 12F, 14, 15B, 18C, 19A, 19F, 22F, 23F, and 33F are detected either by UAD or culture.

PCV13: If any of the following serotypes 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F, and 23F are detected either by UAD or culture

cPCV7 (complimentary seven serotype): If any of the following serotypes 8, 10A, 11A, 12F, 15B, 22F, and 33F are detected by UAD or culture

PCV7: If any of the following serotypes 4, 6B, 9V, 14, 18C, 19F and 23F are detected by UAD or culture

PPV23: if any of the folling serotype 1, 2, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F, 18C, 19F, 19A, 20, 22F, 23F and 33F detected by UAD or culture.

Non-vaccine type (NVT): If none of the serotypes covered in PCV20, PCV13, PCV7, or PPV23 are detected from both UAD and culture.

Non-typeable (NT): A cultured isolate that could not be serotyped. This is a valid result and so will be summarized as a separate line. Note that this is not the same as a missing serotype result.

6. ANALYSES AND SUMMARIES

One focus of this study is description and serotype distribution of *S. pneumoniae,* therefore statistical methods appropriate for these descriptive analyses will be used. Standard summary measures along with the number of participants used in the analysis will be displayed for all analyses. As a general rule, all proportions will be expressed as percentages.

6.1. Primary Endpoints

6.1.1. Primary Endpoint 1 - Effectiveness of PCV13 among immunocompetent participants aged ≥60 years

VE is to be calculated as 1 minus odds ratio for PCV13 vaccination among cases and controls, multiplied by 100%, overall and restricted to immunocompetent participants.

6.1.1.1. Primary Endpoint 1- Main Analysis

Pneumococcal Vaccination Status for Primary Endpoint

Pneumococcal vaccination status is defined as receipt of ≥1 dose of PCV13 or PPV23 and confirmed by electronic registries. Vaccination status will be categorized as: a) never exposed, b) ever exposed. Ever exposed will be subdivided into: i) exposed <5 years, and ii) exposed ≥5 and <10 years. Persons vaccinated ≥10 years will be considered unexposed.

The Madrid vaccination guidelines (**Protocol Table 1**) stipulate the timing of PCV13, and we will exclude following participants from the VE analyses:

- Immunocompetent participants who received PCV13 within one year after receiving PPV23
- Immunocompromising participants who received PCV13 within eight weeks after receiving PPV23
- Any participants who received PCV20

Note: Cases and controls include either radiologically confirmed CAP, not radiologically confirmed, or chest radiology not done. If more than 5% of the CAP cases are not radiologically confirmed, then a separate analysis will be performed selecting cases and controls only from the radiologically confirmed CAP cases.

The characteristics of cases and controls will be compared using the chi-square test.

For the primary endpoint, the cases will be the patients with CAP in whom PCV13 serotypes are identified by any method and the controls will be the patients with CAP without having VT serotype (even the CAPs for which the serotype was not identified). Multivariable logistic regression model (Appendix 3.3), will be applied to estimate odds of exposure to PCV13 among cases versus controls. All VE estimates will be calculated as (1 – odds ratio) × 100%. Both unadjusted and adjusted VE will be estimated. Patients' age, sex, study site and the variables deemed to be potential confounders on the basis of prior knowledge (e.g. nationality to be modeled as Spanish vs. non-Spanish; risk level to be modeled as low vs. at-risk vs. high-risk; BMI category to be modeled as obese vs. overweight vs. normal vs. underweight; PSI score as continuous; current drug abuse; antibiotic use within past 14 days; influenza vaccination within the last year; pneumonia in last year [ordinal variable, 0,1,2,3] and PPV23 receipt in past 5 years) will be included in the multivariable regression model for adjustment. Initially, we will employ bivariate models, and only the variables significant at p<.20 in the bivariate models will be included in the final model for adjustment.

Analysis

If more than 5% of the participants are not radiologically confirmed CAP, then we will perform a sensitivity analysis using participants only with the radiologically confirmed CAP cases.

A second model will be developed excluding immunocompromised persons to evaluate PCV13 VE without those who may not be able to respond to the vaccine.

6.1.1.2. Primary Endpoint 1- Sensitivity Analyses

Several sensitivity analyses will be performed for estimating the VE against PCV13 VT serotype as well as against PPV23 VT serotype *S. pneumoniae*.

Selection of controls using propensity score matching

In the 1st sensitivity analysis, the cases will be the patients with CAP in whom PCV13 serotypes are identified by any method and the controls will be the patients with CAP without having VT serotype (even the CAPs for which the serotype was not identified). The controls will be selected using propensity score matching. A propensity score is the conditional probability of assignment to a treatment given a vector of covariates including the values of all treatment confounders (*vida infra)*. We will select the required number of controls matching by the propensity score of the case. Several types of propensity score matching are available in literature. We will use nearest neighbor matching (without replacement, meaning a control cannot be the control of any other case) within a specified caliper distance (Rosenbaum & Rubin, 1985). The caliper distance will be set to 0.2 of propensity score (i.e., 0.2 standard deviations of whatever distance will be measured), so that matching controls remain within the specified distance of the propensity score. Note that 0.2 of the pooled standard deviation of the logit of the propensity score will eliminate approximately 99% of the bias due to the measured confounders (Austin 2011). The maximum number of controls for a case will be determined based on the exact case-control ratio in the data. If less than the maximum number of controls exists within the set caliper distance of 0.2, then only those controls within the caliper distance will be selected and that case will have less than the maximum number of controls.

We may also perform a sensitivity analysis by selecting controls setting the caliper distance to 0.1 of the propensity score.

Variables to be used to calculate the propensity scores

Variables for calculating the propensity score.

- Patients' age: continuous
- Sex: female=0, male=1
- Nationality: Spanish=0, non-Spanish=1
- Risk level: low-risk=0, at-risk=1, high-risk=2
- PSI score: continuous
- Antibiotic use within past 14 days: no=0, yes=1)
- Influenza vaccination within past one year: yes=0, no=1
- Pneumonia in the last year: ordinal value=0, 1,2,3
- Season of admission: Spring (Mar-May)=0, Summer (Jun-Aug)=1, Autumn (sep-Nov)=2, Winter (Dec-Feb)=3.
- PPV23 receipt in past 5 years: yes=0, no=1
- CAP event occurrence during COVID-19 pandemic: 0=not during COVID-19, 1=During COVID-19 (The study is started in the pandemic period; the variable will be used if we find a non-pandemic period during the study period)

The study site will be exactly matched. Note that the categorical variable will be put in the model like the ordinal values.

The conditional logistic regression will be employed to estimate odds of exposure to PCV13 among these matched cases and controls. The VE estimate will be calculated as (1 – odds ratio) × 100%. We will use a two-sided alpha of 0.05 for all analyses.

