The SafeBoosC phase II clinical trial: an analysis of the interventions related with the oximeter readings

Joan Riera,1,2 Simon Hyttel-Sorensen,3 María Carmen Bravo,1 Fernando Cabañas,4 Paloma López-Ortego,1 Laura Sanchez,1 Marta Ybarra,1 Eugene Dempsey,5 Gorm Greisen,3 Topun Austin,6 Olivier Claris,7 Monica Fumagalli,8 Christian Gluud,9 Petra Lemmers,10 Gerhard Pichler,11 Anne Mette Plomgaard,1 Frank van Bel,10 Martin Wolf,12 Adelina Pellicer1

ABSTRACT
Background The SafeBoosC phase II randomised clinical trial recently demonstrated the benefits of a combination of cerebral regional tissue oxygen saturation (rStO2) by near-infrared spectroscopy (NIRS) and a treatment guideline to reduce the oxygen imbalance in extremely preterm infants.

Aims To analyse rStO2-alarm-related clinical decisions and their heterogeneity in the NIRS experimental group (NIRS monitoring visible) and their impact on rStO2 and SpO2.

Methods Continuous data from NIRS devices and the alarms (area under the curve of the rStO2 out of range had accumulated 0.2%/h during 10 min), clinical data at discrete time points and interventions prompted by the alarms were recorded.

Results Sixty-seven infants had data that fulfilled the requirements for this analysis. 1107 alarm episodes were analysed. The alarm triggered a treatment guideline intervention in 25% of the cases; the type of intervention chosen varied among clinical sites. More than 55% of alarms were not followed by an intervention (‘No action’); additionally, in 5% of alarms the rStO2 value apparently was considered non-reliable and the sensor was repositioned. The percentage of unresolved alarms at 30 min after ‘No action’ almost doubled the treatment guideline intervention (p<0.001). Changes in peripheral oxygen saturation (SpO2), were observed only after treatment guideline interventions.

Conclusions This study shows that 25% of rStO2 alarms were followed by a clinical intervention determined by the treatment guideline. However, the rStO2 and SpO2 returned to normal ranges after the intervention, supporting the notion that decisions taken by the clinicians were appropriate.

Trial registration number ClinicalTrials.gov NCT01590316.

INTRODUCTION
The SafeBoosC phase II randomised clinical trial was a pragmatic medical device trial to obtain evidence-based knowledge on the benefits and harms of cerebral monitoring using near-infrared spectroscopy (NIRS) as a part of routine management of premature infants.1 NIRS-derived absolute regional cerebral tissue oxygen saturation (rStO2) is a composite measure of tissue oxygen saturation across arterial, capillary and venous beds and reflects a balance between cerebral oxygen delivery and consumption.2–3 The aim of the SafeBoosC phase II randomised clinical trial was the stabilisation of cerebral oxygenation by the combined use of rStO2 monitoring and an evidence-based treatment guideline6 as a brain-oriented protection strategy with the goal of reducing fluctuations in oxygen delivery and consumption.7–9

The treatment guideline was developed focusing on interventions that could have an impact on arterial oxygen saturation (SpO2), haemoglobin concentration and cerebral blood flow, which are the main determinants of oxygen delivery to the brain. This treatment guideline suggested potential interventions to be considered by the physicians within the range of the policies currently in place in their units.6

The SafeBoosC phase II randomised clinical trial demonstrated that cerebral oxygenation can be stabilised in extremely preterm infants during the first 3 days after birth with a median 58% reduction in the burden of hypoxia and hyperoxia in the experimental group in comparison with the control group.1 The purpose of this report is to analyse the trial-related clinical decisions and their impact on...
the infants’ condition (rStO₂ and SpO₂) as well as potential heterogeneity among the participating centres with a focus on the optimal design for a large scale, multicentre trial with a patient-relevant outcome.

METHODS
Design and study population
The trial included infants from eight different European countries. It was approved by each hospital’s local research ethics committee, and, where required (Austria, Denmark, France, and Switzerland), by the competent authority responsible for medical devices. Details regarding the trial protocol, clinical characteristics of infants and main outcomes have been reported elsewhere. Recruitment was conducted between June 2012 and December 2013. Briefly, infants of gestational age up to 27 weeks and 6 days, with the possibility to start cerebral NIRS monitoring within 3 h after birth, who had written informed consent, were considered eligible for the trial. By means of a central web-based randomisation system infants were allocated to either the experimental group, with the cerebral NIRS monitoring visible, or the control group with the cerebral NIRS monitoring blinded.

