ACTA PÆDIATRICA
NURTURING THE CHILD

Acta Pædiatrica ISSN 0803-5253

REVIEW ARTICLE

Safe oxygen saturation targeting and monitoring in preterm infants: can we avoid hypoxia and hyperoxia?

Augusto Sola (augustosolaneo@gmail.com)[1,2], Sergio G. Golombek[1,2,3], María Teresa Montes Bueno[4], Lourdes Lemus-Varela[1,5], Claudia Zuluaga[6], Fernando Domínguez[1,7], Hernando Baquero[1,8], Alejandro E. Young Sarmiento[1,9], Diego Natta[1,10], Jose M. Rodriguez Perez[1,11], Richard Deulofeut[12], Ana Quiroga[13], Gabriel Lara Flores[1,14], Mónica Morgues[1,15], Alfredo García-Alix Pérez[16], Bart Van Overmeire[17], Frank van Bel[18]

1.Ibero American Society of Neonatology (SIBEN), Dana Point, CA, USA
2.New York Medical College, Valhalla, NY, USA
3.Maria Fareri Children's Hospital at Westchester Medical Center, Valhalla, NY, USA
4.Nursing Council, SIBEN, Hospital La Paz, Madrid, Spain
5.Hospital de Pediatría del Centro Médico Nacional de Occidente, IMSS, Guadalajara, Jalisco, México
6.Pediatric Ophthalmology, Universidad del Valle, Cali, Colombia
7.Pediatrics, Hospital González Coro, Universidad De La Habana, Habana, Cuba
8.Neonatology Department, Universidad del Norte, Barranquilla, Colombia
9.Neonatal ICU, Hospital Escuela Universitario, Tegucigalpa, Honduras
10.Pediatrics, Hospital Privado de la Comunidad, Mar del Plata, Argentina
11.Stella Maris Hospital, International Neurodevelopment Neonatal Center (CINN), Sao Paulo, Brazil
12.Neonatology, Pediatrix Medical Group North Dallas Practice, Dallas, TX, USA
13.Nursing Council of SIBEN, Universidad Austral, Buenos Aires, Argentina
14.Neonatology, Hospital Ginecología-Obstetricia 4 IMSS, Mexico City, Mexico
15.Pediatrics and Neonatology, Master in Epidemiology, University of Chile, North Campus, Santiago, Chile
16.Hospital Sant Joan de Déu, Pediatrics, Barcelona University, Barcelona, Spain
17.Neonatology, Erasme University, Brussels, Belgium
18.Perinatal Center, University Medical Center Utrecht, Utrecht, The Netherlands

Keywords
Hyperoxia, Oximetry, Oxygen saturation, Premature infant, Retinopathy of prematurity

Correspondence
Augusto Sola, Ibero American Society of Neonatology (SIBEN), 24896 Sea Crest Dr, Dana Point, CA 92629, USA.
Tel: +949 433 8962 |
Fax: +949 276 6466 |
Email: augustosolaneo@gmail.com

Received
19 February 2014; revised 25 April 2014; accepted 13 May 2014.

DOI:10.1111/apa.12692

ABSTRACT

Oxygen is a neonatal health hazard that should be avoided in clinical practice. In this review, an international team of neonatologists and nurses assessed oxygen saturation (SpO_2) targeting in preterm infants and evaluated the potential weaknesses of randomised clinical trials.

Conclusion: SpO_2 of 85–89% can increase mortality and 91–95% can cause hyperoxia and ill effects. Neither of these ranges can be recommended, and wider intermediate targets, such as 87–94% or 88–94%, may be safer.

INTRODUCTION

The toxic nature of excess oxygen has been causing concern for more than 200 years. Despite these concerns, oxygen is the most commonly used drug in neonatal intensive care units (NICUs) around the world. We sometimes fail in our laudable attempts to prevent and avoid hypoxia by unnecessarily exposing newborn infants worldwide to levels of oxygen that are too high and to hyperoxic states. Hyperoxia is caused by healthcare providers, it does not occur in nature, and evolution has not equipped the body to deal with it.

J H Comroe, a pioneer in clinical respiratory physiology, espoused the dictum that 'No drug produces only the effect

Abbreviations
BOOST II, Benefits of Oxygen Saturation Targeting; COT, Canadian Oxygen Trial; FDA, Food and Drugs Administration; FiO_2, Fraction of inspired oxygen; NICU, Neonatal intensive care unit; RCTs, Randomised control trials; ROP, Retinopathy of prematurity; SET, Signal extraction technology; SpO_2, Oxygen saturation; SUPPORT, SUPPORT Study Group of the Eunice Kennedy Shriver NICHD Neonatal Research Network.

