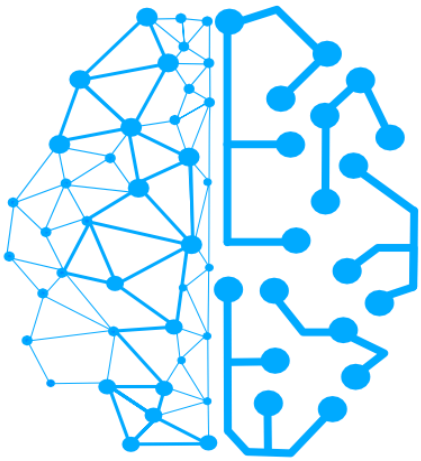
UNIVERSITY OF COPENHAGEN
GRADUATE SCHOOL OF HEALTH AND MEDICAL SCIENCES



SafeBoosC III

SAFEGUARDING THE BRAIN OF OUR SMALLEST CHILDREN

PhD Thesis

Mathias Lühr Hansen

Preparing for the multinational, pragmatic, randomised clinical trial SafeBoosC III

Supervisor: Gorm Greisen

This thesis has been submitted to the Graduate School of Health and Medical Sciences, University of Copenhagen 23rd of February 2021

Name of department: Department of Neonatology, Rigshospitalet

Author(s): Mathias Lühr Hansen

Title and subtitle: Preparing for the multinational, pragmatic, randomised clinical trial SafeboosC III

Topic description: Describing the process of designing, preparing for, and implementing a large-scale pragmatic randomised clinical trial to test if treatment guided by cerebral NIRS monitoring can improve the chances that extremely preterm infants survive without severe brain injury

Supervisor: Gorm Greisen, Professor, Department of Neonatology, Rigshospitalet

Co-supervisor(s)

Simon Hyttel-Sørensen, MD, PhD
Department of Intensive Care, Rigshospitalet

Janus Christian Jakobsen, Chief Physician, MD, PhD
Copenhagen Trial Unit, Rigshospitalet

Submitted on: 23rd of February 2021

Number of study units: 2

Department of Neonatology, Rigshospitalet, Copenhagen University, Copenhagen, Denmark

Copenhagen Trial Unit, Centre for Clinical Interventional Research, Copenhagen University Hospital, Copenhagen, Denmark

Number of characters: 125,225

Acknowledgements

This PhD project was funded by the Elsass Foundation, as part of the central funding for the SafeBoosC-III trial.

First of all, I would like to thank my fiancée Alexandra for supporting me in taking on this position as trial manager for the SafeBoosC-III trial, despite long work hours, phone calls in the middle of the night, interrupted holidays, late night meetings and multiple international travels during the last three years. I was hesitating at first, when Gorm offered me this position, but she convinced me to jump on board – and for that I am grateful.
I would also like to thank Janus Christian Jakobsen and Christian Gluud from Copenhagen Trial Unit. Not only supervising me in my thesis work, but also for educating me in evidence-based medicine. They set a uniquely high standard for clinical research and I am looking forward to continuing learning from them in the future.
Simon Hyttel-Sørensen also deserves thanks for supervising me in my thesis work and for reminding me that all choices comes with a consequence, and when it all comes down, family is first.
Endless thanks to Marie Rasmussen for working with me on the SafeBoosC project and for being so independent and strong-minded. I would have drowned in work if she had not been there. Luckily for me, we will continue our collaboration during the SafeBoosC-III follow-up study (and hopefully many years following).

At last, I would like to thank Gorm Greisen for having given me this possibility. It has been hard and constant work for the last three years, and the learning curve has been steep. I also never thought that I would hold a position with this much communication – 12,625 sent e-mails in 34 months. However, I would never have been without this experience. I am especially grateful for having met so many international colleagues, during the travels with the SafeBoosC project.
I should also thank Gorm for sharing with me, his pragmatic approach to, not only research, but also life in general. It is something that I try to implement in my own mindset. Gorm and I started working together in November 2015 when I was a medical student, and I hope that we will be able to continue collaboration in many years to come.

Contents

Included manuscripts 6

Summary 8

Abbreviations 11

Background 12

Extremely preterm infants, a vulnerable population 12

Brain injury in extremely preterm infants 12

Changes in the cardiopulmonary system during the first days of life 12

Impaired autoregulation of the brain 13

Brain injury and neurodevelopmental impairment 13

Multiple pathophysiological factors involved in brain injury 14

Treatment and monitoring strategies for extremely preterm infants 14

NIRS monitoring 15

Theory and practice 15

Reference values for cerebral StO_2 in preterm infants 16

Different devices differ in absolute StO_2 values 17

Wrongly use of NIRS due to worry of skin marks 17

Use and evidence for the implementation of cerebral NIRS monitoring in a clinical setting 18

Use and evidence in adults and children 18

Use and evidence in neonates 19

SafeBoosC II 20

Design 20

Results 20

Perspective 20

Aims of PhD thesis 21

The SafeBoosC III protocol (paper 1) 22

Methods 22

Trial design 22

Organisation 22

Eligibility 23

Randomisation 23

Blinding 23

Intervention 24

Training and certification 24

Outcomes and follow-up 25

Safety 25

Monitoring 25

Strength and weaknesses 26

The statistical analysis plan for SafeBoosC III (paper 2) 27

Methods/design 27

Sample size 27

Power calculations for exploratory outcomes 28

Level of significance 28

Handling of missing data 29

Statistical analysis 29

Twins and their intracluster correlation – a simulation study 30

Strength and weaknesses **32**

Developing and testing the feasibility of a multilingual web-based training and certification program for SafeBoosC III (paper 3) **33**
Methods/design **33**
SafeBoosC web-based training and certification program 33
Training module on NIRS 34
Survey 34
Data analysis 35
Results **35**
Close-ended questions 35
Open-ended questions – thematic analysis 36
Strength and weaknesses **38**

Exploring if keeping the cover on an INVOS neonatal sensor to avoid contact between the infants' skin and the sensors adhesive layer affects the StO_2 measurements (paper 4) **39**
Methods/design **40**
Statistical analysis 41
Results **43**
Strength and weaknesses **44**

Overall discussion and perspective **45**
Principal findings and status on the trial **45**
Pragmatic trials – evidence designed to alter clinical practice **46**
Delays and alterations due to obstructions **49**
Future considerations **52**
Conclusion **53**

Dansk resumé (Summary in Danish) **55**

References **59**

Paper and manuscript appendix **71**

Included manuscripts

Paper 1

Cerebral near-infrared spectroscopy monitoring versus treatment as usual for extremely preterm infants: a protocol for the SafeBoosC randomised clinical phase III trial

M Hansen, A Pellicer, C Gluud, E Dempsey, J Mintzer, SH Sørensen, A Heuchan, C Hagmann, E Ergenekon, G Dimitriou, G Pichler, G Naulaers, G Cheng, H Guimaraes, J Tkaczyk, K Kreutzer, M Fumagalli, O Claris, P Lemmers, S Fredly, T Szczapa, T Austin, JC Jakobsen, G Greisen

Trials 20;811, p 1-11, 2019

Paper 2

Detailed statistical analysis plan for the SafeBoosC III trial: a multinational randomised clinical trial assessing treatment guided by cerebral oxygenation monitoring versus treatment as usual in extremely preterm infants

M Hansen, A Pellicer, C Gluud, E Dempsey, J Mintzer, SH Sørensen, AM Heuchan, Cornelia Hagmann, G Dimitriou, G Pichler, G Naulaers, G Cheng, A Vilan, J Tkaczyk, K Kreutzer, M Fumagalli, O Claris, S Fredly, T Szczapa, T Lange, JC Jakobsen, G Greisen

Trials 20;746, p 1-12, 2019

Paper 3

Pilot test of an online training module on near-infrared spectroscopy monitoring (NIRS) for the randomised clinical trial SafeBoosC III

M Hansen, M Rasmussen, S Rubin, A Pellicer, G Cheng, X Xu, Y Zhaoqing, V Zoffmann, G Greisen

Trials 21;356, p 1-13, 2020

Paper 4

Not removing the glossy white cover from adhesive INVOS neonatal sensors affects the oxygenation measurements

M Hansen*, D Ostojic*, S Kleiser, G Greisen, M Wolf

Accepted for publication in Advances in Experimental Medicine and Biology on 30th of November 2019

Related publications referred to in the thesis

The clinical effects of cerebral near-infrared spectroscopy monitoring (NIRS) versus no monitoring in children and adults: a protocol for a systematic review with meta-analysis and Trial Sequential analysis

M Hansen, SH Sørensen, JC Jakobsen, C Gluud, E Kooi, J Mintzer, WP de Boode, M Fumagalli, AA Allen, T Alderliesten, G Greisen

Submitted to Systematic Review Journal (in peer-review), available at https://www.researchsquare.com/article/rs-93110/v1

Extremely preterm infant admissions within the SafeBoosC-III consortium during the COVID-19 lockdown: an international, observational study

M Rasmussen, **M Hansen** et al. (45 additional authors from the SafeBoosC-III consortium)

Submitted to Frontiers in Pediatrics (in peer-review), available at https://www.medrxiv.org/content/10.1101/2020.10.02.20204578v1

Summary

Background

Due to an immature cardiopulmonary system, extremely preterm infants are likely to experience low systemic blood flow during the first days of life. Combined with a poor cerebral autoregulation, this may cause inadequate cerebral blood flow, leading to an increased risk of brain injury, and subsequently death. Early neonatal intensive care involves complex treatment therapy guided by extensive monitoring. While monitoring of the respiratory system is well established, the possibilities to continuously monitor the circulatory status is poor. Furthermore, and end-organ monitor for the brain is missing. Cerebral near-infrared spectroscopy (NIRS) could be the next monitoring tool in the neonatal intensive care unit, to solve these issues. It is a non-invasive technology, that utilises near-infrared light to monitor cerebral tissue oxygenation and has been shown to correlate well with both central venous tissue oxygenation and cardiac output. So far, no sufficiently powered trial has shown a clinical benefit of utilising cerebral NIRS monitoring, as an additional tool to guide intensive care therapy in extremely preterm infants. Despite this, the clinical uptake is growing. As with all interventions, there is a risk of harm, including damage to the lungs, eyes and skin. Therefore, a large-scale trial providing real-world evidence on the benefits and harms of cerebral NIRS monitoring in extremely preterm infants is needed. It is therefore, that the multinational, pragmatic, randomised clinical trial, Safeguarding the Brains of our smallest Children (SafeBoosC) III, is now being conducted.

The aims of this thesis are to describe the process of designing and preparing for the conduct of SafeBoosC III, initiate the trial and randomise the first infants. This includes

1) Designing a feasible and pragmatic trial and publish the protocol before randomisation of the first participant (paper 1)
2) Designing a statistical analysis plan with consideration to the pragmatic study design, the large prevalence of twins in the population and the importance of not only statistical but also clinically relevant results. The aim was also to publish the statistical analysis plan before any data analysis (paper 2)
3) Piloting an online training module on NIRS monitoring, as a measure to evaluate the feasibility of developing a multilingual web-based training and certification program for the trial (paper 3)

4) Exploring if keeping the cover on an INVOS neonatal sensor to avoid contact between the infants' skin and the sensors adhesive layer, affects the tissue oxygen saturation measurements (paper 4)

Results

The trial is open-label with a two-parallel arm design. Infants randomised to the experimental arm receives treatment guided by cerebral NIRS during the first 72 hours of life. Treatment is based on an evidence-based treatment guideline and shall be initiated when cerebral oxygenation drops below a predefined hypoxic threshold. Infants randomised to the control arm receives treatment and monitoring as usual, i.e. no cerebral NIRS monitoring. Follow-up is conducted at 36 weeks of postmenstrual age or discharge to home. The primary outcome is a composite of either death or survival with severe brain injury, as detected on cranial ultrasound scans. The protocol was submitted to the Trials journal (trialsjournal.biomedcentral.com) four months before the first infant was randomised and published after 71 infants had been randomised. Up until the 18th of February 2021, the protocol has been approved in 74 local ethical committees across 18 countries.

A detailed statistical analysis plan has been published before the first data analysis. The sample size calculation showed that 1600 extremely preterm infants must be randomised, in order to detect a 22% relative risk difference for the composite outcome, between the experimental and control arm. The primary analysis will include the intention-to-treat population. Mixed-model logistic regression or linear regression will be used in the primary analysis, dependent on the outcome type, adjusting for stratification variables. Twin intra-class correlation coefficient (ICC) will not be included in the primary analysis. Instead, a sensitivity analysis will evaluate the potential effect on outcome correlation within twin couples. A simulation study to quantify the potential effect of a high twin correlation showed, that if the ICC and twin proportion is high, the coverage of the primary analysis will decrease, although minor. For the alternative generalised estimation equation (GEE) analysis, a high proportion of twins and a high ICC showed, as expected, a decrease in power. However, this was also minor.

For the pilot of the training module on NIRS, a total of 81 of out 100 invited staff members entered the training module and completed the survey. Median time to completion was 15 minutes and the median number of questions used to complete the module was seven. The academic level of both the quiz and learning material was found appropriate by most staff members (93% and 85% respectively). Furthermore, 90% agreed that the module was relevant as

a preparation for using NIRS. The thematic analysis detected important issues including technical problems, lack of clarity as well as a discrepancy between learning material and quiz. These results have been used to improve the development of the additional training modules including a revision of the piloted module on NIRS. As of the 18th of February 2021, the training program is available in six languages and more than 1300 staff members have started training. The overall satisfaction rate is high.

The blood-lipid phantom study showed that, despite the relationship between the INVOS neonatal sensor with cover and OxiplexTS was linear ($r^2 = 0.999$), the cover decreased the hypoxic threshold by more than 3% (60.3% with cover and 63.8% without cover). Furthermore, it also influenced the linear equation, with the INVOS neonatal sensor with cover being more sensitive to tissue oxygenation changes than without the cover: $StO2_{\_INVOS\_cover}=1.133*StO2_{\_ISS}+7.1$ as opposed to $StO2_{\_INVOS\_nocover}=1.103*StO2_{\_ISS}+12.0$.

Trial status

The first infant was randomised in Copenhagen on the 27th of June 2019 and since then, 65 hospitals across China, Europe, India and the US have been opened for randomisation. As of 18th of February 2021, 819 infants have been randomised and with the present randomisation rate, the trial is expected to complete recruitment in January 2022.

To conclude

1) A pragmatic protocol has been designed and published, sufficient of size and external validity to answer if treatment guided by cerebral NIRS monitoring during the first days of life, improves clinical outcomes in extremely preterm infants. The number of randomisations and active hospitals proves the feasibility of the study design. The next step will be to complete the trial and thereafter, evaluate if potential improvement persists into early childhood.
2) A statistical analysis plan focusing on the intention-to-treat population, taking into consideration the importance of both statistical and clinical significance as well as the large proportion of twins and their intra-class correlation coefficient, has been designed and published before any data analysis. The simulation study showed that a potential correlation within twin couples for the primary outcome, will have minimal effect on the trial results.
3) Based on the pilot study results, as well as the high participation and satisfaction rate of the SafeBoosC web-based training program, the described method of developing and

implementing an online training program for an international trial, despite language barriers, limited resources and potential difference in clinical practice between participating sites, has proven feasible. The results can be used for future trialists intending to plan and prepare for multicentre trials.

4) If the cover is kept on the INVOS neonatal sensor during cerebral NIRS monitoring, the hypoxic threshold will differ by more than 3%. It is plausible that other clinical practices to minimise sensor adhesiveness will affect StO_2 measurements as well. Clinicians must be aware of this, both inside and outside the SafeBoosC-III setting. Further studies evaluating how such practices affects the StO_2 measurements are needed.

Abbreviations

CI	confidence interval
CONSORT	consolidated standards of reporting trials
FiO_2	fraction of inspired oxygen
GEE	generalised estimation equation
HHb	deoxygenated haemoglobin
ICC	intra-class correlation coefficient
MARSI	medical adhesive-related skin injury
NIRS	near-infrared spectroscopy
O_2Hb	oxygenated haemoglobin
SafeBoosC	safeguarding the brains of our smallest children
SPIRIT	standard protocol items: recommendations for interventional trials
StO_2	tissue oxygen saturation
tHb	total haemoglobin

Background

Extremely preterm infants, a vulnerable population

In developed countries, more than 50,000 infants are born before 28 weeks of gestational age every year (1). They are classified as extremely preterm infants (2) and are in high risk of dying or surviving with complications (3), including impaired neurodevelopmental (4,5). Currently, around 25% of extremely preterm infants die (3) and among those who survive, 20% will suffer from neurodevelopmental impairments such as cerebral palsy, cognitive- and motor developmental delay, blindness or hearing impairment (5). In a population of 50,000, this corresponds to more than 10,000 deaths each year, and 10,000 survivors with neurodevelopmental impairment. Living with neurodevelopmental impairment is difficult and comes with increased health care costs, due to frequent visits to the family care physician as well as the use of hospital services. Furthermore, these children often needs special education and social service support throughout their childhood, thus adding to the economic burden (6,7). Since survival rates are stable in this population, perhaps even increasing for the smallest infants (8), and since life expectance is increasing in general, this is an issue that will only grow in the future (9).

Brain injury in extremely preterm infants

Changes in the cardiopulmonary system during the first days of life

During the first days of an infant's life, the respiratory and circulatory system transitions from a foetal to a neonatal state (10,11). This is especially an issue for extremely preterm infants, since their immature organs struggles during this process (7,12). Due to a combination of hypoxic pulmonary vasoconstriction and the foramen ovale, as well as the arterial duct causing right-to-left shunting of blood, the perfusion of the lungs is minimal in foetal life (10,11,13). When the infant is born, the lungs are ventilated and the oxygen levels in the blood rises, which leads to an increased constriction of the arterial vessels and thereby an increase in blood pressure, but a decrease in pulmonary vascular resistance. This will cause a reorientation of blood flow, with blood shunting from left to right through the arterial duct. In order to avoid low systemic blood flow, cardiac output must be increased (14). For a term, healthy infant this is not a problem, since the duct normally closes quickly and the mature myocardium can increase stroke volume sufficiently (14). Preterm infants, however, often fails to close the foetal communication channels completely. At the same time, the immature myocardium cannot sufficiently increase

the stroke volume. Therefore, the transition can cause periods of low systemic blood flow and hypotension, followed by impaired oxygen delivery to essential organs (10,11).

Impaired autoregulation of the brain

Pressure autoregulation is the body's capability to maintain a stable blood flow to specific organs, when the systemic blood flow fluctuates due to pressure changes. In order to maintain a stable organ blood flow through such periods, organs can either dilate or constrict its arterial blood vessels (15). For vital organs such as the brain, the pressure autoregulation is robust, and a stable organ blood flow can be maintained, despite large fluctuations in systemic blood pressure (15). However, this compensatory mechanism has a capacity limit. When the blood pressure exceeds or drops below a certain threshold, the organ will no longer be able to compensate, and organ blood flow will follow systemic blood pressure. This is known as the 'pressure passive state'. Multiple factors can alter the robustness of the pressure autoregulation and decrease its capacity, including immaturity (15).

Studies have shown that the capacity of the cerebral pressure autoregulation is limited in the immature brain of preterm infants, especially amongst the most immature and sickest (16–20). Since low systemic blood flow is common in this population during the first days of life (10,11), and since cerebral autoregulation is impaired in the immature brain, this population is in great risk of experiencing low cerebral blood flow (10,11), thereby causing cerebral hypoxia (20), which can lead to cerebral haemorrhages and brain injury (21–24). For the most severe cases, it can lead to death (8) and for survivors, a life with neurodevelopmental impairment (25).

Brain injury and neurodevelopmental impairment

In neonatal care, cranial ultrasound is a widely used neuroimaging tool, to diagnose brain injuries in preterm infants (26,27). The correlation between a normal ultrasound scan and a normal neurodevelopment is rather good, with a positive predictive value of 94% (95% confidence interval (CI) 92% to 96%) for a normal neuromotor development and 82% (95% CI 79% to 85%) for a normal cognitive development (28). One of the most common types of acute brain injury in extremely preterm infants is intraventricular haemorrhage (29). It can be diagnosed with cranial ultrasound, often within the first days of life (30) and can be divided into four grades (grade I – IV) (29). Severe intraventricular haemorrhage is defined as grade III or IV haemorrhages (27) and is present in approximately every sixth extremely preterm infant (3). It is debated if IVH grade I and II increases the risk of neurodevelopmental impairment (28,31) but

for grade III and IV, the risk of an abnormal neuromotor development is certain and estimated to 26% (95% CI 13% to 25%) and 53% (95% CI 29 to 76%) respectively (28). A few weeks after occurrence of a severe intraventricular haemorrhage, survivors may develop ventricular dilatation (32,33). This has a risk of 22% (95% CI 17% to 28%) for abnormal neuromotor development (28). For those developing hydrocephalus, the risk is even higher (28). Another severe brain injury, often detected within the first days of life in preterm infants, is cerebellar haemorrhage. Studies have reported findings of cerebellar haemorrhage on 1-7% of cranial ultrasound scans in preterm infants (25,34). Cerebellar haemorrhage is strongly associated with abnormal neuromotor and cognitive development, with a positive predictive value of 71% (95% CI 42% to 90%) (28). Periventricular leukomalacia is an ischemic white matter injury with focal and diffuse necrosis. In some cases, the focal necrosis develops into a cystic form, known as cystic periventricular leukomalacia (35). De Vries et al found a positive predictive value for cerebral palsy following cystic periventricular leukomalacia of 77% (95% CI 59% to 89) (36) and O'Shea et al found moderate to severe mental developmental impairment in 40% of extremely preterm infants, with previous cystic periventricular leukomalacia (25). Brain atrophy at term age has also been shown to increase the risk of neurodevelopmental impairment in preterm infants. A study by Horsch et al showed that preterm infants with brain atrophy on a cranial ultrasound scan, scored significantly lower in neurodevelopmental tests when compared to infants without brain atrophy (37).

Multiple pathophysiological factors involved in brain injury

The exact pathophysiology of brain injury in preterm infants is complex and many factors are involved; as mentioned above, changes in systemic blood flow i.e. low cardiac output (10), disturbances in systemic blood pressure (38) and ductal steal (39) can result in the disturbance of cerebral blood flow followed by brain injury. Another important factor is unintentional hyperventilation during mechanical ventilation, causing hypocapnia and thereby cerebral vasoconstriction and low cerebral blood flow (40,41). Furthermore, anaemia has been shown to affect cerebral oxygenation as well (42).

Treatment and monitoring strategies for extremely preterm infants

Management of extremely preterm infants during the first days of life, includes complex intensive care therapy based on extensive monitoring. Continuous positive airway pressure or mechanical ventilation is used to support the respiratory function together with surfactant

administration (43), and intravenous fluids, vasopressors or inotropes are used as hemodynamic support to treat hypotension (44). If deemed necessary, a persistent duct can also be treated with conservative or pharmaceutical management and for the sickest infants, it is sometimes closed with surgical ligation (45). Nutritional care, including both parenteral and enteral nutrition as well as treatment to stabilise blood glucose levels, is also part of the routine care of extremely preterm infants (46,47). Monitoring of the respiratory system is well established, especially during mechanical ventilation (48), but also due to the possibility of continuous transcutaneous partial pressure of carbon dioxide monitoring and continuous arterial pulse oximetry (49,50). For the circulatory system however, the only continuous monitoring routinely used is blood pressure and heart rate (51), and so far, there is no evidence proving that low blood pressure alone is a predictor for worsened neurological outcome (52). Furthermore, it is also doubtful whether treatment based on blood pressure thresholds improves neurological outcome (44). Neonatal echocardiography is a promising tool, to non-invasively assess haemodynamic instability and guide therapeutic interventions in preterm infants. However, the technique requires extensive training and experience, and it is not a continuous monitoring method (51). Additionally, an end-organ monitor for the brain, taking into consideration the multiple factors affecting cerebral blood flow and oxygenation is lacking. NIRS is a non-invasive, continuous monitoring technique that makes it possible to monitor cerebral tissue oxygenation and hence, cerebral blood flow (53). Since cerebral tissue oxygenation has been shown also to correlate well with cardiac output (54) and systemic blood pressure (19,20), it is suggested that cerebral NIRS monitoring can function as an indirect continuous monitor of cardiac output and a direct end-organ monitor for the brain, to guide intensive care therapy (51,55–57).

NIRS monitoring

Theory and practice

NIRS utilises near-infrared light in the wavelength spectra of 700-1000 nano meters. In practice, a NIRS sensor is placed on the area of the body, where you want to monitor the tissue oxygenation. The sensor consists of a light source and several light detectors. When the light is sent from the source and into the tissue, some of the light will be absorbed and some will be scattered back to the sensor and light detectors. It is the proportion of oxygenated [O_2Hb] and deoxygenated haemoglobin [HHb] in the underlying tissue, that determines the amount of light, that is either absorbed in the tissue or scattered back to the sensor. NIRS relies on the same principles as arterial pulse oximetry, but where pulse oximetry only measures O_2Hb in pulsating

arterial blood, NIRS measures both O_2Hb, HHb and total haemoglobin [tHb = O_2Hb + HHb] (53). This is possible, since the near-infrared light is differently absorbed by HHb and O2Hb; both molecules absorbs light at 800 nm, but at 760 nm, light is mainly absorbed by HHb which allows for the calculation of HHb and O_2Hb concentrations, respectively. These concentrations can be used to calculate tissue oxygenation (StO_2) i.e., the fraction of oxygenated haemoglobin in the vascular bed; [StO_2 = O_2Hb/tHb]. Thus, NIRS monitoring reflects the balance of oxygen delivery and oxygen consumption in the tissue (53,58,59). Since NIRS measures all blood components in the underlying tissue, including arterial, capillary and venous blood, tissue oxygenation is 75-80% venous weighted (60). This also means, that NIRS values tend to be lower than arterial pulse oximetry values (60). There seems to be a fairly good correlation between cerebral StO_2 and invasively measured central venous saturation, but with wide limits of agreements, especially for lower tissue oxygenation values. This is probably due to both inter- and intrasubject variance of the cerebral venous-arterial ratio (61–63).

NIRS monitoring can be applied to multiple parts of the body, dependent on the purpose. Usually, the sensor is placed either at the forehead to monitor cerebral tissue oxygenation, abdomen to monitor mesenteric tissue oxygenation, lower back to monitor renal tissue oxygenation or at a peripheral muscle to monitor muscle tissue oxygenation (61,64). So far, it has mostly been used to assess cerebral tissue oxygenation (65–67).

Reference values for cerebral StO_2 in preterm infants

Previous studies have shown that it is difficult to define and estimate ‘normal’ tissue oxygenation values, which complicates the clinical use of NIRS (68,69); if you do not know which StO_2 values to aim for, when trying to stabilise cerebral tissue oxygenation, it complicates the treatment. For preterm infants, Van Bel et al. (personal communication) previously estimated statistical normal cerebral StO_2 values by monitoring 390 preterm infants (gestational age below 32 weeks) in Wilhelmina Children’s Hospital in Utrecht, during the first three days of life. They used the INVOS 4100/5100 device with the small adult sensor and found a statistical normal StO_2 range from 55% to 85% (±2 standard deviations from the mean). It is these upper and lower reference values, that defined the hypoxic thresholds for interventions in the SafeBoosC-II trial (see **SafeBoosC II**) (70). Since the conduct and publication of SafeBoosC II, the same author group have extended their work, and published a paper on StO_2 reference values for 999 preterm infants (69).

Different devices differ in absolute StO_2 values

It is evident from both in vivo and in vitro studies, that different NIRS devices differ in absolute StO_2 values (68,71–73). This also means that the 'normal' StO_2 reference values defined in Utrecht are only valid for the INVOS small adult sensor (69). Therefore, in order to use other NIRS devices and sensors to guide treatment, one would have to either a) define 'normal' StO_2 reference values for the specific device and sensor or b) be able to determine StO_2 values for the specific device and sensor, corresponding to the reference StO_2 values found with the INVOS small adult sensor in Utrecht (69).

As a preparation for the SafeBoosC-II trial (see under '**SafeBoosC II**'), several NIRS devices were compared in-vivo to the INVOS small adult sensor on an adult forearm (72,73). Only devices where absolute StO_2 values as well as the dynamic range was within 5% of the INVOS small adult sensor, were declared eligible to be used in the trial (70). Since this method can only evaluate, whether a specific device has similar or different absolute StO_2 values and dynamic ranges as compared to the INVOS small adult sensor, it limits the number of devices that could be used for SafeBoosC II. Furthermore, in-homogeneity of the underlying tissue as well as physical alterations over time raises some methodological issues (72).

Another way to calibrate NIRS devices against INVOS with the small adult sensor, is by using an in-vitro blood lipid phantom (74). Such a setup provides the possibility of determining StO_2 values for all NIRS devices and sensors that corresponds to specific StO_2 values with the INVOS small adult sensor (71,74).

Wrongly use of NIRS due to worry of skin marks

Due to the immature and fragile skin of extremely preterm infants, great care must be taken to avoid skin injuries consequently to treatment and daily care in the NICU (75). Medical adhesive-related skin injuries (MARSIs) is a broad term for skin injuries, caused by contact with adhesive surfaces (76). It is a common cause of skin injury in the NICU, with epidermal stripping being the most frequent MARSI seen in preterm infants (76).

The majority of NIRS sensors use adhesiveness to secure the sensors position and therefore, monitoring with adhesive sensors in extremely preterm infants may cause skin problems. In the SafeBoosC-II trial (see under '**SafeBoosC II**'), skin marks were seen in one out of ten babies (16 marks in total). It was suspected that the skin marks occurred in the babies with the poorest circulation, due to heat from the sensor or pressure as a consequence to excessive fixation (70).

As a preventive measure to avoid heat and pressure injuries, it was recommended to reposition the sensor every four hours. However, repetitive removal and application of an adhesive sensor can disrupt the skin barrier and cause MARSIs (76). Based on personal communication, it has been reported that some neonatologists tend to minimise adhesiveness in different ways. Some neonatologists repetitively attach and remove the sensor from a clean surface, before initiation of monitoring, using see-through band-aid between the sensor and skin, while others keep the paper cover on the sensor during monitoring. Although it would be unexpected, there is a possibility that such measures could affect the StO_2 measurements, hence alter the hypoxic intervention threshold.

Use and evidence for the implementation of cerebral NIRS monitoring in a clinical setting

Use and evidence in adults and children

Despite limited evidence, cerebral NIRS monitoring is being used in multiple clinical settings in the adult and paediatric population. Particularly in the intensive care unit (77) and in the surgical theatre as a perioperative monitoring tool during aortic and cardiac surgery (66,67,78). According to our knowledge, no randomised clinical trials have evaluated the clinical effect cerebral NIRS monitoring in the intensive care setting. For perioperative monitoring, more than twenty trials have been published, but none have included children. In 2017, a systematic review with meta-analysis based on ten randomised clinical trials including 1466 adult patients, evaluated the clinical effect of cerebral NIRS monitoring during cardiac surgery (67). The meta-analysis showed no difference in mortality rate or the prevalence of strokes, between participants with or without cerebral NIRS monitoring during surgery. However, the number of deaths and strokes were low in both groups and the included trials were in high risk of bias. Furthermore, no sequential method was used to control the risks of type I and type II errors (79). The authors conclude that for now, the existing evidence does not suggest a clinical benefit of cerebral NIRS monitoring during cardiac surgery, and that addition large pragmatic trials with low risk of bias and clinical important endpoints are necessary (67). In 2018, a Cochrane review evaluating the clinical effects of perioperative cerebral NIRS monitoring across all types of surgery, had a similar conclusion. The authors concluded that the number of clinical events and quality of the published trials were too low to properly assess the effect of perioperative cerebral NIRS monitoring. However, the review was based on fifteen trials and at the time point, twelve additional trials awaited classification and eight were still ongoing (66). Furthermore, all point estimates in the included trials were trending towards a positive effect of cerebral NIRS

monitoring on neurological outcome (66). As a new systematic review with meta-analysis and Trial Sequential Analysis, evaluating the clinical effects of cerebral NIRS monitoring across all clinical settings and within all age groups is ongoing, the available evidence will soon be updated (80).

Use and evidence in neonates

In 2018, Hunter et al. published the results from an international survey study evaluating the use of NIRS monitoring in neonatal intensive care units (NICU) across Asia, North America and Australia. Of the 235 NICUs that responded to the questionnaire, 85 owned a NIRS device and 69 used it in a clinical setting. In 59 of the 150 NICUs who did not own a NIRS device, the reason was limited evidence on a clinical benefit (65). Thus, it seems that implementation of NIRS in the NICU is growing, despite the lack of evidence showing a positive benefit on clinical relevant endpoints (81).

According to our knowledge, only two randomised clinical trials have evaluated the use of cerebral NIRS monitoring to guide intensive care therapy in neonates, both in preterm infants; COSGOD II and SafeBoosC II. The COSGOD II feasibility trial assessed, if cerebral NIRS monitoring to guide respiratory and supplemental oxygen support during resuscitation immediately after birth, could reduce the burden of cerebral hyperoxia and hypoxia in preterm infants. Thirty babies were randomised to visible NIRS, thus receiving treatment guided by cerebral NIRS monitoring immediately after birth in the delivery room, and thirty babies were randomised to blinded NIRS and thereby treatment and monitoring as usual. The study showed that the burden of cerebral hyperoxia and hypoxia was halved in the NIRS-visible arm. Mortality was also reduced in the NIRS-visible arm, while the incidence of retinopathy of prematurity and bronchopulmonary dysplasia was increased. However, the event numbers were small (82). A larger trial, COSGOD III, with the clinical relevant endpoint, survival without severe brain injury, is now being conducted (83). Outside of the delivery room, the only published trial evaluating the use of cerebral NIRS monitoring among preterm infants is the SafeBoosC-II feasibility trial (70,81).

SafeBoosC II

Design

SafeBoosC II was a randomised clinical feasibility trial, that evaluated if treatment guided by cerebral NIRS monitoring could reduce the time where the brain was either 'blue' (hypoxic) or 'over-pink' (hyperoxic). As an addition to the usual monitoring, infants in the experimental arm also received cerebral NIRS monitoring during the first three days of life. If the StO_2 value dropped below a predefined hypoxic threshold or increased above a predefined hyperoxic threshold, treatment should be considered to try and normalise StO_2 (70). As mentioned previously, it was the upper and lower StO_2 reference values derived from analysing monitoring data on the 390 preterm infants from Utrecht, that defined the hypoxic and hyperoxic treatment thresholds, i.e. 55% and 85% respectively (personal communication). To support the clinicians, an evidence-based treatment guideline was developed and published, including suggestions on possible interventions when the StO_2 were out of range (84). Infants in the control group received blinded NIRS monitoring as well as standard monitoring and treatment. In total, 166 babies were randomised across eight European NICUs. The primary outcome was the burden of hypoxia and hyperoxia during the intervention period. Secondary outcomes were the clinical parameters mortality, brain injuries, retinopathy of prematurity, bronchopulmonary dysplasia and necrotising enterocolitis (70).

Results

The burden of hypoxia and hyperoxia was reduced to less than half in the experimental arm, which was mainly due to less hypoxia. The burden of hyperoxia was overall low and did not differ between the two arms. For the clinical outcomes, there was a lower mortality rate and a reduced incidence of severe brain injury in the experimental arm. There was also an increased incidence of bronchopulmonary dysplasia and retinopathy of prematurity in the experimental arm (70). However, SafeBoosC II was not powered to detect a relevant difference on clinical outcomes (85). Although a large number of the administered NIRS-based interventions were aimed at increasing cardiac output, the majority of the interventions were an increase in the fraction of inspired oxygen (FiO_2) (86).

Perspective

While the SafeBoosC-II trial revealed that it is possible to stabilise cerebral tissue oxygenation during the first days of life, it is uncertain whether this translates into a better clinical outcome.

Despite a promising decrease in the incidence of severe brain injury and death in the experimental arm (70), the difference was insignificant and the power too low to reject the possibility of type II errors (87). On the other side, the prevalence of lung and eye damage were increased in the experimental arm. If cerebral NIRS monitoring leads to unnecessary increases in FiO_2, it might potentially increase the risk of bronchopulmonary dysplasia and retinopathy of prematurity (88). It is also possible that the monitoring itself can cause skin injuries. At last, if NIRS is implemented but not beneficial, it will cause unnecessary disturbance and stress to the infants as well as a waste of staff time and increased costs. Thus, there is a possibility that cerebral NIRS monitoring might be harmful and it would be unfortunate, if NIRS was incorporated as a part of routine monitoring in the NICU, without proper evidence on its positive effect on clinical outcomes. As described previously, the use of NIRS in the NICU is increasing (65). Therefore, a sufficiently powered trial, evaluating the benefits and harms of treatment guided by cerebral NIRS monitoring on a clinically relevant end point is needed.

For such a trial to influence decision-makers and alter clinical practice, it must be able to answer the question; “will a general implementation of cerebral NIRS monitoring in routine clinical practice, overall benefit or harm extremely preterm infants?”. To reflect a general implementation in routine practice across multiple countries, the design of the trial must mimic a ‘real-world’ scenario, in order to obtain ‘real-world’ evidence. Such a trial is labelled a pragmatic trial (89).

It is therefore, that the pragmatic, multinational, randomised clinical trial SafeBoosC III is now being conducted.

Aims of PhD thesis

The aims of this thesis are to describe the process of designing and preparing for the conduct of SafeBoosC III, initiate the trial and randomise the first infants. This includes

1) Designing a feasible and pragmatic trial and publish the protocol before randomisation of the first participant (paper 1)
2) Designing a statistical analysis plan with consideration to the pragmatic study design, the large prevalence of twins in the population and the importance of not only statistical but also

clinically relevant results. The aim was also to publish the statistical analysis plan before any data analysis (paper 2)

3) Piloting an online training module on NIRS monitoring, as a measure to evaluate the feasibility of developing a multilingual web-based training and certification program for the trial (paper 3)
4) Exploring if keeping the cover on an INVOS neonatal sensor to avoid contact between the infants' skin and the sensors adhesive layer, affects the StO_2 measurements (paper 4)

The SafeBoosC III protocol (paper 1)

Methods

The purpose of the SafeBoosC-III trial is to evaluate the benefits and harms of cerebral NIRS monitoring combined with an evidence-based treatment guideline (84) in extremely preterm infants, for the first 72 hours of life. The hypothesis is that the intervention will increase survival without severe brain injury at 36 weeks of postmenstrual age (90).

Trial design

It is an international, multicentre, investigator-initiated, open-label, pragmatic randomised clinical phase III trial with two arms – experimental and control. In total, 1600 extremely preterm infants will be randomised from an expected 50 NICUs across 20 countries. The countries include Austria, Belgium, China, Czech Republic, Denmark, France, Germany, Greece, India, Ireland, Italy, Norway, Poland, Portugal, Switzerland, Spain, Turkey, Ukraine, USA and United Kingdom. In a previous funding application for SafeBoosC III, the 93 NICUs partaking, had admission rates between 15 and 90 extremely preterm infants per year and overall, 3000 each year. Thus, it is feasible to randomise 1600 babies across 50 hospitals within a two-year period.

The trial is designed in line with the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) guideline for protocols (91).

Organisation

The trial organisation is built on three levels. The executive committee consists of the trial manager, coordinating investigator, two representatives from Copenhagen Trial Unit and four

national coordinators. They are responsible for the day-to-day management of the trial. The steering committee consists of all national coordinators (one from each country) as well as the trial manager, coordinating investigator and the two representatives from Copenhagen Trial Unit. The steering committee meets online every second month, to discuss the progress of the trial and take important decisions if needed. For each participating NICU, there is also a principal investigator, who is responsible for the preparation and conduct of the trial in his/her local NICU.

Eligibility

In order for an infant to be included, it must be born before 28 weeks of postmenstrual age and the parents must have signed a consent form, unless 'opt-out' or deferred consent is used (92,93). If it has been decided not to conduct full life support or if it is not possible to start NIRS monitoring within six hours, the infant must be excluded (90).

Randomisation

Randomisation is computer-based and central by Copenhagen Trial Unit. The allocation ratio is 1:1, to either the experimental or control arm, and the sequence is with a varied block size, generated by a computer. This is done to conceal the allocation sequence for investigators. Infants are stratified by centre (NICU) and gestational age (lower gestational age (below 26 weeks) versus higher gestational age (26 weeks or higher)) (90).

Blinding

The nature of the intervention makes it difficult, if not impossible, to blind parents and clinical staff to the study arms (see under 'Strength and weaknesses'). Therefore, parents and clinicians are not blinded. The assessment of mortality status is also not blinded, but assessment and diagnosis of severe brain injury is. Principal investigators are responsible for developing a local blinding procedure, describing how the assessment of severe brain injuries is conducted blinded. The procedure must be approved centrally by the trial manager and it must also be demonstrated to the Good Clinical Practice monitor person, before the NICU can start enrolling infants. All local blinding procedures are available at www.safeboosc.eu. The data managers, statisticians and the investigators drawing conclusions will be blinded to the study arms (94).

Intervention

Infants allocated to the experimental arm starts cerebral NIRS monitoring within six hours from birth and optimally, as soon as possible. After 72 hours, the sensor is removed. During this time period, the NIRS monitoring serves as an extra tool to modify and adjust intensive care therapy (90). To support the clinicians in choice of action when the StO_2 drops below a predefined device-specific hypoxic threshold, the evidence-based SafeBoosC treatment guideline developed for SafeBoosC II is used (84). Since hyperoxic events were few in SafeBoosC II and the interventions administered had little effect on this (70), SafeBoosC III does not include treatment due to hyperoxic StO_2 values. As in SafeBoosC II, 55% with the INVOS small adult sensor is used as the hypoxic threshold. However, SafeBoosC III is a pragmatic trial, aiming for a generic result with a high external validity (real-world evidence) and therefore, as many different NIRS devices as possible shall be used in the trial. Therefore, as a preparation for the trial, the hypoxic threshold corresponding to 55% with the INVOS small adult sensor has been defined for a number of NIRS devices and sensors commercially approved for use in neonates, in a blood lipid phantom (74). So far, the hypoxic thresholds for six of the nine approved devices have been published (71,74). Papers including the hypoxic thresholds for the additional three devices will be submitted for publication but are, for now, available on www.safeboosc.eu.