In the 2nd sensitivity analysis (PPV23 Naïve population), the cases will be the hospitalized CAP in whom PCV13 serotypes are identified by any method and had not received PPV23 in the last five years from the date of visit; the controls (also PPV23 Naïve population), will be selected using the propensity score matching as described above. A conditional logistic regression will be employed to estimate odds of exposure to PCV13 among cases versus controls in the PPV23 naïve population. The VE estimate will be calculated as (1 – odds ratio) × 100%. We will use a two-sided alpha of 0.05 for all analyses.

In the 3rd sensitivity analysis (Suzuki Method), cases are the hospitalized patients with CAP in whom VT serotypes are identified by any method including UADs, routine culture of blood, pleural fluid, or other high-quality respiratory tract specimen, and the controls are the patients in whom pneumococcus is not detected by any method (culture, UAD, or Binax), similar to a recently published article (Suzuki et al, 2017). The controls will be selected using the propensity score matching as described above. A conditional logistic regression will be employed to estimate odds of exposure to PCV13 among cases versus controls. The VE estimate will be calculated as (1 – odds ratio) × 100%. We will use a two-sided alpha of 0.05 for all analyses.

In the 4th sensitivity analysis (Broome/Indirect cohort method), cases are the patients with CAP in whom VT serotypes are identified by any method including UADs, routine culture of blood, pleural fluid, or other high-quality respiratory tract specimen, and the controls are the patients having a NVT pneumococcal CAP, i.e. the patients in whom pneumococcus is detected from blood/pleural/high-quality sputum culture, Binax or UAD1/2 positive, and VT serotypes are not identified from culture or UAD1/2. The controls will be selected using the propensity score matching as described above. A conditional logistic regression will be employed to estimate odds of exposure to PCV13 among cases versus controls. The VE estimate will be calculated as (1 – odds ratio) × 100%. We will use a two-sided alpha of 0.05 for all analyses.

Adjustment of Bias in the Indirect Cohort (Broome Method)

Since the Broome method is subject to bias caused by PCV13 vaccination increasing the chance of NVT carriage in vaccinated compared to unvaccinated individuals, through reduction in VT carriage and replacement by NVT carriage (Andrews et al, 2011), we will pursue adjustment for the bias due to carriage replacement using the following formula if pneumococcal colonization data is available for the UAD control population:

$$VE_{Broome} = 1 - \frac{1-VE}{P_{n|v} \div P_{n|u}} \quad \text{..... equation (1)}$$

where, VE_{Broome} is observed VE by the indirect cohort method, $P_{n|u}$ is the probability of NVT carriage in an unvaccinated individuals; $P_{n|v}$ is the probability of NVT carriage in a vaccinated individuals.

If we assume complete serotype replacement at any point in time, i.e. overall carriage rates remain stable and equal in vaccinated and unvaccinated individuals then the equation (1) becomes:

$$VE_{Broome} = 1 - \frac{1-VE}{1+VE_c P_u \div (1-P_u)} \quad \text{..... equation (1)}$$

where, VE_c is VE against carriage, P_u is the proportion of carriage that is VT in the unvaccinated group. This formula will be used along with estimates for VE_c and P_u to assess the possible bias in estimating VE.

6.2. Secondary Endpoints

6.2.1. Secondary Endpoint 1 - Proportion of CAP events by *S pneumoniae* serotypes

Distribution of *S. pneumoniae* serotypes identified by blood, high-quality respiratory cultures, or a serotype-specific UAD assay among the CAP population with culture and/or UAD1/2 testing are done, and among all CAP participants with a pneumococcus identified.

6.2.1.1. Secondary Endpoint 1 - Main Analysis

We will calculate the proportion of participants infected with any of these 20 serotypes (1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F, 23F, 8, 10A, 11A, 12F, 15B, 22F, or 33F) as well as for each serotype among all CAP population 60 years or older having their culture and/or UAD test done with the 95% CIs for proportion. Groupings for PCV13 and PCV20 will be included.

The calculations will also be made among the *Sp*+ subpopulation that includes all CAP population in whom a pneumococcus identified by culture, UAD, or Binax.

The subgroup analysis will be done by risk status (high-risk, at-risk, and low-risk), age group (60-74 years,75-84 year and ≥85 years), and by COVID-19 status (COVID-19 positive and COVID-19 negative).

6.2.2. Secondary Endpoint 2 - Proportion of respiratory pneumococcal carriage among the CAP population

Proportion of CAP cases where the pneumococcal carriage is identified from saliva by culture or by PCR, divided by all CAP cases where the participants had a valid saliva specimen test result.

6.2.2.1. Secondary Endpoint 2 - Main Analysis

The numerator in this analysis will be the participants with pneumococcal carriage identified from saliva by culture or PCR by reference laboratory. The denominator will be the CAP population and who will have valid carriage results from the reference lab. We will calculate the proportion of participants with pneumococcal carriage of two target genes (lytA and *piaB*) detected from saliva among this population.

The calculations will also be made by risk status (high-risk, at-risk, and low-risk) and by age group (60-74 years,75-84 years and ≥85 years).

6.2.3. Secondary Endpoint 3 - Proportion of participants having *S. pneumoniae* with underlying risk among the CAP population

Proportion of participants with a medical risk condition (at-risk and high-risk) among all CAP population as well as among participants with pneumococcus isolated from blood, high-quality respiratory cultures, a serotype-specific urinary antigen detection (UAD) assay or BinaxNow.

6.2.3.1. Secondary Endpoint 3 - Main Analysis

The following calculations will be made, overall and by risk condition:

- Proportion of participants (with 95% CI for proportion) with at-risk medical condition among all CAP population

- Proportion of participants (with 95% CI for proportion) with high-risk medical condition among all CAP population

- Proportion of participants (with 95% CI for proportion) with at-risk medical condition among the *Sp*+ subpopulation

- Proportion of participants (with 95% CI for proportion) with high-risk medical condition among the *Sp*+ subpopulation

6.2.4. Secondary Endpoint 4 - Proportion of participants with an *S. pneumoniae* with antibiotic resistance

Proportion of *S. pneumoniae* isolates with antibiotic susceptibility identified by SOC testing, overall and by susceptibility type. Proportion with multidrug resistance will also be calculated.

6.2.4.1. Secondary Endpoint 4 - Main Analysis

The analysis will be conducted including participants *Sp*+ by culture (based on testing of isolates by reference lab or local lab) among all CAP population with information on antibiotic susceptibility results. If the antibiotic susceptibility data is missing or not done for a participant and for a drug, then the participant will be excluded from the analysis for that particular drug. For participants whose *Sp* isolates were tested for antibiotic susceptibility at the Reference Lab, the results of the Reference Lab will be used in the analysis rather than the local clinical laboratory results. For participants who are detected as *Sp*+ via culture method but without serotype information ('Not Serotyped'), the participants will also be included in the analysis if antibiotic susceptibility results are available from the reference lab or the local clinical lab.

