Statistical Analysis Plan

ALL-IN-META-BCG-CORONA

Title:	Anytime Live and Leading Interim* meta-analysis of the impact of Bacillus Calmette-Guérin vaccination in healcare workers during the SARS-CoV-2 pandemic
Project partners:	UMCU / CWI / UvA
SAP version - date:	V2.0 of 19 September 2022 Updated to include the new definition of participants at risk and of Covid-19 infection following the advisory committee meeting on April 23rd 2021 to also include detection using lung CTs and antigen tests (see 2021-05_Newsletter_3_ALL-IN-META-BCG-CORONA section **Trial event definitions** for this decision). Any omissions from v1.0 in this v2.0 of the SAP are crossed instead of deleted.

SAP authors:

Coordinating investigator	C.H. (Henri) van Werkhoven Julius Center for Health Sciences and Primary Care University Medical Center Utrecht Heidelberglaan 100, 3584 CX Utrecht, the Netherlands E-mail: c.h.vanwerkhoven@umcutrecht.nl
Trial statisticians	Judith ter Schure Machine Learning group Centrum Wiskunde & Informatica (CWI) (the national research institute for mathematics and computer science in the Netherlands) Science Park 123, 1098 XG Amsterdam, the Netherlands E-mail: j.a.ter.schure@cwi.nl Alexander Ly Machine Learning group Centrum Wiskunde & Informatica (CWI) (the national research institute for mathematics and computer science in the Netherlands) Science Park 123, 1098 XG Amsterdam, the Netherlands E-mail: alexander.ly@cwi.nl

* ALL-IN meta-analysis stands for Anytime Live and Leading INterim meta-analysis. This statistical methodology provides meta-analysis of interim data that can be updated at Anytime – so after each new observed data point – Live – so without the need to prespecify when to look – and Leading – because the evidence in the meta-analysis can inform whether individual trials should be stopped or expanded.

This Statistical Analysis Plan proposes to perform ALL-IN meta-analysis for time-to-event data, using the Safe logrank test and Anytime-valid hazard ratio confidence sequences. The analysis is literally ALL-IN: any new trial result can be included, even if the decision to start or expand a trial was based on the data already available. For details, see: https://projects.cwi.nl/safestats/

Signature page

Title:	Anytime Live and Leading Interim meta-analysis of the impact of Bacillus Calmette-Guérin vaccination in healthcare workers during the SARS-CoV-2 pandemic
Document release - date:	Statistical analysis plan V2.0 of 19 September 2022

Approval page *V1.0 signed on 17-06-2020, signatures not available in web version*

Henri van Werkhoven **Coordinating principle investigator** Julius Center for Health Sciences and Primary Care UMC Utrecht Date: Signature:	**Judith ter Schure** **Trial statistician** Machine Learning group CWI Date: Signature: **Alexander Ly** **Trial statistician** Machine Learning group CWI Date: Signature:

Table of contents

List of abbreviations

ALL-IN:	Anytime Live and Leading INterim
BCG:	Bacillus Calmette-Guérin
RCT:	Randomised Controlled Trial
RR:	Relative Risk
HR:	Hazard Ratio

1 Overview of the study

Since the start of the SARS-CoV-2 pandemic, several investigators across the world have set up RCTs of BCG vaccination among healthcare workers. The hypothesis is that the aspecific immune response to BCG provides indirect protection against COVID-19 or against a severe course of COVID-19. Until today, over 10 trials are ongoing or have completed. This is great, because if BCG is effective, more trials will ensure that the putative effect will be found, probably earlier and with a more precise estimate of the magnitude of the effect. However, with multiple parallel trials there is an increased risk of false-positive findings, due to multiple testing, which could endanger the completion of the other trials.

If BCG does not prevent COVID-19 (or not to a meaningful extent), each trial needs to limit its chance to incorrectly find superiority: the type-I error or false-positive rate which is normally set at 5%. However, if 10 trials are performed, the chance that at least one of them will be false-positive increases to 40%. Normally, this is not so much of a problem: the majority of trials will not be positive and in the end the conclusion is that there is no effect. However, in this pandemic setting, the collateral damage of a single false-positive finding can be substantial.

We expect that most trials, if not all, will perform interim analyses. The benefit of this is that superiority – if present – will be found earlier, which is important during the pandemic. However, the downside is that, if there is no effect, it is still quite likely that at least one trial will be positive. Since this is found during an interim analysis, this positive trial will most likely be the first to release results. In response to a (possibly false) positive result, investigators of all other trials will need to decide whether or not to stop their trial for ethical reasons and offer BCG to the control group. Also, it may be decided to offer BCG to COVID-19 risk groups, causing global shortage of BCG. All of this might be in vain if the result acted on is indeed a false-positive. Pooling the evidence across all trials avoids such a scenario. Moreover, prospectively pooling the accruing evidence across all trials has the advantage of increased power, making it likely that a conclusion is reached earlier than in a single trial, when the effect exists.

Normally it will take months to collect and analyse the data of all trials and perform a meta-analysis. In this meta-analysis, central collection of selected data needed for the meta-analysis will already be performed while the trials are ongoing. Data will be analysed on a regular basis using statistical tests and stringent criteria that control false-positive rates. Results will be released as soon as a pooled statistically significant effect is observed or an individual trial yields a statistically significant effect, or when all trials have been completed.