Infants allocated to the control arm receives monitoring and treatment as per local clinical routine (90).

Training and certification

To comply with the Good Clinical Practice requirement stating that staff members participating in a clinical trial "should be qualified by education, training and experience to perform his or her respective task(s)" (95), and to provide a balanced, practical and realistic level of training and introduction for physicians and nurses partaking in the trial, a multilingual web-based training and certification program has been developed and is available for clinical staff. Certification rates within each NICU is reported online at www.safeboosc.eu and the overall certification rate will be included in the main publication (90).

Outcomes and follow-up

The primary outcome is a composite of death or survival with severe brain injury. The severe brain injury diagnosis is based on assessment of all available routine cranial ultrasound scans up until the age of follow-up, and is defined as one or more of the following brain injuries; intraventricular haemorrhage grade III/IV, cystic periventricular leukomalacia, post-haemorrhagic ventricular dilatation, cerebellar haemorrhage and cerebral atrophy (90).

The exploratory outcomes confide the major neonatal morbidities seen within this population, including bronchopulmonary dysplasia, retinopathy of prematurity, necrotising enterocolitis, late-onset sepsis and a 'major neonatal morbidities count' of bronchopulmonary dysplasia, retinopathy of prematurity and severe brain injury (96).

All diagnoses are set as per local criteria in each NICU. Outcomes are evaluated at 36 weeks of postmenstrual age or at the date of discharge, as written in the infants' records, expect for severe brain injury, which is based on assessment of the available cranial ultrasound scans (90).

Safety

Predefined severe adverse reactions (SAR) with suspected direct relationship to the NIRS monitoring, are reported at the end of the intervention period, i.e. 72 hours after birth. This includes 'physical mishaps associated with managing the oximeters and sensors' as well as 'critical mismanagement based on data from the cerebral NIRS monitoring'. Severe adverse events consist of the exploratory outcomes, as well as each component of the primary outcome and are reported at follow-up, i.e. 36 weeks of postmenstrual age (90). The Data Monitoring and Safety Committee evaluates on safety during interim analyses, the first to be conducted after one third of the participants have been randomised.

Monitoring

Central monitoring is conducted by the Trial Manager, Copenhagen Trial Unit and the Coordinating investigator. The purpose is to identify NICUs with missing data, predefined data quality deficiencies as well as noteworthy data deviations. If identified, the Trial Manager will contact the relevant principal investigator, so that appropriate action is taken (90).

Local monitoring is conducted by a Good Clinical Practice monitor person, appointed by the local principal investigator in each NICU. The Good Clinical Practice monitor person conducts the monitoring according to the SafeBoosC-III GCP monitor plan (90).

Monitor plans as well as monthly monitoring reports are uploaded and publicly available at www.safeboosc.eu.

Strength and weaknesses

As described previously under 'Blinding', clinical staff and parents are not blinded to the study arms. Clinical staff could be blinded by using a sham instrument, showing random values. However, adjustment of intensive care therapy based on false monitoring values could be harmful. Furthermore, this is also a pragmatic trial and therefore, the intervention should be compared to routine practice and not a shame/placebo instrument (97). Blinding of parents is theoretically possible, since the monitoring only lasts for 72 hours and hereafter, there will be no difference between the two study arms. In practice however, it is difficult seen from an ethical perspective, to separate the infant and parents for such a long time period, especially if the infant is severely ill. The lack of group allocation concealment for parents and clinical staff introduces a risk of bias (98). Lack of blinded outcome assessment can also lead to assessment bias and potentially, an overestimation of the interventions effect, especially on subjective outcomes such as severe brain injury diagnosed by cranial ultrasound (98–101). It is therefore of highly importance, that assessment of subjective outcomes is done blinded (98). As described under 'Blinding', this will also be the case in SafeBoosC III. On the contrary, mortality status will not be assessed blinded. Even though there is a potential for outcome assessment bias, regarding whether an infant dies before or after 36 weeks of postmenstrual age, the number of infants that dies around this time point is minor (102) and therefore, it is not expected that this potential bias will affect the trial results. Also, a meta-analysis by Perner et al. has shown, that unblinded assessment of mortality status does not introduce any significant assessment bias in intensive care trials (103).

The pragmatic design also raises some potential weaknesses. In pragmatic trials, outcome assessment should primarily rely on routine data, in order to avoid unnecessary interference with the 'real-life' clinical setting (104). This will also be the case in SafeBoosC III, where outcome assessment will be done by reviewing the infants' clinical records, as well as the

routine cranial ultrasound scans. However, since the scanning protocol, frequency and number of cranial ultrasound scans vary between NICUs, there is a risk that some brain injuries will be underdiagnosed. Cerebellar haemorrhage is best visualised through the posterior or mastoid fontal (105) and since these views are not a part of the routine cranial ultrasound scanning protocol in all NICUs (106), some events might be missed. Cerebral atrophy is often diagnosed in later cranial ultrasound scans around term age (37) and therefore, NICUs that do not conduct later scans might miss such events. This could eventually lead to an overall reduced incidence of severe brain injury in both study arms. However, since the prevalence of severe brain injury used in the sample size calculations for SafeBoosC III, is based on the diagnosis of intraventricular haemorrhage grade III or IV and cystic periventricular leukomalacia alone, the potential missed events should not affect the statistical power of the trial. Furthermore, the number of missed events should be distributed evenly between the two study arms, hence the randomisation.

In SafeBoosC III, cranial ultrasound scans are assessed locally rather than centrally, as was done in SafeBoosC II. This is a potential cause of concern, since variance in the interpretation of the scans, could be expected between NICUs. However, a previous retrospective analysis of cranial ultrasound data from a clinical trial on preterm infants, showed high accuracy and reliability of local readings compared to central readings for severe brain injuries (107). Furthermore, the web-based training program available to staff members, includes a module on cranial ultrasound and diagnosis of severe brain injury, in order to increase data quality and decrease inter-variance between local readers (108).

The statistical analysis plan for SafeBoosC III (paper 2)

Methods/design

Sample size

The sample size calculation for the SafeBoosC-III trial is based on the incidence of the primary outcome, i.e. death or survival with severe brain injury at 36 weeks of postmenstrual age, in the SafeBoosC-II trial. In SafeBoosC II, 34% either died or survived with a severe brain injury in the control arm, while in the experimental arm, the incidence was 26% (70). Assuming similar incidences in the SafeBoosC III trial, 1600 infants must be randomised – 800 to each arm – in order to demonstrate such a reduction, at a 5% alfa level and with 90% power (94).

Power calculations for exploratory outcomes

Since the sample size is based on the primary outcome, separate power calculations have been conducted for the exploratory outcomes (table 1). Details on the calculations can be found in paper 2 (94).

Table 1 (94). Power calculations for the explorative outcomes.

Outcome	Assumption on prevalence in background population (%)	Assumption on risk increase or decrease (%)	Power%
Major neonatal morbidities (109)	0.62 (0.8)	20	87
Bronchopulmonary dysplasia (110)	40	20	89
Retinopathy of prematurity (81)	13	30	68
Late-onset sepsis (3)	40	20	91
Necrotising enterocolitis (81)	11	17	23

All calculations are based on a 5% alfa level. Major neonatal morbidities are presented as a mean count (0-3) with a standard deviation. Literature references for the assumptions on prevalence and risk increase or decrease are presented in the table.

Level of significance

A five-step procedure suggested by Jakobsen et al., will be used to assess both statistical as well as clinical significance and thereby draw a conclusion on the trial results (111). At first, the p-value and confidence intervals will be calculated and reported for all outcomes. For the primary outcome, a p-value of less than 0.05 will be considered statistically significant, since there is only one primary outcome and since a 5% risk of type 1 error was accepted in the sample size calculation (94). Despite a chosen statistical threshold of 0.05, the confidence interval around the point estimate will be heavily considered, when interpreting the results from the primary analysis (112). Although the exploratory outcome analysis will only be hypothesis generating, the results will still be taken into consideration when interpreting the results and drawing the overall conclusion. Secondly, in

order to interpret the SafeBoosC-III results in light of previous trial results, the Bayes factor will be calculated and reported for the primary outcome (113). Thirdly, at each interim analysis (first analysis being done after one third of the infants have been randomised), Lan-DeMets monitoring boundaries will be used to adjust the confidence intervals and p-value and thus, decide whether the trial should be terminated early (114). This adjustment during the interim analyses is important, in order to avoid a false rejection of the null-hypothesis due to an insufficient sample size (115). The fourth step regarding adjustment of the p-value, for declaring statistical significance due to multiple testing (116) is not relevant in SafeBoosC III, since there is only one primary outcome and since additional outcomes only are considered hypothesis generating. At last, clinical significance will be assessed, based on the number needed to treat (94).

Handling of missing data

Based on a previous publication by Jakobsen et al. on multiple imputation, it will be evaluated whether it is valid to ignore missing data in the outcome analyses (117). If it is not, a best-worst and worst-best case scenario will be presented, to assess the potential effect which the missing data might have on the results. When calculating the best-worst case scenario, it will be assumed that all infants in the experimental arm, with missing data for the primary outcome, have had a beneficial outcome, i.e. survival without severe brain injury. For infants in the control arm with missing data on the primary outcome, a harmful outcome (death or survival with severe brain injury) will be assumed. For the worst-base case scenario, the assumption will be the opposite for the two trial arms (117). To identify any imbalance between the two arms due to missing data, two comparisons of explanatory variables will be conducted; one comparison where all infants are present, including those with missing data on the primary outcome, and one comparison where only infants with a reported primary outcome is included (118). The results of these interpretations will also be taken into consideration when drawing the conclusion (94).

Statistical analysis

The primary analysis will be an intention-to-treat analysis for all outcomes, including all randomised infants, regardless of missing data, adherence to the intervention and loss to follow-up. This is done to maintain the comparability of baseline characteristics, obtained due to the randomisation sequence (118). Dichotomous outcomes will be analysed using mixed effect logistic regression and count data using mixed-effect linear regression, both with robust standard errors. The stratification variables ‘group allocation’, ‘site’ and ‘gestational age’ will be included in all

outcome analyses (119–121); ‘site’ as a random effect and ‘gestational age’ and ‘group allocation’ as fixed effects (94).
Two sensitivity analysis will be conducted. The first sensitivity analysis accounting for the possible correlation between twins is described in ‘Twins and their intra-cluster correlation’. The second sensitivity analysis is per-protocol, only including participants who had no missing data, were not lost to follow-up, and adhered to the intervention. Adherence to the intervention is defined as more than 58 hours (80% of the intervention period) of continuous cerebral NIRS monitoring during the first 72 hours of life. Premature stoppage of NIRS monitoring within the first 72 hours of life due to withdrawal of life support or death is not considered inadherence to the intervention.
As a secondary analysis, the data will be analysed, using the random-effects meta-analysis (94,122).

Reporting will be made on both the primary and explorative outcomes. However, as mentioned previously, the conclusion will be based on the primary analysis of the primary outcome. The additional analyses will be used, in order to discuss and reflect on the results from the primary analysis.

Twins and their intracluster correlation – a simulation study
Almost one third of the study population in SafeBoosC III will be twins (70) and since outcomes within twin pairs might be correlated, this raises a potential statistical problem (123).
Randomisation of twins, or triplets, also raises an ethical issue since previous research has shown that parents of multiples, and also adult multiples, prefer randomisation to the same arm in neonatal trials (124). In the SafeBoosC-III trial, twins or triplets are randomised to the same study arm (94).

The correlation of outcomes within a twin ‘cluster’ can be evaluated by calculating the intracluster correlation (ICC) (123). In SafeBoosC II, the size of the twin ICC for the primary outcome was unimportant. Other studies have reported a twin ICC for death before discharge to 0.00 (95% CI -0.04 to 0.02) and for intraventricular haemorrhage grade 3 or 4 to -0,01 (95% CI -0.05 to 0.01), which is negligible (125). Based on these results, twins will be analysed as independent observations. However, since there is a theoretical possibility, that outcomes will be correlated within a twin cluster, a sensitivity analysis will be conducted, taking this uncertainty into account (126). This analysis will be a generalised estimation equation (GEE) analysis, where an exchangeable covariance matrix with site, using a logit link, will be used. Additionally, the twin ICC for the primary outcome will be calculated and reported (94).

To quantify the theoretical consequence of outcome correlation within twin clusters, i.e. the potential impact on study power as well as the coverage of the confidence intervals (123), a simulation study was conducted (94). In the simulation study, the primary analysis for the primary outcome (naïve analysis in table 2) was compared to the GEE sensitivity analysis. Opposed to the primary analysis, the GEE analysis takes outcome correlations within twin clusters into account. In total, 10,000 trials were simulated, all of them with a similar sample size and incidence of the primary outcome as used in the sample size calculation for SafeBoosC III. The proportion of twins and twin ICC was set to vary between the simulations (see table 2).

Table 2 (94)**.** Simulation study to assess power and coverage probabilities of confidence intervals of the primary outcome.

ICC	Proportion of twins	Power of naive analysis	Power of GEE analysis	Coverage probability of naive analysis	Coverage probability of GEE analysis
0	0.1	0.91	0.91	0.95	0.95
0	0.2	0.91	0.91	0.95	0.95
0	0.3	0.90	0.90	0.95	0.95
0	0.4	0.90	0.90	0.95	0.95
0.01	0.1	0.91	0.91	0.95	0.95
0.01	0.2	0.90	0.90	0.95	0.95
0.01	0.3	0.90	0.90	0.95	0.95
0.01	0.4	0.91	0.90	0.95	0.95
0.03	0.1	0.90	0.90	0.95	0.95
0.03	0.2	0.91	0.91	0.95	0.95
0.03	0.3	0.90	0.90	0.95	0.95
0.03	0.4	0.90	0.90	0.95	0.95
0.13	0.1	0.90	0.90	0.95	0.95

ICC	Proportion of twins	Power of naive analysis	Power of GEE analysis	Coverage probability of naive analysis	Coverage probability of GEE analysis
0.13	0.2	0.90	0.89	0.94	0.95
0.13	0.3	0.90	0.89	0.94	0.95
0.13	0.4	0.90	0.88	0.94	0.95
0.2	0.1	0.90	0.90	0.95	0.95
0.2	0.2	0.90	0.89	0.94	0.95
0.2	0.3	0.90	0.88	0.94	0.95
0.2	0.4	0.89	0.87	0.94	0.95

The results of the simulation showed that if both the ICC and the proportion of twins are low, coverage of confidence intervals and power remain unchanged for both analyses. If the ICC and the twin proportion is high, only the coverage of the GEE sensitivity analysis will remain unchanged. However, the decrease in coverage for the primary analysis is minimal. At a high ICC and a high twin proportion, the power is also decreased for the GEE analysis. This is expected, as this takes into account the non-independence of outcomes between twins; when the ICC is high, twins cannot be seen as independent observations and thereby, the 'effective' sample size will be reduced. When the effective sample size is reduced, the analysis has less statistical power (123). However, the decrease in statistical power was minimal (94).

Strength and weaknesses

When conducting multiple analysis on a primary outcome, there is a potential for both type I and type II errors (127). If a significant difference between the two arms, on any of the primary outcome analysis is defined as being sufficient to declare superiority, and the alfa-level is not decreased, the risk of type I errors increases. Conversely, if significance on all primary outcome analysis is needed to declare superiority, there is a possibility that some analysis will be underpowered and thus, increase the chance of type II errors. To avoid type I and II errors in SafeBoosC III, only one primary analysis of the primary outcome will be conducted. The additional analysis will be defined as sensitivity and secondary analyses.

It is a major strength of the analysis plan, that the issue of twin ICC has been taken into account. Furthermore, the impact on the results, which can be expected if the twin ICC is high, is minimal. Of weaknesses, it is worth pointing out that only three of the five explorative outcomes are sufficiently powered at a 5% alfa-level, and if conclusions should be drawn on the basis of these outcomes, the significance level would have be to decreased using Bonferonni adjustments (128), which would decrease their power even more. Therefore, no conclusions will be made based on the exploratory outcome analyses.

Developing and testing the feasibility of a multilingual web-based training and certification program for SafeBoosC III (paper 3)

As reports on the development and implementation of online training programs for international randomised clinical trials are sparse and since the SafeBoosC-III trial group is inexperienced within the field, it was decided to pilot the first online training module in the SafeBoosC-III web-based training program, focusing on NIRS monitoring. This was done to test if the intended setup was feasible, despite challenges such as limited resources, difference in clinical practice between countries and potential language barriers. The results of this pilot study have been used to improve the development of the additional modules in the SafeBoosC-III web-based training program, as well as the piloted module on NIRS monitoring.

Methods/design

The principal investigator from five NICUs across China (n=3), Denmark (n=1) and Spain (n=1) were asked to invite ten nurses and ten doctors to participate, i.e. a total of 100 participants. In order to obtain generalisable results, both staff members with and without previous NIRS experience were invited. The nurses and doctors were asked to complete the online training module on cerebral NIRS monitoring, followed by an online survey to evaluate their experience (108).

SafeBoosC web-based training and certification program

The training program includes, besides the module on NIRS monitoring, four additional modules: 1) introduction to the trial, 2) SafeBoosC treatment guideline, 3) cranial ultrasound and severe brain injury diagnosis, and 4) Good Clinical Practice in SafeBoosC III. Each module is built up by a learning part and a quiz. The quiz recognises prior learning, meaning that correct answers will get

you through faster. The learning material and quiz in each module is based on a number of predefined learning objectives (108).

The quiz is case-based, and the questions are supposed to reflect clinical situations that could occur during the conduct of SafeBoosC III. The questions are multiple choice and there is often multiple correct answers, which should all be ticked off in order to pass a question. Once answers to a question are submitted, staff members are presented to explanations of both wrong and correct answers. A module is passed when one question for each predefined learning objective is answered correctly. Thus, staff members are presented to new questions (or previous questions answered wrong) from the pool of questions, until the criterion is met (108).

Training module on NIRS

The training module on NIRS is based on four learning objectives and included initially 11 questions, focusing on 1) principles of NIRS monitoring, 2) practicalities including application and fixation of the sensor and repositioning, 3) risk of skin marks and 4) interpretation of StO_2 values. Since SafeBoosC III is an international trial, the level of English varies between participating countries and staff groups. Therefore, the web-based training program has to be available in multiple languages. However, translation of content while managing to uphold the academic level, can be difficult. Especially if translation must be done by local investigators, due to limited economic resources. To test the feasibility of translating the content locally and train staff members in their native language, the training module on NIRS was piloted in three different languages (English, Chinese and Spanish) and the content was translated by the local principal investigators (108).

Survey

The online survey was also translated into native languages by the local principal investigators. It consisted of fifteen close-ended questions, with answers on a three to four step Likert scale as well as seven open-ended questions with free text-answers. The themes of the survey were 1) performance, 2) learning material, 3) quiz material, 4) interface, and 5) preparation to use NIRS in a clinical context (108).

Data analysis

Descriptive statistics were used to present quantitative data from close-ended questions, while answers to the open-ended questions underwent thematic analysis (129).

Results

Overall, 81 of the 100 invited doctors and nurses completed the online survey (fifty from China, 16 from Spain and 15 from Denmark) and 46 (57%) of the participants had previous experience with NIRS (108).

Close-ended questions

Table 3 outlines the number of responders who answered either 'agree/strongly agree' or 'appropriate' to the close-ended questions, as well as data on the participants performance. The results are presented for the total group and for each nationality. In paper three, additional tables are presented, stratifying answers according to experience level and clinical position (108).

The median time spent to complete the module was 15 minutes and the median number of questions was seven (table 3). In total, 85% of staff members agreed, that the academic level of the learning material was appropriate. For the quiz, it was 93%. Overall, 90% of staff members also agreed that the NIRS module was relevant as a preparation for using the NIRS device. However, almost one third (30%) did not agree, that the learning material was sufficient to complete the quiz and more than one third (40%), did not agree that the number of answering possibilities per question was appropriate. Furthermore, almost one third (30%) did not agree, that the NIRS module was stable and did not crash (108).

Table 3 (108). Time in minutes and number of quiz questions used to complete the module, and number of responding participants who answered either 'agree'/'strongly agree' or 'appropriate' to the questions, regarding the design of the module. Data stratified by country.

Question	Denmark	Spain	China	Total
Performance				
Minutes to complete module, median [range]	14 [7-30] (11/15)[#]	10 [1-60] (13/16)[#]	20 [2-420] (46/50)[#]	15 [1-420] (70/81)[#]

Number of questions to complete module median [range]	7 [6-20] (5/15)[#]	4 [4-12] (13/16)[#]	8 [4-50] (43/50)[#]	7 [4-50] (61/81)[#]
Learning material				
Academic level of learning material appropriate, n/N (%)	14/15 (93)	15/16 (94)	40/50 (80)	69/81 (85)
Learning material sufficient to complete quiz*, n/N (%)	3/12 (25)	13/16 (81)	39/50 (78)	55/78 (70)
Quiz				
Academic level of quiz appropriate*, n/N (%)	14/15 (93)	15/16 (94)	46/50 (92)	75/81 (93)
Number of answering possibilities per question appropriate, n/N (%)	6/15 (40)	9/16 (56)	34/50 (68)	49/81 (60)
Quiz questions clinically relevant and up-to-date*	13/14 (93)	15/16 94)	49/50 (98)	77/80 (96)
Interface				
The NIRS module was stable and did not crash, n/N (%)**	6/15 (40)	9/15 (60)	42/50 (84)	57/80 (71)
Preparation for using NIRS				
Relevant to prepare for using the NIRS device*	13/15 (87)	12/15 (80)	47/50 (94)	72/80 (90)

*Pooling of the answers 'agree' or 'strongly agree'.
**Yes' to the statement.
#=number of responders answering the specific question under the total number of overall responders completing the online survey

Open-ended questions – thematic analysis

The thematic analysis of the answers to the open-ended questions, revealed four main themes with additional sub-themes (figure 1), including 1) discrepancy between learning material and quiz, 2) lack of clarity within the course, 3) technical issues and 4) unsolicited positive comments (figure 1).

Regarding discrepancy, it was especially commented that the learning material was insufficient, too simple and not detailed enough to answer all the quiz questions.

"For someone who know[s] little or nothing about the topic, the introduction material is not sufficient enough to answer the quiz questions" Doctor (108)

"No introduction to how you prepared for NIRS monitoring, so it was pure guessing – you have no idea whether you need to calibrate/shave/wash or something else (prior monitoring), if you have not been told forehand" Nurse (108)

Some participants conversely stated that it was the questions in the quiz, that were simply too difficult.

"The content is too hard to understand" Nurse (108)

Regarding lack of clarity, participants stated that the module structure was unclear, especially the feed-back mechanism, multiple answer possibilities and the case-question setup.

"I think the quality of learning is increased if there are more questions with fewer answer possibilities. The purpose is learning and I think this could be heightened if one is presented with more questions with lesser answer possibilities…." Doctor (108)

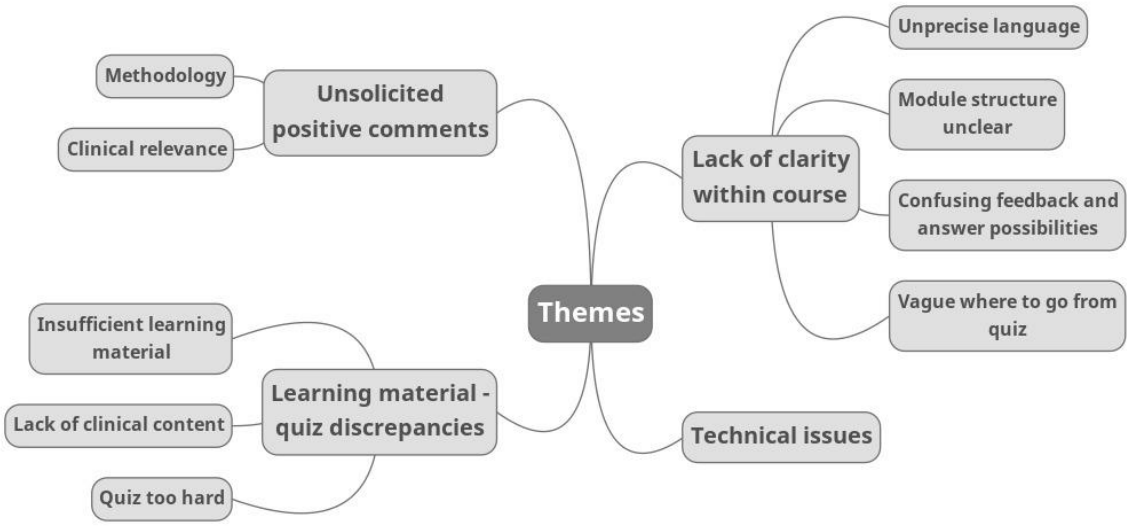
"It would be nice if one could learn something by answering wrong, hence that you could use the box that pops up after you answer incorrectly to see what was the correct answer." Nurse (108)

Regarding technical issues, the majority of comments stated that the training module crashed or simply 'froze' during training.

"If you do it, you will be stuck, you can not finish it, what the hell" Doctor (108)

"The page hangs on some occasions and does not allow to advance. When there is an incorrect answer, it loops in and you must restart the questionnaire to get out of there" Doctor (108)

Figure 1. Main themes and sub-themes as revealed by thematic analysis of open-ended question answers.



Strength and weaknesses

Since the training module was translated and tested in multiple languages, and since participants both included experienced and inexperienced staff members, the generalisability of the results is high. The methodological design where both open- and closed-ended questions were included in the survey, made it possible to get a deeper and wider insight into the participants' experience, thus revealing important points for improvement.

As described in the results section, this pilot study revealed weaknesses in the module, requiring room for improvement. The three main critique points were 1) questions having too many answer possibilities, 2) the learning material was inadequate to answer quiz questions, and 3) the technical problems delayed completion and left a significant number of staff members frustrated. All of these points have been taken into consideration and implemented in the development of the additional training modules, but also during a revision of the NIRS module.

As part of the pragmatic setup, translation from English to Spanish and Chinese was conducted by local principal investigators and due to limited resources and time, checking of translation quality and precision was not possible. This is a weakness of the study, since quality and precision of translations could potentially affect satisfaction rates as well as performance.

Although the survey completion rate was reasonable (81 out of 100 participants), it is unknown whether all 81 staff members completed the training module, since this cannot be tracked on the platform where the piloting was hosted. Conversely, there is also a possibility that the 19 participants who did not answer the survey, entered the training module without completing it. This is also a weakness of the study, since this uncertainty regarding responders and non-responders introduces the risk of bias; If the 19 participants who did not answer the survey actually started the training module, but did not complete it because it was too difficult, or perhaps experienced technical problems, their missing answers would cause an overestimation of the training modules feasibility.

One of the disadvantages of web-based training versus on-site training, is that it is difficult to assess if training was conducted individually, or whether it was done in smaller groups. If training was done in groups or two-and-two, the participants could have supported each other, thus not revealing the true individual performance and satisfaction. However, you could argue that since SafeboosC III is a pragmatic trial, the training setup and evaluation of this, should also reflect how training and certification will be conducted and achieved, outside of the pilot setting.

Exploring if keeping the cover on an INVOS neonatal sensor to avoid contact between the infants' skin and the sensors adhesive layer affects the StO_2 measurements (paper 4)

As mentioned previously, some clinicians tend to minimise sensor adhesiveness before initiation of monitoring, in order to avoid exposing the fragile skin of extremely preterm infants to the adhesive layer on the sensor. One of the reported clinical practices, is to keep the paper cover on the adhesive NIRS sensors and thereby, creating a barrier between the adhesive sensor and the skin (personal communication). In previous blood-lipid phantom studies, NIRS sensors have been calibrated for the SafeBoosC-III trial without the cover on adhesive sensors. Since it is possible that the cover affects StO_2 measurements and thereby also the SafeBoosC hypoxic threshold, it is relevant to evaluate, if not removing the cover affected the StO_2 measurements, particularly the SafeBoosC-III hypoxic threshold (130).

Methods/design

The design of the blood-lipid phantom is explained in a previous publication by Kleiser et al (74). An image of the phantom setup can be seen in figure 2a. In this blood-lipid phantom study, the INVOS Infant-Neonatal Sensor together with the INVOS 5100C oximeter (Medtronic, Inc. Minneapolis, MN, USA), was tested with and without the sensors glossy white paper cover, against a reference oximeter, the Oxiplex TS (ISS, Inc., Champaign, IL, USA). In the blood-lipid phantom, recently expired human erythrocyte bags were mixed together with Intralipid, in order to obtain a 40 µM haemoglobin concentration and a scattering coefficient of 5.3 cm^{-1} (at 834 nm) (130). Before performing the first deoxygenation, an INVOS sensor with cover (S1), an INVOS sensor without cover (S2), and the Oxiplex TS, were attached to the windows of the phantom (figure 3). The sensor without cover (S2) was functioning as a control sensor, as the hypoxic threshold and sensitivity of the INVOS Infant-Neonatal Sensor has already been determined in previous blood-lipid phantom studies (63%) (74). A total of four deoxygenations were performed, changing the position of the INVOS sensors as well as cover-status between deoxygenations: between deoxygenation one and two, S1 and S2 changed position. Between deoxygenation two and three, the cover was removed from S1 so that both sensors were without cover. Between deoxygenation three and four, the two sensors were returned to their initial positions (figure 3) (130).

Figure 2a and 2b. Image of the general phantom setup (2a) and difference in absolute StO_2 values between the S1 and S2 sensor, during deoxygenation two in this experiment (2b).

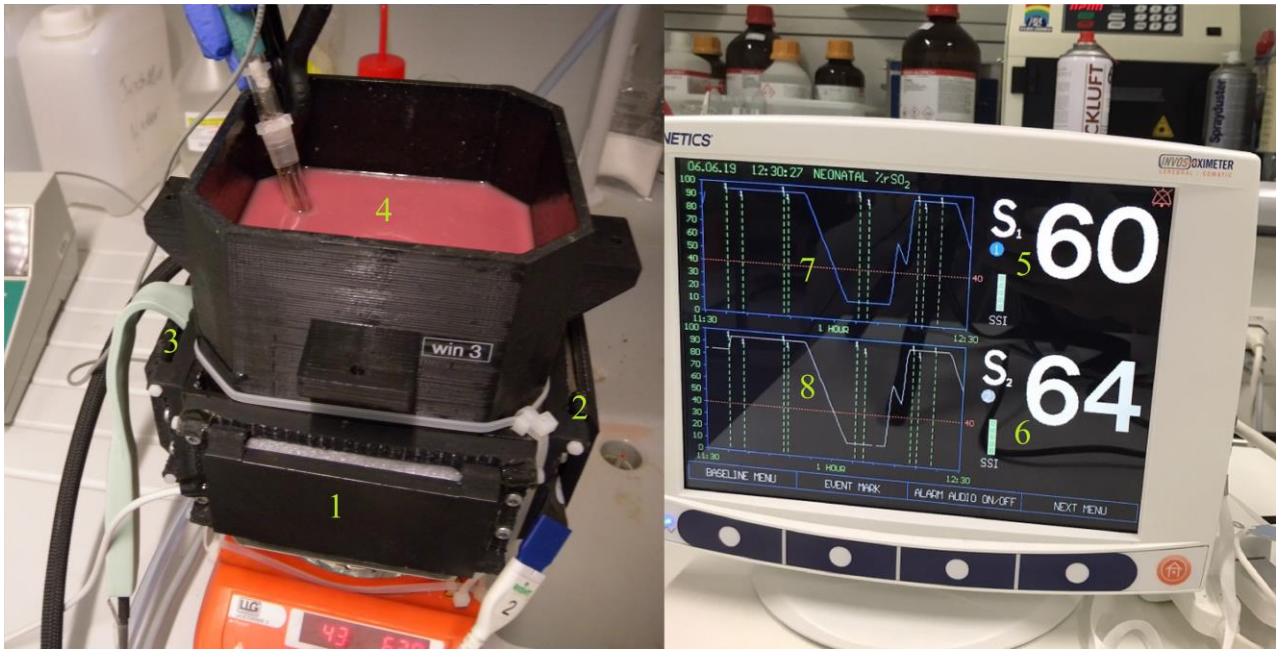


Figure 2a: 1,2 and 3 = phantom windows for NIRS sensors, 4 = blood-lipid mix in the phantom. Figure 2b: 5 and 6 = StO_2 values for S1 with cover and S2 without cover at specific timepoint during deoxygenation two. 7 and 8 = time-dependent changes in StO_2 for S1 and S2, during deoxygenation one and two.

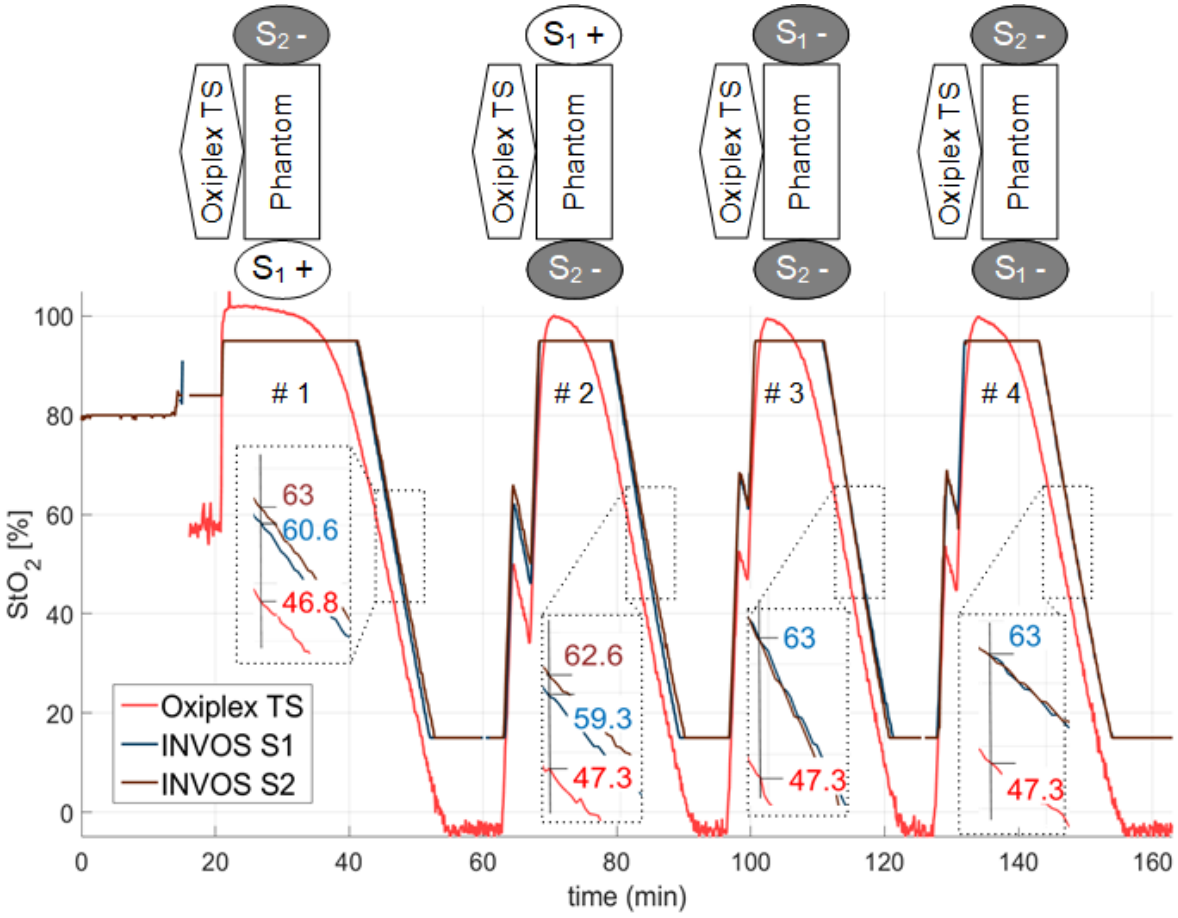
Statistical analysis

As a primary analysis, linear regression was used to determine the coefficients for linear fit between S1 with cover and Oxiplex TS (alfa being the sensors sensitivity to StO_2 changes) (figure 4). The previous blood-lipid phantom studies showed that an StO_2 of 47% for the Oxiplex TS, corresponds to 55% StO_2 for the INVOS small adult SomaSensor (74) and to 63% for the INVOS Infant-Neonatal sensor (130). The equation presented in the previous publication by Kleiser et al. (74), enabled us to calculate the StO_2 for the S1 (and S2) sensor with and without the cover, corresponding to 47% with the Oxiplex TS, i.e. the SafeBoosC hypoxic threshold

$$StO_2(INVOS) = a*StO_2(OxiplexTS) + b,\ a = slope,\ b = intercept$$

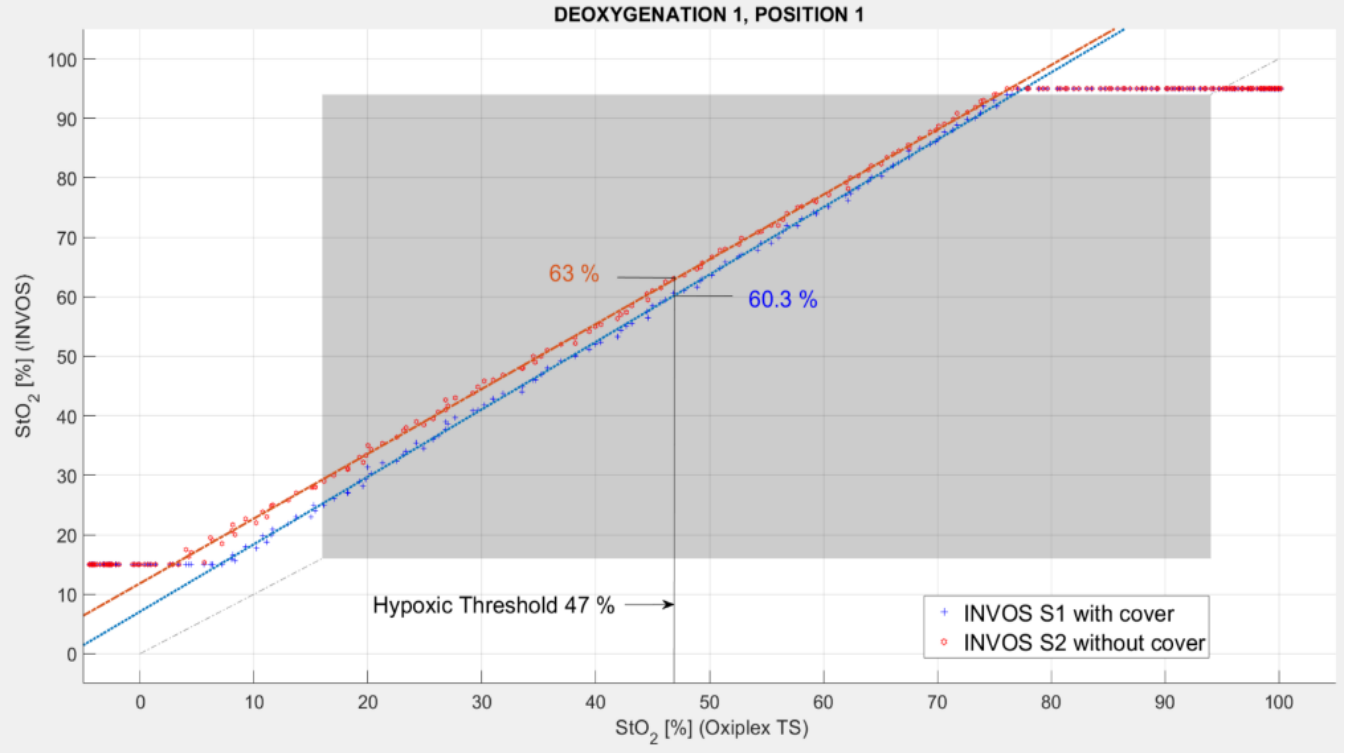
Secondly, sensitivity and hypoxic threshold for S1 during deoxygenation one and two where the cover was on, was compared to values for deoxygenation three and four, where the cover was off. As a control measure, sensitivity and hypoxic threshold were determined for S2, in order to make sure that these values were consistent with previous publications (74). To evaluate on inter-variance between sensors, values for S1 and S2 during deoxygenations three and four, where the cover was off for both sensors, were compared. To evaluate on sensor positioning, values for S1 during deoxygenation one was compared to values for deoxygenation two (130). Data was processed in MATLAB (MathWorks, Natick, MA, USA), as was done in previous blood-lipid phantom studies (71,74).

Figure 3 (130)**.** Overview of the four deoxygenations as well as sensor positions/status in the phantom



In the lower half of the figure, the four deoxygenations are illustrated as a time series with StO_2 values. In the upper half, the phantom including sensor positions during each deoxygenation is illustrated. S1 = INVOS neonatal sensor with cover during the first two deoxygenations, S2 = control sensor, i.e. INVOS neonatal sensor without cover. '+' means with cover, '-' means without cover.

Figure 4 (130)**.** INVOS neonatal sensors and Oxiplex TS linear correspondence during deoxygenation one



The blue curve represents the linear correlation of oxygenation values between S1 and Oxiplex TS. The red curve represents the linear correlation of oxygenation values between S2 and Oxiplex TS. The StO_2 values for S1 and S2, corresponding to 47% with Oxiplex TS is illustrated as well.

Results

In figure 2b, the difference in absolute StO_2 values between S1 (with cover) and S2 during deoxygenation two is illustrated. In figure 3, the four deoxygenations as well as sensor repositioning, and cover status is illustrated. A linear relationship was found between Oxiplex TS and S1, with and without cover (R^2 = 0.999). However, the S1 sensor was more sensitive to StO_2 changes when the cover was on, as can be seen by comparing the different coefficients for the linear fits in table 4 (a=1.133 for S1 with cover during deoxygenation one compared to a=1.103 for S1 without cover during deoxygenation four) (130). The hypoxic threshold was also dependent on cover status, with the S1 sensor with cover having a hypoxic threshold of 60.3% and 63.8% without cover (table 4). Repositioning of the sensor had little effect on these values (table 4). The inter-sensor variation was also negligible (table 4) (130).