The following drugs: Penicillin, Amoxicillin, Cefotaxime, Erythromycin, Clindamycin, Tetracycline, and Levofloxacin are included in the antimicrobial susceptibility test. The antibiotic susceptibility will be categorized as: Susceptible, Resistant, and Multidrug Resistant and the proportion of each of the categories for each of the drugs will be calculated.

6.3. Other Endpoints

6.3.1. Exploratory Endpoint 1 – Rate of CAP events in the surveillance region

Incidence rates for hospitalized CAP and *Sp*+CAP (both overall and radiologically confirmed) obtained from screening log among individuals ≥60 years in surveillance region.

6.3.1.1. Exploratory Endpoint 1 - Main Analysis

The incidence rate of CAP (IR_{CAP}) will be calculated as

$$IR_{CAP} = \frac{no.\,of\ CAP\ cases\ in\ the\ screening\ log \div no.\,of\ years\ of\ surveillance}{total\ population\ of\ the\ hospitals'\ geographical\ catchment\ \ area} \times 100{,}000$$

The no. of years of surveillance will be calculated as
(date of last person 1st visit – date of the start of the surveillance period +1)/365.25

Note that the date of start of surveillance will be obtained from the Clinical Scientist of this study based on the data that it is confirmed that all pneumonia cases are being captured on the screening log.

We will compute distribution of the following variables between cases obtained from eCRF (enrolled patients) and Screening Log (not included in the eCRF/not enrolled) to evaluate the representativeness of the enrolled cases:

- Age
- Sex
- Nationality
- Radiologically confirmed CAP
- ICU Status

We will perform chi-square or Fisher's exact test, as appropriate, for the categorical variable and Student's T-test or Wilcoxon-Mann-Whitney test, as appropriate, for the continuous variables to evaluate the differences between the two populations (enrolled versus not enrolled).

6.3.2. Exploratory Endpoint 2 – Differences in VE between PPV23 and PCV13

The difference in VEs of PCV13 and PPV23 against respective serotypes overall and among immunocompetent persons only.

6.3.2.1. Exploratory Endpoint 2 - Main Analysis

In this analysis, the VE of PCV13 against VT (PCV13) pneumococcal CAP will be obtained from the primary analysis as described in Section 6.1.1.1. The same method will be used to estimate the VE of PPV23 against VT (PPV23) pneumococcal CAP. The crude difference will then be calculated as:

VE of PCV13 against VT (PCV13) pneumococcal CAP – VE of PPV23 against VT (PPV23) pneumococcal CAP

To evaluate the differences of the effectiveness of these two vaccines, the ORs from both the models that used conditional logistic regression using the propensity matching cases and controls, will be tested using the likelihood ratio and/or z-score test as described in Appendix 3.4.

6.3.3. Exploratory Endpoint 3 – Compare serotype results for carriage testing by specimen type

Percent agreement in serotype results for each serotype among the CAP population who have at least two of the three following specimen results available: a) high-quality respiratory specimen (SOC), b) UAD1/2 (study procedure), and c) saliva.

6.3.3.1. Exploratory Endpoint 3 - Main Analysis

The percentages of specimens with serotype results (for each serotype) by method of detection will be determined dividing the number of specified serotypes identified by the total number of specimens tested by the detection method.

Cohen's kappa will be used to describe the agreement between the detection methods. Each comparison will involve two specimen types only (i.e, UAD versus saliva, UAD versus respiratory culture, Saliva versus respiratory culture). The kappa term ranges from -1 to 1. If the agreement is less than what would be expected by chance than it will be negative. The following labels are assigned to the corresponding ranges of kappa strength: <0 is poor agreement, 0.00-0.20 is considered slight agreement, 0.21-0.40 is fair agreement, 0.41-0.60 is moderate agreement, substantial 0.61-0.80 is substantial agreement, and 0.81-1.00 is perfect agreement.

Besides the kappa statistic, concordant and discordant pairs will be counted among these three methods, and the McNemar's test, a nonparametric method, will be used to assess the level of imbalance in the discordant status between specified serotype identified in each of these tests.

A venn diagram will be produced to represent the concordant and discordant pairs among these three methods.

6.3.4. Exploratory Endpoint 4 – Incidence Rate of CAP cases with specific risk condition

The incidence rate of CAP residents with a specific risk conditions will be calculated using the number of CAP events, identified in the screening log, multiplied by the proportion of CAP population with specific risk condition(s), divided by the estimated number of persons ≥60 years in surveillance region with the same specific risk condition(s), in aggregate for at-risk and high-risk conditions and for individual risk groups.

6.3.4.1. Exploratory Endpoint 4 - Main Analysis

The incidence rate among the high-risk population (IR_{hr}) will be calculated as

$$IR_{hr} = \frac{(no.of\ cases\ in\ the\ high\ risk\ category \div no.of\ years\ of\ surveillance)}{total\ high\ risk\ population\ in\ the\ hospitals'geographical\ catchment\ area} \times 100{,}000$$

The incidence rate among the at-risk population (IR_{ar}) will be calculated as

$$IR_{ar} = \frac{(no.of\ cases\ in\ the\ at\ risk\ category \div no.of\ years\ of\ surveillance)}{total\ at\ risk\ population\ in\ the\ hospitals'geographical\ catchment\ \ area} \times 100{,}000$$

The no. of years of surveillance will be calculated as

(date of last participant first visit – date of the start of surveillance period +1)/365.25

If the data do not provide complete years (e.g. 1 year, 2 years, etc.), then the no. of years of surveillance will be adjusted for seasonality, if the data on seasonality are available for the sites.

Note that we will apply the proportion of CAP events in a specific risk group among enrolled participants to account for the number of CAP events in specific risk group among the unenrolled patients. The risk population will be obtained by multiplying the frequency of a given risk factor obtained from other sources (e.g., UAD control subjects or real-world data analyses) by

denominator counts for that age group from local government sources. **This analysis will be performed if such data are made available.**

6.3.5. COVID-19 Related Endpoint 1 – Distribution of the CAP cases by COVID-19 positive and negative

Percentage of CAP cases with pneumococcal infection identified by any means (UAD, BinaxNOW, or bacterial culture) associated with a positive COVID-19 test compared to percentage of CAP cases with pneumococcal infection identified by any means (e.g., UAD, BinaxNOW, bacterial culture) associated with a negative COVID-19 test, overall and stratified by age, sex, and pneumococcal risk status.

