Statistical Analysis Plan

BCG-PRIME

Title:	BACILLUS CALMETTE-GUÉRIN VACCINATION TO PREVENT SERIOUS RESPIRATORY TRACT INFECTION AND COVID-19 IN VULNERABLE ELDERLY – AN ADAPTIVE RANDOMIZED CONTROLLED TRIAL
Sponsor:	University Medical Center Utrecht (UMCU)
SAP version - date:	V1.0 of 13 January 2021

SAP authors:

Sponsor's representative, Coordinating principle investigator	Marc J.M. Bonten, MD, PhD Dpt. of Medical Microbiology University Medical Center Utrecht Heidelberglaan 100, 3584 CX Utrecht, the Netherlands Tel: +31-75-57676 E-mail: m.j.m.bonten@umcutrecht.nl
Trial statistician / Coordinating investigator	Cornelis H. van Werkhoven, MD, PhD Julius Center for Health Sciences and Primary Care University Medical Center Utrecht Heidelberglaan 100, 3584 CX Utrecht, the Netherlands Tel: +31-88-75-68171 E-mail: c.h.vanwerkhoven@umcutrecht.nl

Signature page

Title: BACILLUS CALMETTE-GUÉRIN VACCINATION TO PREVENT SERIOUS RESPIRATORY TRACT INFECTION AND COVID-19 IN VULNERABLE ELDERLY – AN ADAPTIVE RANDOMIZED CONTROLLED TRIAL

Document release - date: Statistical analysis plan V1.0 of 13 January 2021

Approval page

Marc. J.M. Bonten, MD, PhD **Coordinating principle investigator** Dpt. of Medical Microbiology University Medical Center Utrecht Date: 13-01-2021 Signature: 	**Cornelis H. van Werkhoven, MD, PhD** **Trial statistician** Julius Center for Health Sciences and Primary Care University Medical Center Utrecht Date: 13-01-2021 Signature: 

Table of contents

1 List of abbreviations

ADL:	Activities of Daily Living
BCG:	Bacillus Calmette-Guérin
CI:	Confidence interval
COVID-19	Corona Virus Disease 2019: Disease associated with infection with SARS-CoV-2 (please refer to the study definition in section 3.3)
crRTI	Clinically relevant respiratory tract infection
SARS-CoV-2	Severe acute respiratory syndrome coronavirus-2: The virus that causes COVID-19
eCRF:	electronic Case Report Form
ICU:	Intensive Care Unit
IQR:	Inter Quartile Range
RCT:	Randomised Controlled Trial
SD:	Standard deviation
UMCU:	University Medical Center Utrecht
VAS:	Visual Analogue Scale

2 Overview of the study

This is a copy of the protocol summary, version 2.1, 12-Nov-2020:

Rationale: On March 11 2020 the World Health Organization (WHO) declared the coronavirus (SARS-CoV-2) outbreak a pandemic. Worldwide, the number of confirmed cases continues to rise, leading to significant morbidity and mortality. In the Netherlands, although the incidence is currently low due to social distancing measures, recurrence of infections is expected once measures are going to be lifted. Although individuals of any age can acquire SARS-CoV-2, adults of middle and older age are at highest risk for developing severe COVID-19 disease. Moreover, recent reports demonstrate that mortality rates rise significantly among patients 60 years and older. Therefore, strategies to prevent SARS-CoV-2 infection or to reduce its clinical consequences in vulnerable populations are urgently needed. Bacille Calmette-Guérin (BCG) vaccine not only protects against tuberculosis, but also induces protection against various respiratory infections, including those with a viral etiology. We hypothesize that BCG vaccination reduces clinically relevant respiratory tract infections requiring medical intervention, including COVID-19, in vulnerable elderly.

Objective: To determine the impact of BCG vaccination on the incidence of clinically relevant respiratory infections or COVID-19 in vulnerable elderly.

Study design: An adaptive multi-center double-blind randomized placebo-controlled trial.

Study population: 5,200 to 7,000 vulnerable elderly, defined as ≥60 years of age and being discharged from hospital in the last 6 weeks, having a chronic somatic comorbidity, undergoing major surgery, or attending a thrombosis care service. Patients with contraindications to BCG vaccination as stipulated in the Summary of Product Characteristics (SPC) and patients with a history of COVID-19 will be excluded.