Table 4 (130)**.** Overview of the four deoxygenations

Sensor	Position	Deoxygenation	Cover	a	b	StO_2 % Oxiplex TS 47%
S1	1	1	+	1.133	7.067	60.324
S1	2	2	+	1.144	6.708	60.468
S1	2	3	-	1.086	12.020	63.047
S1	1	4	-	1.103	11.992	63.836
Sensor	Position	Deoxygenation	Cover	a	b	StO_2 % Oxiplex TS 47%
S2	2	1	-	1.089	11.823	63.013
S2	1	2	-	1.092	12.466	63.797
S2	1	3	-	1.152	8.410	62.548
S2	2	4	-	1.105	12.173	64.124

This table provides an overview of sensor positioning as well as cover status (+ with and - without cover) during the four deoxygenations. It also includes linear coefficients (a=slope, b=intercept, StO_2(INVOS) = a*StO_2(OxiplexTS) + b) and hypoxic thresholds corresponding to 47% StO_2 with Oxiplex TS (55% with INVOS small adult sensor).

Strength and weaknesses

Using in vitro blood-lipid phantoms to evaluate NIRS oximeters and sensors, carries the advantage of being able to control the optical properties, such as scattering coefficient, haemoglobin concentration and the distance between the sensor and the phantom soluble, which makes it possible to ensure homogeneity between experiments and thus, a high reproducibility of the results (74). In this blood-lipid phantom study, the hypoxic thresholds for both sensors when the cover was taken off, were similar to what is previously reported for the INVOS neonatal sensor, in blood-lipid phantom studies (74,131) (threshold range for S1 without cover was 63% to 63.8%, and for S2 62.5% to 63.8%), thereby confirming the high reproducibility of the blood-lipid phantom (74). Furthermore, the in vitro setup makes it possible to test how the NIRS oximeters and sensors react to StO_2 changes, across a wide range of oxygenation values (74). Such validations would not be possible in vivo, since exposing human brains to low cerebral oxygenation for a longer time period can be harmful and therefore, ethically wrong. Evaluating how the NIRS oximeters and sensors react to StO_2 changes in the lower range, is especially valuable in the SafeBoosC setup, since the intervention focuses on hypoxic borderline values (70,90).

A major strength of this specific blood-lipid phantom study and setup was the change of sensor position as well as cover status, which allowed us to quantify both the repositioning effect, as well

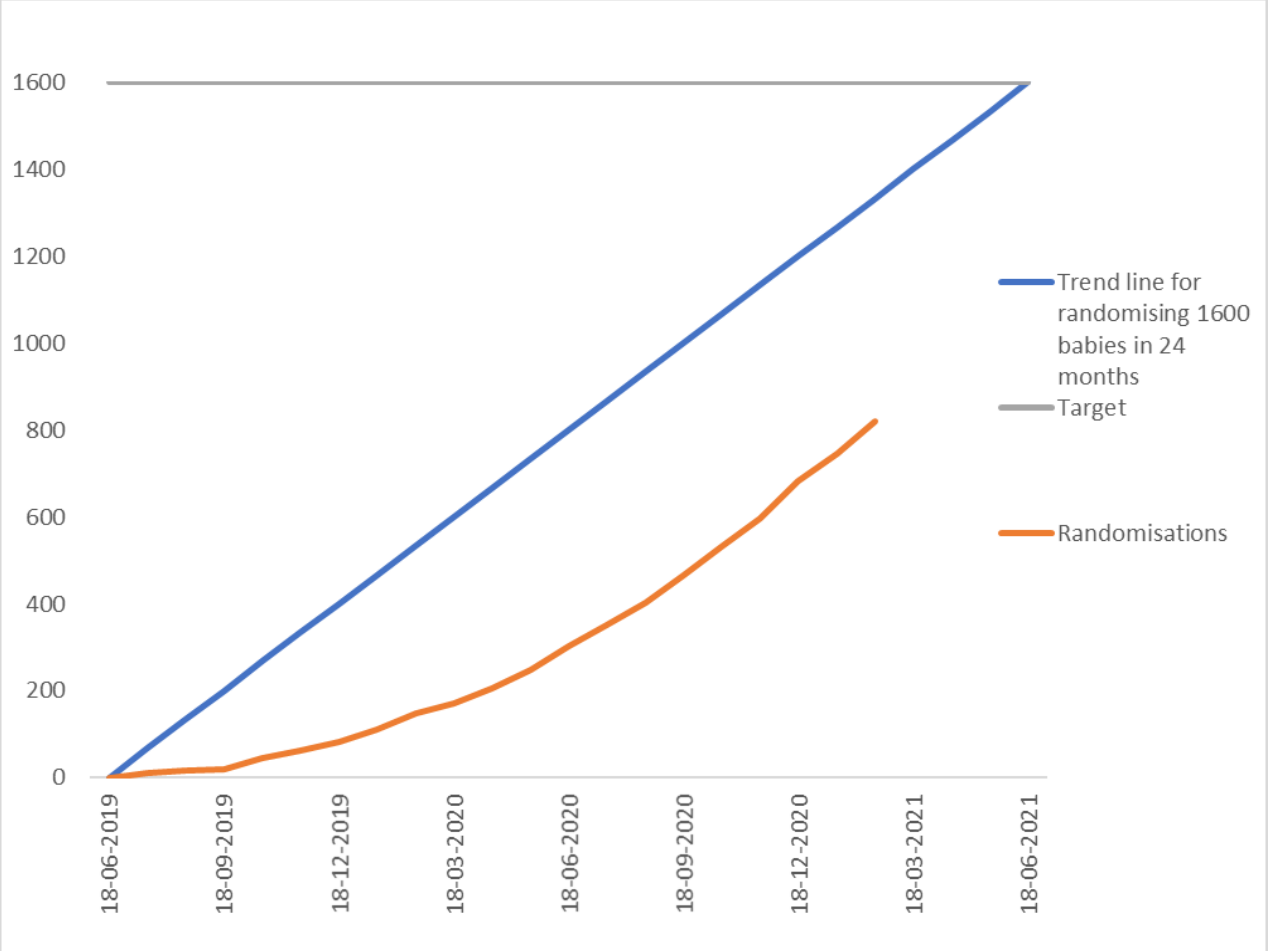
as inter-variance between sensors. As mentioned above, both had minor effect on the hypoxic threshold and sensitivity (table 4).

Overall discussion and perspective

Principal findings and status on the trial

A design paper, based on the full SafeBoosC-III protocol was submitted to Trials journal (trialsjournal.biomedcentral.com) on 22nd of February 2019, i.e. four months before the first infant was randomised, and accepted 4th of December 2019 (90). At this time point, only 71 infants had been randomised. Up until the 18th of February 2021, the protocol has been approved in 74 local ethical committees across 18 countries. The first infant was randomised in Copenhagen on the 27th of June 2019 and since then, 65 hospitals across China, Europe, India and the US have been opened for randomisation and as of the 18th of February 2021, a total of 819 infants have been randomised. Figure 5 shows the randomisation rate in the trial.

Figure 5. Randomisation rate in SafeBoosC III.



The 65 hospitals actively randomising have been opened for randomisation on a continuous basis, as they have completed the relevant trial preparations. This is also reflected in figure 5, where the randomisation rate is steadily increasing, as more hospitals are opened for randomisation. with the present randomisation rate, the trial is expected to complete recruitment in January 2022.

The detailed statistical analysis plan was also published prior to any data analysis (94). The twin simulation study showed that when the twin proportion and ICC are high, the primary analysis has a lower coverage. It also revealed, as expected, that the GEE analysis had a lower power if the twin proportion and ICC is high, due to a reduced effective sample size. However, both decrease in coverage and loss of power were minimal and unlikely to influence the trial results (94). The pilot study on the training module on NIRS monitoring, showed that the setup was feasible in an international, multilingual setting despite limited resources and difference in clinical practice; it was possible to complete the module within reasonable time, the academic level was appropriate and the majority found the content to be of high clinical relevance (108). The major critique points revealed from the quantitative and qualitative analyses, have been used to improve the NIRS module itself as well as the additional training modules. As of the 18th of February 2021, the SafeBoosC web-based training and certification program (trialeducation.info) is available in six different languages, more than 1300 staff members are participating in the training, and the satisfaction rate is high. Among the 274 responders to the evaluation and feedback form, 84% rates the overall quality of the online training program as ‘good’ or ‘very good’, 94% ‘agree’ or ‘strongly agree’ that the program is relevant for clinical practice, and that they know more about NIRS than they did before, respectively. Furthermore, 93% would recommend the training program to colleagues (unpublished data).

The blood-lipid phantom study showed that, despite the relationship between the INVOS neonatal sensor with cover and OxiplexTS was linear ($r^2 = 0.999$), the cover decreased the hypoxic threshold by more than 3% (60.3% with cover and 63.8% without cover). Furthermore, it also influenced the linear equation, with the INVOS neonatal sensor with cover being more sensitive to oxygenation changes than without the cover: $StO_{2\_INVOS\_cover}=1.133*StO_{2\_ISS}+7.1$ as opposed to $StO_{2\_INVOS\_nocover}=1.103*StO_{2\_ISS}+12.0$ (130).

Pragmatic trials – evidence designed to alter clinical practice

Randomised clinical trials are considered among the highest levels of evidence, when evaluating the benefits and harms of an intervention (132). Therefore, results from well-designed

randomised clinical trials should also be available and part of the decision-making when clinical practice is altered (133), preferably in the form of a meta-analysis including the results from available trials (134). Unfortunately, many new interventions are being implemented without proper evidence, including implementation of new technologies (133). Furthermore, the evidence available is often suboptimal for decision-making on changing routine practice, since most trials are of explanatory design, meaning they are designed to test if the intervention is beneficial or harmful under optimal conditions (104). Such trials are often conducted with experienced staff, on a selective participant group, with strict interventional instructions (135). The comparator is also often placebo and thereby, not comparing the intervention to usual practice (135). Furthermore, the primary end-point in such trials is often a surrogate or biological outcome, instead of a clinical or patient-relevant outcome (135). Explanatory trials have a high internal validity and are highly reproducible, but unfortunately lack external validity and generalisability, which are important when results from a trial are being used to predict the effect of an intervention, when it is implemented in the 'real world' (97). Therefore, results from explanatory trials do not necessarily mirror the interventions effect, if it is implemented into routine practice (104). If the trial is conducted at highly expertise sites with experienced investigators, the trial could potentially overestimate benefits and underestimate harms, as compared to the use of the intervention at a less experienced site (89). The SafeBoosC-II trial was more of explanatory design and was conducted in sites with experienced investigators. Thus, there is a possibility that SafeBoosC II have overestimated the potential benefit and underestimated potential harms, as compared to implementation of the intervention in a wide patient group across multiple hospitals, with various clinical practice. Moreover, a positive result on a surrogate endpoint as was used in SafeBoosC II, does not necessarily transfer to a positive result on a patient relevant, clinical end point. Therefore, before implementing a new intervention into routine practice, including treatment guided by cerebral NIRS monitoring, real-world evidence data should be obtained. This can be done through a pragmatic randomised clinical trial (89). A pragmatic trial is designed with the purpose of evaluating how the intervention works, if it was implemented into routine clinical practice, i.e. a 'real-world scenario'. This means that pragmatic trials should not test whether the intervention works under optimal conditions administered by experts, since that does not reflect a real-world scenario. Instead, a pragmatic trial needs to estimate how the intervention works, if it is administered by an average physician, supported by an average nurse, in an average hospital.

The PRECIS tool by Thorpe et al. is created to help trialists differentiate between explanatory and pragmatic trial design, when designing and planning a trial (135). Pragmatic trial criteria includes, among others, the following; minimal eligibility criteria, no constricted guidelines on how to use the intervention, treatment-as-usual or best alternative is used for comparison, follow-up based on routine data, clinical relevant primary outcome and using the 'intention-to-treat' population for the primary analysis (135). The intention is minimal intervention in routine practice. Although pragmatic versus explanatory trial designs is a continuum and that SafeBoosC III could have been more pragmatic in multiple aspects, it is overall of pragmatic design. Participants in SafeBoosC-III are included from both experienced and inexperienced hospitals across three different continents, thus ensuring a heterogenous study group. The instructions on how to implement and use the intervention is also very flexible; although the intervention demands an assessment of the clinical status and a possible intervention from the treatment guideline, when the StO_2 value drops below the predefined hypoxic threshold, the decisions taken on how (and if) to intervene, is completely dependent on the individual physicians' clinical evaluation. The clinician should use the evidence-based treatment guideline to adjust the cardiorespiratory support. However, the treatment guideline is not a flowchart, but simply a list of interventional suggestions (84). While it is important to avoid, that doctors and nurses participating in the trial are experts in the intervention, it is still relevant to ensure that they have some knowledge on the intervention as well as the trial conduct itself. This is relevant, not only for the safety of the trial participants, but also to give a practical estimate of the interventions effect, when implemented into routine care. If a new drug or technology is implemented in normal routine care, some introduction and training will be given to staff members before implementation. Therefore, the SafeBoosC web-based training and certification program has been developed, in order to provide a practical and realistic level of training and introduction to the trial (108). In SafeBoosC III, the comparator to the intervention is also treatment and monitoring as usual and not placebo. Furthermore, only data collected as per usual routine is used, the primary outcome is of direct clinical relevance and the statistical primary analysis relies on the intention-to-treat population (90).

However, being of pragmatic design is not enough for a trial to potentially change clinical practice. The effect of the intervention must also be clinically relevant. As previously stated under 'Sample size calculation', the SafeBoosC III-trial aims to detect an absolute risk reduction of 7.5% on the primary outcome (94). If the results of SafeBoosC III shows such a difference, it

means that implementation of the intervention in high-income countries, where the population at risk is approximately 50,000 per year, will result in an additional 4000 survivors without severe brain injury each year. It also corresponds to a number needed to treat of 15, and since the cost of one NIRS sensor is approximately 1000 dkk, the cost of saving one infant from death or a life with potential handicap, is 15,000 dkk.

Based on the pragmatic design of the SafeBoosC III trial reflecting a real-world scenario, as well as the size of the trial and potential health and economic impact, it is reasonable to state that the results will be both highly generalisable and have potential to effect clinical practice.

Of course, pragmatic trials cannot stand alone. Explanatory trials are still important, especially when evaluating the feasibility of an intervention, as was done in SafeBoosC II (70). In some scenarios, post-hoc analyses of the results from a pragmatic trial will require additional explanatory trials, in order to investigate these results deeper (136). A scenario in the SafeBoosC-III trial could be that the primary analysis showed no overall benefit of the intervention, but that the planned secondary random-effects meta-analysis (94) showed a high heterogenicity between centres and potential benefit of the intervention in some centres. Such results would require additional studies to explain this. If the SafeBoosC-III trials showed an overall benefit of the intervention, the next could be to design and conduct an explanatory trial with multiple arms using different interventional thresholds. This is one of the weaknesses of pragmatic trials which must be taken into account.

Delays and alterations due to obstructions

Despite that the trial has proven feasible in its present pragmatic design and that infants are now being randomised with an acceptable rate, there have been a number of obstructions during the preparation phase, significantly delaying randomisations as well as altering the trial design and conduct. In this section, the most significant obstructions will be outlined and discussed.

The SafeBoosC consortium behind the SafeBoosC-II trial, have tried to obtain funding for a phase-III trial since 2015, but it was not until March 2018 that funding was obtained from the Elsass foundation. However, despite applying for money to cover both central and local trial costs, only money for central trial costs were granted. This included salary to the trial manager (author of this thesis), money for Copenhagen Trial Unit to handle data management and

statistical analysis, yearly investigator meetings in Copenhagen, development of the web-based training program and project coordination. The money for local trial costs, was supposed to cover work hours for the principal investigator, NIRS monitors and sensors, local Good Clinical Practice monitoring and patient insurance. Despite the lack of local funding, the trial was initiated, meaning that hospitals wanting to participate had to find a way to cover their local expenses. Parallel to establishing the trial organisation, additional funding applications were sent out, in order to try and cover the local expenses. In particular, medical device companies producing and selling NIRS devices were contacted, in order to seek out the possibilities for industry funding. However, only a few were interested and eventually, all the major companies declined to support the trial. Especially the Medtronics company (Minneapolis, MN, USA), producing the INVOS devices, showed initial interest and encouraged submission of an application to cover the remaining trial costs. However, after months of dragging out their response, they declined the application. One can only speculate, why the medical device companies were not interested in collaborating. If the trial shows a clinical benefit of cerebral NIRS monitoring, the clinical uptake of NIRS might progress rapidly. On the other hand, if the trial shows no clinical benefit, or even harm, it might be detrimental for the future sale of NIRS devices, and since clinical uptake is already growing without the SafeBoosC-III trial (65), initiation of the SafeBoosC-III trial might not be in the interest of the industry.

So far, the lack of local funding has been solved by collaboration between individual hospitals and two medical device companies, as well as obtaining national grants. In Europe and India, multiple NICUs are collaborating with the smaller medical device company Oxyprem (Zürich, Switzerland), who mainly supports with monitors and sensors. In China, monitors, sensors, Good Clinical Practice Monitoring and patient insurance is covered by the Chinese medical device company Enginmed (Suzhou, China). Some countries, including Spain, Belgium and Turkey, have obtained national grants to cover local trial costs. Others have had less luck. In the United Kingdom, more than ten hospitals initially planning to participate in the trial, have had to withdraw due to the lack of funding to cover a national trial coordinator. In the United States, a large number of hospitals are also highly dependent on local funding in order to participate. They sent a US funding application to the Cerebral Palsy Alliance Research Foundation in November 2019, but due to the COVID-19 epidemic, the response got significantly delayed. It was not until the end of 2020, that they were informed that the funding had come through. Thus, as of February 2021, a group of US hospitals are now working hard to catch up and start recruiting. In

France, there was also total ‘radio-silence’ during 2019 and therefore, France hospitals are not expected to participate. It is unknown whether this is due to the lack of funding, to cover local costs.

The lack of funding has not only caused a delay, but it has altered the trial design in a more pragmatic direction. Initially, the trial was designed to include 1) central reading of cranial ultrasound scans by an expert panel instead of local reading, 2) pre-planned cranial ultrasound scans instead of routine scans 3) three follow-up times until 36 weeks of postmenstrual age instead of two and 4) extensive data collection including real-time entry of vital signs and administered interventions. However, with limited funding, such a setup was simply not feasible. Therefore, the design was altered to the present design so that participation in the trial requires minimal extra work for investigators. Furthermore, the Good Clinical Practice plan was also scaled so that it only included the minimal requirements to be in line with the Declaration of Helsinki. As described previously, the pragmatic alterations have weaknesses, but on the other hand is more in line with a pragmatic study design, thus strengthening the generalisability of the results (135).

Other obstructions that significantly delayed the trial conduct included 1) writing up the bi-lateral collaboration agreement between Copenhagen and each NICU due to the European General Data Protection Regulations, 2) development of the web-based training and certification program and 3) the COVID-19 pandemic.

During the almost worldwide lockdown that was initiated in March 2020 due to the COVID-19 pandemic (137), the randomisation rate decreased, despite that new hospitals were started to actively randomise on a continuous basis. The impact on the trial was even clearer when evaluating on the performance of the randomising hospitals; in January and February, each hospital had, on average, randomised 1,24 and 1,37 infants respectively. In March, this number decreased to 0,67 and increased only slowly during Spring. As the SafeBoosC-III steering group discussed this development and the potential mechanisms behind it, the picture was not clear. Some hospitals had to put all research on hold due to the pandemic, while others continued. There were also reports from some national coordinators that almost no extremely preterm infants were admitted to their hospitals, while others reported the opposite, or no difference. The SafeBoosC-III consortium therefore decided to evaluate whether the number of admitted

extremely preterm infants had actually changed in hospitals participating in SafeBoosC-III, during the lockdown. The principal investigators were asked to report the number of extremely preterm infant admissions to their hospital during the three most rigorous lockdown months, and the number of extremely preterm infant admissions during the same three months in 2019. Across the 46 hospitals participating in the study, 457 extremely preterm infants were born during the three most rigorous lockdown months in 2020, and 428 in the same months in 2019 (p=0.33, Chi-square test for 1x2 tables). Furthermore, regional differences were insignificant as well (138). It is therefore still uncertain what caused the sudden decrease; However, it is fair to assume that it is a combination of known (minor reduction in extremely preterm birth rates, temporarily termination of clinical research in hospitals) and unknown factors.

The initial aim was to recruit all 1600 infants over a 24 months period but due to the delays, only 574 infants have been randomised during the first 19 months. However, the randomisation rate has been good for the last six months and since 20 additional hospitals are actively preparing to start randomising as well, we are confident that the present randomisation rate will continue, perhaps even increase. Therefore, the trial is expected to complete recruitment by January 2021. Due to the decentralised funding, the SafeBoosC-III trial is resilient to delays, so this delay is not a thread to the trial.

Future considerations

In the SafeBoosC-III trial, the primary outcome is assessed at 36 weeks of postmenstrual age, thereby reflecting brain injuries occurring in the neonatal period (90). However, as previously mentioned, not all infants with a severe brain injury on cranial ultrasound scans in the neonatal period suffers from long-term neurologic complications (25,28–30). Even more important, some infants who have no brain injuries on cranial ultrasound scans in the neonatal period are being diagnosed with neurodevelopmental impairment later in life. In a study by O'Shea et al, it was found that 23% of infants with no ultrasound lesion in the neonatal period, had delayed mental development and 26% had delayed psychomotor development at two years of age, as defined by the Bayley test (25). Therefore, it is relevant to detect whether a potential beneficial effect on brain injury also persists into childhood.

There is also a potential risk of harm that cannot be detected at the follow-up time of 36 weeks postmenstrual age. This has been seen in previous neonatal trials. A Cochrane review evaluating the use of early systemic postnatal corticosteroids to prevent bronchopulmonary dysplasia in

preterm infants, found that the intervention facilitated earlier extubation and decreased the risk of bronchopulmonary dysplasia up until 36 weeks of life. However, when data from long-term follow up were analysed, it revealed an increased risk of adverse neurological outcome including cerebral palsy (139). This demonstrates the importance of later follow-up after randomised clinical trials, to assess potential harmful effect. Thus, it is also relevant to evaluate if the SafeBoosC intervention has any long-term harmful effects.

Following the SafeBoosC-II trial, 115 infants were followed up at 24 months of corrected age, in order to evaluate the interventions' effect on neurodevelopment. The analyses showed no significant difference between the experimental and control arm. However, the study was not powered to detect a relevant difference on clinical outcomes (140). Since we intend to randomise 1600 babies in the SafeBoosC III trial, there is a potential for a much greater follow-up and therefore, a possibility to obtain sufficient power for a meaningful test of the interventions effect on neurodevelopment later in life, and to assess any potential harms. To conclude, the next step for the SafeBoosC project following the SafeBoosC-III trial, will be to conduct the two-year follow-up study.

Conclusion

To conclude

1) A pragmatic protocol has been designed and published, sufficient of size and external validity to answer if treatment guided by cerebral NIRS monitoring during the first days of life, improves clinical outcomes in extremely preterm infants. The number of randomisations and active hospitals proves the feasibility of the study design. The next step will be to complete the trial and thereafter, evaluate if potential improvement persists into early childhood.
2) A statistical analysis plan focusing on the intention-to-treat population, taking into consideration the importance of both statistical and clinical significance as well as the large proportion of twins and their ICC, has been designed and published before any data analysis. The simulation study showed that a potential correlation within twin couples for the primary outcome, would have minimal effect on the trial results.
3) Based on the pilot study results as well as the high participation and satisfaction rate of the overall SafeBoosC web-based training program, the described method of developing and implementing an online training program for an international trial despite language barriers,

limited resources and potential difference in clinical practice between participating sites, has proven feasible. The results can be used for future trialists intending to plan and prepare for multicentre trials.

4) If the cover is kept on the INVOS neonatal sensor during cerebral NIRS monitoring, the hypoxic threshold will differ by more than 3%. It is plausible that other clinical practices to minimise sensor adhesiveness will affect StO_2 measurements as well. Clinicians must be aware of this, both inside and outside the SafeBoosC-III setting. Further studies evaluating how such practices affects the StO_2 measurements are needed.

Dansk resumé (Summary in Danish)

Grundet et immaturt kardiopulmonalt system, er ekstremt for tidligt fødte børn i risiko for at opleve nedsat systemisk blodcirkulation i de første dage efter fødslen. Kombineret med en dårlig cerebral autoregulering, kan dette medføre insufficient blodtilførsel til hjernen og dermed øge risikoen for hjerneskade og efterfølgende død. Tidlig neonatal intensiv terapi involverer kompleks behandling, vejledt af ekstensiv monitorering. Monitorering af respirationen er veletableret, men mulighederne for kontinuerligt at monitorere cirkulationen er dårlig. Ydermere findes der ikke et redskab til at monitorere hjernen. Cerebral nær-infrarød spektroskopi (NIRS) kunne være det næste monitoreringsredskab til at løse disse problemer, på de neonatale intensivafdelinger. Det er en ikke-invasiv teknologi, som anvender nær-infrarødt lys til at monitorere hjernevævets iltning, og det har vist sig, at målingerne korrelerer godt med både central venøs iltning og hjertets pumpeevne. Frem til nu, er der ikke noget lodtrækningsforsøg med sufficient statistisk styrke, som har vist en klinisk fordel ved cerebral NIRS-monitorering som et ekstra redskab, til at justere intensivterapien hos ekstremt for tidligt fødte børn. På trods af dette, øges brugen af NIRS. Som ved alle andre interventioner, er der en risiko for skade, i dette tilfælde især på lunger, øjne og hud. Derfor er der behov for et stort randomiseret klinisk forsøg, som kan frembringe 'real world' evidens og belyse fordele og ulemper ved cerebral NIRS-monitorering hos ekstremt for tidligt fødte børn. Det er derfor, at det multinationale, pragmatiske, randomiserede kliniske forsøg, 'Safeguarding the Brains of our smallest Children (SafeBoosC) III' nu bliver udført.

Formålet med denne afhandling er at beskrive processen med at designe og forberede SafeBoosC-III forsøget, initiere forsøget og randomisere de første børn. Dette inkluderer

1) At designe et gennemførligt og pragmatisk forsøg, og publicere protokollen før det første barn bliver randomiseret (artikel 1).
2) At designe en statistisk analyseplan som tager højde for det pragmatiske forsøgsdesign, den store prævalens af tvillinger i populationen samt vigtigheden af ikke kun statistisk, men også klinisk relevante resultater. Målet var også at publicere den statistiske analyseplan før den første dataanalyse (artikel 2).
3) At pilotere et online træningsmodul omkring NIRS-monitorering, for at evaluere mulighederne for at udvikle et større, gennemførligt online træningsprogram på flere sprog (artikel 3)

4) Eksplorere om dét at beholde papirsdækkenet på en INVOS neonatal sensor, for at undgå kontakt mellem barnets hud og sensorens klæbende overflade, påvirker iltmålingen (artikel 4)

Resultater

SafeBoosC-III er open-label og med to parallelle forsøgsgrupper. Børn randomiseret til eksperimentalgruppen modtager behandling vejledt af cerebral NIRS-monitorering i de første 72 timer af livet. Behandlingen er baseret på en evidensbaseret behandlingsvejledning, og skal initieres når den cerebrale iltning falder under en prædefineret hypoksisk tærskel. Børn randomiseret til kontrolgruppen modtager behandling og monitorering som vanligt, dvs. ingen cerebral NIRS-monitorering. Børnene følges op når de er nået 36 ugers postmenstruel alder eller når de udskrives til hjemmet. Det primære effektmål er død eller overlevelse med svær hjerneskade detekteret ved cerebral ultralyd. Protokollen til SafeBoosC-III blev indsendt til Trials tidsskrift (trialsjournal.biomedcentral.com) fire måneder før det første barn blev randomiseret, og den blev publiceret efter 71 børn var blevet randomiseret. Frem til den 18. Februar 2021, er protokollen godkendt af 74 lokale etiske komiteer på tværs af 18 lande.

En detaljeret statistisk analyseplan er blevet publiceret inden den første dataanalyse. Sample size beregningen viste at det er nødvendigt at randomisere 1600 ekstremt for tidligt fødte børn, såfremt man vil detektere en 22% relativ risiko forskel mellem eksperimental- og kontrolgruppen, for det primære effektmål. Den primære analyse vil blive baseret på intention-to-treat populationen. Mixed-model logistisk regression eller lineær regression vil blive anvendt i den primære analyse, afhængigt af effektmålet. Analysen vil blive justeret for stratificeringsvariablerne. Tvillinge intraclass korrelationskoefficienten (ICC) vil ikke blive inkluderet i den primære analyse. I stedet vil der blive foretaget en sensitivitetsanalyse, for at belyse den potentielle effekt af korrelationen mellem tvillinger, på effektmålet. Et simuleringsforsøg viste, at hvis ICC'en og proportionen af tvillinger er høj, vil dækningen af den primære analyse falde, dog kun minimalt. For den alternative generalised estimation equation (GEE) analyse, vil en høj tvillingeproportion og en høj ICC, som forventet, medføre et fald i statistisk styrke. Dette var dog også minimalt.

Af de 100 læger og sygeplejerske som blev inviteret til at deltage i piloteringen af det online NIRS træningsmodul, var der 81 som begyndte på træningsmodulet og besvarede det online spørgeskema. Mediantiden for at gennemføre modulet var 15 minutter og medianantallet af spørgsmålet, som blev brugt for at gennemføre modulet, var syv. De fleste deltagere fandt det akademiske niveau passende, for både quizzen og læringsmaterialet (93% og 85% respektive). Ydermere, var 90% af

deltagerne enige i at modulet var relevant som forberedelse til at bruge NIRS. Den tematiske analyse afslørede vigtige problemer, herunder tekniske vanskeligheder, generel uklarhed, samt diskrepans imellem læringsmaterialet og quizzen. Disse resultater er blevet brugt til at forbedre udviklingen af de resterende moduler til træningsprogrammet, samt til at forberede det piloterede NIRS modul. Per den 18. februar 2021, er træningsprogrammet tilgængeligt på seks forskellige sprog og mere end 1300 læger og sygeplejersker er begyndt at træne. Den overordnede tilfredshed er høj.

Blod-lipid fantomforsøget viste, at selvom forholdet mellem INVOS neonatal sensor med dækken og OxiplexTS var lineær ($r^2 = 0.999$), så sænkede dækkenet den hypoksiske tærskel med mere end 3% (60.3% med dækken og 63.8% uden dækken). Ydermere påvirkede dækkenet også den lineære sammenhæng, idet INVOS neonatal sensor med dækken er mere sensitiv for ændringer i iltningen, end når dækkenet er taget af: $StO_{2\_INVOS\_dækken}=1.133*StO_{2\_ISS}+7.1$ i forhold til $StO_{2\_INVOS\_intetdækken}=1.103*StO_{2\_ISS}+12.0$.

Forsøgsstatus

Det første barn blev randomiseret i København den 27. juni 2019. Siden da er 65 hospitaler på tværs af Kina, Europa, Indian og USA blevet åbnet for randomisering. Per den 18. februar 2021 er der blevet randomiseret 819 børn. Med den nuværende randomiseringsrate, forventes det at alle børn er rekrutteret til forsøget senest januar 2022.

Konklusion

1) Et pragmatisk forsøg, beskrevet ved protokollen, af sufficient størrelse og med høj ekstern validitet er blevet publiceret, for at forsøge at svare på om behandling vejledt af cerebral NIRS-monitorering i de første dage efter fødslen, forbedrer kliniske effektmål hos ekstremt for tidligt fødte børn. Antallet af randomiseringer og aktive hospitaler beviser, at SafeBoosC-III er gennemførligt. Næste skridt vil være at gennemføre forsøget og herefter undersøge, om en potentiel fordel også kan måles senere i livet.
2) En statistisk analyseplan med fokus på 'intention-to-treat' populationen, og som tager højde for vigtigheden af både statistisk og klinisk signifikans samt tvillingeproblematikken, er blevet designet og publiceret inden den første dataanalyse. Simuleringsforsøget viste at en potentiel korrelation imellem tvillinger, med hensyn til det primære effektmål, vil have minimal effekt på resultaterne.

3) Baseret på resultaterne fra piloteringen samt den høje deltagelses- og tilfredshedsrate i SafeBoosC-III webbaseret træningsprogram, vurderes det at den beskrevne metode til at udvikle og implementere et online træningsprogram for et international forsøg er gennemførlig. Dette på trods af sprogbarriere, få ressourcer og potentielle forskellige i kliniske praktisk, de deltagende afdelinger imellem. Det forventes at resultaterne kan bruges af andre forskere, som også planlægger og udfører multinationale, randomiserede kliniske forsøg.
4) Hvis dækkenet bliver siddende på INVOS neonatale sensor under cerebral NIRS-monitorering, vil den hypoksiske tærskel ændres med mere end 3%. Det er rimeligt at antage, at andre lignende kliniske praksis for at mindske sensorens klæbeevne, også vil påvirke iltmålingerne. Det er vigtigt at klinikere er opmærksomme på dette, både indeni og uden for SafeBoosC-III forsøget. Yderligere studier, som evaluerer hvordan lignende kliniske praksis påvirker iltmålingerne er nødvendige.

References

1. Blencowe H, Cousens S, Oestergaard MZ, Chou D, Moller A-B, Narwal R, et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *Lancet* 2012; 379:2162–72.
2. Howson C, Kinney M, Lawn J, eds. March of Dimes, PMNCH, Save the Children, WHO. Born Too Soon: The Global Action Report on Preterm Birth. Geneva: World Health Organization; 2012.
3. Stoll BJ, Hansen NI, Bell EF, Shankaran S, Laptook AR, Walsh MC, et al. Neonatal Outcomes of Extremely Preterm Infants From the NICHD Neonatal Research Network. *Pediatrics* 2010; 126:443–56.
4. Rogers EE, Hintz SR. Early neurodevelopmental outcomes of extremely preterm infants. *Semin Perinatol* 2016; 40:497–509.
5. Adams-Chapman I, Heyne RJ, DeMauro SB, Duncan AF, Hintz SR, Pappas A, et al. Neurodevelopmental Impairment Among Extremely Preterm Infants in the Neonatal Research Network. *Pediatrics* 2018; 141:e20173091.
6. Petrou S, Sach T, Davidson L. The long-term costs of preterm birth and low birth weight: results of a systematic review. *Child Care Health Dev* 2001; 27:97–115.
7. Behrman R, Butler AS. Preterm birth: Causes, Consequences and Prevention. Institute of Medicine. Washington, D.C.: National Academies Press; 2007.
8. Stoll BJ, Hansen NI, Bell EF, Walsh MC, Carlo WA, Shankaran S, et al. Trends in Care Practices, Morbidity, and Mortality of Extremely Preterm Neonates, 1993-2012. *JAMA* 2015; 314:1039.
9. Kontis V, Bennett JE, Mathers CD, Li G, Foreman K, Ezzati M. Future life expectancy in 35 industrialised countries: projections with a Bayesian model ensemble. *Lancet* 2017; 389:1323–35.
10. Kluckow M. Low systemic blood flow and pathophysiology of the preterm transitional circulation. *Early Hum Dev* 2005; 81:429–37.
11. Kluckow M. The Pathophysiology of Low Systemic Blood Flow in the Preterm Infant. *Front Pediatr* 2018; 6:1–6.
12. Ward RM, Beachy JC. Neonatal complications following preterm birth. *BJOG An Int J Obstet Gynaecol* 2003; 110:8–16.

13. Ghanayem NS, Gordon JB. Modulation of pulmonary vasomotor tone in the fetus and neonate Nancy. *Respir Res* 2001; 2:139–44.
14. Rudolph AM. The Changes in the Circulation After Birth. *Circulation* 1970; 41:343–59.
15. Meng L, Wang Y, Zhang L, McDonagh DL. Heterogeneity and Variability in Pressure Autoregulation of Organ Blood Flow. *Crit Care Med* 2019; 47:436–48.
16. Boylan GB, Young K, Panerai RB, Rennie JM, Evans DH. Dynamic Cerebral Autoregulation in Sick Newborn Infants. *Pediatr Res* 2000; 48:12–7.
17. Soul JS, Hammer PE, Tsuji M, Saul JP, Bassan H, Limperopoulos C, et al. Fluctuating Pressure-Passivity Is Common in the Cerebral Circulation of Sick Premature Infants. *Pediatr Res* 2007; 61:467–73.
18. Greisen G. Autoregulation of cerebral blood flow in newborn babies. *Early Hum Dev* 2005; 81:423–8.
19. Munro MJ, Walker AM, Barfield CP. Hypotensive Extremely Low Birth Weight Infants Have Reduced Cerebral Blood Flow. *Pediatrics* 2004; 114:1591–6.
20. Wong FY, Silas R, Hew S, Samarasinghe T, Walker AM. Cerebral oxygenation is highly sensitive to blood pressure variability in sick preterm infants. *PLoS One* 2012; 7:e43165.
21. Bel F, Bor M, Stijnen T, Baan J, Ruys JH. Aetiological Rôle Of Cerebral Blood-Flow Alterations In Development And Extension Of Peri-Intraventricular Haemorrhage. *Dev Med Child Neurol* 1987; 29:601–14.
22. Perlman JM, Goodman S, Kreusser KL, Volpe JJ. Reduction in Intraventricular Hemorrhage by Elimination of Fluctuating Cerebral Blood-Flow Velocity in Preterm Infants with Respiratory Distress Syndrome. *N Engl J Med* 1985; 312:1353–7.
23. Volpe JJ. Brain Injury in the Premature Infant: Neuropathology, Clinical Aspects, Pathogenesis, and Prevention. *Clin Perinatol* 1997; 24:567–87.
24. Alderliesten T, Lemmers PMA, Smarius JJM, van de Vosse RE, Baerts W, van Bel F. Cerebral Oxygenation, Extraction, and Autoregulation in Very Preterm Infants Who Develop Peri-Intraventricular Hemorrhage. *J Pediatr* 2013; 162:698–704.e2.
25. O'Shea TM, Kuban KCK, Allred EN, Paneth N, Pagano M, Dammann O, et al. Neonatal Cranial Ultrasound Lesions and Developmental Delays at 2 Years of Age Among Extremely Low Gestational Age Children. *Pediatrics* 2008; 122:e662–9.
26. Ment LR, Bada HS, Barnes P, Grant PE, Hirtz D, Papile LA, et al. Practice parameter: Neuroimaging of the neonate: Report of the Quality Standards Subcommittee of the American Academy of Neurology and the Practice Committee of the Child Neurology

Society. *Neurology* 2002; 58:1726–38.

27. de Vries LS, Benders MJNL, Groenendaal F. Imaging the premature brain: ultrasound or MRI? *Neuroradiology* 2013; 55:13–22.
28. Nongena P, Ederies A, Azzopardi D V, Edwards AD. Confidence in the prediction of neurodevelopmental outcome by cranial ultrasound and MRI in preterm infants. *Arch Dis Child - Fetal Neonatal Ed* 2010; 95:F388–90.
29. Papile L-A, Burstein J, Burstein R, Koffler H. Incidence and evolution of subependymal and intraventricular hemorrhage: A study of infants with birth weights less than 1,500 gm. *J Pediatr* 1978; 92:529–34.
30. Guzzetta F, Shackelford GD, Volpe S, Perlman JM, Volpe JJ. Periventricular intraparenchymal echodensities in the premature newborn: critical determinant of neurologic outcome. *Pediatrics* 1986; 78:995–1006.
31. Novak CM, Ozen M, Burd I. Perinatal Brain Injury. *Clin Perinatol* 2018; 45:357–75.
32. Murphy BP. Posthaemorrhagic ventricular dilatation in the premature infant: natural history and predictors of outcome. *Arch Dis Child - Fetal Neonatal Ed* 2002; 87:F37–41.
33. Cooke R, Rennie J, Weindling M, Whitelaw A, Stewart A, Hope P, et al. Randomised trial of early tapping in neonatal posthaemorrhagic ventricular dilatation. Ventriculomegaly Trial Group. *Arch Dis Child* 1990; 65:3–10.
34. Zayek MM, Benjamin JT, Maertens P, Trimm RF, Lal C V., Eyal FG. Cerebellar hemorrhage: a major morbidity in extremely preterm infants. *J Perinatol* 2012; 32:699–704.
35. Volpe JJ. Brain injury in premature infants: a complex amalgam of destructive and developmental disturbances. *Lancet Neurol* 2009; 8:110–24.
36. de Vries LS, van Haastert I-LC, Rademaker KJ, Koopman C, Groenendaal F. Ultrasound abnormalities preceding cerebral palsy in high-risk preterm infants. *J Pediatr* 2004; 144:815–20.
37. Horsch S, Muentjes C, Franz A, Roll C. Ultrasound diagnosis of brain atrophy is related to neurodevelopmental outcome in preterm infants. *Acta Paediatr* 2005; 94:1815–21.
38. Miall-Allen VM, De Vries LS, Whitelaw AG. Mean arterial blood pressure and neonatal cerebral lesions. *Arch Dis Child* 1987; 62:1068–9.
39. Lemmers PMA, Benders MJNL, DAscenzo R, Zethof J, Alderliesten T, Kersbergen KJ, et al. Patent Ductus Arteriosus and Brain Volume. *Pediatrics* 2016; 137:e20153090.
40. Perlman JM. White matter injury in the preterm infant: an important determination of

abnormal neurodevelopment outcome. *Early Hum Dev* 1998; 53:99–120.