6.3.5.1. COVID-19 Related Endpoint 1 - Main Analysis

The following calculations will be made for overall and stratified by age group, sex, and pneumococcal risk status among COVID-19 subpopulations (test positive or negative), i.e., all enrolled population with a valid results from COVID-19 viral testing. The following calculations will be made:

- Proportion (95% CI for proportion) of CAP events with pneumococcus among COVID-19 test positive cases
- Proportion (95% CI for proportion) of CAP events with pneumococcus among COVID-19 test negative cases

These will also be calculated for the following subgroups: age group (60-74 years, 75-84 years and ≥85 years), sex, and pneumococcal risk status.

6.3.6. COVID-19 Related Endpoint 2 – Proportion of CAP patients who experienced severe clinical outcome(s) between patients with or without COVID-19 coinfection with *Sp*

These analyses will assess differences in the severe clinical outcomes between patients with or without laboratory-confirmed COVID-19 coinfection with *Sp* identified by any means (e.g., UAD, Binax, and bacterial culture) and separately by penumococcal carriage.

6.3.6.1. COVID-19 Related Endpoint 2 - Main Analysis

The patients with COVID-19 will be divided into two groups:

- those with coinfection with *Sp* identified by any means (e.g., UAD, Binax, bacterial culture)
- Those without coinfection with *Sp* identified by any means (e.g., UAD, Binax, bacterial culture)

The severe clinical outcomes are:

- Admitted to ICU (yes vs. no)

- PaO2/FiO2 ratio <250 mmHg (yes vs. no)

- Respiratory rate >30 breaths/min (yes vs. no)
- Died (yes vs. no)
- Received mechanical ventilation (yes vs. no)
- Diagnosed as ARDS (yes vs. no)
- Any one of the above (yes. vs. no)

The difference in the severe clinical outcomes between the two groups will be evaluated using Chi-square or Fisher's Exact test, as appropriate.

The analysis will also be done stratifying by age group (60-74 years,75-84 years and ≥85 years), sex, and pneumococcal risk status.

6.3.7. COVID-19 Related Endpoint 3 - Effectiveness of 13vPnC against severe COVID-19 disease

VE will be calculated as 1 minus the odds ratio for PCV13 vaccination among cases and controls, multiplied by 100, overall and among immunocompetent persons only.

6.3.7.1. COVID-19 Related Endpoint 3 - Main Analysis

In this analysis:

- The cases are CAP cases with COVID-19 with severe clinical outcomes
- The controls are CAP cases with COVID-19 without severe clinical outcomes

The controls will be selected by matching by the propensity score as described above. We will then develop six models as follows:

Model 1: Cases: COVID-19+ve with any severe outcome; Controls: COVID-19+ve without a severe outcome

Model 2: Cases: COVID-19+ve and admitted to ICU; Controls: COVID-19+ve and not admitted to ICU

Model 3 : Cases: COVID-19+ve and PaO2/FiO2 ratio <250 mmHg; Controls: COVID-19+ve and PaO2/FiO2 ratio ≥250 mmHg

Model 4: Cases: COVID-19+ve and Respiratory rate >30 breaths/min; Controls: COVID-19+ve and Respiratory rate ≤30 breaths/min

Model 5: Cases: COVID-19+ve and died; Controls: COVID-19+ve and alive.

Model 6: Cases: COVID-19+ve and received mechanical ventilation; Controls: COVID-19+ve and did not receive mechanical ventilation

Model 7: Cases: COVID-19+ve and diagnosed as ARDS; Controls: COVID-19+ve and not diagnosed as ARDS

For each of the models, we will employ a conditional logistic regression model and will estimate odds of exposure to PCV13 among cases versus controls. The VE estimate will be calculated as (1 – odds ratio) × 100%. We will use a two-sided alpha of 0.05 for all analyses.

These models will also be replicated for immunocompetent (absence of immunocompromising conditions, i.e., high-risk conditions) persons only.

6.3.8. COVID-19 Related Endpoint 4 – Frequency distribution of pneumococcal carriage among participants with and without current/recent COVID-19 infection

Frequency distribution of serotypes for pneumococcal carriage among participants with or without laboratory-confirmed COVID-19 infection, overall and stratified by age, sex, and pneumococcal risk status.

6.3.8.1. COVID-19 Related Endpoint 4 - Main Analysis

In this analysis the CAP population will be divided into two groups:

- CAP cases associated with COVID-19 infection
- CAP cases without COVID-19 infection

Distribution of serotype for the pneumococcal carriage between these two groups will be evaluated. Chi-square or Fisher's exact tests, as appropriate will be used to evaluate the difference between the two groups (cases and controls as defined herein).

6.3.9. Additional RSV Related Endpoint 1 – Proportion of CAP events where RSV was identified from Saliva

The proportion of the RSV cases among CAP population will be calculated for which a valid saliva specimen test result for RSV is available.

6.3.9.1. Additional RSV Related Endpoint 1 - Main Analysis

In this analysis the denominator will be CAP population whose a valid saliva specimen test result for RSV is available. Proportion (95% CI for proportion) of RSV infection and other respiratory viruses among the CAP population will be calculated. The proportion will be calculated overall and stratified by age, sex, and pneumococcal risk status.

6.4. Subset Analyses

Age group, sex, and nationality will be used in the subset analyses. Age at enrolment will be calculated in years as integer ((Date of enrolment/visit - Date of Birth + 1) /365.25). Sex and nationality at enrolment will be used as reported on the CRF. For purposes of analysis, age at enrolment will be broken into the following categories: 60-74 years and 75 years or older.

6.5. Baseline and Other Summaries and Analyses

Data will be summarized based on the all enrolled population, showing the number and percent of participants consented, enrolled, completed study and withdrawn from study (including reasons for withdrawal). The number and percent of participants will also be displayed by individual sites within all enrolled, per-protocol, CAP and RAD+CAP associated with *Sp*+ populations. In addition, the number and percent of participants within each analysis subpopulation will be reported. The following four analyses pertain to information collected at enrollment:

6.5.1. Demographic Data

Demographic data will include age, sex, nationality, weight, height, BMI (to be calculated, if not already derived in the data set) and previous enrollment in the study (Yes/No). Weight, height, and BMI will be analyzed as continuous variables, and all others as categorical variables.

6.5.2. Risk factors and Significant Medical History

Most of the risk factors are to be analyzed as categorical variables with three categories as present, absent or unknown (recorded as Yes, No and Unknown, respectively on the CRF). Significant medical history will be analyzed as recorded on the CRF (using pre-filled conditions). For each condition, the number and percent of participants in each category, namely, present, absent, unknown and not assessed (recorded as Yes, No, Unknown and Not Assessed, respectively on the CRF) will be displayed. Based on their medical history and risk factors, participants will be grouped into 'at-risk' and 'high-risk' subgroups for analysis. We will analyze per-protocol population breakdown by *Sp*+ versus *Sp*-.