Software analysing procedure and study workflow
In the experimental group the alarm was activated when the area under the curve (AUC) of the rStO₂ out of normal range, defined as rStO₂ between 55% and 85%, had accumulated 0.2%h during the past 10 min. The units are expressed as %h the AUC and generated the alarm. The software also showed that listed in the treatment guidelines was implemented, it was documented as ‘Other’. If after careful evaluation, the attending staff judged that there was no reason to believe that the infant was compromised, then the decision was ‘No action’.

A dataset was recorded every 6 h according to the protocol and included heart rate, peripheral oxygen saturation (SpO₂), mean arterial blood pressure, mean airway pressure, inspired oxygen fraction, and pH, blood gases, lactate, base excess, haemoglobin and glycaemia if available. These data were also requested at the start of the alarm and 30 min later (first review); in case the alarm was not resolved (AUC still above 0.2%h in the last 10 min) within the first 30 min a second review dataset was also requested 30 min after the first review. The bedside software programme displayed the quality of the NIRS values provided by the device; this prevented missing values and false readings that could lead to inappropriate interventions. The various cerebral oximeters used in the study were comparable in terms of absolute value, repeatability and sensitivity to changes in oxygenation on the adult, human forearm.

Local staff was trained as appropriate to supervise all participants during the intervention. To minimise skin irritation related to the device it was recommended to move the NIRS sensor to a different location at the skull as often as every 4 h to avoid damage from heat and/or pressure. Each NIRS sensor reposition was also documented.

Data analysis
Each site downloaded data to the data management centre after completion of the participant’s intervention period. Using a Matlab (MathWorks, Natick, Massachusetts, USA) script created for this purpose, NIRS files, event files (decisions and sensor reposition information) and dataset were preprocessed. Data were reviewed manually to check data consistency.

Sensor reposition that occurred within the 10 min period following a previous sensor site reposition was considered a single event. In addition, NIRS sensor reposition occurring within the 10 min period that followed an alarm in which the decision taken was ‘No action’ was analysed separately; this was considered as an intervention that was termed ‘Alarm-prompted reposition’. In these cases the alarm apparently was judged to be the consequence of poor sensor–skin contact yielding to non-reliable warning. Alarms occurring within the 15 min period following sensor repositioning were discarded.

Statistics
The data were analysed using the statistical software SAS V9.2 (Cary, North Carolina, USA). The quantitative data are given as means (SD or 95% CI) or median (IQR) and the qualitative data as counts and percentages. Mixed-model analysis was used to adjust the effect of repeated measures per centre or total population. To avoid the effect of variations in study duration among infants results are expressed as ‘per time unit’.

In addition to rStO₂, the peripheral SpO₂ was chosen as the physiological variable to evaluate the impact of warnings and...
eventual clinical decisions on the infants’ status. The peripheral SpO2 value was calculated as the averaged mean value entered at the time of the alarm and at 30 and 60 min of warning. The rStO2 value was derived from the averaged mean of all samples obtained during 1 min at the time datasets were entered. The unresolved alarm by time according to type of decision taken by clinicians was represented using the Kaplan and Meier method; differences were assessed by means of the log-rank test.

RESULTS

Among the 86 infants included in the SafeBoosC trial in the NIRS experimental group,1 67 (78%) had data that fulfilled the requirements for this analysis. The remaining 19 infants were excluded due to inappropriate use of the study software. Infants were recruited in eight different centres, with a variable number of infants in the experimental group per centre (ranging from 2 to 18). The mean duration of NIRS recordings during the intervention period was 70.24 (SD 9.22) hours without differences between centres. The clinical characteristics of the study population are detailed in table 1.

Periodical repositioning per protocol of the NIRS sensor, as suggested in the protocol, varied according to centre (p<0.001), with a median of 0.18 (IQR 0.19) repositions per hour (ranges 0.0–0.28). No differences in sensor reposition according to NIRS device were found.

A total of 1107 alarm episodes (cerebral hypoxia, 874; cerebral hyperoxia, 233) were analysed from the 67 infants. The decision taken in relation to alarm varied according to centre (p=0.001) (table 2). The alarm triggered a treatment guideline intervention in 25% of the total alarm episodes (figure 1). In most treatment guideline interventions the choice was a change in FiO2, consisting of an increase in FiO2 in 177 (88.1%) episodes and a decrease in FiO2 in 24 (11.9%) episodes.