Intervention: Participants will be randomized between intracutaneous administration of BCG vaccine (Danish strain 1331) or placebo (0.1ml 0.9% NaCl) in a 1:1 ratio.

Main study parameters/endpoints: The trial has an adaptive primary endpoint. Based on accrual of the two endpoints, the primary endpoint will be either (a) COVID-19 or (b) clinically relevant respiratory tract infection (RTI) requiring medical intervention, potentially including COVID-19 episodes. The other will be declared secondary endpoint. Other secondary endpoints include: all SARS-CoV-2 infections (including asymptomatic infections), influenza infection, RTI (all infections regardless of medical intervention), RTI-related hospital admission, COVID-19 related hospital admission, pneumonia, mental, physical and social functioning, serious adverse events and adverse events, and death.

Nature and extent of the burden and risks associated with participation, benefit and group relatedness: Based on previous experience in published and ongoing randomized controlled trials in adults and elderly, the risks of BCG vaccination are considered negligible. Vaccination could cause local pain, suppuration and scarring at the site of injection. The size of the trial will ensure rapid availability of results that can inform policy-making during the ongoing pandemic and may be of benefit for participants that received no intervention, as well as individuals who did not participate in the trial.

3 Sample size

This is an event driven trial. To demonstrate a 50% reduction of the incidence of the primary endpoint with a two-sided alpha of 0.05, 67 endpoints are needed for 80% power and 90 events for 90% power.

Two candidate primary endpoints are selected: COVID-19 and clinically relevant RTI. For the primary endpoint COVID-19 we consider 80% acceptable given the high need for a fast answer to whether BCG is efficacious. For clinically relevant RTI 90% power is selected. The final primary endpoint will be selected based on the accrual of number of endpoints.

Given the urgent medical need, we consider a sample size of 7,000 to provide the optimal balance between number of participants and expected duration of the trial, while a sample size of 5,200 would be minimally acceptable. For a detailed sample size calculation, please refer to the protocol.

4 Definitions

4.1 Treatment groups

BCG: participants randomised to receive BCG

Placebo: participants randomised to receive placebo

4.2 Analysed populations

The primary anlysis population is the intention-to-treat (ITT) population in which participants will be analysed according to the random allocation.

Subgroups:

We will explore the presence of interaction between BCG and other vaccines, in particular influenza and pneumococcal vaccine.

4.3 Study definitions

COVID-19 disease is defined as respiratory tract symptoms and/or systemic symptoms in combination with detection of SARS-CoV-2 in a respiratory sample.

SARS-CoV-2: severe acute respiratory syndrome coronavirus-2. The virus that causes COVID-19 disease.

crRTI: clinical symptoms of a respiratory tract infection in combination with the need for medical intervention. (See protocol for a detailed definition.)

5 General statistical considerations

5.1 General statistical methodology

The methodology presented in this statistical analysis plan takes into account methods planned in the protocol, specifies them, and completes them. The analysis will be performed in the statistical software R.(1) The version of R and R libraries used will be reported.

5.2 Handling of missing data

For analyses of primary or secondary endpoints with missing data, multiple imputed data analysis will be performed using Multivariate Imputation by Chained Equations; complete data analysis will be performed as a sensitivity analysis to check for inconsistencies.(2) Missing outcome data will not be imputed.

5.3 Descriptive statistics

According to the variable type, descriptive statistics will be:

- Quantitative criteria: number of observations (N), mean, standard deviation (SD), median and inter-quartile range (IQR), as applicable.
- Qualitative criteria: number of observations (N), absolute frequency (n) and relative frequency (%). Percentages will be calculated on the number of participants with documented data.

5.4 Inferential statistics

Effect estimates for primary and secondary endpoints will be reported with 95% CI. Effect estimates with a 95% CI not including the zero-effect will be considered statistically significant.