6.5.3. Prior vaccination

Prior vaccination with PCV13 and/ or PPV23 will be summarized as binary (Yes/No) variables (with numbers and percentages). Participants will be divided into ever vaccinated and never vaccinated for each vaccine. The timing of last dose will be presented by year since last vaccination (e.g., <1 year, 1 year, 2 years, etc.) The timing (in years) will be calculated as int((date of visit – date of last dose +1)/365.25). A table with PCV13 exposure in rows and PPV23 exposure in columns will be developed to show the overlap. Proportion with influenza vaccination in the last year will also be displayed. Breakdowns will be provided for final diagnosis of CAP versus final diagnosis of non-CAP.

6.5.4. Culture Data

The number and percentage of cultures performed as well as the number and percentage of pathogens identified will be summarized as frequency distributions for the entire study population.

Summaries will also be calculated for the specimen source (e.g., blood, pleural fluid, sputum, etc.) as well as for the various organisms within specimen source. Note that a participant may have multiple specimen sources and/or organisms isolated. Organism type (pathogen, for *Sp* only) will also be summarized individually for each culture source as well as combined across all sources. This analysis will be performed for the microbiology culture subpopulation. Breakdowns will be provided for final diagnosis of CAP versus final diagnosis of non-CAP.

6.5.5. Viral Testing Data

The number and percentage of viral tests performed as well as the number and percentage of viral pathogens identified will be summarized as frequency distributions for all participants. Note that a participant may have multiple viral pathogens isolated. Pathogen type will also be summarized individually. This analysis will be performed for the viral testing subpopulation within all enrolled population. Breakdown will be provided for final diagnosis of CAP versus final diagnosis of non-CAP. Within CAP and non-CAP, breakdown will also be provided for *Sp*+ versus *Sp*-.

6.5.6. Antimicrobial Susceptibility

Antimicrobial susceptibility will be performed on the microbiology culture subpopulation within the CAP population. Only participants with *Sp* isolates (which are confirmed as a pathogen on the CRF) from cultures can contribute to this analysis and participants with responses of 'Not Done' for any antimicrobial test will be excluded. For each *Sp* isolate and a given antimicrobial test, a categorical variable will denote whether the isolate is susceptible, intermediate or resistant as recorded on the CRF. Frequencies and percentages of these three categories will be calculated first for each antimicrobial test overall (across all serotypes) and then for each antimicrobial test grouped further by PCV20, the 7 serotypes in PCV20 but not in PCV13, and non-PCV20 serotypes. Percentages will be calculated out of the total number of *Sp* isolates within each serotype.

The following antimicrobials will be evaluated: penicillin, amoxicillin, cefotaxime, erythromycin, clindamycin, tetracycline, and levofloxacin.

6.5.7. Lung Imaging (Chest X-ray/CT/Other) and Final Diagnosis

The number and percent of participants with chest X-ray or chest CT or other diagnostic method performed will be displayed. Among the participants with imaging results, the number and percent of participants with abnormal results will be calculated. This analysis will be performed on the all per-protocol, CAP and RAD+CAP populations. Breakdowns will be provided for final diagnosis of CAP versus not final diagnosis of CAP.

6.5.8. Mortality

The number and percentage of participants who died during the study will be summarized along with a 95% confidence interval for the percentage. For participants who died, time to death from study enrolment will be calculated as (Date of death – Visit 1 date + 1). The number and percentage of participants who died within 30 days of study enrolment will also be summarized with a 95%

CI. We will also summarize the data using the participants who died within 180 days of the study enrolment

Mortality rates will be summarized for CAP, RAD+CAP, *Sp*+CAP and COVID-19+CAP populations. As per the definition of mortality rate, the numerator will be the number of deaths (within each population, respectively), while the denominator would be the total number of participants with that condition. The method of analysis will be performed as recorded in Section 5.2.1 along with a 95% CI.

6.6. Safety Summaries and Analyses

All analyses of safety data will be based on the all enrolled population. Frequencies and percentages for reported AEs, SAEs (overall and Pfizer drug related) and RRIs will be calculated. CIs will not be required. A listing of all AEs reported by the participants will be created showing all the data from the AEs CRF.

7. INTERIM ANALYSES

We will not perform the interim analysis of the primary endpoints, i.e. not evaluting effectiveness of the vaccine. Instead, we will perform descriptive analyses on an *ad hoc* basis throughout the study period, which will include basic characteristics of the study population, overall vaccine coverage among the study population (NOT by the case and control status), disease incidences, and serotype distribution among the study population.

8. REFERENCES

1. Andrews N, Waight PA, Borrow R, Ladhani S, George RC, et al. (2011) Using the indirect cohort design to estimate the effectiveness of the seven valent pneumococcal conjugate vaccine in England and Wales. PLoS ONE 6(12): e28435.

2. Austin PC. An Introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behav Res. 2011; 46(3): 399–424.

3. Bartlett JG, Breiman RF, Mandell LA, File TM. Community-acquired pneumonia in adults: guidelines for management. Clin Infect Dis. 1998;26:811–838.

4. Daly LB, Bourke GJ, Mcgilvray J. *Interpretation and use of medical statistics*, 4th Ed. Oxford: Blackwell Scientific Publications, 1991.

5. Daly LB. Simple SAS macros for the calculation of exact binomial and Poisson confidence limits. *Comp. Biol. Med.* 1992; **22**:351-361.

6. Collett D. *Modelling Binary Data.* London: Chapman & Hall; 1991.

7. España PP MR, Torres A, Fernández-Villar JA, Marimón JM, Martínez de la Fuente AP, et al. Differences in the burden of all-cause CAP due to PCV13 serotypes based in the different use of the vaccine by regions in Spain (the CAPA study). ECCMID; April;

Madrid, Spain 2018. Available from: https://www.escmid.org/escmid_publications/escmid_elibrary/?q=Espa%C3%B1a&id=2173&L=0&x=0&y=0&tx_solr%5Bsort%5D=relevance%2Basc&tx_solr%5Bfilter%5D%5B0%5D=main_filter_eccmid%253Atrue&tx_solr%5Bfilter%5D%5B1%5D=pub_dat%253A201801010000-201812312359. Accessed: June 26, 2019.

8. Kimberly M. Shea, John Edelsberg, Derek Weycker, Raymond A. Farkouh, David R. Strutton, and Stephen I. Pelton "Rates of pneumococcal disease in adults with chronic medical conditions" *Open Forum Infect Dis* 2014 : ofu024v1-ofu024

9. Murray PR, Washington JA. Microscopic and baceriologic analysis of expectorated sputum. Mayo Clin Proc. 1975 Jun;50(6):339-44. PMID: 1127999.

10. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A Simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996; 49(12):1373-9.

11. Sousa D, Justo I, Domínguez A, Manzur A, Izquierdo C, Ruiz L, et al. Community acquired pneumonia in immunocompromised older patients: incidence, causative organisms and outcome. Clin Microbiol Infect. 2013;19(2):187-92.