41. Greisen G, Vannucci RC. Is Periventricular Leucomalacia a Result of Hypoxic-Ischaemic Injury? *Biol Neonate* 2001; 79:194–200.
42. Dani C, Pratesi S, Fontanelli G, Barp J, Bertini G. Blood transfusions increase cerebral, splanchnic, and renal oxygenation in anemic preterm infants. *Transfusion* 2010; 50:1220–6.
43. Beltempo M, Isayama T, Vento M, Lui K, Kusuda S, Lehtonen L, et al. Respiratory Management of Extremely Preterm Infants: An International Survey. *Neonatology* 2018; 114:28–36.
44. Dempsey EM, Barrington KJ. Treating hypotension in the preterm infant: when and with what: a critical and systematic review. *J Perinatol* 2007; 27:469–78.
45. Prescott S, Keim-Malpass J. Patent Ductus Arteriosus in the Preterm Infant. *Adv Neonatal Care* 2017; 17:10–8.
46. McNelis K, Fu TT, Poindexter B. Nutrition for the Extremely Preterm Infant. *Clin Perinatol* 2017; 44:395–406.
47. Kao LS, Morris BH, Lally KP, Stewart CD, Huseby V, Kennedy KA. Hyperglycemia and morbidity and mortality in extremely low birth weight infants. *J Perinatol* 2006; 26:730–6.
48. Schmalisch G. Basic principles of respiratory function monitoring in ventilated newborns: A review. *Paediatr Respir Rev* 2016; 20:76–82.
49. Hochwald O, Borenstein-Levin L, Dinur G, Jubran H, Ben-David S, Kugelman A. Continuous Noninvasive Carbon Dioxide Monitoring in Neonates: From Theory to Standard of Care. *Pediatrics* 2019; 144:e20183640.
50. Bergon-Sendin E, Perez-Grande C, Lora-Pablos D, De la Cruz Bertolo J, Moral-Pumarega MT, Bustos-Lozano G, et al. Auditing of Monitoring and Respiratory Support Equipment in a Level III-C Neonatal Intensive Care Unit. *Biomed Res Int* 2015; 2015:1–9.
51. El-Khuffash A, McNamara PJ. Hemodynamic Assessment and Monitoring of Premature Infants. *Clin Perinatol* 2017; 44:377–93.
52. Alderliesten T, Lemmers PMA, van Haastert IC, de Vries LS, Bonestroo HJC, Baerts W, et al. Hypotension in Preterm Neonates: Low Blood Pressure Alone Does Not Affect Neurodevelopmental Outcome. *J Pediatr* 2014; 164:986–91.
53. Wolf M, Greisen G. Advances in Near-Infrared Spectroscopy to Study the Brain of the Preterm and Term Neonate. *Clin Perinatol* 2009; 36:807–34.

54. Moran M, Miletin J, Pichova K, Dempsey E. Cerebral tissue oxygenation index and superior vena cava blood flow in the very low birth weight infant. *Acta Paediatr* 2009; 98:43–6.
55. Wolf M, Naulaers G, van Bel F, Kleiser S, Greisen G. A Review of near Infrared Spectroscopy for Term and Preterm Newborns. *J Near Infrared Spectrosc* 2012; 20:43–55.
56. Toet MC, Lemmers PMA. Brain monitoring in neonates. *Early Hum Dev* 2009; 85:77–84.
57. Greisen G, Leung T, Wolf M. Has the time come to use near-infrared spectroscopy as a routine clinical tool in preterm infants undergoing intensive care? *Philos Trans R Soc A Math Phys Eng Sci* 2011; 369:4440–51.
58. Marin T, Moore J. Understanding Near-Infrared Spectroscopy. *Adv Neonatal Care* 2011; 11:382–8.
59. Mancini DM, Bolinger L, Li H, Kendrick K, Chance B, Wilson JR. Validation of near-infrared spectroscopy in humans. *J Appl Physiol* 1994; 77:2740–7.
60. Watzman HM, Kurth CD, Montenegro LM, Rome J, Steven JM, Nicolson SC. Arterial and Venous Contributions to Near-infrared Cerebral Oximetry. *Anesthesiology* 2000; 93:947–53.
61. Tortoriello TA, Stayer SA, Mott AR, McKenzie ED, Fraser CD, Andropoulos DB, et al. A noninvasive estimation of mixed venous oxygen saturation using near-infrared spectroscopy by cerebral oximetry in pediatric cardiac surgery patients. *Pediatr Anesth* 2005; 15:495–503.
62. Weiss M, Dullenkopf A, Kolarova A, Schulz G, Frey B, Baenzigar O. Near-infrared spectroscopic cerebral oxygenation reading in neonates and infants is associated with central venous oxygen saturation. *Pediatr Anesth* 2005; 15:102–9.
63. Daubeney PEF, Pilkington SN, Janke E, Charlton GA, Smith DC, Webber SA. Cerebral oxygenation measured by near-infrared spectroscopy: Comparison with jugular bulb oximetry. *Ann Thorac Surg* 1996; 61:930–4.
64. McNeill S, Gatenby JC, McElroy S, Engelhardt B. Normal cerebral, renal and abdominal regional oxygen saturations using near-infrared spectroscopy in preterm infants. *J Perinatol* 2011; 31:51–7.
65. Hunter CL, Oei JL, Suzuki K, Lui K, Schindler T. Patterns of use of near-infrared spectroscopy in neonatal intensive care units: international usage survey. *Acta Paediatr* 2018; 107:1198–204.

66. Yu Y, Zhang K, Zhang L, Zong H, Meng L, Han R. Cerebral near-infrared spectroscopy (NIRS) for perioperative monitoring of brain oxygenation in children and adults. *Cochrane Database Syst Rev* 2018:CD01947.
67. Serraino GF, Murphy GJ. Effects of cerebral near-infrared spectroscopy on the outcome of patients undergoing cardiac surgery: a systematic review of randomised trials. *BMJ Open* 2017; 7:e016613.
68. Dullenkopf A, Frey B, Baenziger O, Gerber A, Weiss M. Measurement of cerebral oxygenation state in anaesthetized children using the INVOS 5100 cerebral oximeter. *Pediatr Anesth* 2003; 13:384–91.
69. Alderliesten T, Dix L, Baerts W, Caicedo A, van Huffel S, Naulaers G, et al. Reference values of regional cerebral oxygen saturation during the first 3 days of life in preterm neonates. *Pediatr Res* 2016; 79:55–64.
70. Hyttel-Sørensen S, Pellicer A, Alderliesten T, Austin T, Van Bel F, Benders M, et al. Cerebral near infrared spectroscopy oximetry in extremely preterm infants: phase II randomised clinical trial. *BMJ* 2015; 350:1–11.
71. Kleiser S, Ostojic D, Andresen B, Nasseri N, Isler H, Scholkmann F, et al. Comparison of tissue oximeters on a liquid phantom with adjustable optical properties: an extension. *Biomed Opt Express* 2018; 9:86.
72. Hyttel-Sorensen S, Sorensen LC, Riera J, Greisen G. Tissue oximetry: a comparison of mean values of regional tissue saturation, reproducibility and dynamic range of four NIRS-instruments on the human forearm. *Biomed Opt Express* 2011; 2:3047.
73. Hyttel-Sorensen S, Hessel TW, Greisen G. Peripheral tissue oximetry: comparing three commercial near-infrared spectroscopy oximeters on the forearm. *J Clin Monit Comput* 2014; 28:149–55.
74. Kleiser S, Nasseri N, Andresen B, Greisen G, Wolf M. Comparison of tissue oximeters on a liquid phantom with adjustable optical properties. *Biomed Opt Express* 2016; 7:2973.
75. Johnson DE. Extremely Preterm Infant Skin Care. *Adv Neonatal Care* 2016; 16:S26–32.
76. Lund C. Medical Adhesives in the NICU. *Newborn Infant Nurs Rev* 2014; 14:160–5.
77. Green MS, Sehgal S, Tariq R. Near-Infrared Spectroscopy: The New Must Have Tool in the Intensive Care Unit? *Semin Cardiothorac Vasc Anesth* 2016; 20:213–24.
78. Thiele RH, Shaw AD, Bartels K, Brown CH, Grocott H, Heringlake M, et al. American Society for Enhanced Recovery and Perioperative Quality Initiative Joint Consensus Statement on the Role of Neuromonitoring in Perioperative Outcomes: Cerebral Near-

Infrared Spectroscopy. *Anesth Analg* 2020; 131:1444–55.

79. Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC Med Res Methodol* 2014; 14:120.
80. Hansen ML, Hyttel-Sørensen S, Jakobsen JC, Gluud C, Kooi EMW, Mintzer J, et al. The clinical effects of cerebral near-infrared spectroscopy monitoring (NIRS) versus no monitoring in children and adults: a protocol for a systematic review with meta-analysis and Trial Sequential Analysis. *Syst Rev* 2020. Available from: https://www.researchsquare.com/article/rs-93110/v1
81. Hyttel-Sorensen S, Greisen G, Als-Nielsen B, Gluud C. Cerebral near-infrared spectroscopy monitoring for prevention of brain injury in very preterm infants. *Cochrane Database Syst Rev* 2017:CD011506.
82. Pichler G, Urlesberger B, Baik N, Schwaberger B, Binder-Heschl C, Avian A, et al. Cerebral Oxygen Saturation to Guide Oxygen Delivery in Preterm Neonates for the Immediate Transition after Birth: A 2-Center Randomized Controlled Pilot Feasibility Trial. *J Pediatr* 2016:73–8.
83. Pichler G, Baumgartner S, Biermayr M, Dempsey E, Fuchs H, Goos TG, et al. Cerebral regional tissue Oxygen Saturation to Guide Oxygen Delivery in preterm neonates during immediate transition after birth (COSGOD III): An investigator-initiated, randomized, multi-center, multi-national, clinical trial on additional cerebral tissue. *Trials* 2019; 20:1–10.
84. Pellicer A, Greisen G, Benders M, Claris O, Dempsey E, Fumagally M, et al. The SafeBoosC Phase II Randomised Clinical Trial: A Treatment Guideline for Targeted Near-Infrared-Derived Cerebral Tissue Oxygenation versus Standard Treatment in Extremely Preterm Infants. *Neonatology* 2013; 104:171–8.
85. Hyttel-Sorensen S, Austin T, van Bel F, Benders M, Claris O, Dempsey E, et al. A phase II randomized clinical trial on cerebral near-infrared spectroscopy plus a treatment guideline versus treatment as usual for extremely preterm infants during the first three days of life (SafeBoosC): study protocol for a randomized controlled tria. *Trials* 2013; 14:120.
86. Riera J, Hyttel-Sorensen S, Bravo MC, Cabañas F, López-Ortego P, Sanchez L, et al. The SafeBoosC phase II clinical trial: an analysis of the interventions related with the oximeter readings. *Arch Dis Child - Fetal Neonatal Ed* 2016; 101:F333–8.

87. Akobeng AK. Understanding type I and type II errors, statistical power and sample size. *Acta Paediatr* 2016; 105:605–9.
88. Darlow BA, Morley CJ. Oxygen Saturation Targeting and Bronchopulmonary Dysplasia. *Clin Perinatol* 2015; 42:807–23.
89. Ford I, Norrie J. Pragmatic Trials. *N Engl J Med* 2016; 375:454–63.
90. Hansen ML, Pellicer A, Gluud C, Dempsey E, Mintzer J, Hyttel-Sørensen S, et al. Cerebral near-infrared spectroscopy monitoring versus treatment as usual for extremely preterm infants: A protocol for the SafeBoosC randomised clinical phase III trial. *Trials* 2019; 20:1–11.
91. Chan A-W, Tetzlaff JM, Gotzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013; 346:e7586.
92. Gale C, Hyde MJ, Modi N. Research ethics committee decision-making in relation to an efficient neonatal trial. *Arch Dis Child - Fetal Neonatal Ed* 2017; 102:F291–8.
93. Woolfall K, Frith L, Gamble C, Gilbert R, Mok Q, Young B. How parents and practitioners experience research without prior consent (deferred consent) for emergency research involving children with life threatening conditions: a mixed method study. *BMJ Open* 2015; 5:e008522.
94. Hansen ML, Pellicer A, Gluud C, Dempsey E, Mintzer J, Hyttel-Sorensen S, et al. Detailed statistical analysis plan for the SafeBoosC III trial: a multinational randomised clinical trial assessing treatment guided by cerebral oxygenation monitoring versus treatment as usual in extremely preterm infants. *Trials* 2019; 20:746.
95. EMA/CHMP/ICH/135/1995. Guideline for good clinical practice E6 (R2). European Medicines Agency. 2016.
96. Holsti A, Serenius F, Farooqi A. Impact of major neonatal morbidities on adolescents born at 23-25 weeks of gestation. *Acta Paediatr* 2018; 107:1893–901.
97. Zuidgeest MGP, Goetz I, Groenwold RHH, Irving E, van Thiel GJMW, Grobbee DE. Series: Pragmatic trials and real world evidence: Paper 1. Introduction. *J Clin Epidemiol* 2017; 88:7–13.
98. Savović J, Turner RM, Mawdsley D, Jones HE, Beynon R, Higgins JPT, et al. Association between Risk-of-Bias Assessments and Results of Randomized Trials in Cochrane Reviews: The ROBES Meta-Epidemiologic Study. *Am J Epidemiol* 2018; 187:1113–22.
99. Hróbjartsson A, Emanuelsson F, Skou Thomsen AS, Hilden J, Brorson S. Bias due to lack

of patient blinding in clinical trials. A systematic review of trials randomizing patients to blind and nonblind sub-studies. *Int J Epidemiol* 2014; 43:1272–83.

100. Hróbjartsson A, Thomsen ASS, Emanuelsson F, Tendal B, Rasmussen JV, Hilden J, et al. Observer bias in randomized clinical trials with time-to-event outcomes: Systematic review of trials with both blinded and non-blinded outcome assessors. *Int J Epidemiol* 2014; 43:937–48.
101. Hrobjartsson A, Thomsen ASS, Emanuelsson F, Tendal B, Hilden J, Boutron I, et al. Observer bias in randomized clinical trials with measurement scale outcomes: a systematic review of trials with both blinded and nonblinded assessors. *Can Med Assoc J* 2013; 185:E201–11.
102. Patel RM, Kandefer S, Walsh MC, Bell EF, Carlo WA, Laptook AR, et al. Causes and Timing of Death in Extremely Premature Infants from 2000 through 2011. *N Engl J Med* 2015; 372:331–40.
103. Anthon CT, Granholm A, Perner A, Laake JH, Møller MH. No firm evidence that lack of blinding affects estimates of mortality in randomized clinical trials of intensive care interventions: a systematic review and meta-analysis. *J Clin Epidemiol* 2018; 100:71–81.
104. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ* 2015; 350:h2147.
105. Fumagalli M, Parodi A, Ramenghi L, Limperopoulos C, Steggerda S. Ultrasound of acquired posterior fossa abnormalities in the newborn. *Pediatr Res* 2020; 87:25–36.
106. Rodie M, Brunton A, Heuchan AM, Watt A. https://www.clinicalguidelines.scot.nhs.uk/ggc-paediatric-guidelines/ggc-guidelines/neonatology/cranial-ultrasound-a-guideline-for-the-performance-of-routine-cranial-uss-for-preterm-infants/. Cranial ultrasound: a guideline for the performance of routine cranial USS for preterm infants. 2018.
107. Hintz SR, Slovis T, Bulas D, Van Meurs KP, Perritt R, Stevenson DK, et al. Interobserver Reliability and Accuracy of Cranial Ultrasound Scanning Interpretation in Premature Infants. *J Pediatr* 2007; 150:592-596.e5965.
108. Hansen ML, Rasmussen MI, Rubin S, Pellicer A, Cheng G, Xu X, et al. Pilot test of an online training module on near-infrared spectroscopy monitoring for the randomised clinical trial SafeBoosC-III. *Trials* 2020; 21:356.
109. Schmidt B, Roberts RS, Davis PG, Doyle LW, Asztalos E V., Opie G, et al. Prediction of Late Death or Disability at Age 5 Years Using a Count of 3 Neonatal Morbidities in Very

Low Birth Weight Infants. *J Pediatr* 2015; 167:982–986.e2.

110. Davidson L, Berkelhamer S. Bronchopulmonary Dysplasia: Chronic Lung Disease of Infancy and Long-Term Pulmonary Outcomes. *J Clin Med* 2017; 6:4.
111. Jakobsen JC, Gluud C, Winkel P, Lange T, Wetterslev J. The thresholds for statistical and clinical significance – a five-step procedure for evaluation of intervention effects in randomised clinical trials. *BMC Med Res Methodol* 2014; 14:34.
112. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. *Nature* 2019; 567:305–7.
113. Goodman SN. Introduction to Bayesian methods I: measuring the strength of evidence. *Clin Trials J Soc Clin Trials* 2005; 2:282–90.
114. Demets DL, Lan KKG. Interim analysis: The alpha spending function approach. *Stat Med* 1994; 13:1341–52.
115. Guyatt GH, Briel M, Glasziou P, Bassler D, Montori VM. Problems of stopping trials early. *BMJ* 2012; 344:e3863.
116. Zhang J, Quan H, Ng J, Stepanavage ME. Some statistical methods for multiple endpoints in clinical trials. *Control Clin Trials* 1997; 18:204–21.
117. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Med Res Methodol* 2017; 17:162.
118. Groenwold RHH, Moons KGM, Vandenbroucke JP. Randomized trials with missing outcome data: how to analyze and what to report. *Can Med Assoc J* 2014; 186:1153–7.
119. Kahan BC, Morris TP. Reporting and analysis of trials using stratified randomisation in leading medical journals: review and reanalysis. *BMJ* 2012; 345:e5840.
120. Kahan BC, Morris TP. Improper analysis of trials randomised using stratified blocks or minimisation. *Stat Med* 2012; 31:328–40.
121. Kahan BC. Accounting for centre-effects in multicentre trials with a binary outcome – when, why, and how? *BMC Med Res Methodol* 2014; 14:20.
122. Deeks JJ, Higgins JP, Altman DG. Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al., eds. Cochrane Handbook for Systematic Reviews of Interventions 6.1. Cochrane; 2020. Available from: www.training.cochrane.org/handbook.
123. Gates S, Brocklehurst P. How should randomised trials including multiple pregnancies be analysed? *BJOG An Int J Obstet Gynaecol* 2004; 111:213–9.

124. Bernardo J, Nowacki A, Martin R, Fanaroff JM, Hibbs AM. Multiples and parents of multiples prefer same arm randomization of siblings in neonatal trials. *J Perinatol* 2015; 35:208–13.
125. Yelland LN, Sullivan TR, Collins CT, Price DJ, McPhee AJ, Lee KJ. Accounting for twin births in sample size calculations for randomised trials. *Paediatr Perinat Epidemiol* 2018; 32:380–7.
126. Shaffer ML, Hiriote S. Analysis of time-to-event and duration outcomes in neonatal clinical trials with twin births. *Contemp Clin Trials* 2009; 30:150–4.
127. Li G, Taljaard M, Van den Heuvel ER, Levine MA, Cook DJ, Wells GA, et al. An introduction to multiplicity issues in clinical trials: the what, why, when and how. *Int J Epidemiol* 2017; 46:746–55.
128. Abdi H. The Bonferonni and Šidák Corrections for Multiple Comparisons. In: Salkind N, ed. Encyclopedia of Measurement and Statistics Thousand Oaks (CA): Sage; 2007. p. 1–9.
129. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006; 3:77–101.
130. Hansen ML, Ostojic D, Kleiser S, Greisen G, Wolf M. Not removing the glossy white cover from adhesive INVOS neonatal sensors affects the oxygenation Measurement (manuscript accepted for publication). *Adv Exp Med Biol* 2019.
131. Ostojic D, Kleiser S, Nasseri N, Isler H, Andresen B, Wabnitz H, et al. In Vitro Comparisons of Near-Infrared Spectroscopy Oximeters: Impact of Slow Changes in Scattering of Liquid Phantoms. *Adv Exp Med Biol* 2018; 1072:375–9.
132. Garattini S, Jakobsen JC, Wetterslev J, Bertelé V, Banzi R, Rath A, et al. Evidence-based clinical practice: Overview of threats to the validity of evidence and how to minimise them. *Eur J Intern Med* 2016; 32:13–21.
133. Tunis SR, Stryer DB, Clancy CM. Practical Clinical Trials. *JAMA* 2003; 290:1624–32.
134. Askie L, Offringa M. Systematic reviews and meta-analysis. *Semin Fetal Neonatal Med* 2015; 20:403–9.
135. Thorpe KE, Zwarenstein M, Oxman AD, Treweek S, Furberg CD, Altman DG, et al. A pragmatic–explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *J Clin Epidemiol* 2009; 62:464–75.
136. Patsopoulos NA. A pragmatic view on pragmatic trials. *Dialogues Clin Neurosci* 2011; 13:217–24.
137. Hale T, Webster S, Petherick A, Phillips T, Kira B. Oxford COVID-19 Government

Response Tracker, Blatnavik School of Government. Data use policy: Creative Commons Attribution CC BY standard. 2020.

138. Rasmussen MI, Lühr Hansen M, Pichler G, Dempsey E, Pellicer A, EL-Khuffash A, et al. Extremely preterm infant admissions within the SafeBoosC-III consortium during the COVID-19 lockdown. *medRxiv* 2020:2020.10.02.20204578. Available from: http://medrxiv.org/content/early/2020/10/05/2020.10.02.20204578.abstract
139. Doyle LW, Cheong JL, Ehrenkranz RA, Halliday HL. Early (< 8 days) systemic postnatal corticosteroids for prevention of bronchopulmonary dysplasia in preterm infants. *Cochrane Database Syst Rev* 2017; 10:CD001146.
140. Plomgaard AM, Alderliesten T, van Bel F, Benders M, Claris O, Cordeiro M, et al. No neurodevelopmental benefit of cerebral oximetry in the first randomised trial (SafeBoosC II) in preterm infants during the first days of life. *Acta Paediatr* 2019; 108:275–81.

Paper and manuscript appendix

Trials

STUDY PROTOCOL Open Access

Cerebral near-infrared spectroscopy monitoring versus treatment as usual for extremely preterm infants: a protocol for the SafeBoosC randomised clinical phase III trial

Mathias Lühr Hansen[1*], Adelina Pellicer[2], Christian Gluud[3], Eugene Dempsey[4], Jonathan Mintzer[5], Simon Hyttel-Sørensen[6], Anne Marie Heuchan[7], Cornelia Hagmann[8], Ebru Ergenekon[9], Gabriel Dimitriou[10], Gerhard Pichler[11], Gunnar Naulaers[12], Guoqiang Cheng[13], Hercilia Guimarães[14], Jakub Tkaczyk[15], Karen B. Kreutzer[16], Monica Fumagalli[17,18], Olivier Claris[19], Petra Lemmers[20], Siv Fredly[21], Tomasz Szczapa[22], Topun Austin[23], Janus Christian Jakobsen[3,24,25] and Gorm Greisen[1]

Abstract

Background: Cerebral oxygenation monitoring may reduce the risk of death and neurologic complications in extremely preterm infants, but no such effects have yet been demonstrated in preterm infants in sufficiently powered randomised clinical trials. The objective of the SafeBoosC III trial is to investigate the benefits and harms of treatment based on near-infrared spectroscopy (NIRS) monitoring compared with treatment as usual for extremely preterm infants.

Methods/design: SafeBoosC III is an investigator-initiated, multinational, randomised, pragmatic phase III clinical trial. Inclusion criteria will be infants born below 28 weeks postmenstrual age and parental informed consent (unless the site is using 'opt-out' or deferred consent). Exclusion criteria will be no parental informed consent (or if 'opt-out' is used, lack of a record that clinical staff have explained the trial and the 'opt-out' consent process to parents and/or a record of the parents' decision to opt-out in the infant's clinical file); decision not to provide full life support; and no possibility to initiate cerebral NIRS oximetry within 6 h after birth. Participants will be randomised 1:1 into either the experimental or control group. Participants in the experimental group will be monitored during the first 72 h of life with a cerebral NIRS oximeter. Cerebral hypoxia will be treated according to an evidence-based treatment guideline. Participants in the control group will not undergo cerebral oxygenation monitoring and will receive treatment as usual. Each participant will be followed up at 36 weeks postmenstrual age. The primary outcome will be a composite of either death or severe brain injury detected on any of the serial cranial ultrasound scans that are routinely performed in these infants up to 36 weeks postmenstrual age. Severe brain injury will be assessed by a person blinded to group allocation. To detect a 22% relative risk difference between the experimental and control group, we intend to randomise a cohort of 1600 infants.

(Continued on next page)

* Correspondence: mathias.luhr.hansen@gmail.com
[1]Department of Neonatology, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark
Full list of author information is available at the end of the article

(Continued from previous page)
Discussion: Treatment guided by cerebral NIRS oximetry has the potential to decrease the risk of death or survival with severe brain injury in preterm infants. There is an urgent need to assess the clinical effects of NIRS monitoring among preterm neonates.

Trial registration: ClinicalTrial.gov, NCT03770741. Registered 10 December 2018.

Keywords: Randomised clinical trial, Preterm, Near infrared spectroscopy, Protocol

Background

Every year, approximately 50,000 extremely preterm infants (< 28 weeks postmenstrual age) are born in countries where they routinely will be offered neonatal intensive care [1]. Extremely preterm birth carries a high risk of death or long-term cerebral impairment. With a current mortality of about 25% and a prevalence of psychomotor impairment in approximately 20% of survivors, more than 10,000 will die each year and a further 10,000 will suffer from cerebral palsy or moderate-to-severe cognitive impairment [2–4].

When an infant is born extremely preterm, all organs are immature and vulnerable [5, 6]. This is particularly relevant for the immature brain [7]. Cerebral autoregulation is limited and believed to be fragile in extremely preterm infants [4]. It is hypothesised that large fluctuations in cerebral blood flow may result in cerebral haemorrhage arising from immature blood vessels. These fluctuations in systemic blood flow are common during the transition from foetal to neonatal circulation during the first days of life, thus putting the immature brain in danger [8].

Neonatal brain injury may be diagnosed by cranial ultrasound [9]. The most severe injuries, including grade III or IV intraventricular haemorrhage and the non-haemorrhagic white matter injury cystic periventricular leukomalacia, entail a high probability of death or cerebral palsy [10, 11]. Several pre- and postnatal factors have been shown or are thought to be associated with cerebral injury, including ascending infections [12], insufficient nutrition early in life [13], insufficient blood pressure, cardiac dysfunction, and suboptimal mechanical ventilation [14–16].

Among extremely preterm infants during their first days of life, current practice standards involve multiple parallel interventions, including respiratory and haemodynamic support, intravenous fluids, antibiotics, nutrition, and monitoring of physiological parameters. Despite significant advances in the management of extremely preterm infants over the past three decades, many of these interventions are used with little evidence. Furthermore, an end-organ monitor with sufficient time resolution to guide evidence-based treatment is lacking. Near-infrared spectroscopy (NIRS) has the potential to function in this manner. Cerebral NIRS provides a real-time continuous estimate of the cerebral tissue oxygenation ($rStO_2$), expressed as a percentage. The normal ranges of $rStO_2$ in preterm infants have been determined and change somewhat with gestational age and postnatal age [17].

The evidence on the utility of NIRS monitoring in extremely preterm infants during the first days of life is sparse. Only one previous randomised clinical trial has assessed the effects of cerebral monitoring—the SafeBoosC phase II feasibility trial [18]. This trial showed that NIRS monitoring reduced the burden of cerebral hypoxia to less than half compared with treatment as usual and there were also non-significant trends towards reduced incidence of severe brain injury and reduced mortality in the NIRS group [18]. The clinical interventions used in the NIRS-open group included a significant number with likely beneficial effects on blood oxygen content and transport, blood pressure, cardiac output, and cerebral blood flow [19]. Despite these promising results, it is theoretically possible that NIRS monitoring may cause harm. This includes skin marks from the sensors, inappropriate modifications in cardio-respiratory support based on hypoxic values, and unnecessary infant disturbance due to manipulation of the forehead-based NIRS sensor. Furthermore, the SafeBoosc II trial showed a higher prevalence of bronchopulmonary dysplasia and retinopathy of prematurity in the experimental group. As NIRS devices and sensors are also costly and monitoring confers additional nursing tasks, it would be unfortunate to incorporate NIRS monitoring into standard practice without clear evidence of clinical benefit.

To evaluate the potential benefits and harms of NIRS monitoring, large-scale randomised clinical trials are urgently warranted. Since the intervention is complex—NIRS monitoring itself in addition to evidence-based modification of cardio-vascular support—a pragmatic design is preferable to ensure relevance for routine neonatal intensive care. International participation is additionally necessary to achieve adequate subject numbers and ideally promote generalisability of the results.

Methods/design

This trial will be conducted in compliance with the guidelines of The Declaration of Helsinki in its latest form, the International Conference on Harmonization

Good Clinical Practice guidelines [20], and applicable national regulations and directives. No clinical site will start randomisation before their eligibility has been confirmed and the protocol has been approved by the relevant ethics committee. Any amendments to the protocol will need approval by the Steering Committee and ethical review before being implemented. Written informed consent will be obtained by a qualified physician or nurse connected to the trial, prior to randomisation of any participant, unless the Neonatal Intensive Care Unit (NICU) uses deferred informed consent or prior assent as consent methods (see below). These consent procedures will be approved by local ethics committees or institutional review boards.

Objective

The objective of this trial is to examine the benefits and harms of treatment based on NIRS monitoring compared with treatment as usual (standard monitoring and treatment) to reduce cerebral hypoxia during the first 72 h of life in extremely preterm infants. The hypothesis is that the application of treatment based on NIRS monitoring will decrease a composite outcome of severe brain injury or death at 36 weeks postmenstrual age.

Roles and responsibilities for committees

SafeBoosC III is led by a Steering Committee comprising the coordinating investigator (GG), the national coordinators, and two representatives from the Copenhagen Trial Unit (CG and JCJ). Decisions will be made by a simple majority. The executive committee will be responsible for the day-to-day management and will comprise the coordinating investigator, the trial manager (MLH), co-investigators (AP, GD, JM, SHS), and the two representatives from the Copenhagen Trial Unit (CG and JCJ).

There will be one principal investigator in each department who will be responsible for obtaining ethical approval, organising local Good Clinical Practice monitoring, informing clinical staff members on the web-based training and certification program, recruitment of patients, and data entry into the patient report forms. The Copenhagen Trial Unit will be responsible for randomisation, development of the patient report forms, and central monitoring.

Trial design

This is an investigator-initiated, multinational, randomised, pragmatic phase III clinical trial with a two-parallel group design that will enrol 1600 extremely preterm infants from 20 countries (Austria, Belgium, China, Czech Republic, Denmark, England, France, Germany, Greece, India, Ireland, Italy, Norway, Poland, Portugal, Switzerland, Scotland, Spain, Turkey, USA). A list of all study sites will be available at www.safeboosc.eu. It is an open label trial, but parts will be conducted blinded to the intervention (see the 'Blinding' section).

The trial has been designed according to the SPIRIT guidelines (Fig. 1 and Additional file 1) [21].

Inclusion criteria

The inclusion criteria will be infants born before 28 weeks postmenstrual age and signed parental informed consent unless the NICU has chosen to use 'opt-out' or deferred consent as their consent method.

Exclusion criteria

The exclusion criteria will be no signed parental informed consent (or if the 'opt-out' method is used, lack of a record that the clinical staff have explained the trial and the 'opt-out' consent process to parents and/or a record of the parents' decision to opt-out in the infant's clinical file); decision not to provide full life support; and no possibility to initiate cerebral NIRS monitoring within 6 hours after birth.

Participation in other trials

Participants included in the SafeBoosC III trial can participate in any other study or intervention on the condition that: it does not allow clinical staff access to cerebral oximetry in the control group from inclusion in SafeBoosC III to the end of the intervention period 72 h after birth; and does not exclude a treatment that would be clearly indicated by the SafeBoosC III evidence-based treatment guideline during the intervention period. All partners are encouraged to design ancillary studies and draw on data collected by SafeBoosC III, if not compromising the blinding of assessors or the equipoise of the trial. Ancillary studies must seek approval by the SafeBoosC Steering Committee.

Participant discontinuation and withdrawal

A participant's parents are free to withdraw them from the SafeBoosC III trial at any time, and this will not have any consequences for the infant's further treatment. Reasons for discontinuation, if provided by the parents, will be documented. When possible, the parents will be asked if they will allow their child's data to be used in the analysis.

The attending clinician can withdraw the participant from the trial at any time in case there are safety concerns. Reasons for withdrawal will be documented. There are no pre-specified criteria for discontinuation of participants from the trial. Discontinuation of participants from the trial will not result in replacement with new participants.

Visit description	Consent and randomisation	Intervention period: first 72 hours of life (Cerebral monitoring and treatment according to guideline versus 'treatment as usual')	Follow-up: two possible time points up to week 36+0 postmenstrual age	
Visit code	V0	V1	V2	
Time period	0-6 hours	0-72 hours	Discharge to home	36±1 weeks postmenstrual age
Assessing inclusion and exclusion criteria	X			
Informed consent (can be obtained before birth) *	X			
Allocation to experimental or control group	X			
SARs		X		
SAEs		X	X	
Explanatory variables		X	X	
Exploratory outcomes (NEC, ROP, BPD, Sepsis)			X	X
Severe brain injury (IVH grade III/IV, cPVL, post-haemorrhagic ventricular dilatation, cerebellar haemorrhage, cerebral atrophy)			X	X
All-cause mortality			X	X

Fig. 1 Schedule for enrolment, intervention and assessment, based on the SPIRIT 2013 guidance for protocols of clinical trials. *If approved by the local ethics committee, deferred informed consent or prior informed assent may be sought. Time to ask parents for deferred consent will be decided individually by clinical staff members

Recruitment

In this phase III trial, we have prolonged the enrolment period from 3 hours, as used in SafeBoosC II, to 6 hours after birth, although we recommend that monitoring is started as early as possible to help decision-making when cardio-respiratory support is established. This 6-hour window is similar to what is currently used for another neonatal intervention—therapeutic hypothermia for hypoxic-ischaemic encephalopathy after birth asphyxia [22]. We believe this will make the trial relevant in settings where antenatal transfer to a perinatal centre is used less often, and thereby increase recruitment feasibility without compromising the effect of NIRS monitoring.

Extremely preterm infants are expected to be included at about 50 NICUs in about 20 countries. The 93 units that took part in a previous funding application for the SafeBoosC III trial had rates of admission of between 15 and 90 extremely preterm infants per year. The total admissions were estimated to be 3000 infants per year. We should, therefore, have a good chance of recruiting 1600 participants within 2 years. Sites that expect to enrol at least 15 participants per year within the 2-year recruitment period will take part. Inclusion of new NICUs after the common start date will be done ad hoc, considering expected contributions and time remaining.

Randomisation

Infants will be centrally randomised to either the experimental or control group with a 1:1 allocation ratio at the Copenhagen Trial Unit using a web-based randomisation application. The allocation sequence will be computer-generated with varying block sizes concealed for all investigators, as the web-based program will not release the randomisation until the patient has been included in the trial and stratified by NICU and gestational age group (lower gestational age (< 26 weeks) compared to higher gestational age (≥ 26 weeks)). Twin couples will be randomised to the same group, either intervention or control. In centres where only one or two NIRS devices are available, it may not be possible to include all infants from twin births. Thus, only one of a pair of twins may be included. The sibling enrolled will be the one born last.

Blinding

Due to the nature of the experimental intervention, it is not possible to blind the clinical staff, the infant, or the parents to study group allocation. Outcome assessment of mortality will not be blinded but the mortality data will be checked by Good Clinical Practice via source data verification in all patients. The diagnosis and classification of brain injury along with the entry of these data

into the patient report form will be conducted by an assessor blinded to study group allocation. Data entry procedures will depend on local factors and will be agreed on between the principal investigator at each NICU and the coordinating investigator. The data managers, statisticians, and those drawing conclusions will be blinded to study group allocation. Details on this is described in a report on the statistical analysis plan [23].

Intervention

Experimental group participants will undergo cerebral NIRS monitoring applied as soon as possible after arrival in the NICU and always within 6 hours after delivery and receive treatment based on NIRS monitoring during the first 72 h of life. Treatment will be based on the same evidence-based guideline as used in the SafeBoosC II trial (see below) [24].

The control group participants will not receive any cerebral NIRS monitoring and will be monitored and treated according to local guidelines and clinical practices.

Treatment guideline based on NIRS monitoring

An evidence-based treatment guideline recommending modification of cardio-respiratory support or interventions aiming at increasing blood oxygen transport capacity will be followed in order to maintain cerebral oxygenation above 55% (Additional file 2) [24]. As the SafeBoosC II trial showed a low burden of hyperoxia unaffected by monitoring-based interventions, the SafeBoosC III trial will not target cerebral hyperoxia and therefore the interventions for hyperoxia have been removed from this trial's treatment guidelines. The same SafeBoosC III treatment guideline will be used in all participating centres.

Devices

All commercially available cerebral oximeters that are approved for clinical use in newborns may be used. The aim is to use several different devices to generate results of generic value. There are now seven commercially available devices that are approved for clinical use in different countries: INVOS (Medtronics, Minneapolis, MN, USA); NIRO (Hamamatsu, Hamamatsu City, Japan); Fore-Sight (CAS Medical, Branford, CT, USA); Sensmart (Nonin Medical, Plymouth, MN, USA); O3 (Masimo, Irvine, CA, USA); Egos (Enginmed, Suzhou, China); and Oxyprem 1.4 (Oxyprem, Zürich, Switzerland). The normal range of $rStO_2$ was determined with the INVOS adult sensor [17] and defined the $rStO_2$ thresholds for intervention used in the SafeBoosC II trial. Each eligible device in SafeBoosC III will be compared with the INVOS adult sensor using a blood lipid phantom and device-specific thresholds will be determined [25] before being used in the SafeBoosC III trial.

Training and certification

Clinical staff will be offered a web-based training and certification program consisting of short modules covering the trial rationale, NIRS and monitoring of cerebral oxygenation, the treatment guideline, cerebral ultrasound and classification of brain injury, and Good Clinical Practice (www.safeboosc.eu). The use of these modules and the completion rate will be monitored and reported with the results of the trial. Sites with low compliance may be selected for subgroup analyses.

Trial duration

NIRS monitoring will start within 6 postnatal hours and the intervention will last until 72 h of life. Each participant will be followed up at 36 weeks postmenstrual age.

Explanatory variables

To allow comparisons between intervention groups, additional baseline clinical data will be obtained, including birth weight, gestational age, mechanical ventilation, and use of cardiovascular support. Data will be drawn from clinical records at 72 h of age and 36 weeks postmenstrual age, the same time as the primary and exploratory outcomes are assessed and documented. The majority of these selected variables are usually reported to neonatal network databases such as the Vermont Oxford Network [26].

Outcomes

Primary and exploratory outcomes will be assessed at 36 weeks postmenstrual age as documented in the infants' clinical files. If an infant has been discharged to a step-down unit, data will be sought from that unit, and if this is not possible, data will be used until the date of discharge to the step-down unit. In case the last entry in an infant's clinical file is prior to 36 + 0 weeks postmenstrual age, for example due to discharge home, the date of discharge will be reported in the online patient report form.

The primary outcome is a composite of either death or severe brain injury detected on any one of a series of cranial ultrasound scans that are routinely performed in extremely premature infants. Severe brain injury is defined as grade III or IV intraventricular haemorrhage (IVH), cystic periventricular leukomalacia (cPVL), cerebellar haemorrhage, post-haemorrhagic ventricular dilatation, or cerebral atrophy. The exploratory outcomes will be bronchopulmonary dysplasia (BPD), retinopathy of prematurity (ROP) stage 3+, necrotising enterocolitis (NEC) stage 2 or higher using the modified Bell's staging system and/or focal intestinal perforation, late-onset sepsis (> 72 h after birth) defined as being treated with antibiotics for a minimum of 5 days, and a count of the presence of three major neonatal morbidities (BPD,

ROP, and severe brain injury). All diagnoses, except severe brain injury, are made as per routine in each NICU.

Statistical plan and data analysis

Full details regarding statistical considerations and data analysis are outlined in a separate report [23], which will be published before the analysis phase begins, without knowledge of any data collected.

Sample size

We have calculated our sample size based on the composite primary outcome, with an alpha of 5%, a power of 90%, and a ratio of experimental trial participants to control trial participants of 1:1.

In the 2009 EuroNeoNet report, the mortality among extremely preterm infants was 33% and severe intracranial haemorrhage was observed in 15%. In the SafeBoosC II trial, the proportion of participants with the composite primary outcome was approximately 34% in the control group and 26% in the experimental group [27].

Based on the above, a total of 1600 infants would be required to demonstrate a similar relative risk reduction of 22%, with an alpha of 5%, and a power of 90%.

In SafeBoosC II, the intra-class correlation coefficient (ICC) of the burden of hypoxia within pairs of twins was negligible. The ICC for death before discharge and for intraventricular haemorrhage grade 3 or 4 have previously been estimated to 0.00 (95% confidence interval (CI) − 0.04 to 0.02) and − 0.01 (95% CI − 0.05 to 0.01) [28]. These values correlate to a design effect very close to 1 [28]. Based on this, we have not included twin ICC in the sample size estimation.

Analysis of the primary outcome

The primary outcome analysis will be made on the intention-to-treat population, and we will use mixed-effect logistic regression. 'Site' will be included as a random effect (intercept) and the remaining stratification variables, age and intervention groups, will be included as fixed effects. In addition, we will perform a range of pre-defined sensitivity analyses to inform the interpretation of the results of the primary analysis [23].

Safety

Predefined serious adverse reactions (SAR) will be reported at 72 h after birth and serious adverse events (SAE) will be reported at 36 weeks postmenstrual age. Expedited reporting will not be used. An independent data monitoring and safety committee is established to monitor mortality, neonatal morbidity, and SARs with 'certain' or 'probably/likely' relationships with the cerebral NIRS oximeter and/or the application of the evidence-based treatment guideline or any of its interventions. They include two neonatologists and a biostatistician. The charter for the data monitoring and safety committee has been written prior to the enrolment of trial participants. The trial will not be stopped early because of futility, and Lan-DeMets sequential monitoring boundaries will be used at each interim analysis to assess if thresholds for statistical significance of benefits or harms have been crossed [29]. Only one interim analysis is planned, after one-third of trial participants have been randomised. Additional analyses will be decided by the data monitoring and safety committee members [23]. Based on primarily safety considerations, the data monitoring and safety committee will make recommendations to the steering group to continue, change, hold, or terminate the trial. The recommendations will be guided by the statistical monitoring guidelines, which is defined in the data monitoring and safety committee charter (available from www.safeboosc.eu).

