Central monitoring plan SafeBoosC III

Version	Author(s)	Date	Changes	Approved by
1.0	Mathias Lühr Hansen Gorm Greisen Janus Christian Jakobsen Christian Gluud	25-nov 2019	Initial version	SafeBoosC Steering committee

Purpose

The purpose of the central monitoring plan is to optimise the data quality including 1) minimising missing data, and 2) ensuring that data are entered correctly. The purpose is not safety monitoring, as this is a task for the Data Safety and Monitoring Committee.

Missing data

Regarding missing data there will be zero tolerance. Once a month, site data completion will be calculated and reported for each centre. An overview of the site data completion rate within each site, will be posted in the monthly newsletters. Missing data monitoring will start after 100 babies have been randomised.

Site data completion

Site data completion is defined as all data forms completed before the relative data entry deadline (a) over all data forms that should have been completed (b) and reported in percentage

$$\text{Site data completion rate} = a/b * 100\ \%$$

Each baby included in the trial has five data forms:

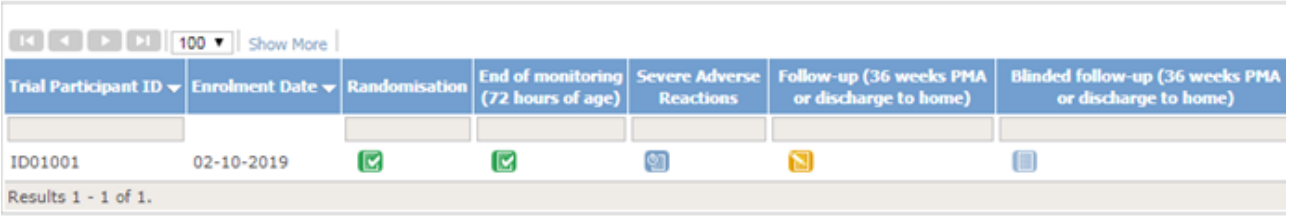
1. Randomisation form (data entry deadline at six hours from birth) (not relevant since always completed)
2. End-of-monitoring form (data entry deadline at 72 hrs of age + 7 days)
3. Severe adverse reactions (not relevant since only very few data entries are expected)
4. Follow-up form (data entry deadline at 36 weeks PMA + four weeks)
5. Blinded follow-up form (data entry deadline at 36 weeks PMA + four weeks)

This means that the total number of available data forms for a site can be calculated as:

Available data forms = number of randomised babies*3

To complete a data form, the investigator must 1) perform all data entries and 2) press “submit”. If a data form has *not* been completed within the data entry deadline, it will be classified as missing data.

Example
Site ID01 started randomising 1/10 2019, and the first included baby (IB01001) was born 2/10 2019, gestational age 27+0 and randomised to the experimental group. The date is now 14/10 2019 and the status in OpenClinica is this:



As for now, only one baby (ID01001) has been randomised at site ID01.

Green means that data entry is completed, yellow that it has begun (but not completed) and blue that it has not begun yet. Let us do a manual site data completion calculation: first we need to establish which dates the different forms should be completed, based on the above information on data entry deadlines. As stated above, randomisation- and Severe Adverse Reaction form is not relevant.

- *End-of-monitoring form: 12/10 2019 (completed)*
- *Follow-up form: 1/1 2020 (begun, data entry deadline has not been exceeded)*
- *Blinded follow-up form: 1/10 2020 (not begun, data entry deadline has not been exceeded)*

This means that, when calculating site data completion 14/10 2019 for this site, only the end-of monitoring form should be accounted for. Thus, the site data completion is 100% (one form is relevant and completed, two is not relevant yet).

If there had been more babies included at site ID01, they should also have been included in the calculation.

Example

If four babies have been randomised and three of them have completed all three data forms within the data entry deadline, but the third one has exceeded data entry deadline for all three data forms, the site data completion rate would be:

9/12 = 75% (9 forms completed out of 12 forms, i.e. three forms per baby).

For practical reasons, these calculations will be done automatically by a computer once a month.

Individual monitoring and reporting of incomplete data forms

Besides calculating and reporting site data completion rates for each site, we will also identify the study ID for babies with missing data including which data forms that are incomplete and have exceeded the data entry deadline and report this information to the local principal investigator.

Correct and realistic data

Primary outcome

Since the trial is pragmatic and is run with a minimum of resources, we will only monitor the primary outcome at the level of individual sites, using a probabilistic approach. We will calculate the prevalence of the primary outcome (death or survival with severe brain injury) for the total number of babies at each site after 10, 15, 20 and 25 babies have reached 36 weeks postmenstrual age and completed data entry for the primary outcome. This will be done once a month, simultaneously with the calculation and reporting of missing data. Correct and realistic data monitoring will begin after 100 babies have reached 36 weeks of postmenstrual age.

If an average site is expected to have a 30% prevalence of the primary outcome [1], we can also calculate an expected 95% confidence intervals after having completed data entry on 5, 10, 15, 20 and 25 babies by using proportion calculations:

https://www.openepi.com/Proportion/Proportion.htm

- 95% CI after 10 babies: 1 to 6 events
- 95% CI after 15 babies: 2 to 8 events
- 95% CI after 20 babies: 3 to 10 events
- 95% CI after 25 babies: 4 to 12 events

This means, that if the primary outcome incidence is outside these confidence intervals, the trial manager will make contact with the responsible principal investigator to 1) inform of the findings, and 2) ask for comments. During this process, however, it is important to keep in mind that 1) this is done without unblinding group assignment, and 2) 1/20 sites will have a primary outcome prevalence outside these confidence intervals, simply by the play of chance.

Exploratory outcomes

These outcomes include retinopathy of prematurity (ROP), bronchopulmonary dysplasia (BPD), necrotising enterocolitis (NEC), late-onset sepsis, and major neonatal morbidities count.

We will not monitor data for major neonatal morbidities count, since it is a count of the individual outcomes severe brain injury, ROP and BPD.

For each exploratory outcome, we will calculate the rate for all babies at all sites where data entry has been completed. This will be calculated once a month, simultaneously with the calculation and reporting of site data completion.

The prevalence will be compared to the expected 95% confidence interval, by using proportion calculations as for the primary outcome:

https://www.openepi.com/Proportion/Proportion.htm

We will use the expected prevalence's stated in the SafeBoosC III statistical analysis plan (REF):

1. Retinopathy of prematurity: 13%
2. Bronchopulmonary dysplasia: 40%
3. Necrotising enterocolitis: 11%
4. Late-onset sepsis: 40%

Example

Six months after randomisation opened, 300 babies have been randomised and 200 have completed data entry for ROP. The rate of ROP is 17% at present, meaning that 34 babies have been registered as having ROP.

If we use proportion calculations and an expected prevalence of 13% retinopathy of prematurity in the population of 200 babies, the expected 95% confidence interval of the incidence is: 18 to 36 babies.

This means that the prevalence of ROP is not higher than the expected incidence interval. Thus, no actions will be taken. If it had been higher or lower, the steering group would have been called in for a meeting on how to proceed.

It is important to clarify that this is not safety monitoring but quality control, thus we will not consider stoppage or termination of the trial, if we find an outcome incidence outside the expected ranges.