12. Suzuki M, Dhoubhadel BG, Ishifuji T, Yasunami M, Yaegashi M, Asoh N, et al. Serotype-specific effectiveness of 23-valent pneumococcal polysaccharide vaccine against pneumococcal pneumonia in adults aged 65 years or older: a multicentre, prospective, test-negative design study. Lancet Infect Dis. 2017;17(3):313-21.

13. Zeger SL, Linag KY. An overview of methods for analysis of longitudinal data. Statistics in medicine 1992, II(14-15):I825-I839

9. APPENDICES

Appendix 1. List of Abbreviations

Abbreviation	Definition
ACIP	Advisory Committee on Immunisation Practices
AE	Adverse Event
AIDS	Acquired Immunodeficiency Syndrome
ANOVA	Analysis Of Variance
BMI	Body Mass Index
CAD	Coronary Artery Disease
CAMHB	Cation-Adjusted Mueller-Hinton broth
CAP	Community Acquired Pneumonia
CAPA	Acronym for an epidemiological study assessing the burden of hospitalised community-acquired pneumonia (CAP) due to *Streptococcus pneumoniae* in adults in Spain using urinary antigen detection (UAD) testing
CAPITA	Community-acquired Pneumonia Immunization Trial in Adults
CDC	United States Centres for Disease Control and Prevention
CI	Confidence Interval
COPD	Chronic Obstructive Pulmonary Disease
COVID-19	Coronavirus Disease 2019
CRB-65	Confusion Urea Respiratory Rate Blood Pressure
CRF	Case Report Form
CT	Computed Tomography
CXR	Chest X-ray
DCT	Data Collection Tool
EC	Ethics Committee
eCRF	Electronic Case Report Form
EUCAST	The European Committee on Antimicrobial Susceptibility Testing
FiO2	Fraction of Inspired
FSFV	First Subject First Visit
HCP	Healthcare Practitioner
HIV	Human Immunodeficiency Virus
ICU	Intensive Care Unit
IEC	Independent Ethic Committee
ID	Identification
IPD	Invasive Pneumococcal Disease
IR	Incidence Rate

Abbreviation	Definition
IRB	Institutional Review Board
LSLV	Last Subject Last Visit
ISCIII	Instituto de Salud Carlos III
NIP	National Immunisation Program
OR	Odds Ratio
PCR	Polymerase Chain Reaction
PCL	Pfizer Central Laboratory
PCV13	13-valent Pneumococcal Conjugate Vaccine
PCV20	20-valent Pneumococcal Conjugate Vaccine
PI	Principal Investigator
PPV23	23-valent Pneumococcal Polysaccharide Vaccine
PSI	Pneumonia Severity Index
RIP	Regional Immunisation Program
RL	Reference Laboratory
RRI	Research Related Injury
RSV	Respiratory Syncytial Virus
SISPAL	"Sistema de Información en Salud Pública y Alimentación" (in Spanish)
SP	Streptococcus pneumoniae, *S. pneumoniae*
SpO2	Oxygen Saturation Level
SOC	Standard Of Care
TND	Test Negative Design
UAD	Urinary Antigen Detection
US	United States
VE	Vaccine Effectiveness
VT	Vaccine Type

Appendix 2. Data Derivation Details

Appendix 2.1. Serotypes

Central laboratories will perform pneumococcal serotyping for all participants with *S. pneumoniae* (*Sp*) isolated by culture.

VT is defined as serotypes included in PCV20, PCV13, PCV7 or the PPV23

PCV7 serotypes: 4, 6B, 9V, 14, 18C, 19F and 23F.

PCV13 serotypes: 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F, and 23F.

PCV20 serotypes: 1, 3, 4, 5, 6A, 6B, 7F, 8, 9V, 10A, 11A, 12F, 14, 15B, 18C, 19A, 19F, 22F, 23F, and 33F.

PPV23 serotypes: 1, 2, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F, 18C, 19F, 19A, 20, 22F, 23F and 33F.

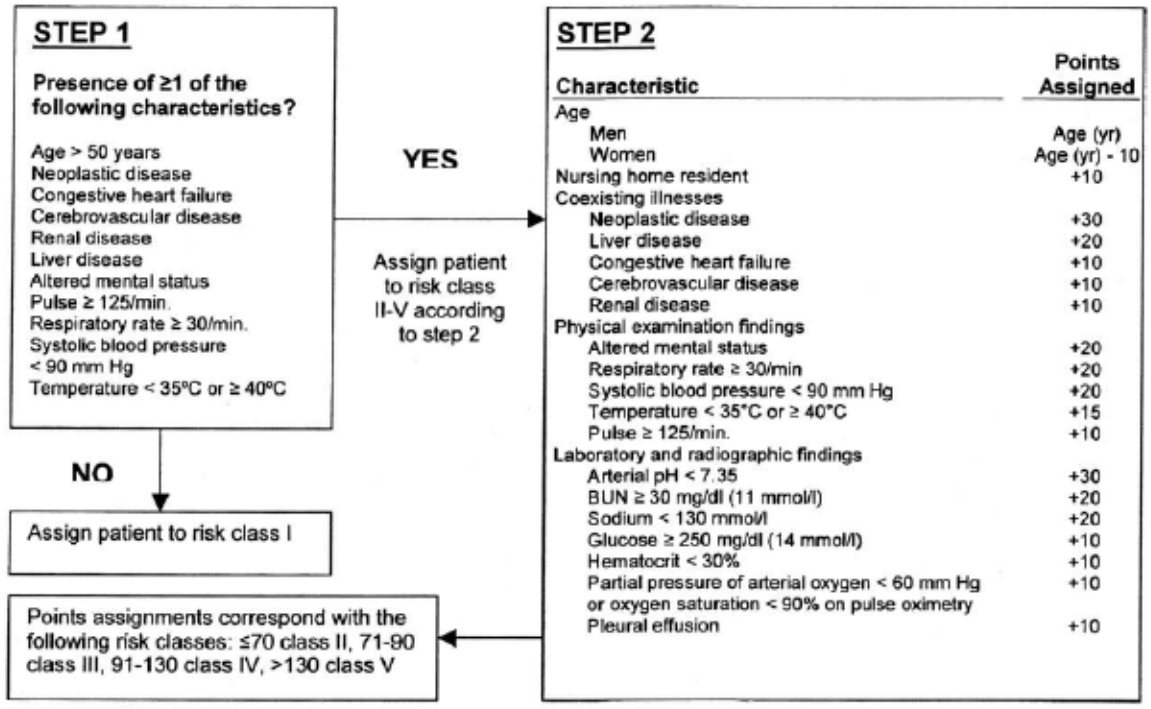
NVT are serotypes that are not included in any of PCV20, PCV13, PCV7, or PPV23.

Non-typeable or Untypeable (UT) is defined as a result which could not be serotyped. This is a valid result and so will be summarized as a separate line. Note that this is not the same as a missing serotype result.