Key notes
- Hyperoxia is a neonatal health hazard that should be avoided in clinical practice.
- International neonatologists and nurses assessed SpO_2 targeting in preterm infants and concluded that SpO_2 targets of 85–89% and 91–95% should not be used due to their potential associations with increased mortality and hyperoxic ill effects, respectively.
- Without knowing the optimal SpO_2 targets, wider intermediate targets, such as 87–94% or 88–94%, may be safer.

for which it is prescribed. It invariably affects other functions and organ systems' (1,2). In 1955, he also stated that 'Pulmonary physiologists understand pulmonary physiology reasonably well. Many doctors and medical students do not' (3–5). It is disappointing that a similar statement could be applied today to neonatal oxygenation physiology. As early as 1939, Comroe demonstrated the deleterious effects of oxygen, when he showed that oxygen provided during asphyxia-induced aortic chemoreceptor discharge caused an abrupt decrease in arterial blood pressure (6). Furthermore, in the early 1950s, he warned us about the adverse effects of oxygen in the lung, with uneven alveolar ventilation during a single breath of 100% oxygen, and on respiratory control and mental status, with mental changes during oxygen therapy (3–5).

Since then, we must constantly remind ourselves that too much of a good thing can be poisonous. Hyperoxaemia can be pernicious, and there is no clinical or pathophysiological condition known to cause it. Therefore, each time undue oxidant stress occurs, it is likely to be caused by excess or unnecessary oxygen administered by healthcare providers or by reperfusion after hypoxaemia.

In 2007, we stated that inappropriate oxygen use is a neonatal health hazard associated with many morbidities and that neonatal exposure to pure oxygen, even briefly, or to pulse oximetry of more than 95% when breathing supplemental oxygen must be avoided as much as possible (1). The highly positive editorial that accompanied our paper (7) said that our manuscript 'should be obligatory reading for everyone working with newborn intensive care' and added that 'we should all take a deep breath and consider Sola et al.'s proposal to stop inducing hyperoxia in any newborn infant as long as we do not know the short- and long-term consequences of such a practice' (7).

Unfortunately, despite many older and recent studies, there is no evidence to demonstrate the perfect oxygen saturation (SpO_2) target for daily neonatal practice, and SpO_2 targeting in preterm infants remains a very controversial worldwide topic. However, findings of recent randomised clinical trials (RCTs) (8–11) and published commentaries and meta-analyses (12,13) have added significant insight and allow for discernment.

While investigators are looking for further improvements, neonatal oxygenation is not well understood by many caregivers, and oxygen is often given at fickle, capricious or whimsical doses, with SpO_2 interpreted erroneously in many newborns. Of course, trying to avoid hyperoxaemia does not equate to permitting hypoxaemia. Through the years, it has become evident that no neonatal bedside care provider leaves an infant unattended if hypoxia is suspected or confirmed. However, the same does not always apply when hyperoxaemia may be present, despite advancing knowledge of cerebral oxygenation, extraction and auto regulation in very preterm infants. It is becoming possible to assess the presence of excess oxygen in the brain, but this is not yet widely used by clinicians. We have been concerned, recently, to learn that high regional cerebral saturation has been potentially associated with peri-intraventricular haemorrhage and with poor outcome at 2 years of age (14).

Regrettably, blindness due to retinopathy of prematurity (ROP) is a serious and epidemic health problem in Latin America and many other regions of the world. At the moment, many neonatal clinical care providers in North America, Central and South America and Europe are somewhat confused about which SpO_2 target should be used. It is likely that changed clinical protocols are, once again, exposing many babies to unnecessary and potentially detrimental hyperoxia. Our current level knowledge connotes that we do not know the ideal SpO_2 target range and cannot formulate best practice in this area. Instead, it may be prudent to avoid and eradicate possible bad practices that could prove detrimental to newborn infants. For example, recommending a low or narrow SpO_2 target range is not realistic in clinical practice and could have ill effects for preterm infants worldwide.