Statistically significant differences were found among centres in relation to the type of intervention chosen (p=0.001). Centres 1, 2, 4, 6 and 7 were more prone to FiO2 adjustment, whereas centres 3 and 5 chose more often a different balance between FiO2 and ventilator settings. The time needed for the alarm to resolve varied according to type of decision taken (p<0.001) (figure 2).

We explored the impact of the decision taken on SpO2. The averaged mean SpO2 at start of the alarm warning was 90.0% (SD 9.6), while it was 92.4% (SD 6.8) at the first (30 min) and 92.9% (SD 6.7) at the second (60 min) review. However, the profile of SpO2 and rStO2 changes showed differences according to the decisions taken in response to alarm (figure 3). Higher FiO2 was associated with lower rStO2 (p=0.057) and

---

**Table 2** Decision taken by clinicians in response to alarm per centre

| Centre | Alarm/h per patient (95% CI) | Total n (%) | FiO2 adjustment n (%)* | No action, n (%) | Other, n (%) | Alarm-prompted reposition, n (%) |
|--------|-----------------------------|-------------|------------------------|-----------------|-------------|---------------------------------|
| 1      | 0.32 (0.07 to 0.55)         | 38 (24)     | 35 (92)                | 100 (63)        | 16 (10)     | 5 (3)                            |
| 2      | 0.39 (0.22 to 0.56)         | 78 (22)     | 65 (83)                | 196 (56)        | 66 (19)     | 13 (4)                           |
| 3      | 0.09 (0.06 to 0.11)         | 30 (26)     | 9 (30)                 | 33 (29)         | 26 (23)     | 26 (23)                          |
| 4      | 0.16 (0.06 to 0.27)         | 31 (42)     | 27 (87)                | 38 (52)         | 4 (5)       | 0 (0)                            |
| 5      | 0.19 (0.12 to 0.28)         | 40 (22)     | 15 (37)                | 115 (65)        | 20 (11)     | 3 (2)                            |
| 6      | 0.33 (0.02 to 0.58)         | 18 (19)     | 16 (89)                | 57 (61)         | 16 (17)     | 2 (2)                            |
| 7      | 0.24 (0.01 to 0.46)         | 37 (37)     | 32 (86)                | 49 (49)         | 9 (9)       | 4 (4)                            |
| 8      | 0.32 (0.12 to 0.90)         | 4 (11)      | 2 (50)                 | 27 (73)         | 2 (5)       | 4 (11)                           |
| Total  | 0.23 (0.18 to 0.29)         | 276 (25)    | 201 (73)               | 615 (56)        | 159 (14)    | 57 (5)                           |

*FiO2 adjustment expressed as percentage within the treatment guideline interventions.
SpO₂ (p<0.001). Table 3 shows rStO₂ and SpO₂ according to the level of supplemental oxygen at the time of dataset.

DISCUSSION

This report focuses on the alarm episodes analyses and the decisions taken in the experimental group (visible NIRS) of the SafeBoosC phase II randomised clinical trial. The alarm warned the clinical staff of situations where the brain could be exposed to hypoxia or hyperoxia, and potential clinical interventions were suggested in a treatment guideline to normalise an out of range rStO₂. However, more than 55% of the warning episodes were not followed by an intervention. Two important aspects need to be considered. First, the infants in the study were monitored by other devices apart from NIRS and only the interventions triggered by NIRS warnings were recorded in the SafeBoosC phase II trial software. Second, the NIRS alarm accumulated the values of rStO₂ obtained over the preceding 10 min, while the other monitoring systems did not (real-time reading and immediate alert system). These facts raise two important limitations of the present study. On the one hand, if compromised cerebral perfusion–oxygenation was suspected, successful treatment interventions may have been driven by other monitoring devices, such as transcutaneous blood gas monitoring, pulse oximetry or others, and been implemented before the NIRS alarm warned the clinicians (not enough time to accumulate an out of range rStO₂). On the other hand, if the intervention triggered by other devices was not rapidly effective and the infant’s impaired condition was maintained for a longer period, the accumulated out of range rStO₂ would have triggered an alarm; in this case, the clinician eventually might have decided not to take any further action because appropriate action had already been implemented some minutes earlier. This may falsely inflate the ‘No action’ response. As there is no alternative monitoring system for the continuous assessment of brain oxygenation the possibility to address these issues is limited. We would like to remark that evidence supporting the treatment guideline interventions is relatively low and cannot be recommended outside a research setting.