Appendix 2.2. Body Mass Index (BMI) Calculation

Body Mass Index (BMI), also known as Quetelet index, is defined as a ratio of body weight in kilograms (kg) to the square of the body height in meters (m). So the unit of BMI is kg/m^2.: The BMI will be classified as a) Underweight (BMI<18.5), b) Normal (BMI:18.5-24.9), c) Overweight (BMI:25.0-29.9), and d) Obese (BMI≥30.0)

Appendix 2.3. Risk class assignment based on pneumonia severity index



Appendix 2.4. Risk Classification

High risk (immunocompromised)
A participant falls in any of the following medical conditions:

- Asplenia
- Cancer/Malignancy, Hematologic
- Cancer/Malignancy, Solid Tumor
- Chronic Kidney Disease
- Human Immunodeficiency Virus (HIV) – AIDS
- Human Immunodeficiency Virus (HIV) – No AIDS
- Immunodeficiency
- Immunosuppressant Drug Therapy
- Multiple Myeloma
- Organ Transplantation

At-risk (immunocompetent)
A participant does not fall in the high-risk category, but having any one of the following medical conditions:

- Alcoholism
- Asthma
- Celiac Disease
- Chronic Liver Disease with Hepatic Failure
- Chronic Liver Disease without Hepatic Failure
- Chronic Neurologic Diseases
- Chronic Obstructive Pulmonary Disease
- Coagulation factor replacement therapy
- Cochlear Implant
- Congestive Heart Failure

- Coronary Artery Disease (CAD)
- CSF Leak
- Diabetes Treated with Medication
- Down syndrome
- Living in a nursing home
- Living in a long-term care facility
- Occupational risk with exposure to metal fumes
- Other Chronic Heart Disease
- Other Chronic Lung Disease
- Other pneumococcal disease risk factors
- Previous Invasive Pneumococcal Disease
- Tobacco smoking (Tobacco/E-Cigarettes)

Low-risk (immunocompetent)
If a participant does not fall in high or at-risk category, then the participant will be classified in the low-risk category. These participants will also be treated as immunocompetent.

Risk factor	Spain's Categorization (Protocol Table 7)	eCRF capture	Characterization for SAP
Age	60 +	DOB and symptom onset date	At risk/ immunocompetent
Alcoholism	immunocompetent	Alcoholism	At risk/ immunocompetent
Chronic heart disease	immunocompetent	Congestive heart or failure or coronary artery disease (cad)/ Other chronic heart disease	At risk/ immunocompetent
Chronic liver disease	immunocompetent	Chronic liver disease without hepatic failure/ Chronic liver disease with hepatic failure	At risk/ immunocompetent

Chronic lung disease	immunocompetent	Chronic obstructive pulmonary disease/ Asthma/ Other chronic lung disease	At risk/ immunocompetent
Congenital or acquired asplenia	immunocompromised	Asplenia	High-risk/ immunocompromised
Cerebrospinal fluid leak	immunocompetent	CSF Leak	At risk/ immunocompetent
Cochlear implant	immunocompetent	Cochlear implant	At risk/ immunocompetent
Congenital or acquired immunodeficiencies	immunocompromised	Immunodeficiency	High-risk/ immunocompromised
Diabetes mellitus	immunocompetent	Diabetes treated with medication (not include diet controlled)	At risk/ immunocompetent
Generalize malignancy	immunocompromised	Cancer/malignancy, solid tumor	High-risk/ immunocompromised
HIV infection	immunocompromised	Human immunodeficiency virus (HIV) – AIDS/ Human immunodeficiency virus (HIV)	High-risk/ immunocompromised
Chronic renal failure	immunocompromised	Chronic kidney disease	High-risk/ immunocompromise
Nephrotic syndrome			
Leukemia	immunocompromised	Cancer/malignancy, hematologic	High-risk/ immunocompromise d
Hodgkin's disease			
Lymphoma			
Multiple myeloma			
Iatrogenic immunosuppression	immunocompromised	Immunosuppressant drug therapy	High-risk/ immunocompromised
Antineoplastic chemotherapy	not specified		
Solid organ transplant	immunocompromised	Organ transplantation	High-risk/ immunocompromised
Bone marrow transplantation			
Occupational risk with exposure to metal fumes	not specified	Occupational risk with exposure to metal fumes	At risk/ immunocompetent
Neurologic diseases (eg, cerebral palsy and seizures)	not specified	Chronic neurologic diseases	At risk/ immunocompetent
Prior IPD	immunocompetent	Previous invasive pneumococcal disease	At risk/ immunocompetent
Smoking	immunocompetent	Smoking or e-cigarettes current on substance abuse CRF	At risk/ immunocompetent

Items in Madrid column of risk classification grid but not in the Table 7 of the Protocol:

Risk factor	Spain's Categorization	eCRF capture	Characterization in SAP
Down syndrome	not specified	Down Syndrome	At risk/ immunocompetent
Celiac disease	not specified	Celiac disease	At risk/ immunocompetent
Cerebrovascular accident (Stroke)	not specified	Chronic neurologic diseases (specified to include this in CCG)	At risk/ immunocompetent
Coagulation factor replacement therapy	not specified	Coagulation factor replacement therapy	At risk/ immunocompetent
Institutionalized in nursing home or LTC facility	not specified	Nursing home or long-term care facility for those with disability or dependency on subject characteristics/risk determinants eCRF	At risk/ immunocompetent

Items on CRF that are not risk factors in Spain/Madrid:

Risk factor	Spain's Categorization	eCRF capture	Characterization in SAP
Other metabolic disease	Not present	Other metabolic disease	Not a risk factor
Diabetes that is diet controlled	Not present	Diabetes that is diet controlled	Not a risk factor
Pneumonia	Not present	Pneumonia	Not a risk factor

Not classified into any risk category

- Other metabolic disease
- Pneumonia
- Diabetes that is diet controlled

Relevant Guidance of Selected Past Medical History from CCG:

Medical history	Comments
Chronic Obstructive Pulmonary Disease	Also known as emphysema.
Asthma	--
Other Chronic Lung Disease	--
Congestive Heart Failure	--
Coronary Artery Disease (CAD)	--
Other Chronic Heart Disease	--
Diabetes That Is Diet Controlled	Diabetes mellitus controlled without use of diabetic medication.