The preterm population is at high risk for SAEs and most adverse events may be of a serious nature with or without relevance to the SafeBoosC III trial intervention. Both groups of the trial are expected to have a high proportion of SAEs. It is therefore neither feasible nor meaningful to record and report all adverse events. Therefore, we have decided only to record and report predefined SAEs and SARs. The SAEs include any event of death, severe brain injury, necrotising enterocolitis, bronchopulmonary dysplasia, retinopathy of prematurity, or sepsis as defined under primary and exploratory outcomes. These predefined SAEs have been chosen since they cover the major neonatal morbidities seen in this study population. The SARs are defined as any adverse reaction related to the trial intervention that results in death, is life-threatening, requires prolongation of existing hospitalisation, results in persistent or significant disability or incapacity, or requires intervention to prevent permanent impairment or damage. This includes physical mishaps associated with managing the oximeter and sensors, such as severe skin damage, critical displacement of endotracheal tubes or endovascular lines, and clinical mismanagement based on cerebral oximetry monitoring data, such as interventions aiming at improving cardiovascular status, respiratory status, and/or oxygen transport.

Data management

All participants' data are protected in accordance with the Danish Act on the processing of personal data and the Danish Health Act. The Copenhagen Trial Unit will provide central, web-based data entry through an online patient report form, in the open-source clinical trial software OpenClinica®. This will handle the inclusion procedure, the documentation of the stratification and randomisation process, the SARs, and the relevant clinical data from enrolled subjects, including primary and

exploratory outcomes and explanatory variables. The data will be entered into the online patient report form directly by the medical staff. Forms for randomisation/inclusion, end-of-monitoring at 72 h of age, and the 36-week follow-up will be created. Data will be stored in accordance with guidelines issued by the Danish Data Protection Agency, from whom approval of the trial will be sought. Only NICU numbers and study numbers will be used to identify participants (i.e. the data kept at Copenhagen Trial Unit is pseudo-anonymised), while lists of study numbers and personal identifying information (e.g. to allow Good Clinical Practice, data cleansing, and later follow-up) will be kept at the NICUs. Six months after the acceptance of the publication that presents the primary outcome, the dataset will be transferred to the Danish data archive. Before transfer, subject study numbers will be removed, NICU numbers will be replaced, sex documentation removed, and birth weight and gestational age recoded into binary variables to minimise the risk of re-identification. Use by other researchers will depend on the permission of the steering group.

The investigators permit trial-related monitoring, audits, and regulatory inspections by providing direct access to the source data and other relevant documents. Trial data will be handled according to regulations of data protection agencies in the respective countries.

Monitoring

Internal monitoring will be conducted by the Copenhagen Trial Unit, who will monitor patient recruitment and quality, completeness, and timeliness of data entry. In case of problems, the principal investigator will be contacted.

External monitoring will be conducted by a Good Clinical Practice person assigned by the principal investigator at each site. The Good Clinical Practice person will perform monitoring according to the monitoring plan, which will is available at www.safeboosc.eu.

Ethical considerations

To obtain evidence-based knowledge on the potential benefit and harms of NIRS-based cerebral monitoring in the clinical management of premature infants, large-scale randomised clinical trials are required. The SafeBoosC II trial served as a feasibility trial for the present large-scale SafeBoosC III trial.

In most NICUs, there is still clinical equipoise regarding the use of NIRS monitoring, meaning there is genuine uncertainty over whether cerebral oximetry monitoring and subsequent monitoring-based treatments are clinically beneficial or harmful. Nevertheless, some NICUs have started to use cerebral oxygenation monitoring as part of routine clinical management. Thus, there might be a limited time-window for this trial, since it may be more difficult to test an intervention that is already in clinical use [30]. Therefore, we aim at a pragmatic trial, rather than doing a proof-of-concept trial first.

Extremely preterm infants demonstrate stress reactions during routine manipulation. Positioning and repositioning of cerebral NIRS sensors can result in such reactions. There are, however, no data to support substantially more risk or discomfort compared with no intervention or compared with current routine care. All interventions proposed in the evidence-based treatment guideline are commonly used in this patient group [21].

'Treatment as usual', defined as treatment according to participating hospital's standard procedures, will be provided to the control group. Also, this will be the care provided to any participant that withdraws consent, in addition to infants who are not included in the trial. Multiple births will be randomised together and undergo allocation to the same study group. This is to avoid parents ascribing differences in their infants' clinical courses and outcomes based on group allocation resulting from participation in this trial.

Publication plan

The trial protocol is registered at ClinicalTrials.gov (NCT03770741) and all versions are available at www.safeboosc.eu. Following trial completion, summary trial data will additionally be entered at www.clinicaltrials.gov. Further summary data of main outcomes will be entered after statistical analyses are conducted. Attempts will be made to publish all results, positive, neutral, as well as negative, in a peer-reviewed international journal. Authorship will be determined according to the International Committee of Medical Journal Editors. An additional requirement is one author per NICU completing at least 30 participants. Ancillary studies with results potentially affecting equipoise with regard to the value of NIRS shall not be published before the main publication of the SafeBoosC III trial. After the publication of trial results, depersonalised individual patient data will be uploaded at Zenodo.

Discussion

In this pragmatic trial, we plan to test the hypothesis that the application of treatment based on cerebral NIRS monitoring in extremely preterm infants will decrease a composite outcome of either death or survival with severe brain injury at 36 weeks postmenstrual age.

A Cochrane systematic review concluded that it is not possible, based on the currently available literature, to determine the specific benefits or harms of NIRS monitoring in extremely preterm infants [27]. The conclusion of this review was that NIRS monitoring should only be used in randomised clinical trials [31]. Despite this, NIRS is routinely used in extremely preterm infants

during the first days of life in numerous NICUs in multiple countries [32]. It is likely that this monitoring approach will become more common as evidence in other patient groups becomes more convincing [33]. Therefore, to prevent a non-evidence-based, large-scale clinical uptake of NIRS monitoring, a robust randomised clinical trial, such as the SafeBoosC III trial, is urgently required.

As described in the 'Blinding' section, it is not possible to blind the clinical staff, the infants, and the parents of infants participating in this trial. This circumstance introduces risks of bias. Several previous studies have shown that inadequate blinding of participants, personnel, and outcome assessors in randomised trials often results in overestimation of treatment effects for a given intervention for all outcome types, including mortality and subjective outcomes such as radiologic image interpretation [34–37]. A meta-epidemiologic study showed a high variability of treatment effect measured on unblinded subjective outcomes, indicating that for trials including subjective outcomes, the magnitude of bias due to lack of blinding is unpredictable [34]. But again, non-blinded trials compared to similar blinded trials showed overestimation of intervention effects [30]. This meta-epidemiologic study included randomised trials across all clinical fields. A meta-analysis, including 361 intensive-care randomised trials, evaluated the effect of adequate blinding on effect estimates of mortality and found no statistical significant difference between blinded and unblinded trials, suggesting that there may be little, if any, effect of adequate blinding on mortality effect estimates in intensive care trials [38]. No meta-epidemiologic studies, meta-analyses, or systematic reviews have evaluated the effect of adequate/inadequate blinding on intervention effects in neonatal randomised trials. In conclusion, previous results suggest there is a risk of biased results due to lack of blinding even on mortality results. The design of the SafeBoosC III trial strives to minimise the risks regarding the primary outcome.

The pragmatic methodology of this trial also has some limitations. Cranial ultrasound-based diagnoses will be performed locally rather than centrally as was done in SafeBoosC II [18]. This may potentially raise concerns in SafeBoosC III since discrepancies between local readers in different centres could be expected. However, when comparing local and central interpretations of cranial ultrasound images in preterm infants in previous clinical trials, the sensitivity and specificity for local interpretations of severe brain injury were quite robust [39]. Furthermore, we have developed a web-based training program for staff members caring for trial participants. Among other topics, this web program includes a cranial ultrasound module for the purpose of decreasing inter-observer variability and heightening data quality.

As in all trial populations of extremely preterm infants, a large number of participants will be twins, which can cause statistical concerns arising from intra-class correlation coefficients (ICC) [28]. We cannot with certainty estimate the ICC for the composite outcome of death or severe brain injury for the present trial. However, the ICC of the burden of hypoxia within pairs of twins in SafeBoosC II was negligible (ICC = 0.027) [27]. Additionally, the twin ICC for pre-discharge death and grade III or IV intraventricular haemorrhage has been estimated in a previous study to 0.00 and – 0.01, which correlates to a negligible design effect [28]. The details of how the twin issue will be statistically accounted for is outlined in the publication of the SafeBoosC III statistical and data analysis plan [23].

The interventions in this trial are complex and rely on a number of separate but interacting components, all relevant for the potential success of the intervention. When NIRS monitors show hypoxic values, neonatologists must evaluate the participant's clinical status by taking additional measures into consideration and deciding on a possible modification of cardio-respiratory support and interventions to increase blood oxygen transport capacity, based on the treatment guideline. This complexity will result in difficulty interpreting specific results, as it cannot be ascertained what exactly causes a potential effect at 36 weeks postmenstrual age. Furthermore, reproducing and generalising complex interventions may be difficult for future clinicians assessing the results of this trial [40]. However, since this is a pragmatic effectiveness trial evaluating outcomes related to NIRS-based cerebral oxygenation monitoring in routine practice and not the specific treatment choices per se, this concern will not affect the purpose of the trial. The Medical Research Council Framework has developed CONSORT guidelines in order to help trialists develop clearly defined and reproducible complex interventions [41, 42]. We believe that the methodology in the SafeBoosC III trial is in agreement with these guidelines, which is a major strength of this trial.

Obtaining prior informed consent from parents of critically ill neonates within the first hours following birth is difficult and may challenge important standards of information delivery, comprehension, competence, and voluntariness [43–45], and can also restrict the population studied with the effect of impairing the generalisability of results. Furthermore, since monitoring of cerebral oxygenation has been used clinically for several years in other patient groups, and now has entered neonatology to a significant degree, the SafeBoosC III trial can be considered comparative effectiveness research rather than a test of an experimental intervention [46]. Therefore, the protocol allows and encourages principal investigators at each NICU to consider and potentially to

seek approval from research ethics boards for one of two other consent forms, i.e. deferred informed consent [43] and prior informed assent (opt-out with enrolment as default) [47]. We believe this offers appropriate flexibility in an international trial in an area where legitimate ethical considerations are in conflict. For this purpose, we have developed parental information sheets specific for each consent method (Additional file 3).

Though extremely preterm infants constitute only 0.5% of all births [1], they represent an extremely high-risk population, and thus their contribution to infant mortality and to the prevalence of cerebral palsy exceeds 10% [48–50]. Accumulating evidence indicates that cerebral hypoxia is a significant cause of mortality as well as brain injury in this population. Thus, monitoring of cerebral oxygenation levels during the first days after birth has the potential to address a significant health problem. Although the overall risk in this population is high, there are many other relevant contributing factors to mortality and brain injury, and thus only a moderate risk reduction can be expected. Therefore, a trial to address this therapeutic question must be large in scope. If the experimental intervention proves successful, we may save 2000 extremely preterm infants or more every year from death or a life with handicap due to brain injury in high-income countries. The ensuing health economics impact may thus be quite robust.

In conclusion, there is an urgent need for a randomised clinical trial to assess the effects of cerebral NIRS monitoring compared with treatment as usual in extremely preterm infants.

Trial status

The protocol is registered at www.clinicaltrials.gov (NCT03770741; registered 10 December 2018). The first infant was enrolled in June 2019 and the anticipated date of study completion is October 2021. Recruitment status can be accessed at www.safeboosc.eu.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s13063-019-3955-6.

Additional file 1. SPIRIT checklist.

Additional file 2. Treatment guideline. A description of the treatment guideline that will be used in the SafeBoosC III trial.

Additional file 3. Parental information sheets and consent form. Templates for parental information sheets for different consent methods and a general consent form.

Additional file 4. WHO trial registration data set.

Abbreviations

BPD: Bronchopulmonary dysplasia; cPVL: Cystic periventricular leukolamacia; ICC: Intra-class correlation coefficients; IVH: Intraventricular haemorrhage; NEC: Necrotizing enterocolitis; NICU: Neonatal intensive care unit; NIRS: Near-infrared spectroscopy; ROP: Retinopathy of prematurity; SAE: Severe adverse events; SAR: Severe adverse reactions

Acknowledgements

We would like to thank the Elsass Foundation, Sven Andersen Foundation and Aage og Johanne Louis-Hansen Foundation for making the conduct of the SafeBoosC III trial possible by covering central trial unit costs including trial coordination and data management. We would also like to thank Marie Isabel Rasmuseen for English language revision of the manuscript.

Authors' contributions

MLH, GG, SHS, JCJ, and CG contributed to the conception and design of the protocol, drafted the main protocol, drafted the manuscript, and will give final approval of the version to be published. AP and ED contributed to the conception and design of the protocol, drafted the main protocol, revised the manuscript critically for important intellectual content, and will give final approval of the version to be published. JM, AMH, CH, EE, GD, GP, GN, GC, HG, JT, KBK, MF, OC, PL, SF, TS, and TA contributed to the conception and design of the protocol, revised the main protocol critically for important intellectual content, revised the manuscript critically for important intellectual content, and will give final approval of the version to be published.

Funding

The sponsor/coordinating investigator, professor of neonatology Gorm Greisen, is the initiator of the SafeBoosC III project. He has no financial interest in the results of the trial, nor in the NIRS devices. The Elsass Foundation supported this work through an unconditional and unrestricted grant of DKK 2,700,000, the Svend Andersen Foundation with an unconditional and unrestricted grant of DKK 1,000,000 and Aage and Johanne Louis-Hansen Foundation with an unconditional and unrestricted grant of DKK 1,000,000. These funding sources had no role in the design of this study and will not have any role during its execution, analyses, interpretation of the data, or decision to submit results. We seek additional local and central funding, including funding from industry sponsors. Such sources will not have any influence on the methodology, data, analysis, reporting, or conclusions of the trial. Furthermore, any participating department can seek local/national support from all sources, including dealers of devices, as long as such sources will not have any influence on the methodology, data, analysis, reporting, or conclusions of the trial.

Availability of data and materials

Not applicable.

Ethics approval and consent to participate

By 31 October 2019, the study has been approved in 57 NICUs from Austria, China, Czech Republic, Denmark, Greece, Ireland, Italy, Norway, Poland, Portugal, Spain, Switzerland, Turkey and the US... No sites will start randomising participants before ethics approval has been granted. Status on ethics approval for all participating sites can be found at www.safeboosc.eu.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Department of Neonatology, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark. [2]Department of Neonatology, La Paz University Hospital, Paseo De La Castellana 261, 28046 Madrid, Spain. [3]Copenhagen Trial Unit, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark. [4]Infant Centre and Department of Paediatrics and Child Health, University College Cork, College Road, Cork, Ireland. [5]Division of Newborn Medicine, Hackensack Meridian Health – Mountainside Medical Center, 1 Bay Ave, Montclair, NJ, USA. [6]Department of Intensive Care, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark. [7]Department of Neonatology, Royal Hospital for Children, 1345 Govan Rd, Glasgow G51 4TF, UK. [8]Department of Neonatology, Children's University Hospital of Zürich, Steinweisstrasse 75, 8037 Zurich, Switzerland. [9]Department of Neonatology,

Gazi University Hospital, Emniyet Mahallesi, Gazeteci Yazar Muammer Yaşar Bostancı Sokak, 06560 Yenimahalle/Ankara, Turkey. [10]NICU, Department of Pediatrics, University General Hospital of Patras, 265 04 Patras, Greece. [11]Department of Pediatrics, Medical University of Graz, Auenbruggerplatz 30, Graz, Austria. [12]Department of Neonatology, University Hospital Leuven, Herestraat 49 Leuven, Belgium. [13]Department of Neonatology, Children's Hospital of Fudan University, 399 Wanyuan Rd, Minhang Qu, Shanghai Shi, China. [14]Department of Neonatology, Centro hospitalar Universitário de São João, Alameda Prof. Hernâni Monteiro, 4200-319 Porto, Portugal. [15]Department of Neonatology, University Hospital Motol, V Uvalu 84, 150 06 Prague 5, Czech Republic. [16]Department of Neonatology, University Children's Hospital Tuebingen, Hoppe-Seyler-Straße 1, 72076 Tuebingen, Germany. [17]Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico Milan, Via della Commenda 12, IT- 20122 Milan, Italy. [18]Department of Clinical Sciences and Community Health, University of Milan, Via Francesco Sforza 35, 20122 Milan, Italy. [19]Department of Neonatology, Hospices Civil De Lyon, 3 Quai des Célestins, 69002 Lyon, France. [20]Department of Neonatology, Wilhelmina Children's Hospital, Lundlaan 6, 3584 EA Utrecht, Netherlands. [21]Department of Neonatology, Oslo University Hospital, Kirkeveien, 166 0450 Oslo, Norway. [22]Department of Neonatology, Poznan University of Medical Sciences, Polna 33, 60-535 Poznań, Poland. [23]Neonatal Intensive Care Unit, Cambridge University Hospitals NHS Foundation Trust, Hills Road, Cambridge CB2 0SW, UK. [24]Department of Cardiology, Holbæk Hospital, Smedelundsgade 60, 4300 Holbæk, Denmark. [25]Department of Regional Health Research, The Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark.

Received: 22 February 2019 Accepted: 4 December 2019
Published online: 30 December 2019

References

1. Blencowe H, Cousens S, Oestergaard MZ, Chou D, Moller A-B, Narwal R, et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. Lancet. 2012;379:2162–72.
2. Stoll BJ, Hansen NI, Bell EF, Shankaran S, Laptook AR, Walsh MC, et al. Neonatal outcomes of extremely preterm infants from the NICHD Neonatal Research Network. Pediatrics. 2010;126:443–56.
3. Adams-Chapman I, Heyne RJ, DeMauro SB, Duncan AF, Hintz SR, Pappas A, et al. Neurodevelopmental impairment among extremely preterm infants in the Neonatal Research Network. Pediatrics. 2018;141: e20173091.
4. Volpe JJ. Brain injury in the premature infant: neuropathology, clinical aspects and pathogenesis. Semin Pediatr Neurol. 1998;5:135–51.
5. Ward RM, Beachy JC. Neonatal complications following preterm birth. BJOG An Int J Obstet Gynaecol. 2003;110:8–16.
6. Behrman R, Butler AS. Preterm birth: Causes, consequences and prevention. Institute of Medicine. Washington, D.C.: National Academies Press; 2007.
7. Stephens BE, Vohr BR. Neurodevelopmental outcome of the premature infant. Pediatr Clin N Am. 2009;56:631–46.
8. Kluckow M. Low systemic blood flow and pathophysiology of the preterm transitional circulation. Early Hum Dev. 2005;81:429–37.
9. Papile LA, Burstein J, Burstein R, Koffler H. Incidence and evolution of subependymal and intraventricular hemorrhage: A study of infants with birth weights less than 1,500 gm. J Pediatr. 1978;92:529–34.
10. Guzzetta F, Shackelford GD, Volpe S, Perlman JM, Volpe JJ. Periventricular intraparenchymal echodensities in the premature newborn: critical determinant of neurologic outcome. Pediatrics. 1986;78:995–1006.
11. Volpe JJ. Brain injury in premature infants: a complex amalgam of destructive and developmental disturbances. Lancet Neurol. 2009;8:110–24.
12. Cordeiro CN, Tsimis M, Burd I. Infections and brain development. Obstet Gynecol Surv. 2015;70:644–55.
13. Prado EL, Dewey KG. Nutrition and brain development in early life. Nutr Rev. 2014;72:267–84.
14. Perlman JM. White matter injury in the preterm infant: An important determination of abnormal neurodevelopment outcome. Early Hum Dev. 1998;53:99–120.
15. Greisen G, Vannucci RC. Is periventricular leucomalacia a result of hypoxic-ischaemic injury? Hypocapnia and the preterm brain. Biol Neonate. 2001;79: 194–200.
16. Perlman M, Volpe J. Are venous circulatory abnormalities important in the pathogenesis of hemorrhagic and/or ischemic cerebral injury? Pediatrics. 1987;80:705–11.
17. Alderliesten T, Dix L, Baerts W, Caicedo A, van Huffel S, Naulaers G, et al. Reference values of regional cerebral oxygen saturation during the first 3 days of life in preterm neonates. Pediatr Res. 2016;79:55–64.
18. Hyttel-Sorensen S, Austin T, van Bel F, Benders M, Claris O, Dempsey E, et al. A phase II randomized clinical trial on cerebral near-infrared spectroscopy plus a treatment guideline versus treatment as usual for extremely preterm infants during the first three days of life (SafeBoosC): study protocol for a randomized controlled tria. Trials. 2013;14:120.
19. Riera J, Hyttel-Sorensen S, Bravo MC, Cabañas F, López-Ortego P, Sanchez L, et al. The SafeBoosC phase II clinical trial: an analysis of the interventions related with the oximeter readings. Arch Dis Child Fetal Neonatal Ed. 2016; 101:F333–8.
20. European Medicines Agency, Committee for Human Medicinal Products. Guideline on Good Clinical Practice E6(R2). 2017.
21. Chan A-W, Tetzlaff JM, Gotzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ. 2013;346:e7586.
22. Davidson JO, Wassink G, van den Heuij LG, Bennet L, Gunn AJ. Therapeutic hypothermia for neonatal hypoxic–ischemic encephalopathy – where to from here? Front Neurol. 2015;6:198.
23. Hansen ML, Pellicer A, Gluud C, Dempsey E, Mintzer J, Hyttel-Sorensen S, et al. Detailed statistical analysis plan for the SafeBoosC III trial: a multinational randomised clinical trial assessing treatment guided by cerebral oxygenation monitoring versus treatment as usual in extremely preterm infants. Trials. 2019;20:746.
24. Pellicer A, Greisen G, Benders M, Claris O, Dempsey E, Fumagally M, et al. The SafeBoosC phase II randomised clinical trial: A treatment guideline for targeted near-infrared-derived cerebral tissue oxygenation versus standard treatment in extremely preterm infants. Neonatology. 2013;104:171–8.
25. Kleiser S, Ostojic D, Andresen B, Nasseri N, Isler H, Scholkmann F, et al. Comparison of tissue oximeters on a liquid phantom with adjustable optical properties: an extension. Biomed Opt Express. 2018;9:86.
26. Vermont Oxford Network. Manual of operations: Part 2 data definitions & infant data forms. 2018.
27. Hyttel-Sørensen S, Pellicer A, Alderliesten T, Austin T, Van Bel F, Benders M, et al. Cerebral near infrared spectroscopy oximetry in extremely preterm infants: phase II randomised clinical trial. BMJ. 2015;350:1–11.
28. Yelland LN, Sullivan TR, Collins CT, Price DJ, McPhee AJ, Lee KJ. Accounting for twin births in sample size calculations for randomised trials. Paediatr Perinat Epidemiol. 2018;32:380–7.
29. Jakobsen JC, Gluud C, Winkel P, Lange T, Wetterslev J. The thresholds for statistical and clinical significance – a five-step procedure for evaluation of intervention effects in randomised clinical trials. BMC Med Res Methodol. 2014;14:34.
30. Greisen G, van Bel F. Equipoise is necessary for randomising patients to clinical trials. Acta Paediatr Int J Paediatr. 2016;105:1259–60.
31. Hyttel-Sørensen S, Greisen G, Als-Nielsen B, Gluud C. Cerebral near-infrared spectroscopy monitoring for prevention of brain injury in very preterm infants. Cochrane Database Syst Rev. 2017;9:CD011506.
32. Hunter CL, Oei JL, Suzuki K, Lui K, Schindler T. Patterns of use of near-infrared spectroscopy in neonatal intensive care units: international usage survey. Acta Paediatr. 2018;107:1198–204.
33. Bevan PJW. Should cerebral near-infrared spectroscopy be standard of care in adult cardiac surgery? Hear Lung Circ. 2015;24:544–50.
34. Savović J, Turner RM, Mawdsley D, Jones HE, Beynon R, Higgins JPT, et al. Association between risk-of-bias assessments and results of randomized trials in Cochrane reviews: The ROBES Meta-Epidemiologic Study. Am J Epidemiol. 2018;187:1113–22.
35. Hrobjartsson A, Thomsen ASS, Emanuelsson F, Tendal B, Hilden J, Boutron I, et al. Observer bias in randomized clinical trials with measurement scale outcomes: a systematic review of trials with both blinded and nonblinded assessors. Can Med Assoc J. 2013;185:E201–11.
36. Hróbjartsson A, Thomsen ASS, Emanuelsson F, Tendal B, Rasmussen JV, Hilden J, et al. Observer bias in randomized clinical trials with time-to-event outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. Int J Epidemiol. 2014;43:937–48.

37. Hróbjartsson A, Emanuelsson F, Skou Thomsen AS, Hilden J, Brorson S. Bias due to lack of patient blinding in clinical trials. A systematic review of trials randomizing patients to blind and nonblind sub-studies. Int J Epidemiol. 2014;43:1272–83.
38. Anthon CT, Granholm A, Perner A, Laake JH, Møller MH. No firm evidence that lack of blinding affects estimates of mortality in randomized clinical trials of intensive care interventions: a systematic review and meta-analysis. J Clin Epidemiol. 2018;100:71–81.
39. Hintz SR, Slovis T, Bulas D, Van Meurs KP. Interobserver reliability and accuracy of cranial ultrasound interpretation in premature infants. J Pediatr. 2007;150:592–6.
40. Campbell M. Framework for design and evaluation of complex interventions to improve health. BMJ. 2000;321:694–6.
41. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ. 2008;337:a2390.
42. Boutron I, Altman DG, Moher D, Schulz KF, Ravaud P. CONSORT statement for randomized trials of nonpharmacologic treatments: A 2017 update and a CONSORT extension for nonpharmacologic trial abstracts. Ann Intern Med. 2017;167:40.
43. Brierley J, Larcher V. Emergency research in children: options for ethical recruitment. J Med Ethics. 2011;37:429–32.
44. Beauchamp TL, Childress JF, Press. OU. Principles of biomedical ethics. 7th ed. Oxford: Oxford University Press; 2012.
45. Mason SA, Allmark PJ. Obtaining informed consent to neonatal randomised controlled trials: interviews with parents and clinicians in the Euricon study. Lancet. 2000;356:2045–51.
46. Gale C, Juszczak E. A paediatrician's guide to clinical trials units. Arch Dis Child Educ Pract Ed. 2016;101:265–7.
47. Gale C, Hyde MJ, Modi N. Research ethics committee decision-making in relation to an efficient neonatal trial. Arch Dis Child Fetal Neonatal Ed. 2017; 102:F291–8.
48. Hafström M, Källén K, Serenius F, Maršál K, Rehn E, Drake H, et al. Cerebral palsy in extremely preterm infants. Pediatrics. 2018;141:e20171433.
49. Roberts G, Anderson PJ, De Luca C, Doyle LW. Changes in neurodevelopmental outcome at age eight in geographic cohorts of children born at 22-27 weeks' gestational age during the 1990s. Arch Dis Child Fetal Neonatal Ed. 2010;95:F90–4.
50. Stanley F, Blair E, Alberman E. Cerebral palsies: epidemiology and casual pathways. Cambridge: Cambridge University Press; 2000.

Publisher's Note

Trials

UPDATE

Open Access

Detailed statistical analysis plan for the SafeBoosC III trial: a multinational randomised clinical trial assessing treatment guided by cerebral oxygenation monitoring versus treatment as usual in extremely preterm infants

Mathias Lühr Hansen[1*], Adelina Pellicer[2], Christian Gluud[3], Eugene Dempsey[4], Jonathan Mintzer[5], Simon Hyttel-Sorensen[1], Anne Marie Heuchan[6], Cornelia Hagmann[7], Gabriel Dimitriou[8], Gerhard Pichler[9], Gunnar Naulaers[10], Guoqiang Cheng[11], Ana Vilan[12], Jakub Tkaczyk[13], Karen B. Kreutzer[14], Monica Fumagalli[15,16], Olivier Claris[17], Siv Fredly[18], Tomasz Szczapa[19], Theis Lange[20,21], Janus Christian Jakobsen[3,22,23] and Gorm Greisen[1]

Abstract

Background: Infants born extremely preterm are at high risk of dying or suffering from severe brain injuries. Treatment guided by monitoring of cerebral oxygenation may reduce the risk of death and neurologic complications. The SafeBoosC III trial evaluates the effects of treatment guided by cerebral oxygenation monitoring versus treatment as usual. This article describes the detailed statistical analysis plan for the main publication, with the aim to prevent outcome reporting bias and data-driven analyses.

Methods/design: The SafeBoosC III trial is an investigator-initiated, randomised, multinational, pragmatic phase III trial with a parallel group structure, designed to investigate the benefits and harms of treatment based on cerebral near-infrared spectroscopy monitoring compared with treatment as usual. Randomisation will be 1:1 stratified for neonatal intensive care unit and gestational age (lower gestational age (< 26 weeks) compared to higher gestational age (≥ 26 weeks)). The primary outcome is a composite of death or severe brain injury at 36 weeks postmenstrual age. Primary analysis will be made on the intention-to-treat population for all outcomes, using mixed-model logistic regression adjusting for stratification variables. In the primary analysis, the twin intra-class correlation coefficient will not be considered. However, we will perform sensitivity analyses to address this. Our simulation study suggests that the inclusion of multiple births is unlikely to significantly affect our assessment of intervention effects, and therefore we have chosen the analysis where the twin intra-class correlation coefficient will not be considered as the primary analysis.

Discussion: In line with the Declaration of Helsinki and the International Conference on Harmonization Good Clinical Practice guidelines, we have developed and published this statistical analysis plan for the SafeBoosC III trial, prior to any data analysis.

(Continued on next page)

* Correspondence: mathias.safeboosc@gmail.com
[1]Department of Neonatology, Rigshospitalet, Copenhagen University Hospital, Blegdamsvej 9, 2100 Copenhagen, Denmark
Full list of author information is available at the end of the article

(Continued from previous page)
Trial registration: ClinicalTrials.org, NCT03770741. Registered on 10 December 2018.

Keywords: Randomised clinical trial, Extremely preterm, Near-infrared spectroscopy, Cerebral oximetry, Statistical analysis plan

Background

Extremely preterm infants carry a high risk of death, with a mortality rate up to 25% [1, 2]. Furthermore, about 20% suffer from long-term neurodevelopmental impairment such as cerebral palsy or low intelligence quotient [2, 3]. Psychomotor impairment is a major cause of reduced quality of life and increased costs of medical care, rehabilitation, and special education in this population [4]. Low intelligence quotient affects all aspects of life. With increasing life expectancy, these combined prematurity-related factors pose a significant problem.

Hypoxia has been associated with mortality and brain injury in the preterm population [5]. In the SafeBoosC II trial, cerebral near-infrared spectroscopy (NIRS) monitoring combined with an evidence-based treatment guideline significantly reduced the burden of hypoxia during the first days of life in preterm infants [6]. There were also trends towards reduced occurrence of severe brain injury and mortality [6]. On the other hand, the incidence of bronchopulmonary dysplasia and retinopathy of prematurity was higher among NIRS-monitored neonates [6]. However, SafeBoosC II was not powered to demonstrate effects on these outcomes; thus, high-certainty evidence of clinical benefit and harm in extremely preterm infants is lacking [7]. We therefore plan a larger phase III trial, SafeBoosC III, powered to demonstrate the potential benefits and harms of treatment based on cerebral NIRS monitoring compared with treatment as usual on patient-centred clinical outcomes. As the SafeBoosC III trial will be conducted in compliance with the Declaration of Helsinki in its latest form and the International Conference on Harmonization Good Clinical Practice guidelines [8], we have developed this detailed statistical analysis plan. We believe this will decrease the risk of outcome reporting bias and data-driven analyses.

Methods/design

Trial overview

SafeBoosC III is an investigator-initiated, open-label, randomised, multinational, pragmatic phase III clinical trial with a parallel group design. The primary objective is to evaluate the benefits and harms of treatment based on cerebral NIRS monitoring during the first 72 postnatal hours in extremely preterm infants [9], compared with treatment and monitoring as usual, to reduce cerebral hypoxia [10]. The hypothesis is that treatment based on NIRS monitoring for extremely preterm infants during the first 72 h of life will result in a reduction in death or severe brain injury assessed at 36 weeks postmenstrual age. We plan to test for superiority of the experimental intervention compared with the control group for only the primary outcome, since exploratory outcomes will only be hypothesis generating (see 'Level of significance'). Infants will be randomised with an allocation ratio of 1:1 to either the experimental group or the control group stratified for neonatal intensive care unit (NICU) and gestational age (lower gestational age (< 26 weeks) compared to higher gestational age (≥ 26 weeks)). Details of the randomisation method are held securely in the statistics master file. Infants in the experimental group will start cerebral NIRS monitoring as close to birth as possible, but at least within 6 h of birth, and receive treatment based on NIRS monitoring during the first 72 h of life (Fig. 1). These treatments will follow an evidence-based treatment guideline [11]. Infants in the control group will not receive cerebral NIRS monitoring and will be monitored and treated according to local guidelines and practices (i.e. treatment as usual). Due to the nature of NIRS, it is difficult to blind the clinical staff or the parents of the trial participants.

Three different consent methods may be used in this trial: prior informed consent (prenatal and postnatal); deferred consent; and prior assent/'opt-out'. The trial will be conducted at more than 50 centres across up to 20 countries (16 European countries, India, China, and the USA), and the protocol will be published in an international peer-reviewed journal [10].

The SafeBoosC III trial is registered at ClinicalTrials.org (NCT03770741) and is compliant with the Declaration of Helsinki in its latest form and with the International Conference on Harmonization Good Clinical Practice. The trial will be approved by relevant authorities, including research ethics boards and data protection agencies, in all participating centres. The progression of the trial can be followed at www.safeboosc.eu. This statistical analysis plan has been written and submitted before randomisation commences and all data analysis for the main publication will be compliant to this plan.

Outcomes

Primary outcome

The primary outcome is a composite of either death or severe brain injury. Severe brain injuries will be defined as grade III or IV cerebral haemorrhage (Papile's

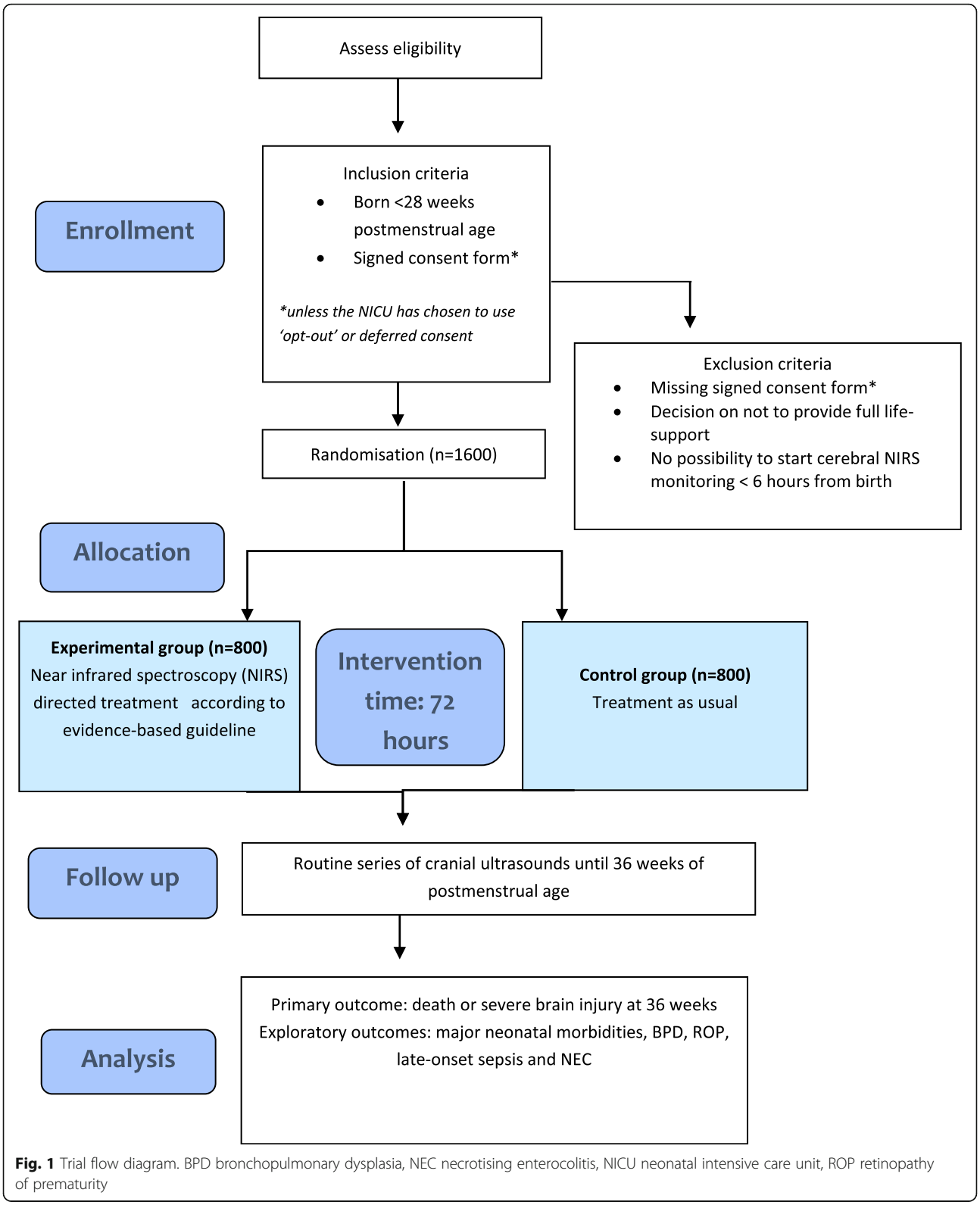


Fig. 1 Trial flow diagram. BPD bronchopulmonary dysplasia, NEC necrotising enterocolitis, NICU neonatal intensive care unit, ROP retinopathy of prematurity

classification) [12], cystic periventricular leukomalacia [2], cerebellar haemorrhage, post-haemorrhagic ventricular dilatation, or cerebral atrophy. These cerebral outcomes will be reported as detected on any one of a series of cranial ultrasound scans that are routinely performed in these infants.

Outcome assessment of mortality will not be blinded, but diagnosis and classification of brain injury and entry of this information into electronic case report forms will be conducted by a clinician blinded to group allocation.

Exploratory outcomes

- A count of the presence of the three major neonatal morbidities associated with neurodevelopmental impairment later in life [13]: bronchopulmonary dysplasia (defined below), retinopathy of prematurity (as defined below), and severe brain injury as defined in the primary outcome (i.e. a value of 0, 1, 2, or 3)
- Bronchopulmonary dysplasia defined as oxygen or ventilator/continuous positive airway pressure requirement at 36 weeks' postmenstrual age
- Retinopathy of prematurity stage 3 and above at any time prior to 36 weeks' postmenstrual age
- Late-onset sepsis (> 72 h after birth) defined as treatment with antibiotics for at least 5 days
- Necrotising enterocolitis stage 2 or higher using the modified Bell's staging system [14] and/or focal intestinal perforation at any time up until 36 weeks' postmenstrual age

Outcome assessment time point

All outcomes will be assessed at 36 weeks postmenstrual age.

Sample size

We have calculated our sample size with an α of 5%, a power of 90%, and a ratio of experimental trial participants to control trial participants of 1:1. The primary outcome is the composite outcome of death or severe brain injury. Sample size calculations were performed for the composite outcome and not for the individual components.

Calculated from the 2009 dataset from the EuroNeoNet project [15] the mortality was 33% and severe intracranial haemorrhage was observed in 15%. In the SafeBoosC II trial, the proportion of trial participants in the control group with the same composite primary outcome was approximately 34% and in the experimental group was 26% [6]. Mortality was 24% in the control group versus 13% in the experimental group and the proportion of infants with severe brain injury was 23% versus 13% [6].

Based on the aforementioned, a total of 1600 infants — 800 infants randomised to the experimental group and 800 infants to the control group — would be required to demonstrate a reduction of the primary outcome from 34.0% to 26.5%, with an α of 5% and a power of 90%. This corresponds to a 22% relative risk reduction or a 7.5% absolute risk reduction. We consider this a clinically relevant and important benefit, since mortality is of direct patient relevance and since surviving infants with severe brain injury (about 25%) are at approximately 40% risk of moderate-to-severe neurodevelopmental impairment [16]. This absolute risk reduction corresponds to a 'number-needed to treat' of 15 infants and, if our null hypothesis is rejected, is likely to influence clinical practice.

Power calculations for exploratory outcomes

For the exploratory outcomes, we have performed power calculations as presented in Table 1.

Assuming a mean major neonatal morbidity count (bronchopulmonary dysplasia, retinopathy of prematurity, and severe brain injury) of 0.62 among extremely preterm infants [17], with a standard deviation of 0.80 and a relative risk increase or decrease of 20% in the experimental group, we will be able to detect this difference between the experimental and control group with 87% power at a 5% significance level (Table 1).

Assuming a 40% prevalence of bronchopulmonary dysplasia among extremely preterm infants [18] and a relative risk decrease or increase of 20% in the experimental group, we will be able to detect this difference between the experimental and control group with 89% power at a 5% significance level (Table 1).

Assuming a 13% prevalence of stage 3 and above retinopathy of prematurity among extremely preterm infants and a relative risk decrease or increase of 30% in the experimental group [7], we will be able to detect this difference between the experimental and control groups with 68% power at a 5% significance level (Table 1).