6 Disposition of participants and conduct of the study

6.1 Participant disposition and size of the analysed populations

The study flowchart will describe:

- Number of subjects screened
- Number of subjects not meeting the eligibility criteria
    - Reasons for non-eligibility
- Number of subjects eligible but not included
    - Reasons for non-inclusion
- Number of participants included and randomised
- Early termination: Number of participants withdrawn, dropouts
    - Reasons for early termination

6.2 Conduct of the study

The following will be reported on the conduct of the study:

- Enrolment period: date First Patient First Visit – date Last Patient First Visit
- Distribution of actual duration of follow-up per participant by treatment arm
- Number and proportion of days with missing outcome measurement by treatment arm (e.g. due to dropout or noncompleted questionnaires)

7 Demographics and other baseline characteristics

7.1 Demographics / baseline characteristics

The following data will be described for included participants by treatment arm and for the total population:

VARIABLE	DERIVATION	STATISTICAL METHOD *
Age (years)	Self-reported at baseline	Median (IQR) or mean (sd) [1]
Gender	Self-reported at baseline: Male/ Female	n/N (%)
History of BCG vaccination	Self-reported at baseline: Yes/No	n/N (%)
Age at BCG vaccination	Self-reported at baseline	Median (IQR) or mean (sd) [1] among those with reported history of BCG vaccination
BMI	Self-reported at baseline	Median (IQR) or mean (sd) [1]
Receipt of influenza vaccination in past season (2019-2020 or 2020-2021 depending on date of enrolment)	Self-reported at baseline: Yes/No	n/N (%)
Receipt of pneumococcal vaccination in past or current season	Self-reported at baseline: Yes/No	n/N (%)
Receipt of other vaccinations in past 12 months	Self-reported at baseline (open question) [2]	n/N (%) for each reported vaccination
History of hypertension	Self-reported at baseline: Yes/No	n/N (%)
History of cardiovascular disease	Self-reported at baseline: Yes/No	n/N (%)
History of stroke	Self-reported at baseline: Yes/No	n/N (%)
History of diabetes mellitus	Self-reported at baseline: Yes/No	n/N (%)
History of COPD	Self-reported at baseline: Yes/No	n/N (%)
History of asthma	Self-reported at baseline: Yes/No	n/N (%)
History of other pulmonary diseases	Self-reported at baseline: Yes/No and open question [2]	n/N (%) for each reported pulmonary disease if deemed relevant
History of moderate to severe renal disease	Self-reported at baseline: Yes/No	n/N (%)
History of solid organ or haematological malignancy	Self-reported at baseline: Yes/No	n/N (%)
History of dementia (diagnosed by physician)	Self-reported at baseline: Yes/No	n/N (%)
History of other chronic diseases	Self-reported at baseline: Yes/No and open question [2]	n/N (%) for each reported disease if deemed relevant
Medication use	Self-reported at baseline	Listing of medication on group level based on ATC code system
Clinical Frailty Scale	Self-reported at baseline	Median (IQR)

History of smoking	Self-reported at baseline. Options are: - Never smoked - Ever smoked - Current smoking	n/N (%) for all categories
Smoking of household member	Self-reported at baseline: Yes/No	
Alcohol use	Self-reported at baseline: Yes/No, quantity in categories	n/N (%) per category
History of allergic rhinitis	Self-reported at baseline: Yes/No	n/N (%)
Other allergies	Self-reported at baseline: Yes/No and free text	n/N (%) per category if deemed relevant
Type of home	Self-reported at baseline: - Independent - Sheltered 11ccommodation - Care home - Nursing home	n/N (%) per category
Presence of household members	Self-reported at baseline: Yes/No	n/N (%)
Education level	Self-reported at baseline, categories	n/N (%) per category
Payed work	Self-reported at baseline: - Yes - Unemployed, searching for a job - Unemployed, not actively searching - Retired	n/N (%) per category
Working environment	Self-reported at baseline: - Patient care - Day-care - Primary or secondary school - Vocational education or highschool - None of the above	n/N (%) per category
Receives domestic help by indication	Self-reported at baseline: Yes/No	n/N (%)
Receives home care	Self-reported at baseline: Yes/No	n/N (%)
Dependence in ADL	Self-reported at baseline: Katz-ADL, Lawton IADL	Median (IQR) or mean (sd) [1]
Quality of life / utilities	Self-reported at baseline, EQ5D3L	Median (IQR) or mean (sd) [1]
Health status	Self-reported at baseline, VAS	Median (IQR) or mean (sd) [1]

* It will not be statistically tested whether baseline characteristics differ by treatment arm.