Diabetes Treated with Medication	Diabetes mellitus treated with any diabetic medication (eg, oral hypoglycemics, insulin, or any others).
Other Metabolic Disease	--
Chronic Neurologic Diseases	Includes chronic neurologic diseases such as epilepsy, cerebral palsies, neuromuscular or cerebrovascular disease (such as history of cerebrovascular accident/stroke).
Chronic Liver Disease without Hepatic Failure	--
Chronic Liver Disease with Hepatic Failure	--
Other Pneumococcal Disease Risk Factors	Please check this if the subject has any disease risk factors other than the one listed in pre-medical history terms in this eCRF.
Chronic Kidney Disease	Includes Chronic Kidney Disease (defined as glomerular filtration rate <60 mL/min), End-Stage Renal Disease (ESRD), nephrotic syndrome, and patients requiring renal dialysis.
Immunodeficiency	Congenital or acquired immunodeficiencies, including complement deficiencies and others.
Human Immunodeficiency Virus (HIV) - AIDS	HIV-infected person with AIDS diagnosis or CD4<200 cells/µl.
Human Immunodeficiency Virus (HIV) - No AIDS	HIV-infected person who has not met AIDS definition.
Cancer/Malignancy, Hematologic	Includes Lymphomas (Hodgkin's disease and others), and Leukemias.
Cancer/Malignancy, Solid Tumor	Includes solid tumors and other generalized malignancies.
Organ Transplantation	Include solid organ, bone marrow transplantation, and hematopoietic stem cell transplantation.
Immunosuppressant Drug Therapy	Any immunosuppressant drug therapy that would have reduced immune function prior to pneumonia onset, including immunosuppressing antineoplastic drug therapy and other chronic treatment with known immunosuppressant medications, including the use of systemic corticosteroids (equivalent of ≥10 mg/day of prednisone) for >14 days within 30 days prior to pneumonia event onset.
Previous Invasive Pneumococcal Disease	Pneumococcal infection from a normally sterile site such as blood, pleural fluid and other normally sterile body fluids, and internal organ/tissue cultures.
Cochlear Implant	--
CSF Leak	Cerebrospinal fluid leak.
Celiac Disease	--
Down Syndrome	--
Asplenia	Includes congenital or acquired asplenia.
Pneumonia	Previous pneumonia events in year prior to admission.
Alcoholism	--

Occupational risk with exposure to metal fumes	--
Coagulation factor replacement therapy	--
Multiple Myeloma	--

Appendix 3. STATISTICAL METHODOLOGY DETAILS

Appendix 3.1. Calculation of Confidence of Interval for the Proportion of Interest

Some of the endpoints of this study may be summarized along with exact 95% confidence intervals, which will be calculated based on the Clopper-Pearson method. The method is based on F distribution. If r equals the number of responses and n equals the number of subjects, then it follows that $p = r / n$ is the estimate of the proportion of responses. An exact 95% confidence interval can then be computed by solving the following 2 equations. For the lower limit p_L, use

$$p_L = \frac{rF_L}{(rF_L + (n-r+1))}$$ and for the upper limit p_U, use $$p_U = \frac{(r+1)F_U}{(n-r)+(r+1)F_U}$$

where F_L is the quantile from the F distribution for α=0.025, with numerator degrees of freedom equal to 2r and denominator degrees of freedom equal to 2(n-r+1). F_U is the quantile from the F distribution for α=0.975, with numerator degrees of freedom equal to 2(r+1) and denominator degrees of freedom equal to 2(n-r). When r equals 0, F_L should be set equal to 1.0 so p_L equals 0. When r equals n, F_U should be set equal to 1.0 so p_U equals 1. The confidence interval using the F distribution is described in Collett (1991) and implemented in SAS PROC FREQ.

Appendix 3.2. Confidence Interval for the Incidence Rates

The incidence rate will be calculated per 100,000 population in that specific age group. Confidence intervals for the incidence rate will be based on the assumption that the actual count of cases arises from a Poisson distribution. The approach used in Daly et al describing the calculation of exact confidence limits will be used. If x denotes the actual count of cases, $(1 - \alpha)$ the confidence limit, x_L the lower confidence limit and x_U the upper confidence limit, then the following formula yields values for x_L and x_U,

When $x > 0$,

$$x_L = 0.5 * \chi^2(\alpha/2;\ 2x) \text{ and } x_U = 0.5 * \chi^2(1-\ \alpha/2;\ 2x+2),$$

When $x = 0$,

$$x_L = 0 \text{ and } x_U = -\ln(1 - \alpha/2).$$

In the above notation, $\chi^2(p;\ n)$ represents the p^{th} quantile of the chi-square distribution with n degrees of freedom and ln(.) represents the natural logarithm function.

Once x_L and x_U are obtained, they are to be expressed in per 100,000 population.

Appendix 3.3. Calculation of odds ratio

To measure the effectiveness of the PCV13 and PPV23 against the target outcomes, the multivariable logistic regression model will be used Coefficient of the vaccine variable in the model will be exponentiated to estimate the odds ratio. To address the effect of the confounding

variables, the variables that were independently associated with the outcome at $p<0.10$ in a bivariate analysis will be selected. When selecting the covariates for the final model, the rule of 10 events per covariate (Peduzzi et al 1996) will be followed. The vaccine effectiveness (VE) will be calculated as

$$VE = 1 - \text{adjusted odds ratio} \times 100.$$

Appendix 3.4. Test for a Difference in Two Odds Ratios

The following steps will be performed to test the difference between the two odds ratio

Calculate the standard error of the log odds ratio use this formula:

$$SE(logOR)=SQRT(1/n1+1/n2+1/n3+1/n4)$$

where n1, n2, n3, n4 are the case-exposed, case-unexposed, control-exposed, and control-unexposed.

To get the p-value for the difference follow these steps:

1. Take the absolute value of the difference between the two log odds ratios. We will call this value δ.
2. Calculate the standard error for δ, SE(δ), using the formula: $SQRT(SE_1^2+SE_2^2)$
3. Calculate the Z score for the test: $z=\delta/SE(\delta)$
4. Calculate the p-value from the z score.

The p-value will be calculated using R or Microsoft Excel using the below formulas.

R: P-value=2*(1-pnorm(Z))
MS Excel: P-value=2*(1-(NORMDIST(Z,0,1,TRUE)))

Appendix 4: Definition of high-quality respiratory specimens

High quality respiratory text specimens will include: pleural fluid, transtracheal aspirate, bronchoalveolar lavage, or sputum meeting Murray/Washington criteria as outlined below:

Murray and Washington Grading System for Sputum

	Epithelial cells per low power field	Leucocytes per low power field
Grade 1	>25	<10

Grade 2	>25	10-25
Grade 3	>25	>25
Grade 4	10-25	>25
Grade 5	<10	>25

Note that the good-quality specimens are those with ≤10 squamous epithelium and >25 leukocytes per low power field

Appendix 5: Conversion SaO2/PaO2

SpO2 (%)	Calculated PaO2 (mmHg)
80	44
81	45
82	46
83	47
84	49
85	50
86	52
87	53
88	55
89	57
90	60
91	62
92	65
93	69
94	73
95	79
96	86
97	96
98	112
99	145

Appendix 6. Charlson Co-morbidity Index