Assuming a 40% prevalence of late-onset sepsis in the control group [1], defined as treatment with antibiotics for at least 5 days, and a 20% relative risk decrease or increase in the experimental group, we will be able to detect this difference between the experimental and control groups with 91.2% power at a 5% significance level (Table 1).

Assuming an 11% prevalence of stage 2 and 3 necrotising enterocolitis among extremely preterm infants and a 17% relative risk decrease or increase in the experimental group, as is the estimate from existing trials [7], we will be able to detect this difference between the experimental and control groups with 23% power at a 5% significance level (Table 1).

Assessment of outcomes and additional clinical variables

There will be three time points for data collection: at randomisation (from 0 to 6 h after birth); at the end of the intervention period (72 h of life); and at 36 weeks postmenstrual age. Data on feasibility will be assessed at randomisation. At the end of the intervention period,

Table 1 Overview of power calculations for exploratory outcomes

Outcome	Assumption on prevalence in background population (%)	Assumption on risk increase or decrease (%)	Power (%)
Major neonatal morbidities	0.62 (0.8)[a]	20	87
Bronchopulmonary dysplasia	40	20	89
Retinopathy of prematurity	13	30	68
Late-onset sepsis	40	20	91.2
Necrotising enterocolitis	11	17	23

For definition of outcomes, see 'Outcomes'. All power calculations have been made with a 5% significance level
[a]Presented as mean count (standard deviation)

data collection will primarily reflect cerebral NIRS monitoring and safety parameters. As mentioned, all outcomes will be assessed at 36 weeks postmenstrual age. Severe brain injury diagnosis and classification data will be collected either by neonatologists assessing all cranial ultrasound scans performed up until 36 weeks postmenstrual age or by reading radiologists' descriptions of these scans. This assessment and data entry will be conducted by a person blinded to group allocation. No long-term follow-up has been formally planned. However, we encourage clinical sites to conduct long-term follow-up, and we have therefore developed an appendix in the protocol (see full protocol at www.safeboosc.eu) describing possible outcomes for later follow-up studies and how these could be conducted. Currently, no protocol for such an ancillary study has been developed.

Explanatory variables

Additional clinical data on trial participants will be drawn from clinical files, in order to compare characteristics between intervention groups. Data will be drawn from clinical records at 72 h of age and 36 weeks postmenstrual age. These data consist of a subset of explanatory variables, with the majority usually being reported to the neonatal network databases, such as Vermont Oxford Network [19]. These data will be presented in a table in the main publication (see Table 2). Tests of statistical significance will not be undertaken for explanatory variables. Categorical data will be summarised by numbers and percentages. Continuous data will be summarised by mean and standard deviation if normally distributed or by median and interquartile range if non-normally distributed.

Safety

We will report the total number of serious adverse reactions, as defined in the protocol [10] for each group, as well as the total number of participants who experienced one or more serious adverse reactions in each group. We will also report the total number of serious adverse events, as defined in the protocol [10] in each group, as well as the number of participants who experienced one or more serious adverse events in each group.

Level of significance

The thresholds for significance will be assessed according to a 5-point procedure, suggested by Jakobsen et al. [20]. We will calculate and report confidence intervals and exact *p*-values for the primary and exploratory outcomes. All confidence intervals presented will be 95% and two-sided. A *p*-value of less than 0.05 will be used as the threshold for statistical significance for our primary outcome, since this value was used as the acceptable risk of type I error in our sample size estimation (see 'Sample size') and since we plan to report on only one primary outcome. However, in our interpretation of the results, we will assess any effect of the experimental intervention according to the point estimate taking into consideration the confidence interval as well as intervention effects on other outcomes [21]. All remaining outcome results will only be considered hypothesis-generating. Since our primary conclusion will be based on one outcome result at one time point, we will limit problems associated with multiple testing, due to multiple outcome comparisons [22].

Secondly, we will calculate and report the Bayes factor [23] for the primary outcome [24]. The Bayes factor is the ratio between the probability of the results given that the null hypothesis (H_0) is true divided by the probability of the results given that the alternative hypothesis (H_A) is true [23]. In the SafeBoosC III trial, the alternative hypothesis is that the treatment effect is the effect that was used for the sample size calculation: a 22% relative risk reduction in the experimental group. By calculating the Bayes factor, we will be able to interpret the results of the primary outcome in relation to former trial results [6].

Thirdly, Lan–DeMets monitoring boundaries will be used to adjust the threshold for statistical significance at each interim analysis to judge whether the trial should be terminated early [25]. This is done in order to avoid a false rejection of the null hypothesis based on insufficient sample sizes [26]. The trial will not be stopped prematurely due to futility. The fourth step in the five-step procedure by Jakobsen et al., regarding adjustment of *p*-values based on multiple testing of the primary outcome, is not applicable to our trial, since we have a single primary outcome [20].

Table 2 Explanatory variables divided by experimental group and control group participants

Variables	Experimental group (*n*)	Control group (*n*)
At randomisation		
Birth weight (g)		
Gestational age (weeks)		
Apgar 1 min (1–10)		
Apgar 5 min (1–10)		
Gender		
Male (%)		
Female (%)		
At 72 h of age		
Age when NIRS monitoring started (h)[a]		N/A
Stopping NIRS monitoring before end of monitoring period (%)[a]		N/A
Parents discontinuing trial participation (%)		
Changes in treatment due to cerebral hypoxia (%)[a]		N/A
Registered cardiovascular support treatment (%)[a]		N/A
Type of NIRS device used[a] INVOS (%) NIRO (%) Fore-Sight (%) Sensmart (%) O3 (%) Egos (%) Oxyprem (%) Other (%)		N/A
Cerebral NIRS monitoring despite being in control group (%)[b]	N/A	
Surfactant therapy (%)		
Severe adverse reactions (%)		
At 36 weeks postmenstrual age		
Major congenital anomaly (%)		
Mechanical ventilation (%)		
Time with mechanical ventilation (days)		
Patent ductus arteriosus (%)		
Weight (g)		
Early cranial ultrasound scan (%)		
Late cranial ultrasound scan (%)		

Data expressed as median (range) for continuous variables, and numbers (percentage) for dichotomous variables
N/A not applicable, *NIRS* near-infrared spectroscopy
[a]Variables only relevant for experimental group participants
[b]Variables only relevant for control group participants

We will take the upper and lower limits of the confidence intervals into consideration when making study conclusions [21]. Clinical significance will be assessed by calculating the number needed to treat based on the absolute risk reduction data. Based on the results from the phase II trial, we expect an absolute risk reduction of 7.5%, which corresponds to a number needed to treat of 15 (see 'Sample size') [6].

Interim analyses

One pre-planned interim analysis will be conducted after one-third of trial participants have been randomised. The timing and prevalence of additional interim analyses will be decided solely by the data monitoring and safety committee members. The data monitoring and safety committee will make recommendations to the steering group to continue, change, hold, or terminate the trial. This recommendation will be based primarily on safety considerations and will be guided by statistical monitoring guidelines, defined in the data monitoring and safety committee charter. The data monitoring and safety committee will be provided with the following data from the Coordinating Data Centre: number of participants randomised, number of participants per intervention group (0,1), number of participants stratified per stratification variable per intervention group (0,1), and number of events (primary outcome, SAEs, and SARs) in the two groups. Based on the evaluations of these outcomes, the

data monitoring and safety committee will decide whether they want further data from the Coordinating Data Centre, and when next to perform analyses of data. Based on the analyses of the safety variables, the data monitoring and safety committee is suggested to use Lan–DeMets sequential monitoring boundaries, based upon a relative risk increase of 100% of mortality from 25% to 50%. For any of the other safety outcomes, the statistical limit to guide its recommendations regarding early termination of the trial for harms is recommended also to be conservative.

Handling of missing data

Missing data will be minimised by performing repeated monitoring of data entry into electronic case report forms. In this way, we will be able to monitor the extent of missing data and intervene if necessary. Hence, we do not anticipate that there will be any significant number of missing values. However, we will consider using multiple imputation and present best–worst and worst–best case scenarios if it is not valid to ignore missing data according to the standards reported by Jakobsen et al. [27]. When using best–worst and worst–best case scenarios, we will assess the potential range of impact of the missing data for the trial results [27]. In the 'best–worst' case scenario, it is assumed that all patients lost to follow-up in the experimental group have had a beneficial outcome, and all those with missing outcomes in the control group have had a harmful outcome [27]. Conversely, in the 'worst–best' case scenario, it is assumed that all patients who were lost to follow-up in the experimental group have had a harmful outcome, and that all those lost to follow-up in the control group have had a beneficial outcome [27].

As recommended, we will describe reasons why outcome data are missing in the main study manuscript [28]. Furthermore, we will compare explanatory variables between all participants randomised to intervention groups (including those with missing outcomes), and also between participants in the intervention groups, where outcomes are reported. This is done to identify imbalances between groups due to missing outcome data [29].

Twins and their intra-cluster correlation

In extremely preterm populations, 30% of births may be twins [6], which poses a potential problem for statistical analyses as the outcomes among pairs of twins are potentially correlated [30]. In the SafeBoosC III trial, multiple birth infants will be randomised as a 'pair' or a 'group' (i.e. all siblings will be allocated to the same intervention group). In centres where only one or two cerebral monitoring devices are available, it may not be possible to include all infants from multiple births. Thus, only one of a pair or only one or two infants of triplets may be included. The sibling(s) enrolled in the trial will be the one(s) born last. In the SafeBoosC II trial, the intra-class correlation coefficient (ICC) of the burden of hypoxia within pairs of twins was negligible. The ICC for various binary outcomes has been estimated in a previous study: ICC for death before discharge was estimated as 0.00 (95% confidence interval (CI) −0.04 to 0.02) and for intraventricular haemorrhage grade 3 or 4 as − 0.01 (95% CI − 0.05 to 0.01) [31]. These values correlate to a design effect very close to 1 [31]. Therefore, in the primary analysis, we will analyse twin data as independent observations. However, due to the possibility that the correlation between the primary outcome within multiple births will interfere with the estimation of the treatment effect [32], and particularly the assessment of estimation uncertainty, we will perform a sensitivity analysis, taking this effect into consideration. This sensitivity analysis will be performed using the generalised estimating equation (GEE) approach utilising an exchangeable covariance matrix with site (NICU) and stratification variables as fixed effects. The results of both primary outcome analyses will be presented and discrepancies between the two analyses discussed in the final publication. Furthermore, we will calculate, report, and discuss the ICC for the primary outcome.

Stratification

We will use site (NICU) and gestational age (lower gestational age (< 26 weeks) compared to higher gestational age (≥ 26 weeks)) as stratification variables in the randomisation. Analyses for all outcomes will be adjusted for these stratification variables [33–35].

Assessment of underlying statistical assumptions

For all regression analyses, we will test for major interactions between each covariate and the intervention variable. We will, in turn, include each possible first-order interaction between included covariates and the intervention variable. For each combination, we will test whether the interaction term is significant and assess the effect size. We will only consider that there is evidence of an interaction if the interaction is statistically significant after Bonferroni-adjusted thresholds (0.05 divided by number of possible interactions) and if the interaction shows a clinically significant effect. If it is concluded that the interaction is significant, we will be presenting an analysis separately for each (e.g. for each site if there is significant interaction between the trial intervention and 'site') and an overall analysis including the interaction term in the model.

Assessment of underlying statistical assumptions for dichotomous outcomes

We will assess whether the deviance divided by the degrees of freedom is significantly larger than 1 to assess for relevant overdispersion, and in this case consider using a maximum likelihood estimate of the dispersion parameter. To avoid analytical problems with either zero events or problems such as all participants dying at a given site, we have only included sites planning to randomise a sufficient number of participants. However, we cannot exclude the risk that some sites might have problems with recruitment. We will, by checking whether the number of participants is larger than 10 (rule of thumb) per site, pool the data from small sites if the number of participants is too low.

Statistical analyses

Analyses will be made on the intention-to-treat population for all outcomes, since this method maintains baseline comparability of the intervention groups [29]. The intention-to-treat population will include all randomised patients, regardless of missing data, lost to follow-up or adherence to the intervention.

In our primary analysis, we will analyse dichotomous outcomes using mixed-effect logistic regression and count data using mixed-effect linear regression with robust standard errors. In all regression models, 'site' will be included as a random effect. The remaining stratification variables (age and intervention groups) will be included as fixed effects. The sensitivity analysis accounting for the possible correlation between twins is described in 'Twins and their intra-cluster correlation'.

As an additional sensitivity analysis, we will perform a per-protocol analysis, only including participants who had no missing data, were not lost to follow-up, and adhered to the intervention. Adherence to the intervention is defined as continuous cerebral oxygenation monitoring during the first 72 h of life or until death.

We will, in a secondary analysis, analyse the results using random-effects meta-analysis [36].

All outcomes will be analysed collectively since the follow-up time is identical.

Data management

The data management plan has been described in the protocol paper [10].

CONSORT flow diagram

The main publication will include a Consolidated Standards of Reporting of Randomised Trials (CONSORT) flow diagram, following the CONSORT 2010 Statement [37]. This will be used to summarise the number of patients who were randomised, allocated to the experimental and control groups, adhered and unadhered to the intervention, lost to follow-up (including parental and physician withdrawal), randomised and included in the primary analysis, and randomised and excluded from the primary analysis.

Withdrawal

Parents will be able to withdraw consent at any time during the trial. However, data on participants up until the day of withdrawal will be used and participants will be part of the intention-to-treat population and analysis.

Blinding of statisticians

All data managers, statisticians, and those drawing conclusions will be blinded to treatment allocation. Two blinded statisticians connected to The Copenhagen Trial Unit will independently perform all statistical analyses and the two statistical reports will be published as supplemental material. Discrepancies between the two reports will be discussed by the Steering Committee of the trial. The two intervention groups will be coded 'A' and 'B'. When comparability between the two independent analyses have been obtained, two abstracts will be written: one assuming 'A' is the experimental group and 'B' is the control group – and one assuming the opposite. After the conclusions have been drawn, blinding will be broken, and the final manuscript will be based on the correct pre-written abstract.

Simulation of twin scenarios

To explore the potential impact of twin correlation, we conducted a simulation study to assess potential impact on power and coverage probabilities of confidence intervals (i.e. does the computed 95% CI contain the true parameter values with 95% probability). We compared the naive analysis (primary analysis of the primary outcome), which ignores twin pairs, to a GEE-based approach which does account for twin correlation. We did this by simulating 10,000 trials with sample size and true parameter values as in the sample size estimation and varied twin probability and ICC. These results are presented in Table 3. This simulation study shows that for a low ICC value or low twin proportion, we can expect both the naive and GEE analyses to have correct coverage and equal power. For a high ICC and a high twin proportion, we can expect the GEE analysis to retain correct coverage, while the naive analysis will have decreased coverage; these differences, however, would be minimal. For high twin proportion and high ICC values, the effective sample size was reduced, which as expected implied that the correct analysis (the GEE) yields a lower power than the intended 90%, albeit only marginally so, and that the coverage for the naïve analysis was a bit too low.

Discussion

This article presents the detailed statistical analysis plan for the SafeBoosC phase III trial. It has been developed and submitted prior to any randomisation or data collection in order to avoid data-driven analyses and outcome reporting bias. Data will be analysed on the intention-to-treat population, and multiple imputations will be used if the proportion of missing data cannot be ignored (see 'Handling of missing data'). An anonymised dataset regarding all outcomes will be uploaded to a public database to be available for other researchers and peers 6 months after acceptance of the study manuscript.

We plan to report on both primary and exploratory outcomes in the main publication, but the conclusion will solely be based on the results of the primary outcome. If the result is statistically insignificant, based on the 5-point procedure by Jakobsen et al. [20], we will conclude that there is no significant difference between the intervention and treatment as usual (see 'Level of significance').

Dealing with multiple analyses

Planning multiple analyses on a primary outcome has the potential to increase the risk of type I errors, due to multiple testing [38]. If it is predefined that a significant difference between the experimental and control groups on any one of the primary outcome analyses is sufficient to declare superiority of a given intervention, one would have to correct for multiple testing by decreasing the α value [22, 39]. On the other hand, planning that all primary outcome analyses must show significant benefit of the intervention to declare superiority has the potential to increase the risk of type II errors, due to insufficiently powered analyses [40]. Hence, by only planning one analysis for the primary outcome and by defining additional analyses as sensitivity analyses (see 'Twins and their intra-cluster correlation'), we have eliminated the type I and type II error-related issues described above. The sensitivity analyses on the primary outcome will only be used to discuss and illustrate the results of the primary analysis.

Strengths

According to our knowledge, SafeBoosC III will provide the largest trial, thus far, evaluating the benefits and harms of treatment guided by cerebral NIRS monitoring – not only in extremely preterm infants [41] but across all patient populations [42].

It is an important strength that both the protocol and statistical analysis plan have been developed and submitted prior to any randomisation or data collection [8, 10]. Furthermore, we have also taken the issue of twins and their intra-cluster correlation into consideration, by

Table 3 Simulation study to assess power and coverage probabilities of confidence intervals of primary outcome

ICC	Proportion of twins	Power of naive analysis	Power of GEE analysis	Coverage probability of naive analysis	Coverage probability of GEE analysis
0	0.1	0.91	0.91	0.95	0.95
0	0.2	0.91	0.91	0.95	0.95
0	0.3	0.90	0.90	0.95	0.95
0	0.4	0.90	0.90	0.95	0.95
0.01	0.1	0.91	0.91	0.95	0.95
0.01	0.2	0.90	0.90	0.95	0.95
0.01	0.3	0.90	0.90	0.95	0.95
0.01	0.4	0.91	0.90	0.95	0.95
0.03	0.1	0.90	0.90	0.95	0.95
0.03	0.2	0.91	0.91	0.95	0.95
0.03	0.3	0.90	0.90	0.95	0.95
0.03	0.4	0.90	0.90	0.95	0.95
0.13	0.1	0.90	0.90	0.95	0.95
0.13	0.2	0.90	0.89	0.94	0.95
0.13	0.3	0.90	0.89	0.94	0.95
0.13	0.4	0.90	0.88	0.94	0.95
0.2	0.1	0.90	0.90	0.95	0.95
0.2	0.2	0.90	0.89	0.94	0.95
0.2	0.3	0.90	0.88	0.94	0.95
0.2	0.4	0.89	0.87	0.94	0.95

GEE generalised estimating equation, *ICC* intra-class coefficient

performing an additional sensitivity analysis to address its potential effect on results (see section on 'Twins and their intra-cluster correlation'). To address the potential impact of twin correlation on our results, we also performed a simulation study, showing that we can expect the potential impact of twin correlation to be minor.

There is genuine evidence that most randomised clinical trials lack external validity, which is an important explanation for why multiple interventions proven beneficial in randomised clinical trials are underused in routine clinical practice [43]. Since SafeBoosC III is an international trial including multiple sites across different countries, limitations to external validity such as different practices between countries and health-care systems seems less of an issue for external validity. Furthermore, the external validity of our results will also be described in the main publication, as recommended in the Consolidated Standards of Reporting of Randomised Trials guidelines [37].

Limitations

Our methodology also has limitations. Only three of the five exploratory outcomes are sufficiently powered (80% power) to show a significant difference between the experimental and control groups, at a 5% significance level. If these were categorised as secondary or additional primary outcomes, we would need to correct for multiple testing by decreasing the α value using Bonferroni adjustments [39]. Therefore, we will not make any clinical conclusions based on these results. However, we believe they are important to report and assess since they represent major neonatal morbidities in our study population [1, 13, 44].

As thoroughly reported in the SafeBoosC III design paper [10], it is difficult to blind clinical staff, the infant, and the parents of the trial participants, which introduces risks of bias [45–48]. This important concern is discussed in detail in our design paper [10].

As recommended in the European Medicines Agency Guidelines on Multiplicity Issues in Clinical Trials, the components of the primary composite outcome (i.e. death and severe brain injury) will be analysed separately [49]. However, interpretation of these sub-analyses will be difficult, since death and severe brain injury as individual outcomes are insufficiently powered to show a real benefit of the intervention.

Trial status

At present, the study protocol has been registered at www.clinicaltrials.gov (NCT 03770741, registered on 10 December 2018) and has been accepted for publication [10]. The first participant was randomised on 27 June 2019. Status on recruitment can be accessed at www.safeboosc.eu.

Statistical analysis plan status

Version 1.0 (8 August 2019). This document has been written based on information available in the protocol paper [10].

Abbreviations

CI: Confidence interval; GEE: Generalised estimating equation; ICC: Intra-class correlation coefficient; NICU: Neonatal intensive care unit; NIRS: Near-infrared spectroscopy

Acknowledgements

The authors thank the trial participants in the SafeBoosC II trial for providing information for the design of the present SafeBoosC III trial. They thank the Elsass Foundation for monetary support.

Roles and responsibilities

MLH has taken the lead in writing the statistical analysis plan. JCJ is the senior statistician. GG is the coordinating investigator.

Authors' contributions

MLH, GG, JCJ, CG, and TL contributed to the conception and design of the statistical analysis plan and drafted the main manuscript. AMH, AP, AV, CH, ED, GD, GP, GN, GC, JM, JT, KBK, MF, OC, SH-S, SF, and TS all made substantial contributions to the process of developing the manuscript and revised the manuscript critically for important intellectual content. All authors will give final approval of the version to be published.

Funding

The sponsor/coordinating investigator, Professor of Neonatology Gorm Greisen, is the initiator of the SafeBoosC III project. He has no financial interest in the results of the trial nor in the NIRS devices. The Elsass Foundation supported this work through an unconditional and unrestricted grant of DKK 2,700,000, Svend Andersen foundation with an unconditional and unrestricted grant of DKK 1,000,000 and Aage og Louis-Hansens foundation with an unconditional and unrestricted grant of DKK 1,000,000. These funding sources had no role in the design of this study and will not have any role during its execution, analyses, interpretation of the data, or decision to submit results. As we seek additional local and central funding, including funding from industry sponsors, such sources will not influence the methodology, data, analysis, reporting, or conclusions of the trial. Furthermore, any participating department can seek local/national support from all sources, including dealers of devices, as long as such sources will not influence the methodology, data, analysis, reporting, or conclusions of the trial.

Availability of data and materials

Not applicable.

Ethics approval and consent to participate

By 15 March 2019, the trial had been approved in 55 neonatal intensive care units in Austria, China, Czech Republic, Denmark, Greece, Ireland, Italy, Norway, Poland, Portugal, Spain, Switzerland and Turkey. No sites will start randomising participants before ethics approval has been granted. Status on ethical approval for all participating sites can be found at www.safeboosc.eu. Written informed consent will be obtained by a qualified physician or nurse connected to the trial prior to randomisation of any participant, unless the neonatal intensive care unit (NICU) uses deferred informed consent or prior assent as consent methods. These consent procedures will be approved by local ethics committees or institutional review boards before use.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Department of Neonatology, Rigshospitalet, Copenhagen University Hospital, Blegdamsvej 9, 2100 Copenhagen, Denmark. [2]Department of Neonatology, La Paz University Hospital, Paseo De La Castellana 261, 28046 Madrid, Spain. [3]Copenhagen Trial Unit, Centre for Clinical Intervention

Research, Rigshospitalet, Copenhagen University Hospital, Blegdamsvej 9, 2100 Copenhagen, Denmark. [4]Infant Centre and Department of Paediatrics and Child Health, University College Cork, College Road, Cork, Ireland. [5]Division of Neonatal-Perinatal Medicine, Mountainside Medical Center, Montclair, NJ, USA. [6]Department of Neonatology, Royal Hospital for Children, 1345 Govan Rd, Glasgow G51 4TF, UK. [7]Department of Neonatology, Children's University Hospital of Zürich, Steinweisstrasse 75, 8037 Zurich, Switzerland. [8]NICU, Department of Pediatrics, University General Hospital of Patras, 265 04 Patras, Greece. [9]Department of Pediatrics, Medical University of Graz, Auenbruggerplatz 30, Graz, Austria. [10]Department of Neonatology, University Hospital Leuven, Herestraat 49, Leuven, Belgium. [11]Department of Neonatology, Children's Hospital of Fudan University, 399 Wanyuan Rd, Minhang Qu, Shanghai Shi, China. [12]Department of Neonatology, Centrohospitalar Universitário de São João, Alameda Prof. Hernâni Monteiro, 4200-319 Porto, Portugal. [13]Department of Neonatology, University Hospital Motol, V Uvalu 84, 150 06 Prague 5, Czech Republic. [14]Department of Neonatology, University Children's Hospital Tuebingen, Hoppe-Seyler-Straße 1, 72076 Tuebingen, Germany. [15]Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico Milan, Via della Commenda 12, 20122 Milan, Italy. [16]Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy. [17]Department of Neonatology, Hospices Civil De Lyon, 3 Quai des Célestins, 69002 Lyon, France. [18]Department of Neonatology, Oslo University Hospital, Kirkeveien, 166 0450 Oslo, Norway. [19]Department of Neonatology, Poznan University of Medical Sciences, Polna 33, 60-535 Poznań, Poland. [20]Section of Biostatistics, Department of Publich Health, Copenhagen University, Øster Farimagsgade 5, Copenhagen K, Denmark. [21]Center for Statistical Science, Peking University, Beijing, China. [22]Department of Cardiology, Holbæk Hospital, Smedelundsgade 60, 4300 Holbæk, Denmark. [23]Department of Regional Health Research, The Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark.

Received: 2 May 2019 Accepted: 25 September 2019
Published online: 19 December 2019

References

1. Stoll BJ, Hansen NI, Bell EF, Shankaran S, Laptook AR, Walsh MC, et al. Neonatal outcomes of extremely preterm infants from the NICHD neonatal research network. Pediatrics. 2010;126:443–56.
2. Volpe JJ. Brain injury in the premature infant: overview of clinical aspects, neuropathology, and pathogenesis. Semin Pediatr Neurol. 1998;5:135–51.
3. Adams-Chapman I, Heyne RJ, DeMauro SB, Duncan AF, Hintz SR, Pappas A, et al. Neurodevelopmental impairment among extremely preterm infants in the neonatal research network. Pediatrics. 2018;141:e20173091.
4. Behrman R, Butler AS. Preterm birth: causes, consequences and prevention. Institute of Medicine. Washington, DC: National Academies Press; 2007.
5. Volpe JJ. Brain injury in premature infants: a complex amalgam of destructive and developmental disturbances. Lancet Neurol. 2009;8:110–24.
6. Hyttel-Sørensen S, Pellicer A, Alderliesten T, Austin T, Van Bel F, Benders M, et al. Cerebral near infrared spectroscopy oximetry in extremely preterm infants: phase II randomised clinical trial. BMJ. 2015;350:1–11.
7. Hyttel-Sørensen S, Greisen G, Als-Nielsen B, Gluud C. Cerebral near-infrared spectroscopy monitoring for prevention of brain injury in very preterm infants. Cochrane Database Syst Rev. 2017;9:CD011506.
8. Guideline on Good Clinical Practice E6(R2). London: European Medicines Agency, Committee for Human Medicinal Products. 2017.
9. World Health Organization. Preterm birth. https://www.who.int/news-room/fact-sheets/detail/preterm-birth. 2018. Accessed 24 Nov 2019.
10. Hansen ML, Pellicer A, Gluud C, Dempsey E, Mintzer J, Hyttel-Sørensen S, et al. Cerebral near-infrared spectroscopy monitoring versus treatment as usual for extremely preterm infants. A protocol for the SafeBoosC phase III randomized clinical trial. Manuscript accepted for publication in Trials. 2019
11. Pellicer A, Greisen G, Benders M, Claris O, Dempsey E, Fumagally M, et al. The SafeBoosC phase II randomised clinical trial: a treatment guideline for targeted near-infrared-derived cerebral tissue oxygenation versus standard treatment in extremely preterm infants. Neonatology. 2013;104:171–8.
12. Papile LA, Burstein J, Burstein R, Koffler H. Incidence and evolution of subependymal and intraventricular hemorrhage: a study of infants with birth weights less than 1,500 gm. J Pediatr. 1978;92:529–34.
13. Holsti A, Serenius F, Farooqi A. Impact of major neonatal morbidities on adolescents born at 23–25 weeks of gestation. Acta Paediatr. 2018;107:1893–901.
14. Bell MJ, Ternberg JL, Feigin RD, Keating JP, Marshall R, Barton L, et al. Neonatal necrotizing enterocolitis. Therapeutic decisions based upon clinical staging. Ann Surg. 1978;187:1–7.
15. Valls-I-Soler A, Carnielli V, Claris O, de la Cruz Bértolo J, Halliday HL, Hallman M, Hummler H, Weindling M; Scientific Steering Committee von EuroNeoStat (siehe Anhang). [EuroNeoStat: a European information system on the outcomes of care for very-low-birth-weight infants (< 1500 g)]. Z Geburtshilfe Neonatol. 2008;212(3):116-8. https://doi.org/10.1055/s-2008-1076802. German. PMID: 18709632.
16. Hintz SR, Barnes PD, Bulas D, Slovis TL, Finer NN, Wrage LA, et al. Neuroimaging and neurodevelopmental outcome in extremely preterm infants. Pediatrics. 2015;135:e32–42.
17. Schmidt B, Roberts RS, Davis PG, Doyle LW, Asztalos EV, Opie G, et al. Prediction of late death or disability at age 5 years using a count of 3 neonatal morbidities in very low birth weight infants. J Pediatr. 2015;167:982–986.e2.
18. Davidson L, Berkelhamer S. Bronchopulmonary dysplasia: chronic lung disease of infancy and long-term pulmonary outcomes. J Clin Med. 2017;6:4.
19. Vermont Oxford Network. Vermont Oxford Network. 2019. https://public.vtoxford.org/. Accessed 24 Nov 2019.
20. Jakobsen JC, Gluud C, Winkel P, Lange T, Wetterslev J. The thresholds for statistical and clinical significance – a five-step procedure for evaluation of intervention effects in randomised clinical trials. BMC Med Res Methodol. 2014;14:34.
21. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. Nature. 2019;567:305–7.
22. Zhang J, Quan H, Ng J, Stepanavage ME. Some statistical methods for multiple endpoints in clinical trials. Control Clin Trials. 1997;18:204–21.
23. Goodman SN. Introduction to Bayesian methods I: measuring the strength of evidence. Clin Trials J Soc Clin Trials. 2005;2:282–90.
24. Wasserstein RL, Schirm AL, Lazar NA. Moving to a world beyond "$p < 0.05$". Am Stat. 2019;73:1–19.
25. DeMets DL, Lan KKG. Interim analysis: the alpha spending function approach. Stat Med. 1994;13:1341–52.
26. Guyatt GH, Briel M, Glasziou P, Bassler D, Montori VM. Problems of stopping trials early. BMJ. 2012;344:e3863.
27. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. BMC Med Res Methodol. 2017;17:162.
28. Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, et al. The prevention and treatment of missing data in clinical trials. N Engl J Med. 2012;367:1355–60.
29. Groenwold RHH, Moons KGM, Vandenbroucke JP. Randomized trials with missing outcome data: how to analyze and what to report. Can Med Assoc J. 2014;186:1153–7.
30. Gates S, Brocklehurst P. How should randomised trials including multiple pregnancies be analysed? BJOG An Int J Obstet Gynaecol. 2004;111:213–9.
31. Yelland LN, Sullivan TR, Collins CT, Price DJ, McPhee AJ, Lee KJ. Accounting for twin births in sample size calculations for randomised trials. Paediatr Perinat Epidemiol. 2018;32:380–7.
32. Shaffer ML, Hiriote S. Analysis of time-to-event and duration outcomes in neonatal clinical trials with twin births. Contemp Clin Trials. 2009;30:150–4.
33. Kahan BC, Morris TP. Reporting and analysis of trials using stratified randomisation in leading medical journals: review and reanalysis. BMJ. 2012;345:e5840.
34. Kahan BC, Morris TP. Improper analysis of trials randomised using stratified blocks or minimisation. Stat Med. 2012;31:328–40.
35. Brennan KC. Accounting for centre-effects in multicentre trials with a binary outcome – when, why, and how? BMC Med Res Methodol. 2014;14:1–11.
36. Deeks JJ, Higgins JPT, Altman DG. Chapter 10: analysing data and undertaking meta-analyses. Draft version (29 January 2019) for inclusion. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. Cochrane handbook for systematic reviews of interventions. London: Cochrane. In Press.
37. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c869.
38. Li G, Taljaard M, Van den Heuvel ER, Levine MA, Cook DJ, Wells GA, et al. An introduction to multiplicity issues in clinical trials: the what, why, when and how. Int J Epidemiol. 2017;46:746–55.

39. Abdi H. The Bonferonni and Šidák corrections for multiple comparisons. In: Salkind N, editor. Encyclopedia of measurement and statistics. Thousand Oaks: Sage; 2007. p. 1–9.
40. Akobeng AK. Understanding type I and type II errors, statistical power and sample size. Acta Paediatr. 2016;105:605–60.
41. Yu Y, Zhang K, Zhang L, Zong H, Meng L, Han R. Cerebral near-infrared spectroscopy (NIRS) for perioperative monitoring of brain oxygenation in children and adults. Cochrane Database Syst Rev. 2018;1:CD010947.
42. Taillefer MC, Denault AY. Cerebral near-infrared spectroscopy in adult heart surgery: systematic review of its clinical efficacy. Can J Anaesth. 2005;52:79–87.
43. Rothwell PM. External validity of randomised controlled trials: "To whom do the results of this trial apply?". Lancet. 2005;365:82–93.
44. Ward RM, Beachy JC. Neonatal complications following preterm birth. BJOG An Int J Obstet Gynaecol. 2003;110:8–16.
45. Savović J, Turner RM, Mawdsley D, Jones HE, Beynon R, Higgins JPT, et al. Association between risk-of-bias assessments and results of randomized trials in Cochrane Reviews: the ROBES Meta-Epidemiologic Study. Am J Epidemiol. 2018;187:1113–22.
46. Hrobjartsson A, Thomsen ASS, Emanuelsson F, Tendal B, Hilden J, Boutron I, et al. Observer bias in randomized clinical trials with measurement scale outcomes: a systematic review of trials with both blinded and nonblinded assessors. Can Med Assoc J. 2013;185:E201–11.
47. Hróbjartsson A, Thomsen ASS, Emanuelsson F, Tendal B, Rasmussen JV, Hilden J, et al. Observer bias in randomized clinical trials with time-to-event outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. Int J Epidemiol. 2014;43:937–48.
48. Hróbjartsson A, Emanuelsson F, Skou Thomsen AS, Hilden J, Brorson S. Bias due to lack of patient blinding in clinical trials. A systematic review of trials randomizing patients to blind and nonblind sub-studies. Int J Epidemiol. 2014;43:1272–83.
49. EMA (European Medicines Agency). Guideline on multiplicity issues in clinical trials. 2017.

Publisher's Note

Hansen et al. Trials (2020) 21:356
https://doi.org/10.1186/s13063-020-4206-6

Trials

RESEARCH Open Access

Pilot test of an online training module on near-infrared spectroscopy monitoring for the randomised clinical trial SafeBoosC-III

Mathias Lühr Hansen[1*], Marie Isabel Rasmussen[1], Snorre Rubin[2], Adelina Pellicer[3], Guoqiang Cheng[4], Xin Xu[5], Yin Zhaoqing[6], Vibeke Zoffmann[7] and Gorm Greisen[1]

Abstract

Background: SafeBoosC-III is an international randomised clinical trial to evaluate the effect of treatment of extremely preterm infants during the first 3 days of life based on cerebral near-infrared spectroscopy (NIRS) monitoring versus treatment and monitoring as usual. To ensure high quality of the trial intervention as well as of patient care, we have developed a multilingual web-based training program to train relevant staff and test their competence. As we enter an under-explored area of e-learning, we have conducted a pilot study on the first of the five modules comprising the web-based training program to test the feasibility of developing such a program for an international trial with limited resources.

Methods: The module in this study focuses on the principles and practice of NIRS monitoring. The pedagogical idea was to integrate training and certification. One-hundred doctors and nurses from five Neonatal Intensive Care Units across China, Spain and Denmark were invited to participate in the pilot study. Upon completion of the NIRS module, participants were invited to evaluate their experience by completing an online survey. Data from closed-ended questions were analysed using descriptive statistics while data from open-ended questions underwent thematic analysis.

Results: In total, 81 of 100 invited staff members entered the training module and completed the online survey. The median time and the number of questions to pass the module was 15 minutes and seven questions, respectively. Most staff found the academic level of the learning material and quiz appropriate (85% and 93% of all staff members, respectively), as well as agreeing that the module was relevant to prepare them to 'use the NIRS device' (90%). Thematic analysis revealed issues such as a discrepancy between learning material and quiz questions, lack of clarity, and technical issues.

Conclusion: We provide evidence of the feasibility of developing a multilingual web-based training program for an international trial, despite challenges such as low budget, language barriers and possibly differences in the clinical training of staff. Exploring the integration of training and certification for international trials, the positive results of this study motivate further developments.

(Continued on next page)

* Correspondence: mathias.luhr.hansen@gmail.com
[1]Department of Neonatology, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark
Full list of author information is available at the end of the article

(Continued from previous page)
Trial registration: ClinicalTrial.gov, NCT03770741. Registered 10 December 2018.

Keywords: SafeBoosC, Randomised clinical trial, Randomized clinical trial, RCT, Extremely preterm, Near-infrared spectroscopy, NIRS, Online training, Web-based training, E-learning

Background

Randomised clinical trials are considered the highest level of evidence when evaluating the effects of a clinical intervention [1]. It is therefore essential that the methodological quality is high. Furthermore, since randomised clinical trials are conducted on human subjects [2], the safety and well-being of participants are of crucial importance. Good Clinical Practice (GCP) is an international standard for designing, conducting, recording and reporting clinical trials involving human subjects, with the purpose of ensuring the safety and well-being of trial subjects as well as high scientific quality [3]. A core principle in GCP is that staff members involved in the trial "should be qualified by education, training and experience to perform his or her respective task(s)" [3]. One way to ensure this is by training the clinical staff [4]. Despite evidence suggesting that training staff members in trial-related tasks has a positive effect on the trial's results [5], the training process is rarely reported [6]. Furthermore, recommendations for specific training requirements for clinical trials are not defined in the standards on GCP by the International Committee on Harmonisation.

To recruit enough participants, large-scale clinical trials often include many centres across multiple countries. This poses the problem of training staff since on-site training is expensive, time-demanding and difficult to standardise. A way to bypass this issue, while preserving the quality of training, is by using e-learning. E-learning is a broad concept describing education facilitated through electronic systems, such as computers or mobile devices [7], and, as such, can be used to ensure standardised delivery of subject matter [8].

E-learning has already proven to be a valuable asset when increasing the competencies of health professionals, in both industrial and developing countries [8–13], and has been proliferating until now with the purpose of medical education at universities, as a counteraction to traditional classroom teachings [14]. It has proven to be a useful tool when harmonizing teachings that are aimed worldwide, across different languages and clinical settings, and has been used with great progress in resource-constrained countries [15].

A recent Cochrane review of e-learning, which it defined as any educational intervention mediated electronically via the internet, was found non-inferior to traditional classroom teaching [8, 16], and in the few published reports on e-learning as preparation for clinical trials, it has been implemented with success [4].

A consensus on a clear definition of e-learning does not exist. Therefore, multiple terms are used as synonyms for e-learning, including internet-based learning, web-based learning and training, computer-assisted instructions and computer-based learning and training [8, 17–19]. For the purpose of this study, we will use the term 'web-based training' (see the 'Web-based training and certification program' section).

The SafeBoosC-III trial

SafeBoosC-III is a randomised clinical trial investigating the benefits and harms of treatment based on cerebral near-infrared spectroscopy (NIRS). The hypothesis is that treatment based on NIRS monitoring during the first 72 h of life of extremely preterm infants will result in a reduction of severe brain injury and death at 36 weeks postmenstrual age. Sixteen-hundred infants born with a gestational age below 28 weeks and admitted to more than 50 neonatal intensive care units across 20 different countries will be randomised. Infants in the experimental group will receive treatment guided by cerebral NIRS monitoring during the first 72 h of life, while infants in the control group will receive treatment and monitoring as usual. The protocol of the SafeBoosC-III trial is registered at www.ClinicalTrials.gov, NCT 03770741 (07.12.2018).

When working at a clinical department, it is often expected that you are familiar with routine practices. If you are not, learning will often happen through supervision by more experienced colleagues familiar with the interventions. However, when an intervention trial is rolled out in a clinical department, only a few staff members may be familiar with the intervention. Thus, training in trial-related procedures is necessary, not only for the safety of trial participants but also to give a relevant and practical estimate on the effect of the intervention in routine practice. This is done through a pragmatic trial such as SafeBooSC-III, where the purpose is to test the effect of a given intervention in a real-world setting [20], i.e. what effect can be expected by implementing a specific intervention in a broad patient group, in a large number of departments. Therefore, in order to estimate the potential effect of implementing an intervention in routine practice, staff on-site should be trained in the intervention before the trial takes off [3, 20]. However, if

the trial ought to reflect how the intervention works in a real-world scenario, the level and intensity of staff training should reflect this, meaning that you do not want to train your staff members to an expert level prior to trial initiation since this does not reflect a 'real-world' scenario and you will not thus get generalizable trial results. To provide a practically realistic level of introduction and training for the SafeBoosC-III trial, we have developed a multilingual online training program to train relevant staff and test their competence.

As we enter an under-explored area of e-learning, we have conducted a pilot study on the first of the five modules comprising the web-based training program to test the feasibility of developing such a program on limited resources for an international trial. We expected that results from this pilot study could be used to enhance and support further development of the web-based training and certification program for SafeBoosC-III, and possibly encourage the use of e-learning when implementing future international clinical trials.

Methods

Fifty nurses and 50 doctors from a total of five neonatal intensive care units across China, Denmark and Spain were invited to participate in the pilot study. In order to ensure that the e-learning tool was appropriate for staff of all levels of experience, the responsible investigators within each of the five participating neonatal intensive care units (AP, GC, MLH, XX, ZY) invited staff members with and without prior NIRS experience to participate. All neonatal intensive care units participating in this pilot study are planning to participate in SafeBoosC-III, and all participating staff members are expected to care for babies enrolled in SafeBoosC-III.