The time needed for the alarm to be resolved varied according to the type of the decision taken (figure 2). Sensor repositioning (‘Alarm-prompted reposition’) returned rStO₂ to normal values (55%–85%) within 30 min in 80% of alarms and within 100 min in all alarms. The percentage of alarms not resolved at 30 and 60 min after ‘No action’ were almost double compared with those in which there was an intervention (‘Treatment guideline intervention’). Yet, despite an intervention, 15% of these were unresolved by 60 min. This means that even if the intervention chosen was appropriate, the mechanism involved in improving brain perfusion–oxygenation takes more time to be fully resolved. Finally, in some very ill infants, it is likely that cerebral hypoxia has been resistant to all interventions. It is important to note, however, that data collection may have varied both within and between centres. As it was manually entered, some interventions may not have been properly documented. Furthermore, the 30 and 60 min points are likely to be insufficient to catch the details of multiple interventions. In the design of the study we chose those to avoid overloading the clinical staff with registration and thereby compromising the primary purpose of the trial.

To further evaluate the impact of alarms and decisions we analysed the evolution of SpO₂ and rStO₂ at the time of the alarm and the subsequent reviews according to the type of decision taken by the clinician. As illustrated in figure 3, when the decision was a ‘Treatment guideline intervention’ the averaged mean SpO₂ showed the greatest change, from the lowest mean value at start of warning to a more physiological and stable post-warning average value at first (30 min) and second (60 min) reviews. Interestingly, when the decision was ‘No action’, the averaged mean SpO₂ did not significantly vary, and remained during the whole intervention period within ‘acceptable’ clinical ranges. More obviously, repositioning the NIRS sensor did not change rStO₂. Thus, trends in SpO₂ support the contention that the decisions taken by clinicians were appropriate in most cases. As the average SpO₂ at the time of the rStO₂ alarm was as low as 85%, it is not a surprise that the most common intervention was to increase FiO₂. It is more of a surprise that the FiO₂ was not increased before the rStO₂ alarm as most units target SpO₂ at 90%–95%.

We found a negative association between FiO₂ and both rStO₂ and SpO₂. By stratifying the group into (a) infants in room air, (b) infants on low supplementary oxygen (22%–29%) and (c) infants on high oxygen supplementation (>29%), we found that the impact of FiO₂ on rStO₂ and SpO₂ differed. The healthier the infant was (and so, the lower the supplementary
oxygen needs) the higher the associated rStO2 and SpO2 values were (Table 3). This also supports the notion that other treatment strategies suggested by the treatment guidelines, apart from FiO2 adjustment, contributed to the return of rStO2 to normal ranges, particularly in the sickest infants.

In our experience, the rStO2 alarm defined in this study protocol was rarely used to initiate treatment of a patent ductus arteriosus or guiding glucose intake. The former is more dependent on the variations in neonatal intensive care unit policies with regards to early echocardiographic assessment in these infants.

We observed wide differences in sensor repositioning among the participating centres. This difference was not explained by the brand of NIRS device used but possibly by differences in sensor fixation, registration of repositioning or compliance with the standard operating procedure among centres.

On average four interventions for infants in the experimental group achieved the 58% reduction in the AUC of hyperoxia and hypoxia that we previously reported.1 As previously stated our data collection was not perfect and an in-depth analysis of the complex interaction of cerebral NIRS and the already available monitoring systems is difficult. From the results presented here it seems that cerebral NIRS monitoring did add some information that resulted in additional interventions. This was expected. Treatment was not changed in 75% of alarms. This may seem high, but it would on the other hand be surprising if out of range NIRS always resulted in interventions in such complex clinical scenarios as presented by this study population.

Whether the NIRS monitoring also hindered some interventions is uncertain. Finally, we observed heterogeneity in the number of alarms between centres. We speculate this could be due to differences in clinical staff training on NIRS monitoring among the study participants1 and in the severity of infant conditions. The patient-relevant clinical benefit of this complex intervention remains to be determined.

The absolute values provided by cerebral NIRS have shown poor repeatability and differences between the varied devices.15 16 To deal with this, systematic pretrial testing was conducted using several NIRS instruments.11 17 Only two NIRS devices were used in this study reducing heterogeneity. We have to depend on the absolute values of cerebral tissue oxygenation and on absolute thresholds for intervention.