1. The description method will depend on the normality of the data.

2. Open questions will be categorized by the coordinating investigator.

8 Analysis of the primary endpoint

The trial has an adaptive primary endpoint. Based on accrual of primary endpoints the primary endpoint will be selected as soon as one of the candidate primary endpoints has reached the required number of episodes. The non-selected candidate primary endpoint will be analysed as secondary endpoint.

The two primary endpoints concerned are: (see protocol chapter 7 for definitions)

a) COVID-19 (67 episodes for 80% power)

b) Clinically relevant RTI (90 episodes for 90% power)

Only first episodes per participant will be considered.

The primary study parameter will be analyzed as a time-to-event outcome using a Cox proportional hazards analysis. The randomization date will be considered T=0 and the symptom onset date of the primary endpoint (or in case of COVID-19 the date of positive test result if that occurs earlier) will be considered the event time. Censoring time of participants with no event will be the last date of follow-up. The analysis will be adjusted for the stratification variables (hospital and age category) and for strong prognostic factors*. The result will be reported as the number of episodes per arm, number of patient-years of follow-up per arm, and hazard ratio (HR) with 95% CI.

* For COVID-19 there are no prognostic factors that can be measured at enrolment so only stratification variables will be used. For clinically relevant RTI: presence of the following (groups of) comorbidities: cardiovascular disease or stroke, chronic pulmonary disease (COPD, asthma or other chronic pulmonary disease), diabetes mellitus, solid organ or hematological malignancy, moderate to severe chronic renal disease.

9 Analysis of the secondary endpoints

Please refer to the protocol for definitions of each endpoint.

9.1 Cumulative incidence of SARS-CoV-2 infection regardless of symptomatology

Analysis will be the same as for the primary endpoint COVID-19.

9.2 Cumulative incidence of asymptomatic, mild/moderate, and severe (requiring hospitalisation) SARS-CoV-2 infection.

Analysis will be the same as for the primary endpoint COVID-19.

9.3 Cumulative incidence of influenza infection

Analysis will be the same as for the primary endpoint COVID-19.

9.4 Cumulative incidence of RTI

Analysis will be the same as for the primary endpoint COVID-19.

9.5 Cumulative incidence of medically attended RTI

Analysis will be the same as for the primary endpoint COVID-19.

9.6 Cumulative incidence of RTI related hospital admission

Analysis will be the same as for the primary endpoint COVID-19.

9.7 Cumulative incidence of pneumonia diagnosed by a GP or medical specialist

Analysis will be the same as for the primary endpoint COVID-19.

9.8 Mental, physical and social functioning at 6 months

The change in Katz-Activities of Daily Living score, Lawton Activities of Daily Living score and EQ5D quality of life score between baseline and end of study will be analysed using linear regression. The stratification variables (hospital and age category) will be added as covariates to the model. The results will be reported as differences in change of score with 95% confidence intervals.

9.9 Serious adverse events and adverse events

Frequencies of AEs, SAEs and SUSARs will be described per intervention arm for total number of events, per solicited event, and for unsolicited events. Evemts will be grouped using the MedDRA classification system. Frequencies will be compared using logistic regression analysis and reported as OR with 95% CI.

9.10 Cumulative incidence of all-cause 6-month mortality

Analysis will be the same as for the primary endpoint COVID-19.

10 Analysis of the exploratory endpoints

10.1 Interaction between BCG and other vaccines

Primary and secondary analyses will be stratified for status of vaccination with other vaccines at baseline, in particular influenza and pneumococcal vaccine. If the data allow, an additional analysis will be performed using a time-varying influenza and pneumococcal vaccination status with interaction to determine the impact of receiving these vaccines *after* BCG vaccination.

10.2 Immunological endpoints (conditional on funding)

Analysis of immunological endpoints are not specified at this stage. The analysis will be described in the publication reporting on these endpoints.

11 References

1. R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
2. Van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. J Stat Softw. 2011 Dec;45(3):1–67.