Participants were asked to 1) complete the web-based training module on NIRS monitoring and 2) evaluate it through an online survey.

Web-based training and certification program

The training module is part of a complete web-based training and certification program for the SafeBoosC-III trial, which will be offered to all doctors and nurses involved in the care of trial participants. It consists of five separate modules covering 1) introduction to SafeBoosC-III and the protocol, 2) cerebral NIRS monitoring, 3) SafeBoosC-III treatment guideline, 4) cranial ultrasound imaging and diagnosing of brain injury, and 5) GCP monitoring in SafeBoosC-III. All modules are designed as integrated training and certification modules, with each module consisting of a) learning material and b) a quiz. With the exception of the introduction module, all modules are built over a simplified adaptive framework, meaning that you are led directly to the quiz and will only be prompted to visit the learning materials if your answers to questions are wrong. If you answer all questions correctly, you have shown mastery of the subject matter and will be certified directly. As such the quiz is designed to recognize prior learning, as correct answers will get participants through the modules faster. For less experienced users, the option is given for the user to bypass the quiz and go directly to the learning materials first (Fig. 1).

The content for all modules was developed with the same approach: Initially, a narrative text covering all essential knowledge on the subject was drafted. Based on the narrative, a number of learning objectives were developed, all clearly described according to Bloom's Taxonomy's cognitive domain [21] to specify which degree of mastery the user should show. The narrative was also used to write the learning material for each module. Next, two to four questions related to each learning objective were developed, thereby representing a pool of questions used to build the quiz. The questions strive to be as relevant as possible and were therefore formulated in a case-like manner, with a short description of a clinical situation given for each question, followed by a varying number of response options. Cases reflect clinical situations that could happen during the conduct of SafeBoosC-III. Often, there are several correct response options constituting the format of multiple-choice questions with several-of-many answers. To pass a question, all the correct answers and only correct answers must be ticked (Fig. 2a). Participants complete a module when they have answered one question per learning objective correctly. We could have chosen to require two (or three) correct answers per learning objective, but this would have inflated the volume of questions—and hence the costs—as well as the time to be used by participants. They will be exposed to new questions or re-exposed to questions they have already met from the quiz-pool on a continuous basis until the above criterion is met. Questions in the learning objectives that at any given time are not yet passed are presented in random order. This is done to reduce the risk that participants adopt a fast game, like a 'trial and error' strategy, rather than learning and understanding.

The teaching methodology is case-based and uses immediate detailed feedback, which means that participants will be presented with explanations for right and wrong answers as they go (Fig. 2b). This method has been shown to increase student performance in previous online medical education programs [22].

The complete web-based training and certification program will be hosted in a Moodle virtual learning environment (Moodle Pty Ltd, West Perth, WA, Australia), a commonly used shareware software within online medical training [10, 11, 23]. However, Moodle was not used as the platform for this pilot project due to restricted time. Instead, a direct link with immediate access to the module was used. The platform used in this pilot study was the

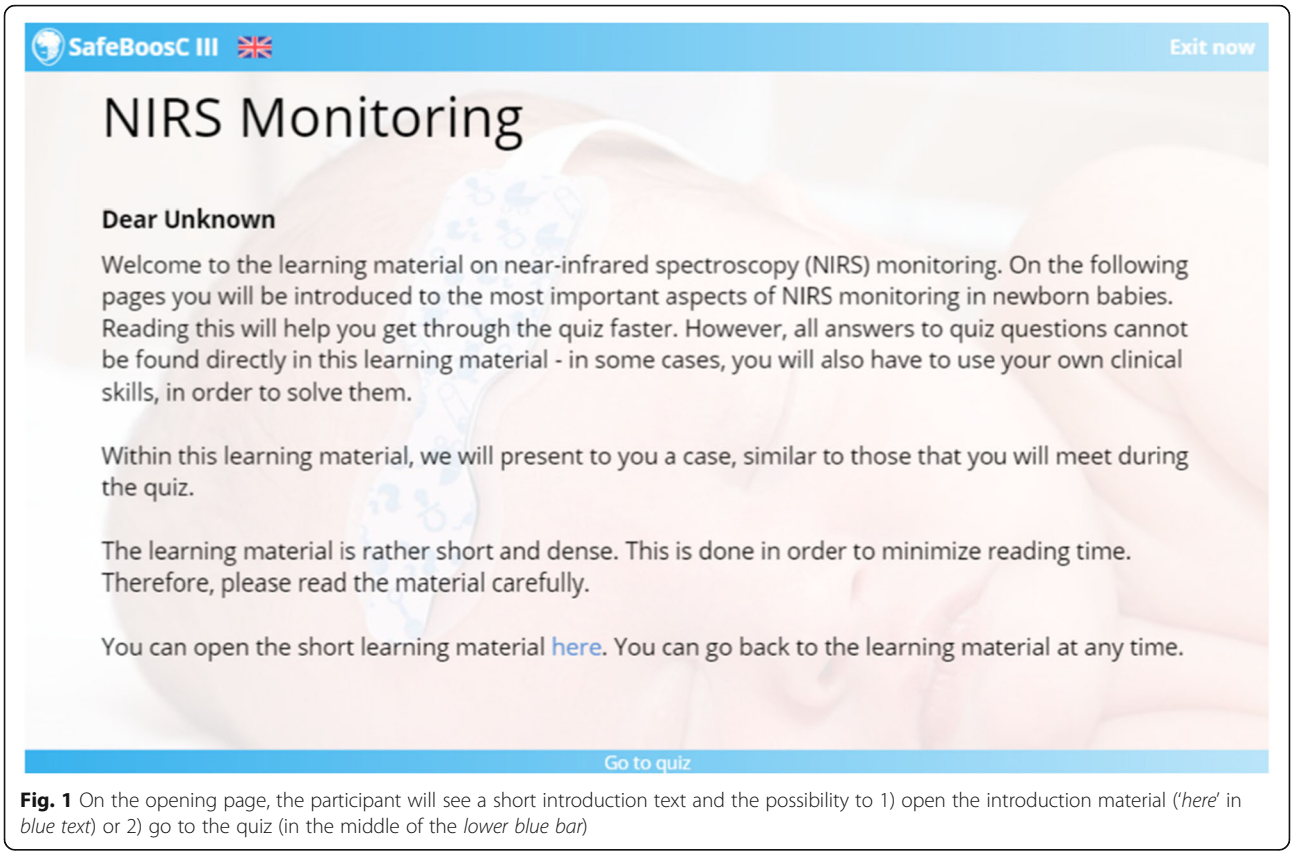


Fig. 1 On the opening page, the participant will see a short introduction text and the possibility to 1) open the introduction material ('*here*' in *blue text*) or 2) go to the quiz (in the middle of the *lower blue bar*)

Capital Region of Denmark's primary platform for e-learning, providing almost 400 different training programs for 40,000 staff members (kursusportalen.plan2learn.dk).

Training module on NIRS

The module on NIRS monitoring which was piloted in this study focuses on the principles of measuring cerebral oxygenation by NIRS, basic device operation, application and fixation of the sensor to the head of the infant, care of sensor and repositioning, the risk of skin marks, interpretation of measured values and the concept of venous-weighted tissue blood oxygenation. It consists of four learning objectives and 11 questions (Table 1). SHS and GG [24–28] took the lead in writing the learning material as well as questions for the quiz. MLH also participated in this process. MIR and SR programmed the training module in the interactive e-learning software Articulate Storyline (Articulate, New York, NY, US) and provided a direct URL link for participants to use. SR is an employee of the Copenhagen University Hospital e-learning section and is mainly responsible for programming all modules for the web-based training and certification program.

Since SafeBoosC-III is a multinational trial, language barriers can pose as a challenge because the content of the web-based training program must be translated to all the languages and still hold an academic level which meets clinical standards without the translation process being too complex. Therefore, in order to test the feasibility of translating the content of the web-based training program and train staff members in local languages, the original English version was translated into both Spanish and Chinese. The translation was done locally by the national coordinators AP in Spain and GC in China, who conducted manual translation of the material from English to Spanish and Chinese, respectively. In China, GC conducted the translation with the aid of the online translation tool 'youdao.com' whenever he was in doubt of the correct translation. Due to limited resources, the quality and precision of the translations were not evaluated by external linguistic experts and no back-translation and comparison with the originals were conducted. Danish participants were trained in the English module and Chinese and Spanish participants in the Chinese and Spanish modules, respectively.

Survey

Upon completion of the NIRS module, participants were asked to evaluate their experience by completing an online survey. For Spanish and Danish participants, the online survey was hosted in Google Analytics (Google LLC, Mountain View, CA, USA), but since Google is blocked in China, a Chinese survey program, Wenjuan (Shanghai

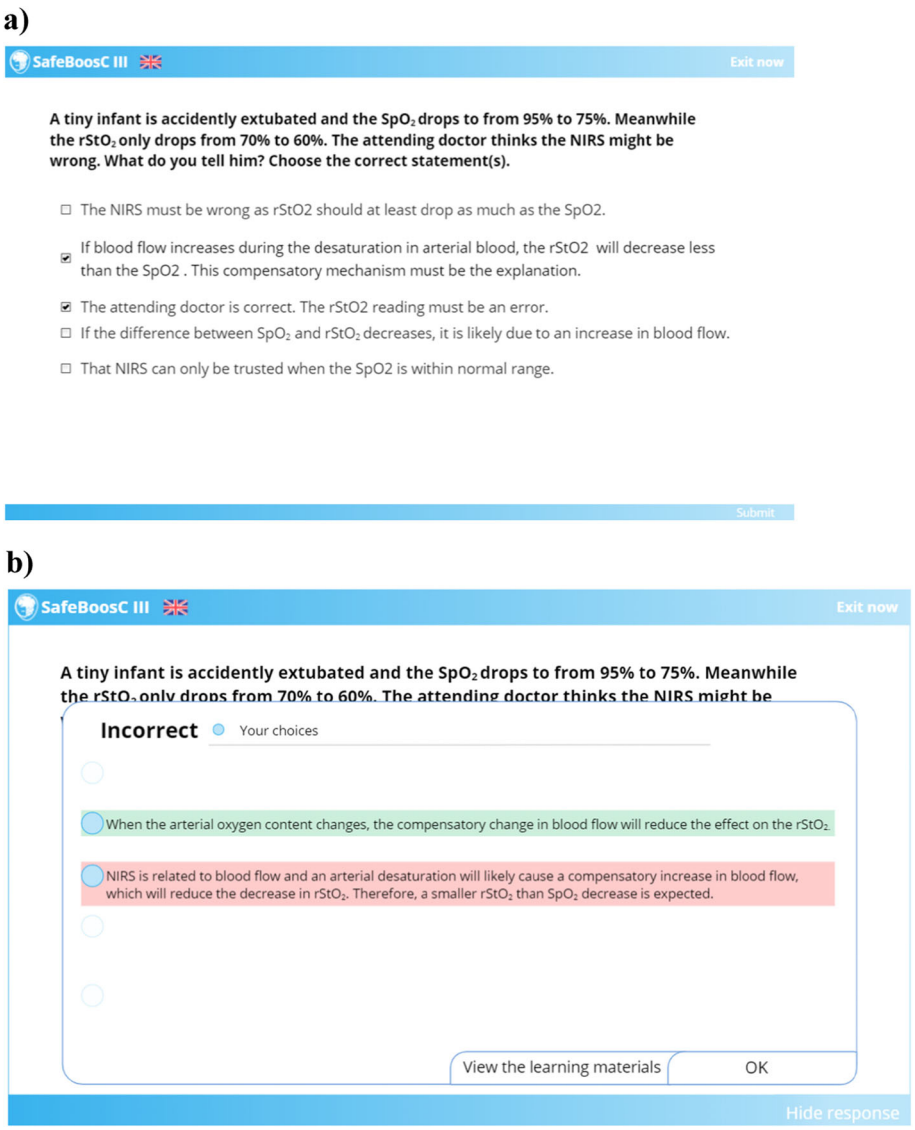


Fig. 2 Example of a question from the NIRS module. **a** The case-text describing the clinical setting and the five answer possibilities; options two and three have been chosen. **b** Explanations to answers are presented; option two was correct, while option 3 was wrong. This means that the question is not passed and the participant will be presented with another question on the same learning objective and have to answer that correctly before completing the quiz

Zhongyan Network, Shanghai, China), was used to host the Chinese survey. Participants completed the survey in local languages. As for the web-based training module, the translation of the survey was done locally by AP and GC. The online survey consisted of 15 closed-ended questions with answers on a three- or four-step Likert scale and seven open-ended questions with free-text answers. The structure and content of the closed-ended questions are based on Wang's principles for e-learner satisfaction [29]. Open-ended questions were added to gain a deeper and more complex understanding of participants' experiences and to clarify potential room for improvement. The survey covered the following themes: 1) performance, 2) learning material, 3) quiz material, 4) interface, and 5) preparation to use NIRS monitoring in a clinical context. MIR and MLH developed the online survey.

Data analysis

Quantitative data from closed-ended questions were analysed using basic descriptive statistics. Analysis of answers to the open-ended questions followed the principles of thematic analysis as described by Braun and Clarke [30]. With an inductive and data-driven approach, an iterative six-step analysis was conducted to

Table 1 Learning objectives and questions for the training module on NIRS monitoring

Learning objective	Question
Point out differences between NIRS tissue oxygenation ($rStO_2$) and pulse oximetry	A father of a very preterm infant asks why the cerebral $rStO_2$ is 65 when the SpO_2 is 94. What do you tell him?
	A baby is pale and mottled and you suspect circulatory failure due to septic shock. You have a hard time getting a signal from the pulse oximeter, but the NIRS gives readings with no apparent problems. Choose the correct statement(s)
Recognise the consequences of rStO2 being a direct measure of cerebral oxygen consumption/ supply balance and indirect measure of cardiac output	A tiny infant is accidentally extubated and the SpO_2 drops from 95% to 75%. Meanwhile, the $rStO_2$ only drops from 70% to 60%. The attending doctor thinks the NIRS might be wrong. What do you tell him?
	A colleague asks for help to understand what rStO2 really measures. Which of the following statements would you include in your explanation
	You care for a baby in the experimental group on the first day of life. Everything has been stable when the rStO2 alarm goes off and shows cerebral hypoxia. No other monitors are sounding an alarm. The ventilator runs normally, the SpO2 is stable around 92% and the mean arterial blood pressure is stable around 28 mmHg. As you look into the incubator to check the cerebral oximeter sensor you see that he has been bleeding from the umbilicus. It is a large spot on the linen and in the diaper and you estimate that the volume may be 10 ml. Could that be the explanation for the cerebral hypoxia?
Know the elements in starting up NIRS monitoring and interpret values during monitoring	You move the sensor to the other side of the forehead of a sick preterm infant as part of routine care. The parents notice that $rStO_2$ is about 7 percentage points higher in the new position. What answers can you give them?
	$rStO_2$ drops suddenly to 40%. What would you do? Please prioritise the following actions from first to last
	You have to start up monitoring cerebral oxygenation. Which of the following actions would you **not** do?
Know the side effects of NIRS monitoring	The parents ask if there are side effects to the near-infrared light used by the oximeter. Choose which statements you may include in your explanation
	You take over the care of a baby in the SafeBoosC trial. He is in the experimental group. Gestational age is 24 weeks and he is mechanically ventilated with high pressures and on high dose pressor (dopamine 15 microgram/ kg/min) and yet the mean arterial blood pressure is only 24 mmHg. The situation, however, has been stable for the last 12 h. The cerebral oximeter seems to work well and the rStO2 is 65% (the hypoxic threshold of your oximeter is 58%). Choose what you will do?
	A mother notices a minor mark on the skin after you have moved the pulse oximeter sensor to another position. She is now concerned about the NIRS sensor as well. What answers can you give her?

identify themes across the entire data set [30]. Initially, answers were systematically reviewed and coded. In total, 111 answers were coded into 70 codes, which subsequently were narrowed into 64 codes, based on their similarities (Table 2).

All 64 codes were collated and grouped into seven candidate themes. In order to get a better overview of data, candidate themes were illustrated in mind maps and reviewed in relation to 1) their specific data extracts and 2) across the entire data set. In this process, themes that covered similar aspects were merged, and irrelevant themes were either deleted or re-assembled, which resulted in four final themes.

Ethics approval and consent to participate

According to Danish, Chinese and Spanish laws, survey studies are not considered biomedical research and ethics approval was not therefore required to conduct this study. An information sheet written by MLH and GG explaining the purpose of the pilot project, that no personal data were collected and that all survey answers were recorded and analysed anonymously was distributed to the responsible investigators in each of the five participating neonatal intensive care units (AP, GC, MLH, XX and YZ). The five investigators invited relevant staff members to participate in the pilot study and, based on the information sheet, informed them of the study and data handling. All staff members had the possibility to ask the responsible investigators questions on the study and the possibility to decline participation in this pilot study. Since no personal identifiers were registered on participants, it was impossible to identify the identity of individual survey responders and thus

Table 2 Examples of data extract coding. Narrative to the left and codes to the right

Fine academic level, but some of the questions did not match the introduction material, which was a shame and frustrating (in relation to agreeing/strongly agree that the academic level of the quiz was appropriate)	1. Discrepancy between introduction material and quiz 2. Frustrations 3. Academic level appropriate 4. Introduction material insufficient to answer quiz questions
The question is not related to the learning material. The language is not enough concise and clear	1. Discrepancy between the introduction material and quiz 2. Unprecise language

withdrawal of data was not possible. This design was chosen to protect the participants against employers and responsible investigators back-tracking the performance or survey completion of individuals.

Results

In total, 81 of 100 invited staff members (81%) entered the training module and completed the online survey. Fifty (62%), 16 (20%) and 15 (18%) staff members responded from China, Spain and Denmark, respectively. Of the 81 responders, 41 were doctors (51%) and 40 were nurses (49%). Previous experience with NIRS monitoring was reported by 46 of the 81 responders (57%), including 26 doctors (57%) and 20 nurses (43%) (Table 5). In Denmark, six of the 15 responders had previous experience (40%), in China 25 of 50 (50%), and in Spain 15 of 16 (94%).

Closed-ended questions

Performance

Overall, responders spent a median time of 15 min (range 1 to 420 min) and a median number of seven questions (range 4 to 50 questions) to complete the NIRS module. Spanish responders were faster than both Danish and Chinese (median 10, 14 and 20 min, respectively) and used fewer questions to pass (median 4, 7 and 8, respectively) (Table 3). Doctors were faster than nurses (median 13.5 versus 20 min) and used fewer questions to pass (median 6 versus 9 questions) (Table 4). Responders with NIRS experience were faster than non-experienced (median 13.5 min versus 20 min) and spent fewer questions to pass (median 5.5 versus 8 questions) (Table 5).

Learning material

Overall, 69 of 81 (85%) responders found the academic level of the learning material appropriate and none found it too easy. Of the 12 responders who found the learning material

Table 3 Time used and number of quiz questions used to complete the module and number of responding participants who answered either 'agree'/'strongly agree' or 'appropriate' to the questions regarding the design of the module (data stratified by country)

Question	Denmark	Spain	China	Total
Performance				
Minutes to complete module, median [range]	14 [7–30] (11/15)[c]	10 [1–60] (13/16)[c]	20 [2–420] (46/50)[c]	15 [1–420] (70/81)[c]
Number of questions to complete module median [range]	7 [6–20] (5/15)[c]	4 [4–12] (13/16)[c]	8 [4–50] (43/50)[c]	7 [4–50] (61/81)[c]
Learning material				
Academic level of learning material appropriate, n/N (%)	14/15 (93)	15/16 (94)	40/50 (80)	69/81 (85)
Learning material sufficient to complete quiz[a], n/N (%)	3/12 (25)	13/16 (81)	39/50 (78)	55/78 (70)
Quiz				
Academic level of quiz appropriate[a], n/N (%)	14/15 (93)	15/16 (94)	46/50 (92)	75/81 (93)
Number of answering possibilities per question appropriate, n/N (%)	6/15 (40)	9/16 (56)	34/50 (68)	49/81 (60)
Quiz questions clinically relevant and up-to-date[a]	13/14 (93)	15/16 94)	49/50 (98)	77/80 (96)
Interface				
The NIRS module was stable and did not crash[b], n/N (%)	6/15 (40)	9/15 (60)	42/50 (84)	57/80 (71)
Preparation for using NIRS				
Relevant to prepare for using the NIRS device[a]	13/15 (87)	12/15 (80)	47/50 (94)	72/80 (90)

[a]Pooling of the answers agree or strongly agree
[b]Yes to the statement
[c] Number of responders answering the specific question and the total number of overall responders completing the online survey

Table 4 Time used and number of quiz questions used to complete the module and number of participants who answered either 'agree'/'strongly agree' or 'appropriate' to the questions regarding the design of the module (data stratified by participants' profession)

Question	Doctors	Nurses	Total
Performance			
Minutes to complete module, median [range]	13.5 [2–420] (34/41)[c]	20 [1–420] (36/40)[c]	15 [1–420] (70/81)[c]
Number of questions to complete module, median [range]	6 [4–30] (28/41)[c]	9 [4–50] (33/40)[c]	7 [4–50] (61/81)[c]
Learning material			
Academic level of learning material appropriate, n/N (%)	37/41 (90)	32/40 (80)	69/81 (85)
Learning material sufficient to complete quiz[a], n/N (%)	28/38 (74)	27/40 (68)	55/78 (70)
Quiz			
Academic level of quiz appropriate[a], n/N (%)	38/41 (93)	37/40 (93)	75/81 (93)
Number of answering possibilities per question appropriate, n/N (%)	29/41 (71)	20/40 (50)	49/81 (60)
Quiz questions clinically relevant and up-to-date[a]	39/40 (98)	38/40 (95)	77/80 (96)
Interface			
The NIRS module was stable and did not crash[b], n/N (%)	29/41 (71)	28/39 (72)	57/80 (71)
Preparation for using NIRS			
Relevant to prepare for using the NIRS device[a]	35/40 (88)	37/40 (93)	72/80 (90)

[a]Pooling of the answers agree or strongly agree
[b]Yes to the statement
[c] Number of responders answering the specific question and the total number of overall responders completing the online survey

Table 5 Time used and number of quiz questions used to complete the module and number of participants who answered either 'agree'/'strongly agree' or 'appropriate' to the questions regarding the design of the module (data stratified by participants' previous experience with NIRS monitoring)

Question	Experience	No experience	Total
Performance			
Minutes to complete module, median [range] (n/N)	13.5 [1–420] (40/46)[c]	20 [4–420] (30/35)[c]	15 [1–420] (70/81)[c]
Number of questions to complete module, median [range] (n/N)	5.5 [4–20] (36/46)[c]	8 [4–50] (25/35)[c]	7 [4–50] (61/81)[c]
Learning material			
Academic level of learning material appropriate, n/N (%)	39/46 (85)	30/35 (86)	69/81 (85)
Learning material sufficient to complete quiz[a], n/N (%)	32/44 (73)	23/34 (68)	55/78 (70)
Quiz			
Academic level of quiz appropriate[a], n/N (%)	41/46 (89)	34/35 (97)	75/81 (93)
Number of answering possibilities per question appropriate, n/N (%)	27/46 (59)	22/35 (63)	49/81 (60)
Quiz questions clinically relevant and up-to-date[a]	43/46 (93)	34/34 (100)	77/80 (96)
Interface			
The NIRS module was stable and did not crash[b], n/N (%)	34/45 (76)	23/35 (66)	57/80 (71)
Preparation for using NIRS			
Relevant to prepare for using the NIRS device[a]	41/45 (91)	31/35 (89)	72/80 (90)

[a]Pooling of the answers agree or strongly agree
[b]Yes to the statement
[c] Number of responders answering the specific question and the total number of overall responders completing the online survey

too advanced, ten were from China, one was from Denmark and one was from Spain (Table 3). Eight of 40 (20%) nurses found the learning material too advanced compared to four of 41 (10%) doctors (Table 4). Additionally, no relevant difference was seen between responders experienced in NIRS monitoring and those with no experience (seven of 46 (15%) experienced versus five of 35 non-experienced (14%)) (Table 5). When asked if the introduction material was sufficient to answer quiz questions, 23 of 78 (29%) responders disagreed or strongly disagreed, nine from Denmark, 11 from China and three from Spain (Table 3). More nurses (13 of 41 (32%)) than doctors (10 of 38 (26%)) disagreed or strongly disagreed with this statement (Table 4). Amongst those with NIRS experience, 12 of 44 (27%) disagreed or strongly disagreed compared to 11 of 34 (32%) responders with no previous experience (Table 5).

Quiz

Seventy-five of 81 (93%) responders agreed or strongly agreed that the academic level of questions was appropriate. Of those who disagreed or strongly disagreed, no relevant difference was found between countries (Table 3) or clinical positions (Table 4). Of the six disagreeing or strongly disagreeing on the statement, five were experienced in NIRS monitoring (Table 5). Thirty-two of 81 (40%) responders thought there were too many answer possibilities for each question, primarily nurses (20 of 40 (50%) nurses compared to 12 of 41 (29%) doctors) (Table 4) and Danish responders (nine of 15 (60%) compared to seven of 16 (43%) Spanish and 16 of 50 (32%) Chinese responders) (Table 3). Among those with NIRS

experience, 19 of 46 (41%) thought there were too many answer possibilities compared to 13 of 35 (37%) with no experience (Table 5). When asked if the quiz questions were clinically relevant and up-to-date, 77 of 80 (96%) responders agreed or strongly agreed on this.

Interface

Almost one-third of all responders (23 of 80 (29%)) experienced a crash once or multiple times while accessing the NIRS module . It seemed that the problem was greatest in Denmark and Spain where nine of 15 (60%) and six of 15 (40%) reported experiencing a crash, compared to only 8 of 50 (16%) in China (Table 3). Among doctors and nurses, 12 of 41 (29%) and 11 of 39 (28%) experienced a crash, respectively (Table 4). Eleven of 45 (24%) experienced with NIRS and 12 of 35 (34%) non-experienced responders reported a crash (Table 5).

Preparation for using NIRS

When asked if the module was relevant to prepare staff members to use the NIRS device, 72 of 80 (90%) agreed or strongly agreed on this (13 of 15 (87%) Danish, 12 of 15 (80%) Spanish and 47 of 50 (94%) Chinese responders) (Table 3). No relevant difference was seen between clinical positions (35 of 40 (88%) doctors and 37 of 40 nurses (93%)) or between experience levels (41 of 45 (91%) experienced and 31 of 35 (86%) non-experienced) (Tables 4 and 5).

Open-ended questions

The thematic analysis resulted in four essential themes, accompanied by sub-themes (Fig. 3). The themes were 1) learning material-quiz discrepancies, 2) lack of clarity within course, 3) technical issues and 4) unsolicited positive comments. These four themes elicit key concepts that are essential throughout the data.

Learning material—quiz discrepancies

Some responders ($n = 18$) described a discrepancy between the learning material and the quiz, with several stating that the learning material was insufficient to adequately answer the questions in the quiz:

> "*For someone who know [s] little or nothing about the topic, the introduction material is not sufficient enough to answer the quiz questions*" Doctor

One responder stated that despite being committed and working hard to gain a comprehension of the learning material, they struggled with answering the questions correctly and finishing the course:

> "*Put in a great effort to understand the intro material and I was surprised that I could not answer questions correctly. I did not feel that there was a connection between theory in the introduction material and questions*" Nurse

A few mentioned ($n = 6$) that the learning material was too simple or not detailed enough and was lacking comprehensiveness:

> "*Additional knowledge is needed in the principles and concepts section*" Nurse

As a possible consequence of this discrepancy, some responders ($n = 10$) also expressed that the content of the course was too hard:

> "*The content is too hard to understand*" Nurse
> "*The questions are difficult, and the basic courses are few*" Doctor

Some responders ($n = 23$) also stated that specific clinical content was missing in the learning material, which made it difficult to complete the quiz. A specific concern raised ($n = 14$) was the absence of knowledge regarding the practicality behind the usage and handling of the NIRS device:

> "*Risk of skin marks and side-effects is not described sufficiently in the introduction material*" Doctor

The lack of clinical content left a few responders ($n = 4$) feeling unequipped for answering questions in relation to this:

> "*No introduction to how you prepared for NIRS monitoring, so it was pure guessing—you have no idea whether you need to calibrate/shave/wash or something else (prior monitoring), if you have not been told forehand*" Nurse

Lack of clarity within the course

Language issues were mentioned (n = 6), including that the language was not precise and clear, which made it hard to understand the context of the course. This was voiced by Spanish ($n = 1$) and Chinese ($n = 5$) participants:

> " … *The language is not enough concise and clear*" Doctor

The transparency of the module's structure was also criticised, with a few responders ($n = 7$) stating that the feedback mechanism was hard to figure out:

> "*[The module] did not tell me what my wrong answers I had, and therefore I didn't know what the correct answers were and I couldn't find it in the introduction material*" Nurse

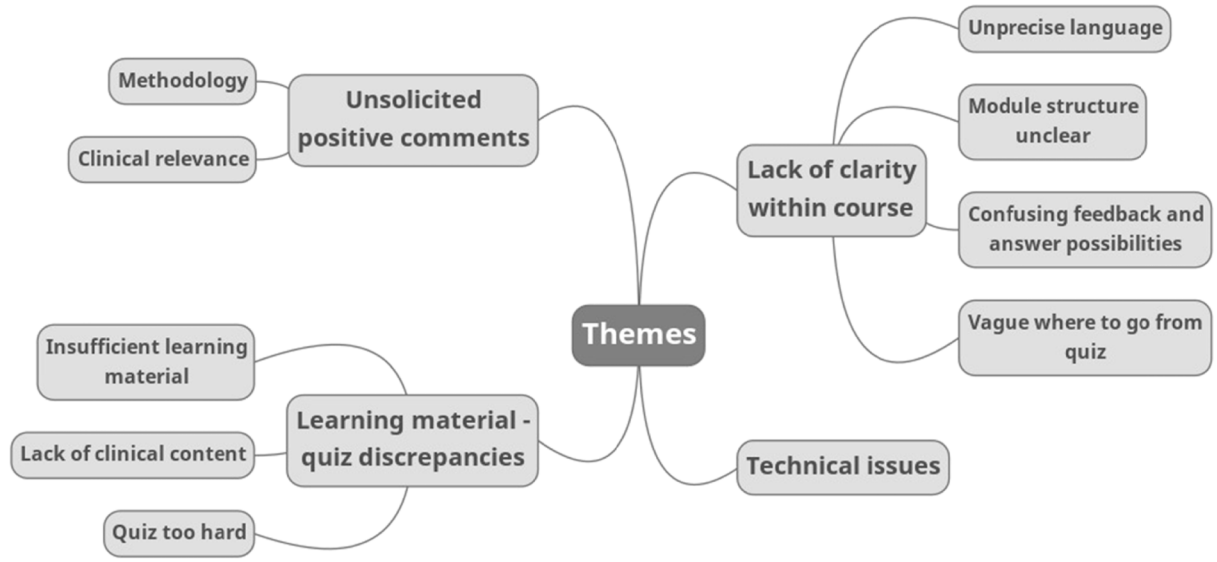


Fig. 3 Inferred themes and sub-themes

In this event, one mentioned that it was hard to learn something from answering incorrectly:

> "*It would be nice if one could learn something by answering wrong, hence that you could use the box that pops up after you answer incorrectly to see what was the correct answer.*" *Nurse*

In specific regards to the module lacking clarity, the deficient explanation of the quiz set-up was described. One responder expressed that it should be stated more explicitly how the module was structured:

> "*Very good, but I was not prepared for a case-setup—and many answer possibilities were not mentioned in the introduction material*" *Nurse*

A few ($n = 4$) respondents stated that having multiple answer possibilities was an issue:

> "*I think the quality of learning is increased if there are more questions with fewer answer possibilities. The purpose is learning and I think this could be heightened if one is presented with more questions with lesser answer possibilities … .*" *Doctor*

Technical issues

Technical issues seemed to be a source of frustration in this course. Responders answered that the module entered into a loop of incorrect questions ($n = 4$), that it crashed ($n = 8$), that the speed was slow ($n = 5$), and that the screen froze ($n = 12$), with one responder describing how it froze three to four times in a row, which caused this person to restart and begin all over again:

> "*If you do it, you will be stuck, you can not finish it, what the hell*" *Doctor*

> "*The page hangs on some occasions and does not allow to advance. When there is an incorrect answer, it loops in and you must restart the questionnaire to get out of there*" *Doctor*

The accessibility also seemed to be a problem. A few ($n = 4$) experienced that they could not easily navigate between the quiz and learning material without losing answers or facing a module crash, which in some cases led to a failure to finish the quiz:

> "*Problems when some question is incorrect: it does not allow one to advance, in spite of reviewing the material and you must leave the page*" *Doctor*

Unsolicited positive comments

Despite the open-ended questions being focused on clarifying any critique points of the module as well as potential improvements, some responders ($n = 12$) also commented on the positive aspects of the module. Some applauded the clinical relevancy and fitness for clinical use:

"*Suitable for application of clinical*" *Doctor*

Others were positive towards the method of learning:

"*I really like the methodology in this e-learning course … " Doctor*

Some were also generally positive such as:

"*Just right, very good*" *Doctor*
"*Super topic*" *nurse*
"*Very helpful*" *nurse*
"*Relatively friendly*" *nurse*

Discussion

This pilot study of a module on cerebral NIRS monitoring for the SafeBoosC-III web-based training and certification program shows that it is possible to complete the module within a reasonable time frame, that the academic level is appropriate and that clinical relevance is high, irrespective of previous experience, clinical position or nationality.

In order to prepare for practical use, training must include clinically relevant scenarios. In the SafeBoosC web-based training and certification program, training cases are based on real-life scenarios and written by clinically experienced neonatologists and experts in the field [24, 25, 31–33], thereby making it possible to merge wide clinical experience and up-to-date literature within the field.

The external validity of our results is high [34] since the training module was tested in three different countries (Denmark, Spain and China), across two continents (Europe and Asia). Furthermore, we invited participants both with- and without previous experience on NIRS monitoring to participate. This was done to evaluate whether previous experience affected performance and comprehension. In SafeBoosC-III, the level of NIRS experience will vary between departments; thus, knowledge on feasibility of the certification and training program dependent on previous experience level is important. By using both closed- and open-ended questions, we were able to gain a wider and deeper understanding of the participants' experience, which revealed important strengths and limitations of our design.

Translation of the training module was done manually by AP and GC (see "Methods" section) without any external translation support. Due to limited resources, we were not able to assess the quality and precision of the translations from English to Spanish and Chinese and were therefore not able to determine whether the quality and precision of the translations affected the difference in performance parameters and satisfaction rates. Despite a reasonable participation rate with 81 of 100 participants completing the online survey, we do not know for certain if all 81 responders completed the module. When looking at performance data (time to completion and number of questions to completion), 77 of the 81 responders had entered data for at least one of these parameters. However, two of the 77 responders commented that they did not complete the module, despite entering data on performance. Thus, we do not find data entry on performance parameters reliable as a measure of module completion. If we ought to rely on comments from responders, a total of five commented that they did not complete the module, primarily due to technical issues. Furthermore, we do not know whether the 19 participants who did not answer the survey still entered the training module but, due to unknown reasons, refrained from participating in the survey. Theoretically, it is possible that some of the 19 participants have been training in the module but gave up before completion and therefore did not answer the survey. Due to restricted time, it was not possible for us to host the piloting in Moodle, which would have made it possible to track completion rates.

When looking at performance data, the ranges of estimates are wide, with an upper limit of 50 questions for 'number of questions to completion' and 420 min for 'time to completion' (Table 3). However, only one responder answered 50 questions, one answered 30 questions and the remaining 79 responders answered using 20 questions or less. Regarding 'time to completion', six responders reported that they spent 420 min in the module, but only between 20 and 8 questions. They were all from the same country. The module automatically tracks time spent in the module, and when you reach completion, it will report the total time until completion. We suspect, therefore, that the six responders have had the module open throughout a 7-hour period but only trained part of the time. The remaining 75 responders spent 60 min or less.

Despite that web-based training provides a platform to train large numbers of staff across multiple countries, it also has the disadvantage of not knowing exactly how training was conducted locally. Since we could not monitor training on-site, we do not know whether responders trained in groups instead of individually, or how much they supported each other in completing the module. This could potentially affect performance data as well as answers to the questionnaire, thereby decreasing the validity of our results. However, using the SafeBoosC web-based training and certification program for group training instead of individual training may also be the case outside this pilot study; thus, this might depict how training and certification will be conducted when it is implemented externally. The module is structured and built for individual training, but group training may encourage discussions regarding the learning topics and therefore an increased learning opportunity.

By evaluating participants' experience only through an online survey, we have potentially missed out important information and thereby the foundation for additional improvement. A semi-structured interview [35] with randomly selected responders from each country might have given us a deeper understanding of responses. However, due to restricted time and limited budget, this was not possible to conduct.

Despite positive responses, this piloting also revealed room for improvement, as described in the "Results" section. Major critique points included 1) too many answer possibilities, 2) inadequate correlation between learning material and quiz, and 3) too many technical problems hindering completion of the module. All of these points have been taken into consideration when we revised this specific module as well as designed the additional modules for the SafeBoosC-III web-based training and certification program; at first, we revised and edited all question within the modules so that only one or two correct answer possibilities persisted, in contrast to the previous design where some questions held up to five correct answers. We also scaled the questions so that the maximum number of answer possibilities was limited to five, as opposed to previously having up to ten possibilities. This condensing of the questions required that we split some of the complex questions into two or more focused questions, thereby creating a larger pool of questions per learning objective. Secondly, for each module, we cross-checked the content of the learning material with questions in the quiz in order to identify inconsistencies. If such were found, relevant content was added to the learning material in order to ensure adequate coverage of the learning material.

Regarding the technical problems, a new IT consultant identified several errors in the coding that caused crashes and other technical difficulties. These were corrected and implemented in all the training modules.

As of today, the full-scale SafeBoosC web-based training and certification program is hosted in Moodle on a commercial platform and up-and-running. It has been translated into Chinese, Turkish, Spanish, German and French using a similar approach as was done for this pilot study. So far, more than 500 staff members have trained, and we have not yet received complaints related to any of the major critique points revealed in the pilot phase, including technical errors.

We hope that our reporting of developing and implementing web-based training for a pragmatic, multinational study will encourage other trialists to take a similar approach despite limited funding, as well as reporting on the process for the benefit of peers.

Conclusion

We believe that we provide evidence of the feasibility of developing a multilingual web-based training program for an international trial, despite challenges such as low budget, language barriers and possibly differences in the clinical training of staff. Exploring the integration of training and certification for international trials, the positive results of this study motivate further developments.

Abbreviations

GCP: Good Clinical Practice; NIRS: Near-infrared spectroscopy; SafeBoosC: Safeguarding the brains of our smallest children; USD: United States dollars

Acknowledgements

We sincerely thank all doctors and nurses from participating neonatal intensive care units who took their time to complete the web-based training module on NIRS as well as answered our online survey. We also thank Simon Hyttel-Sørensen (Department of Neonatology, Rigshospitalet, Denmark) for writing content for the module on NIRS monitoring.

Resources

Since funding is limited, the neonatologists agreed on writing the content for the training modules free of charge. They are academic partners of the SafeBoosC project. Writing content for all five modules took approximately 100 h. Expenses to Copenhagen University Hospital e-learning section was 19,000 USD and covered consultancy on pedagogic methodology as well as programming. Furthermore, we outsourced the tasks of reprogramming the modules in different languages through Fiverr (www.fiverr.com), an online freelance service market place, where online tech professionals can offer their services. Reprogramming one module in one new language is about 100 USD. Local translation has been estimated to approximately 4 hours per module.

Authors' contributions

MLH and MIR contributed to the development of the online training module on NIRS monitoring, contributed to the conception and design of this study, collected and analysed data, drafted the manuscript and will give final approval of the version to be published. SR contributed to the development of the online training module on NIRS monitoring, contributed to the conception and design of this study, drafted the manuscript and will give final approval of the version to be published. AP and GC translated content for the module on NIRS monitoring into Spanish and Chinese respectively, recruited participants, collected data, revised the manuscript critically for important intellectual content and will give final approval of the version to be published. XX and YZ recruited participants, collected data, revised the manuscript critically for important intellectual content and will give final approval of the version to be published. VZ contributed to the conception and design of this study, drafted the manuscript and will give final approval of the version to be published. GG contributed to the development of the online training module on NIRS monitoring, contributed to the conception and design of this study, drafted the manuscript and will give final approval of the version to be published.

Funding

The sponsor/coordinating investigator, Professor of Neonatology Gorm Greisen, is the initiator of the SafeBoosC-III project. He has no financial interest in the results of the trial, nor in the NIRS devices. The Elsass Foundation, Aage and Johanne Louis-Hansen Foundation and Svend Andersen Foundation supported this work through unconditional and unrestricted grants of DKK 2,700,000, DKK 1,000,000 and DKK 1,000,000, respectively. These funding sources had no role in the design of this study and will not have any role during its execution, analysis, interpretation of the data or decision to submit results.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

According to Danish, Chinese and Spanish laws, survey studies are not considered biomedical research and ethics approval is not therefore required to conduct this study. No personal data were collected on the participants, except for the self-reported data on clinical position and nationality. Prior to study initiation, participants were informed about this.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Neonatology, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark. [2]Section for Learning Technology, Center for HR, Kildegårdsvej 28, 2900 Hellerup, Capital Region of Denmark, Denmark. [3]Department of Neonatology, La Paz University Hospital, Paseo de La Castellana 261, 28046 Madrid, Spain. [4]Department of Neonatology, Children's Hospital of Fudan University, 399 Wanyuan Rd, Minhang Qu, Shanghai Shi, China. [5]Department of Neonatology, Key laboratory of Neonatal Diseases, Xiamen Children's Hospital, 10 Zhenhai Road, Xiamen, Fujian province, China. [6]Department of Neonatology, Dehong People's Hospital of Yunnan Province, 13 Yong Han Street, Dehong Mangshi City, Yunnan province, China. [7]Copenhagen University the department of Public Health, Section for Health Services Research, Rigshospitalet, JMC, Department of Women's and Children's Health, Tagensvej 22, 2200 Copenhagen, Denmark.