CONCLUSIONS

This post hoc analysis of the alarm episodes and their subsequent management in the experimental group of the SafeBoosC II randomised clinical trial revealed that three quarters of all alarms triggered were not followed by an active intervention at the bedside, suggesting that the treating clinician was content with the infant’s overall status despite a cerebral oxygen outside

| Table 3 Cerebral (rStO2) and peripheral (SpO2) oxygen saturation according to level of inspired supplemental oxygen (FiO2) |
|-----------------|-----------------|-----------------|
| FiO2 value (%)  | rStO2% (SD) (n) | SpO2% (SD) (n)  |
| 21              | 64.9 (15.9) (580) | 94.5 (5.2) (711)* |
| 22--29          | 67.3 (14.8) (376)† | 92.8 (4.6) (447) |
| ≥30             | 63.8 (11.8) (149)† | 92.0 (4.1) (180) |

*Differences between 21% and the other two FiO2 strata (p<0.001).
†Differences between 21% and 22%–29% FiO2 strata (p=0.001).
of the monitoring range. However, 25% of the alarms triggered a treatment guideline intervention, averaging four interventions per infant, supporting the strength of this non-invasive monitoring system. Limitations of the study design, however, restrict the generalisability of the results. For a large-scale trial of monitoring of cerebral oxygenation with patient-relevant outcomes study staff must be adequately trained in the application of NIRS and the pathophysiological basis for the choice of intervention.

Author affiliations
1 Department of Neonatology, La Paz University Hospital, Madrid, Spain
2 Center for Biomedical Technology, Technical University, Madrid, Spain
3 Department of Neonatology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark
4 Biomedical Research Foundation, La Paz University Hospital-IDiPAZ, Madrid, Spain
5 Department of Paediatrics and Child Health, University College Cork, Cork, Ireland
6 Roos Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK
7 Department of Neonatology, Hospital Femme Mère Enfants, Bron, France
8 NICU, Fondazione IRCCS Ca’ Granda Ospedale Maggiore Policlinico, Milan, Italy
9 Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark
10 University Medical Center Utrecht, Wilhelmina Children’s Hospital, Utrecht, Netherlands
11 Department of Pediatrics, Medical University of Graz, Graz, Austria
12 Biomedical Optics Research Laboratory, Division of Neonatology, University Hospital Zurich, Zurich, Switzerland

Acknowledgements
All authors critically reviewed the revised versions of the paper and approved the final draft for submission. All the authors had access to the data and contributed substantially to the submitted report. There were no financial relationships with any organisations that might have an interest in the submitted work in the previous 3 years and no other relationships or activities that could appear to have influenced the submitted work.

Contributors
JR, SH-S, AP, TA, FvB, DC, ED, MF, CG, PL, AMP, GP, MIW and GG conceived and designed the study. JR designed the boxes for blinding the device displays and designed and maintained the NIRS data collection software. JR, SH-S, AP, TA, FvB, DC, ED, MF, CG, PL, GP, AMP, MCB, FC, PLO, LS, MY and GG assisted with acquisition of the data. JR, MCB, FC, PLO, LS, MY and AP analysed the data. JR and AP wrote the initial draft. All authors critically reviewed revised versions of the paper and approved the final draft for submission. All the authors had access to the data and contributed substantially to the submitted report. JR and AP are the guarantors.

Funding
This work was supported by an unconditional and unrestricted grant from the Danish Council for Strategic Research (DKK 11 100 105) and the SAMID network (RD08/0072/0018 and RD12/0026/0004). The funder had no role in the design, conduct, or analysis of the trial.

Competing interests
None declared.

Patient consent
Obtained.

Ethics approval
This study was approved by the local ethics committees: clinical research ethics committee of the Cork Teaching Hospitals (EMC 3 (03/07/12)); NRES Committee East of England—Cambridge South (12/EE/0329); Ethikkommission Medizinische Universität Graz (24-261 ex (11/12)); De Videnskabelses Komiteen—Region Hovedstaden (H-4-2012-028); CFP Süd-Ext III—Lyon (HCU 2012.728); Direzione Scientifica-Comitato di Etica—The IRCCS Fondazione Ca’ Granda Ospedale Maggiore Policlinico (549/12-all.3); Medisch Ethische Toetsingscommissie—Universitair Medisch Centrum Utrecht (WAG/cr/12/036774); and Comité Etico de Investigación Clínica del Hospital Universitario “La Paz” de Madrid (15/2011). All participants gave informed consent before taking part in the trial.

Provenance and peer review
Not commissioned; externally peer reviewed.

Data sharing statement
Dataset and alarm values are available from the corresponding author on request. Consent was not obtained but the presented data are anonymised and risk of identification is low.

Open Access
This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/