2 Definitions

2.1 Treatment groups

BCG: participants randomised to receive BCG

Control: participants randomised to receive no BCG (either no intervention or placebo)

2.2 Analysed populations

Participants will be analysed according to the intention-to-treat (ITT) principle, i.e. according to the random allocation.

~~**Subgroups:**~~

~~Participants with documented COVID-19 (time-to-recovery analysis as secondary analysis)~~

~~Participants with COVID-19 related hospitalization (time-to-recovery analysis as secondary analysis)~~

2.3 Study definitions

~~Documented COVID-19 disease is defined as PCR-based detection of SARS-CoV-2 in a respiratory sample.~~

Documented COVID-19 disease is defined as infection with (1) PCR-based detection of SARS-CoV-2 in a respiratory sample, (2) detection by lung CT, or (3) detection by antigen test.

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

Participants at risk

Participants are considered at risk of Covid-19 infection and hospitalization from the date of randomization to the date of either a Covid-19 infection/hospitalization, the end of follow-up, loss to follow-up or date of Covid-19 specific vaccination. So follow-up time is censored at the date of Covid-19 specific vaccination for some of the participants and infections after Covid-19 vaccination are not considered as events, and neither are reinfections.

3 General statistical considerations

3.1 General statistical methodology

The analysis will be performed in the statistical software R.[1] The R-library safestats will be used for most inferential statistics. The version of R and R-libraries used will be reported.

3.2 Handling of missing data

Due to the nature of the analysis, missing data are not expected, except in terms of loss-to-follow-up (e.g. withdrawal of participants). Participants with incomplete follow-up will be assumed to have noninformative censoring at the end of their follow-up.

3.3 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. Incidence density will be reported (e.g., number of events per 1000 follow-up days) as appropriate.
- Kaplan Meier survival curve for time-to-event endpoints.

3.4 Inferential statistics

Safe tests will be used for inferential statistics. Safe tests are designed to control type-I error rates, while maximising flexibility in timing and frequency of interim analyses and allowing for the continuation of data collection beyond the planned trial duration. An explanation of the Safe logrank test for time-to-event data is provided here: https://projects.cwi.nl/safestats/.

3.5 Interim analyses and publication of the results

Interim analyses of the co-primary endpoint will be performed frequently. Results of the interim analyses will not be published or reported to the trial investigators except in case of one of the following three events:

1. The interim meta-analysis of one or both of the co-primary endpoints yields a statistically significant effect (see section 5 for definition).
2. One of the trials ends with a conclusion. In this case, unveiling the meta-analysis results serves the purpose of confirming or refuting the trial results in order to facilitate an informed decision by all trial investigators and policy makers. This event can happen multiple times as long as trials are ongoing.
3. All trials have been completed and no new trials are being started.

4 Description of included RCTs

We will describe the following characteristics of the included RCTs:

Meta-data:

- Country/countries in which the trial has enrolled participants
- Health care setting
- Start date of randomization, date of last follow-up

Participant data:

- Number of participants included and randomised to BCG and control (N)
- Number and percentage of participants with incomplete follow-up for BCG and control (N, %)
- Distribution of actual duration of follow-up per participant for BCG and control (Median, IQR)

5 Analysis of the primary endpoints

Both primary endpoints will be analysed as time-to-event outcome. A Safe logrank test will be used stratified by trial and hospital. The time-scale for the analysis will be calendar time, setting the date of first enrolment of the first trial to t = 0 for all. As a consequence, participants being enrolled later will be left-truncated. This is to ensure that the epidemic wave is on the same time scale for all participants in the same site, thus making the hazard function more homogeneous.

5.1 Cumulative incidence of COVID-19 disease, weight: 10%

The Safe test for superiority is designed to be a one-sided test for Hazard Ratios (HR) *less than* 1 in favour of the treatment, with alpha = 0.25%, and optimised for a minimal clinically relevant HR of 0.8.

The Safe test for harm is designed to be a one-sided test for HR *greater than* 1 in favour of the control, with alpha = 0.25%, and optimised for a minimal clinically relevant HR of 1.25.

Results will be reported as a descriptive HR with a 95% two-sided Safe confidence sequence. We will also report a 99.5% two-sided Safe confidence sequence for consistency with the chosen alpha for statistical significance.

5.2 Cumulative incidence of hospital admission due to COVID-19 disease, weight: 90%

The Safe test for superiority is designed to be a one-sided test for HR *less than* 1 in favour of the treatment, with alpha = 2.25%, and optimised for a minimal clinically relevant HR of 0.7.

The Safe test for harm is designed to be a one-sided test for HR *greater than* 1 in favour of the control, with alpha 2.25%, and optimised for a minimal clinically relevant HR of 1.43.

Results will be reported as a descriptive HR with a 95% two-sided Safe confidence sequence. We will also report a 95.5% two-sided Safe confidence sequence for consistency with the chosen alpha for statistical significance.

6 Analysis of the secondary endpoints

~~The following secondary endpoints will be analysed and reported together with the primary endpoints when results are published:~~

~~6.1 Time-to-recovery from documented COVID-19 disease~~

~~6.2 Time-to-recovery from hospital admission due to COVID-19 disease~~

7 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. https://projects.cwi.nl/safestats/