Received: 3 October 2019 Accepted: 25 February 2020
Published online: 23 April 2020

References

1. Garattini S, Jakobsen JC, Wetterslev J, Bertelé V, Banzi R, Rath A, et al. Evidence-based clinical practice: Overview of threats to the validity of evidence and how to minimise them. Eur J Intern Med. 2016;32:13–21.
2. Nardini C. The ethics of clinical trials. Ecancermedicalscience. 2014;8:1–9.
3. CHMP. Guideline for good clinical practice E6(R2) by the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Eur Med Agency. 2016;6:1–70.
4. Walker R, Morris DW, Greer TL, Trivedi MH. Research staff training in a multisite randomized clinical trial: Methods and recommendations from the stimulant reduction intervention using dosed exercise (STRIDE) trial. Addict Res Theory. 2014;22:407–15.
5. Targum SD. Evaluating rater competency for CNS clinical trials. J Clin Psychopharmacol. 2006;26:308–10.
6. Mulsant BH, Kastango KB, Rosen J, Stone RA, Mazumdar S, Pollock BG. Interrater reliability in clinical trials of depressive disorders. Am J Psychiatry. 2002;159:1598–600.
7. Clark RC, Mayer RE. E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedialearning. 3rd ed. San Francisco: Wiley; 2011.
8. Ruiz JG, Mintzer MJ, Leipzig RM. The impact of e-learning in medical education. Acad Med. 2006;81:207–12.
9. Palmer RC, Samson R, Triantis M, Mullan ID. Development and evaluation of a web-based breast cancer cultural competency course for primary healthcare providers. BMC Med Educ. 2011;11:59.
10. Brown M, Bullock A. Evaluating PLATO: postgraduate teaching and learning online. Clin Teach. 2014;11:10–4.
11. Antonoff MB, Verrier ED, Allen MS, Aloia L, Baker C, Fann JI, et al. Impact of Moodle-based online curriculum on thoracic surgery in-training examination scores. Ann Thorac Surg. 2016;102:1381–6.
12. Thukral A, Sasi A, Chawla D, Datta P, Wahid S, Rao S, et al. Online Neonatal Training and Orientation Programme in India (ONTOP-IN)--the way forward for distance education in developing countries. J Trop Pediatr. 2012;58:486–90.
13. Pelayo-Alvarez M, Perez-Hoyos S, Agra-Varela Y. Clinical effectiveness of online training in palliative care of primary care physicians. J Palliat Med. 2013;16:1188–96.
14. Oliveira AC, Mattos S, Coimbra M. Development and assessment of an e-learning course on pediatric cardiology basics. JMIR Med Educ. 2017;3:e10.
15. Frehywot S, Vovides Y, Talib Z, Mikhail N, Ross H, Wohltjen H, et al. E-learning in medical education in resource constrained low- and middle-income countries. Hum Resour Health. 2013;11:4.
16. Vaona A, Banzi R, Kwag KH, Rigon G, Cereda D, Pecoraro V, et al. E-learning for health professionals. Cochrane Database Syst Rev. 2018;5:12.
17. Cook DE, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Internet-based learning in the health professions. JAMA. 2014;300:1181–96.
18. Cook DA. Web-based learning: pros, cons and controversies. Clin Med (Northfield Il). 2007;7:37–42.
19. Benjamin SE, Tate DF, Bangdiwala SI, Neelon BH, Ammerman AS, Dodds JM, et al. Preparing child care health consultants to address childhood overweight: A randomized controlled trial comparing web to in-person training. Matern Child Health J. 2008;12:662–9.
20. Ford I, Norrie J. Pragmatic trials. N Engl J Med. 2016;375:454–63.
21. Bloom BS, Krathwohl DR, Masia BS. Taxonomy of educational objectives. the classification of educational goals : Cognitive domain handbook 1. New York: Longman; 1984.
22. Wojcikowski K, Kirk L. Immediate detailed feedback to test-enhanced learning: An effective online educational tool. Med Teach. 2013;35:915–9.
23. Reis LO, Ikari O, Taha-Neto KA, Gugliotta A, Denardi F. Delivery of a urology online course using moodle versus didactic lectures methods. Int J Med Inform. 2015;84:149–54.
24. Hyttel-Sørensen S, Pellicer A, Alderliesten T, Austin T, Van Bel F, Benders M, et al. Cerebral near infrared spectroscopy oximetry in extremely preterm infants: phase II randomised clinical trial. BMJ. 2015;350:1–11.
25. Greisen G, Andresen B, Plomgaard AM, Hyttel-Sørensen S. Cerebral oximetry in preterm infants: an agenda for research with a clear clinical goal. Neurophotonics. 2016;3:31407.
26. Hyttel-Sorensen S, Kleiser S, Wolf M, Greisen G. Calibration of a prototype NIRS oximeter against two commercial devices on a blood-lipid phantom. Biomed Opt Express. 2013;4:1662–72.
27. Hyttel-Sørensen S, Greisen G, Als-Nielsen B, Gluud C. Cerebral near-infrared spectroscopy monitoring for prevention of brain injury in very preterm infants. Cochrane Database Syst Rev. 2017;9:CD011506.
28. Hyttel-Sorensen S, Hessel TW, la Cour A, Greisen G. A comparison between two NIRS oximeters (INVOS, OxyPrem) using measurement on the arm of adults and head of infants after caesarean section. Biomed Opt Express. 2014;5:3671.
29. Wang Y-S. Assessment of learner satisfaction with asynchronous electronic learning systems. Inf Manag. 2003;41:75–86.
30. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol. 2006;3:77–101.
31. Pellicer A, Greisen G, Benders M, Claris O, Dempsey E, Fumagally M, et al. The SafeBoosC phase II randomised clinical trial: A treatment guideline for targeted near-infrared-derived cerebral tissue oxygenation versus standard treatment in extremely preterm infants. Neonatology. 2013;104:171–8.
32. Plomgaard AM, Hagmann C, Alderliesten T, Austin T, Van Bel F, Claris O, et al. Brain injury in the international multicenter randomized SafeBoosC phase II feasibility trial: Cranial ultrasound and magnetic resonance imaging assessments. Pediatr Res. 2016;79:466–72.
33. Riera J, Hyttel-Sorensen S, Bravo MC, Cabañas F, López-Ortego P, Sanchez L, et al. The SafeBoosC phase II clinical trial: an analysis of the interventions related with the oximeter readings. Arch Dis Child Fetal Neonatal Ed. 2016;101:F333–8.
34. Rothwell PM. External validity of randomised controlled trials: "To whom do the results of this trial apply?". Lancet. 2005;365:82–93.
35. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care. 2007;19:349–57.

Publisher's Note

Not Removing the Glossy White Cover from Adhesive INVOS Neonatal Sensors Affects the Oxygenation Measurement

Mathias Lühr Hansen[a,#], **Daniel Ostojic**[b,#], **Stefan Kleiser**[b], **Gorm Greisen**[a], **Martin Wolf**[b]

[a]Department of Neonatology, Rigshospitalet, University of Copenhagen, Denmark

[b]Biomedical Optics Research Laboratory (BORL), Department of Neonatology, University Hospital Zürich, University of Zürich, Switzerland

[#]Joint first-authorship

Abstract The randomized clinical trial, SafeBoosC III, evaluates the effect of treatment guided by cerebral tissue oximetry monitoring in extremely preterm infants. Treatment should be considered, when cerebral oxygen saturation (StO_2) drops below a predefined hypoxic threshold. This threshold value differs between different brands of instruments. To achieve high external validity, in this pragmatic trial all commercially available cerebral tissue oximeters have been accepted, provided their specific hypoxic threshold value has been determined in phantom studies. Since most companies produce sensors with an adhesive surface on the patient-contacting side, in the phantom studies these sensors were applied according to the specifications, i.e. the glossy cover was removed from the sensor. However, since the skin of preterm infants is particularly fragile, some neonatologists keep this cover on the adhesive sensors, to avoid the risk of skin injury when removing the sensor. Therefore, the aim of this study was to determine whether keeping this cover on, leads to different StO_2 values. To evaluate the effect of the cover, we performed multiple deoxygenations in a blood-lipid phantom and compared an INVOS neonatal sensor (Medtronic), with and without the cover, to a reference oximeter (OxiplexTS, ISS). As expected, the relationship of the StO_2 between the INVOS neonatal sensor and OxiplexTS was linear ($r^2 = 0.999$) with and without cover, but the cover influenced the linear equation: $StO_{2\_INVOS\_cover}=1.133*StO_{2\_ISS}+7.1$ as opposed to $StO_{2\_INVOS\_nocover}=1.103*StO_{2\_ISS}+12.0$. Furthermore, the hypoxic SafeBoosC III threshold differed as well: 60.3% with cover and 63.8% without cover. In conclusion, keeping the adhesive cover on an INVOS neonatal sensor results in lower measured values. At the hypoxic threshold this is more than 3% (from 60.3% to 63.8%) and therefore, if clinicians keep the cover on the sensor, they need to be aware of this difference.

1 Introduction

Near-infrared spectroscopy (NIRS) is a non-invasive technology to monitor tissue oxygenation, by using near-infrared light [1]. It provides a value of tissue oxygen saturation (StO_2), expressed as the ratio of oxygenated to total haemoglobin in the tissue under the sensor [1]. It is possible to monitor StO_2 of various tissues, but so far, it has mostly been used to assess cerebral StO_2 [2–4]. In neonatology, despite lacking clear evidence of positive effects on patient-relevant outcomes [5], clinical usage of cerebral NIRS monitoring is growing [1, 6]. SafeBoosC-III is a randomised clinical trial evaluating the effect of treatment guided by cerebral NIRS monitoring in extremely preterm infants [7]. In SafeBoosC III, treatment should be considered when the cerebral StO_2 drops below a predefined hypoxic threshold. The primary end-point of the trial is survival without severe brain injury. Since SafeBoosC III is a pragmatic trial aiming for a high external validity, all NIRS devices approved for clinical use in newborns are allowed in the trial. However, previous studies have shown that different NIRS devices differ in absolute StO_2 values [8, 9], i.e. the value of the hypoxic threshold differs between different brands of instruments. A previous study has found the StO_2 normal range in preterm babies to be between 55% and 85% with the INVOS adult SomaSensor [10]. For the purpose of SafeBoosC, hypoxic thresholds corresponding to 55% with the INVOS adult SomaSensor, have been determined in a blood lipid phantom for most commercially available NIRS devices [11, 12]. The majority of NIRS sensors are produced with an adhesive surface on the patient-contacting side to make skin attachment easier. However, since the skin of extremely preterm infants is fragile [13], some clinicians within and outside the SafeBoosC organization have previously kept the cover on the adhesive sensors during NIRS monitoring outside of SafeBoosC context, to avoid the risk of skin injury when the sensor is repositioned or removed (personal communication). In the phantom studies [11, 12], sensors were tested according to the specifications, without the cover. Although unexpected, it is possible that the cover affects the measurement of StO_2 and hence the hypoxic threshold. The aim of this study was to evaluate the effect of not removing the cover from an adhesive sensor when measuring cerebral StO_2.

2 Methods

We compared the INVOS 5100C with Cerebral/Somatic Oximetry Infant-Neonatal sensor (Medtronic, Inc. Minneapolis, MN, USA), with and without cover, to the Oxiplex TS (ISS, Inc., Champaign, IL, USA) as a reference oximeter [11, 12]. The phantom design and setup was previously described [11]. Human blood from recently expired human erythrocyte concentration bags and Intralipid were added, to obtain a haemoglobin concentration of 40 μM and a scattering coefficient (at 834

nm) of 5.3 cm^{-1}. Both parameters remained unchanged throughout the measurement. We performed a total of four deoxygenations, where the sensors response to oxygenation changes were measured simultaneously. For deoxygenation one, two INVOS sensors—one with a cover (S1) and one without (S2)—and the Oxiplex TS were attached to the phantom windows (Fig. 1). The INVOS sensor without cover (S2) served as a control measure of the hypoxic threshold and sensitivity to oxygenation changes, since we already know these values from previous experiments [11]. Before deoxygenation two, the two INVOS sensors (S1 and S2) had the window switched (Fig. 1) to assess a potential effect of sensor positioning. Before deoxygenation three, the cover was removed from S1, i.e., both sensors were without a cover (Fig. 1). Before deoxygenation four, the INVOS sensors (S1 and S2) were once again switched back to their starting positions (Fig. 1) and remained without the cover. The primary analysis was a linear regression between the INVOS neonatal sensor with cover (S1) and the Oxiplex TS (Fig. 2). Finally, we compared the sensitivity and hypoxic threshold for S1 during deoxygenation one and two (with cover) and deoxygenation three and four (without cover). The hypoxic threshold and sensitivity for the control sensor (S2) was calculated as well, to ensure that the values were consistent and similar to those previously reported [11]. To quantify inter-sensor variation measurements during deoxygenations three and four, with both sensors lacking a cover, were considered (Table 1). Data processing methods were the same as previously used in [11, 12].

3 Results

Fig. 1 shows the time series for the four deoxygenations. The relationship between S1 and Oxiplex TS was linear ($R^2 = 0.999$) both with and without the cover. Table 1 shows the coefficients for the linear fits, i.e., the sensitivity to oxygenation changes, as well as the hypoxic thresholds. The sensitivity to oxygenation changes differed depending on the cover. The hypoxic thresholds were determined for $StO_2 = 47\%$ for the OxiplexTS, which corresponds to 55% with the INVOS Adult SomaSensor [11]. Based on the linear fits (Tab. 1) and the threshold of the OxiplexTS of 47%, we calculated the hypoxic threshold for S1. This threshold depended on the cover: with the cover, it was 60.3% and without the cover, 63.8%. The repositioning of the sensors only had a minor effect on these values (Table 1). Also, the inter-sensor variation was small (Table 1). These data show that phantom measurements are a highly reproducible way of testing NIRS instrumentation.

4 Discussion

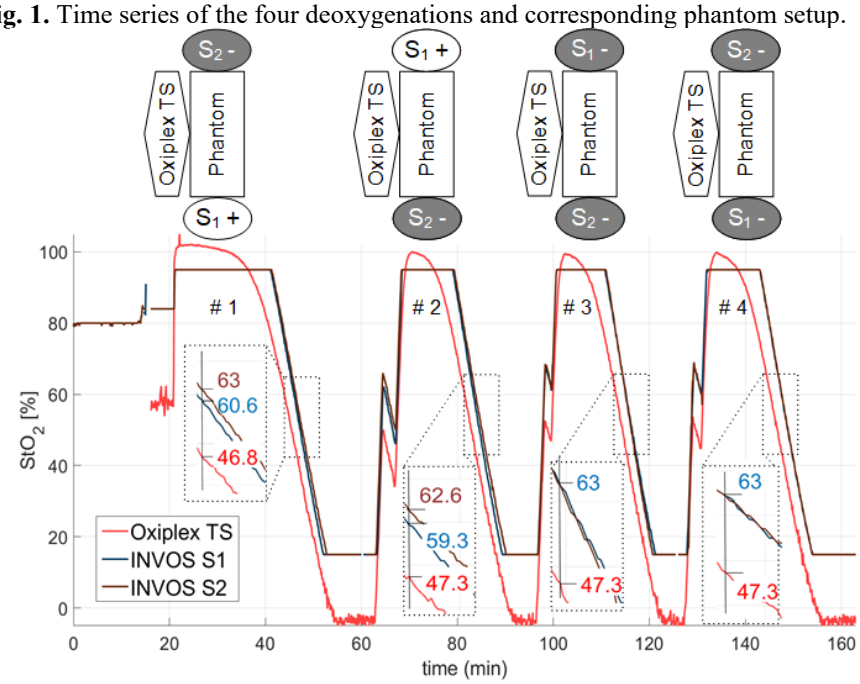
The main finding of this study is that keeping the glossy white cover on an INVOS neonatal sensor during cerebral oxygenation monitoring decreases the hypoxic threshold by more than 3%. This is an unexpected finding. Furthermore, although the relationship to the reference oximeter, Oxiplex TS [12] is linear, the slope also deviates by ~3%. How can this surprising result be explained? Since the algorithm of the INVOS neonatal sensor is not published, we can only speculate why the absolute StO_2 values are lower when the cover is on. We suggest the following possible reasons: a) light is channeling through the adhesive (light piping) when the cover is taken off, i.e., the cover reduces light piping by the adhesive, b) the cover's scattering is higher at low wavelengths, and/or c) the very high scattering of the cover (~60 cm^{-1}) behaves optically like a mirror and creates a boundary condition more similar to an infinite geometry. An explanation must take into account that, at high StO_2, the difference between the measurement with and without the cover is smaller. This indicates that mainly the lowest wavelengths of the INVOS are affected and means that the slope of intensity versus distance at this wavelength must be steeper with a cover. A more infinite boundary condition (option c) would lead to a decrease in the steepness of the slope and is, therefore, unlikely. A higher scattering at lower wavelength (option b) might increase the slope at a lower wavelength. However, when we measured the scattering of the cover, the scattering was strong (~60/cm, approximately ten times higher than the scattering of tissue), but no wavelength dependency was visible. It must be kept in mind that this high scattering was at the detection limit of our spectrometer. Thus, a more likely explanation seems to be the light channeling effect. A major strength in this study is the experimental setup. By performing multiple deoxygenations, changing sensor position as well as cover status, we were able to quantify the effect of sensor position as well as inter-sensor variance with high precision. Both had little effect on sensitivity and oxygenation values (Tab. 1). Furthermore, we used a dynamic in-vitro phantom, which has the advantage of controllable optical properties, minimal variations, and the possibility to test the device's reactions to oxygenation changes for a wide range of StO_2 [11]. This would not be possible if the experiments had been performed in-vivo in neonates. Frequency domain NIRS, such as OxiplexTS, is acceptable as a reference, since it accurately measures the total concentrations of O_2Hb (oxyhaemoglobin) and HHb (deoxyhaemoglobin). Additionally, it enables monitoring of variations in scattering [14]. In conclusion, keeping the adhesive cover on an INVOS neonatal sensor results in lower measured values. At the hypoxic threshold this is more than 3% (from 60.3% to 63.8%) and therefore, if clinicians keep the cover on the sensor, they need to be aware of this difference.

Table 1. Overview of the four deoxygenations

Sensor	Position	Deoxygenation	Cover	a	b	StO_2 % Oxiplex TS 47%
S1	1	1	+	1.133	7.067	60.324
S1	2	2	+	1.144	6.708	60.468
S1	2	3	-	1.086	12.020	63.047
S1	1	4	-	1.103	11.992	63.836
S2	2	1	-	1.089	11.823	63.013
S2	1	2	-	1.092	12.466	63.797
S2	1	3	-	1.152	8.410	62.548
S2	2	4	-	1.105	12.173	64.124

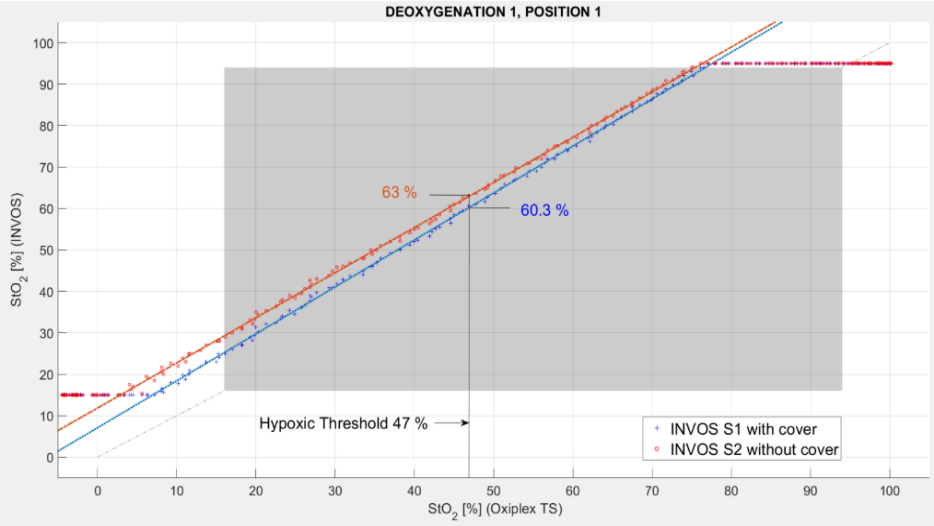
The table includes position in the phantom setup, status on cover (+ with and - without cover), linear coefficients (a=slope, b=intercept, StO_2(INVOS) = a*StO_2(OxiplexTS) + b) and hypoxic thresholds corresponding to an StO_2 of 47% with Oxiplex TS (corresponding to the hypoxic threshold of 55% as defined by the INVOS with the small adult sensor).

Fig. 1. Time series of the four deoxygenations and corresponding phantom setup.



For each deoxygenation, there is a time series (bottom) of the StO_2 values in synchronized time, and an illustration of the phantom setup (top). S1 = INVOS neonatal sensor with cover initially, S2 = INVOS neonatal sensor without cover (control sensor). '+' = with cover, '-' = without cover.

Fig. 2. INVOS neonatal sensor and Oxiplex TS linear correspondence with and without cover



References

1. Green MS, Sehgal S, Tariq R (2016) Near-Infrared Spectroscopy: The New Must Have Tool in the Intensive Care Unit? Semin Cardiothorac Vasc Anesth 20:213–224
2. Taillefer MC, Denault AY (2005) Cerebral near-infrared spectroscopy in adult heart surgery: systematic review of its clinical efficacy. Can J Anaesth 52:79–87
3. Dullenkopf A, Frey B, Baenziger O, et al (2003) Measurement of cerebral oxygenation state in anaesthetized children using the INVOS 5100 cerebral oximeter. Pediatr Anesth 13:384–391
4. Hyttel-Sørensen S, Pellicer A, Alderliesten T, et al (2015) Cerebral near infrared spectroscopy oximetry in extremely preterm infants: phase II randomised clinical trial. BMJ 350:1–11
5. Hyttel-Sørensen S, Greisen G, Als-Nielsen B, Gluud C (2017) Cerebral near-infrared spectroscopy monitoring for prevention of brain injury in very preterm infants. Cochrane Database Syst Rev 1–25
6. Hunter CL, Oei JL, Suzuki K, et al (2018) Patterns of use of near-infrared spectroscopy in neonatal intensive care units: international usage survey. Acta Paediatr 107:1198–1204
7. Hansen ML, Pellicer A, Gluud C, et al (2019) Cerebral near-infrared spectroscopy monitoring versus treatment as usual for extremely preterm infants. A protocol for the SafeBoosC phase III randomized clinical trial. Manuscr Submitt Publ
8. Hessel TW, Hyttel-Sorensen S, Greisen G (2014) Cerebral oxygenation after birth - a comparison of INVOS® and FORE-SIGHT™ near-infrared spectroscopy oximeters. Acta Paediatr Int J Paediatr 103:488–493
9. Dix LML, van Bel F, Baerts W, Lemmers PMA (2013) Comparing near-infrared spectroscopy

devices and their sensors for monitoring regional cerebral oxygen saturation in the neonate. Pediatr Res 74:557–563

10. Hyttel-sørensen S, Austin T, Bel F Van, et al (2013) Clinical use of cerebral oximetry in extremely preterm infants is feasible. Dan Med J A4533
11. Kleiser S, Nasseri N, Andresen B, et al (2016) Comparison of tissue oximeters on a liquid phantom with adjustable optical properties. Biomed Opt Express 7:2973
12. Kleiser S, Ostojic D, Andresen B, et al (2018) Comparison of tissue oximeters on a liquid phantom with adjustable optical properties: an extension. Biomed Opt Express 9:86
13. Greisen G, Leung T, Wolf M (2011) Has the time come to use near-infrared spectroscopy as a routine clinical tool in preterm infants undergoing intensive care? Philos Trans R Soc A Math Phys Eng Sci 369:4440–4451
14. Ostojic D, Kleiser S, Nasseri N, et al (2018) In vitro comparisons of near-infrared spectroscopy oximeters: Impact of slow changes in scattering of liquid phantoms. Adv Exp Med Biol 1072:375–379

DECLARATION OF CO-AUTHORSHIP

The declaration is for PhD students and must be completed for each conjointly authored article. Please note that if a manuscript or published paper has ten or less co-authors, all co-authors must sign the declaration of co-authorship. If it has more than ten co-authors, declarations of co-authorship from the corresponding author(s), the senior author and the principal supervisor (if relevant) are a minimum requirement.

1. Declaration by	
Name of PhD student	Mathias Lühr Hansen (corresponding author to this paper)
E-mail	Mathias.luehr.hansen@regionh.dk
Name of principal supervisor	Gorm Greisen (senior author to this paper)
Title of the PhD thesis	Preparing for the multinational, pragmatic, randomised clinical trial SafeboosC III

2. The declaration applies to the following article	
Title of article	Cerebral near-infrared spectroscopy monitoring versus treatment as usual for extremely preterm infants: a protocol for the SafeBoosC randomised clinical phase III trial
Article status	
Published ☒ Date: 30. December 2019	Accepted for publication ☐ Date:
Manuscript submitted ☐ Date:	Manuscript not submitted ☐
If the article is published or accepted for publication, please state the name of journal, year, volume, page and DOI (if you have the information).	Trials, 2019, 20, 811, 10.1186/s13063-019-3955-6

3. The PhD student's contribution to the article ***(please use the scale A-F as benchmark)*** Benchmark scale of the PhD-student's contribution to the article **A**. Has essentially done all the work (> 90 %) **B**. Has done most of the work (60-90 %) **C**. Has contributed considerably (30-60 %) **D.** Has contributed (10-30 %) **E.** No or little contribution (<10 %) **F.** Not relevant	**A, B, C, D, E, F**
1. Formulation/identification of the scientific problem	E
2. Development of the key methods	D
3. Planning of the experiments and methodology design and development	D

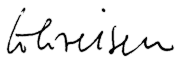
3. The PhD student's contribution to the article ***(please use the scale A-F as benchmark)*** Benchmark scale of the PhD-student's contribution to the article **A**. Has essentially done all the work (> 90 %) **B**. Has done most of the work (60-90 %) **C**. Has contributed considerably (30-60 %) **D.** Has contributed (10-30 %) **E.** No or little contribution (<10 %) **F.** Not relevant	**A, B, C, D, E, F**
4. Conducting the experimental work/clinical studies/data collection/obtaining access to data	B
5. Conducting the analysis of data	F
6. Interpretation of the results	F
7. Writing of the first draft of the manuscript	B
8. Finalisation of the manuscript and submission	A

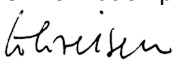
Provide a short description of the PhD student´s specific contribution to the article.[i]

The PhD student has contributed to the continuous development of the protocol for the SafeBoosC-III trial, based on an initial draft from 2015, in collaboration with the SafeBoosC-III steering committee. Based on the final protocol, the PhD student has written the manuscript and submitted it for publication.

4. Material from another thesis / dissertation[ii]	
Does the article contain work which has also formed part of another thesis, e.g. master's thesis, PhD thesis or doctoral dissertation (the PhD student's or another person's)?	Yes: ☐ No: ☒
If yes, please state name of the author and title of thesis / dissertation.	
If the article is part of another author's academic degree, please describe the PhD student´s and the author´s contributions to the article so that the individual contributions are clearly distinguishable from one another.	

5. Signatures of the co-authors[iii]				
	Date	**Name**	**Title**	**Signature**

5. Signatures of the co-authors[iii]				
1.	15-jan 2020	Gorm Greisen (principal supervisor and senior author) THE PUBLICATION HAS MORE THAN TEN CO-AUTHORS	Professor	

6. Signature of the principal supervisor
I solemnly declare that the information provided in this declaration is accurate to the best of my knowledge. Date: 15-jan-2020  Principal supervisor: Gorm Greisen

7. Signature of the PhD student
I solemnly declare that the information provided in this declaration is accurate to the best of my knowledge. Date: PhD student: Mathias Lühr Hansen  Digitalt signeret af Mathias Lühr Hansen Dato: 2021.02.23 12:07:51 +01'00'

Please learn more about responsible conduct of research on the Faculty of Health and Medical Sciences' website.

[i] This can be supplemented with an additional letter if needed.

[ii] Please see Ministerial Order on the PhD Programme at the Universities and Certain Higher Artistic Educational Institutions (PhD Order) § 12 (4):

"Any articles included in the thesis may be written in cooperation with others, provided that each of the co-authors submits a written declaration stating the PhD student's or the author's contribution to the work."

[iii] If more signatures are needed please add an extra sheet.

DECLARATION OF CO-AUTHORSHIP

The declaration is for PhD students and must be completed for each conjointly authored article. Please note that if a manuscript or published paper has ten or less co-authors, all co-authors must sign the declaration of co-authorship. If it has more than ten co-authors, declarations of co-authorship from the corresponding author(s), the senior author and the principal supervisor (if relevant) are a minimum requirement.

1. Declaration by	
Name of PhD student	Mathias Lühr Hansen (corresponding author to this paper)
E-mail	Mathias.luehr.hansen@regionh.dk
Name of principal supervisor	Gorm Greisen (senior author to this paper)
Title of the PhD thesis	Preparing for the multinational, pragmatic, randomised clinical trial SafeboosC III

2. The declaration applies to the following article	
Title of article	Detailed statistical analysis plan for the SafeBoosC III trial: a multinational randomised clinical trial assessing treatment guided by cerebral oxygenation monitoring versus treatment as usual in extremely preterm infants
Article status	
Published ☒ Date: 19. December 2019	Accepted for publication ☐ Date:
Manuscript submitted ☐ Date:	Manuscript not submitted ☐
If the article is published or accepted for publication, please state the name of journal, year, volume, page and DOI (if you have the information).	Trials, 2019, 20, 746, 10.1186/s13063-019-3756-y

3. The PhD student's contribution to the article ***(please use the scale A-F as benchmark)*** Benchmark scale of the PhD-student's contribution to the article **A**. Has essentially done all the work (> 90 %) **B**. Has done most of the work (60-90 %) **C**. Has contributed considerably (30-60 %) **D.** Has contributed (10-30 %) **E.** No or little contribution (<10 %) **F.** Not relevant	**A, B, C, D, E, F**
1. Formulation/identification of the scientific problem	E
2. Development of the key methods	D
3. Planning of the experiments and methodology design and development	D

3. The PhD student's contribution to the article ***(please use the scale A-F as benchmark)*** Benchmark scale of the PhD-student's contribution to the article **A**. Has essentially done all the work (> 90 %) **B**. Has done most of the work (60-90 %) **C**. Has contributed considerably (30-60 %) **D.** Has contributed (10-30 %) **E.** No or little contribution (<10 %) **F.** Not relevant	**A, B, C, D, E, F**
4. Conducting the experimental work/clinical studies/data collection/obtaining access to data	B
5. Conducting the analysis of data	F
6. Interpretation of the results	F
7. Writing of the first draft of the manuscript	B
8. Finalisation of the manuscript and submission	A

Provide a short description of the PhD student´s specific contribution to the article.[i]

The PhD student has contributed to the continuous development of the statistical analysis plan for the SafeBoosC-III trial, based on an initial draft from 2015, in collaboration with the Theis Lange (statistican at Copenhagen University), Copenhagen Trial Unit and Gorm Greisen. The PhD student has been the lead in writing the full statistical analysis plan including all aspects of publication.

4. Material from another thesis / dissertation[ii]	
Does the article contain work which has also formed part of another thesis, e.g. master's thesis, PhD thesis or doctoral dissertation (the PhD student's or another person's)?	Yes: ☐ No: ☒
If yes, please state name of the author and title of thesis / dissertation.	
If the article is part of another author's academic degree, please describe the PhD student´s and the author´s contributions to the article so that the individual contributions are clearly distinguishable from one another.	

5. Signatures of the co-authors[iii]				
	Date	**Name**	**Title**	**Signature**

5. Signatures of the co-authors[iii]				
1.	15-jan 2020	Gorm Greisen (principal supervisor and senior author) THE PUBLICATION HAS MORE THAN TEN CO-AUTHORS	Professor	

6. Signature of the principal supervisor
I solemnly declare that the information provided in this declaration is accurate to the best of my knowledge. Date:15-jan-2020 Principal supervisor: Gorm Greisen

7. Signature of the PhD student
I solemnly declare that the information provided in this declaration is accurate to the best of my knowledge. Date: PhD student: Mathias Lühr Hansen Digitalt signeret af Mathias Lühr Hansen Dato: 2021.02.23 12:19:44 +01'00'

Please learn more about responsible conduct of research on the Faculty of Health and Medical Sciences' website.

[i] This can be supplemented with an additional letter if needed.

[ii] Please see Ministerial Order on the PhD Programme at the Universities and Certain Higher Artistic Educational Institutions (PhD Order) § 12 (4):

"Any articles included in the thesis may be written in cooperation with others, provided that each of the co-authors submits a written declaration stating the PhD student's or the author's contribution to the work."

[iii] If more signatures are needed please add an extra sheet.

DECLARATION OF CO-AUTHORSHIP

The declaration is for PhD students and must be completed for each conjointly authored article. Please note that if a manuscript or published paper has ten or less co-authors, all co-authors must sign the declaration of co-authorship. If it has more than ten co-authors, declarations of co-authorship from the corresponding author(s), the senior author and the principal supervisor (if relevant) are a minimum requirement.

1. Declaration by	
Name of PhD student	Mathias Lühr Hansen
E-mail	Mathias.luehr.hansen@regionh.dk
Name of principal supervisor	Gorm Greisen
Title of the PhD thesis	Preparing for the multinational, pragmatic, randomised clinical trial SafeboosC III

2. The declaration applies to the following article	
Title of article	Pilot test of an online training module on near-infrared spectroscopy monitoring for the randomised clinical trial SafeBoosC-III
Article status	
Published ☒ Date: 23. April 2020	Accepted for publication ☐ Date:
Manuscript submitted ☐ Date:	Manuscript not submitted ☐
If the article is published or accepted for publication, please state the name of journal, year, volume, page and DOI (if you have the information).	Trials, 2020, 21, 356, 10.1186/s13063-020-4206-6

3. The PhD student's contribution to the article *(please use the scale A-F as benchmark)* Benchmark scale of the PhD student's contribution to the article A. Has essentially done all the work (> 90 %) B. Has done most of the work (60-90 %) C. Has contributed considerably (30-60 %) D. Has contributed (10-30 %) E. No or little contribution (<10 %) F. Not relevant	A, B, C, D, E, F
1. Formulation/identification of the scientific problem	C
2. Development of the key methods	C
3. Planning of the experiments and methodology design and development	C

3. The PhD student's contribution to the article *(please use the scale A-F as benchmark)* Benchmark scale of the PhD-student's contribution to the article A. Has essentially done all the work (> 90 %) B. Has done most of the work (60-90 %) C. Has contributed considerably (30-60 %) D. Has contributed (10-30 %) E. No or little contribution (<10 %) F. Not relevant	A, B, C, D, E, F
4. Conducting the experimental work/clinical studies/data collection/obtaining access to data	C
5. Conducting the analysis of data	B
6. Interpretation of the results	C
7. Writing of the first draft of the manuscript	B
8. Finalisation of the manuscript and submission	B
Provide a short description of the PhD student's specific contribution to the article.[i] The PhD student has contributed considerably or more to all aspects of this publication, from getting the idea, designing the study, gathering the study group, organising data collection, collecting data, analysing data, writing the manuscript as well as the submission and publication process.	

4. Material from another thesis / dissertation[ii]	
Does the article contain work which has also formed part of another thesis, e.g. master's thesis, PhD thesis or doctoral dissertation (the PhD student's or another person's)?	Yes: ☐ No: ☒
If yes, please state name of the author and title of thesis / dissertation.	
If the article is part of another author's academic degree, please describe the PhD student's and the author's contributions to the article so that the individual contributions are clearly distinguishable from one another.	

5. Signatures of the co-authors[iii]				
	Date	Name	Title	Signature
1.	15-Jan 2020	Gorm Grelsen	MD, Professor	GGreisen

5. Signatures of the co-authors[ii]				
2.	17. jan 2021	Marie Isabel Rasmussen	Medical student	
3.	20. jan. 2021	Snorre Rubin	Digital e-learning consultant	
4.		Adelina Pellicer	MD, Head of department	PELLICER MARTINEZ ADELINA - 19984140S
5.		Guoqiang Cheng	MD, Head of department	Guoqiang Cheng
6.		Xin Xu	MD	Xin Xu
7.		Yin Zhaoqing	MD	Yin Zhaoqing
8.		Vibeke Zoffman	Nurse, Professor	Vibeke Zoffmann

6. Signature of the principal supervisor

I solemnly declare that the information provided in this declaration is accurate to the best of my knowledge.

Date: 15-jan-2020

Principal supervisor: Gorm Grelsen

7. Signature of the PhD student

I solemnly declare that the information provided in this declaration is accurate to the best of my knowledge.

Date:

PhD student: Mathias Lühr Hansen

Digitalt signeret af Mathias Lühr Hansen
Dato: 2021.02.18 12:06:11 +01'00'

Please learn more about responsible conduct of research on the Faculty of Health and Medical Sciences' website.

[i] This can be supplemented with an additional letter if needed.

[ii] Please see Ministerial Order on the PhD Programme at the Universities and Certain Higher Artistic Educational Institutions (PhD Order) § 12 (4):

"Any articles included in the thesis may be written in cooperation with others, provided that each of the co-authors submits a written declaration stating the PhD student's or the author's contribution to the work."

"If more signatures are needed please add an extra sheet.

DECLARATION OF CO-AUTHORSHIP

The declaration is for PhD students and must be completed for each conjointly authored article. Please note that if a manuscript or published paper has ten or less co-authors, all co-authors must sign the declaration of co-authorship. If it has more than ten co-authors, declarations of co-authorship from the corresponding author(s), the senior author and the principal supervisor (if relevant) are a minimum requirement.

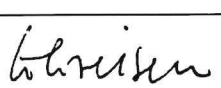
1. Declaration by	
Name of PhD student	Mathias Lühr Hansen
E-mail	Mathias.luehr.hansen@regionh.dk
Name of principal supervisor	Gorm Greisen
Title of the PhD thesis	Preparing for the multinational, pragmatic, randomised clinical trial SafeboosC III

2. The declaration applies to the following article	
Title of article	Not Removing the Glossy White Cover from Adhesive INVOS Neonatal Sensors Affects the Oxygenation Measurement
Article status	
Published ☐	Accepted for publication ☒ Date: 30-November 2019
Manuscript submitted ☐ Date:	Manuscript not submitted ☐
If the article is published or accepted for publication, please state the name of journal, year, volume, page and DOI (if you have the information).	Advances in Experimental Medicine and Biology Vol. 1269, Edwin M. Nemoto et al. (Eds): Oxygen Transport to Tissue XLII, Chapter 56

3. The PhD student's contribution to the article ***(please use the scale A-F as benchmark)*** Benchmark scale of the PhD-student's contribution to the article **A.** Has essentially done all the work (> 90 %) **B.** Has done most of the work (60-90 %) **C.** Has contributed considerably (30-60 %) **D.** Has contributed (10-30 %) **E.** No or little contribution (<10 %) **F.** Not relevant	**A, B, C, D, E, F**
1. Formulation/identification of the scientific problem	D
2. Development of the key methods	C
3. Planning of the experiments and methodology design and development	C

3. The PhD student's contribution to the article ***(please use the scale A-F as benchmark)*** Benchmark scale of the PhD-student's contribution to the article **A.** Has essentially done all the work (> 90 %) **B.** Has done most of the work (60-90 %) **C.** Has contributed considerably (30-60 %) **D.** Has contributed (10-30 %) **E.** No or little contribution (<10 %) **F.** Not relevant	**A, B, C, D, E, F**
4. Conducting the experimental work/clinical studies/data collection/obtaining access to data	C
5. Conducting the analysis of data	E
6. Interpretation of the results	C
7. Writing of the first draft of the manuscript	B
8. Finalisation of the manuscript and submission	B
Provide a short description of the PhD student´s specific contribution to the article.[i] The PhD student has contributed in all aspects of this study, except for the Mathlab data analysis, which was conducted by Daniel Ostojic.	

4. Material from another thesis / dissertation[ii]	
Does the article contain work which has also formed part of another thesis, e.g. master's thesis, PhD thesis or doctoral dissertation (the PhD student's or another person's)?	Yes: ☐ No: ☒
If yes, please state name of the author and title of thesis / dissertation.	
If the article is part of another author's academic degree, please describe the PhD student´s and the author´s contributions to the article so that the individual contributions are clearly distinguishable from one another.	

5. Signatures of the co-authors[iii]				
	Date	**Name**	**Title**	**Signature**
1.	15-jan 2020	Gorm Greisen	MD, Professor	

5. Signatures of the co-authors[iii]				
2.		Daniel Ostojic	Engineer, PhD	Dr. Daniel Ostojic
3.		Stefan Kleiser	Engineer, PhD	
4.		Martin Wolf	Engineer, Professor	

6. Signature of the principal supervisor

I solemnly declare that the information provided in this declaration is accurate to the best of my knowledge.

Date: 15-jan- 2020
Principal supervisor: Gorm Greisen

7. Signature of the PhD student

I solemnly declare that the information provided in this declaration is accurate to the best of my knowledge.

Date:

PhD student: Mathias Lühr Hansen

Digitalt signeret af Mathias Lühr Hansen
Dato: 2021.01.20 10:04:52 +01'00'

Please learn more about responsible conduct of research on the Faculty of Health and Medical Sciences' website.

---

[i] This can be supplemented with an additional letter if needed.

[ii] Please see Ministerial Order on the PhD Programme at the Universities and Certain Higher Artistic Educational Institutions (PhD Order) § 12 (4):

"Any articles included in the thesis may be written in cooperation with others, provided that each of the co-authors submits a written declaration stating the PhD student's or the author's contribution to the work."

[iii] If more signatures are needed please add an extra sheet.