The aim of this review is to provide an additional view, supported by objective data, on this complex and unresolved issue. We present salient points from recent RCT's and highlight their relevant differences and the likely reasons for those differences. This enables us to summarise and clarify points on the value on SpO_2 targeting for clinical practice. Finally, we also include some clinical concepts regarding SpO_2 monitoring and comments on why some RCTs may be flawed, based on the views presented by Ioannidis et al. (15). Our objective is that newborn infants should not be exposed to hypoxia at all and that potentially damaging hyperoxia should also be avoided as much as possible during daily neonatal clinical practice.

EVIDENCE FROM THE SUPPORT, BOOST II AND COT

This section looks at five major randomised controlled trials that were carried out as masked, prospective evaluations comparing two SpO_2 target ranges – 85–89% vs 91–95% – in a large number of infants of <28 weeks of gestation. They are the US SUPPORT Study Group of the Eunice Kennedy Shriver NICHD Neonatal Research Network (SUPPORT), the Benefits of Oxygen Saturation Targeting (BOOST II) trial, which covered the UK, Australia and New Zealand and the Canadian Oxygen Trial (COT) (8–11). There were five RCTs, but the neonatal outcomes of the three BOOST II trials were pooled and reported as a single study. We understand that this was a deviation from the original plan and was prompted by the interim subgroup analysis. This review refers to the findings of the studies as they have been published to date (8–11). All these RCTs compared the results between infants randomly assigned to the lower SpO_2 target (85–89%) and the higher intention to treat (91–95%) target. The targets studied were narrow, and no other SpO_2 ranges or targets were compared.

In all the trials, the oximeters were modified to display and store SpO_2 levels that were either 3% higher or lower than the true values. True SpO_2 values were displayed if the measured SpO_2 values were <84% or more than 96%. During two trials (10,11), the algorithms of the SpO_2

monitors were changed, both algorithms functioning within specifications approved by the US Food and Drugs Administration (FDA). The primary outcome variables were similar among the trials and comprised the composite outcome of the death or disability rate at 18 or 24 months. In the SUPPORT trial, the primary outcome of the short-term part of the trial was a composite of severe ROP, death before discharge from the hospital or both. The primary composite outcome for the longer-term analysis was death before assessment at 18–22 months or neurodevelopmental impairment at 18–22 months of corrected age.

SUMMARY OF THE KEY FINDINGS REPORTED BY THE RCTS

The two groups were separated and analysed by randomised assignment and intention to treat. However, there were varying degrees of SpO_2 overlap between the two groups, with variable percentages of time when the infants assigned to the SpO_2 target ranges of 85–89% were in the 91–95% range and vice versa. The overlap was not the same for all studies, but in all the studies, the low-target-range group had median SpO_2 readings at, or beyond, the upper limit of this range. This meant that the separation between the SpO_2 readings was around 3–4%, instead of 6% as intended. Additionally, the amount of time spent with SpO_2 below 85% in the lower-target group and with SpO_2 above 95% in the higher-target group was not the same when we compared the findings published by the RCTs. This is not at all unexpected. It has been well described previously that, in clinical practice, achieved versus intended SpO_2 is very variable and depends on several factors.

In the SUPPORT study, an SpO_2 target of 85–89% was associated with increased mortality at the time of discharge, compared with the higher target of 91–95% (19.9% vs 16.2%; $p = 0.045$). In the BOOST trial, recruitment was stopped early when an interim analysis showed an increased death rate at 36 weeks in the group with lower oxygen saturation. The authors analysed pooled data and reported hospital discharge outcomes (11) and found that mortality was higher in the lower-target group (23.1% vs 15.9%, relative risk: 1.45 95% confidence interval 1.15–1.84; $p = 0.002$). It is unclear why the death rate was higher in the lower-target group and the best estimate of the effect of SpO_2 on mortality from these trials is unknown. Other interventions that influence oxygen targeting may also influence mortality (11). The COT (10) reported findings at 18 months, after the ascertainment of the primary outcome variable, and reported no difference in mortality between the lower and higher targets. No study has shown that targeting any other lower range apart from 85 to 89%, for example 85% to 93% or 94%, has been associated with increased mortality. On the other hand, several descriptive clinical studies showed decreased rates of ROP, together with decreased exposure to oxygen, ventilator days and incidence of bronchopulmonary dysplasia, without any association with mortality, when SpO_2 targets were 85–94% or similar (16–25).

In the SUPPORT and BOOST II trials, SpO_2 targets of 91–95% were associated with a higher incidence of severe ROP compared with SpO_2 levels of 85–89%, at the same time points mentioned above. Therefore, if clinicians could only chose one of the ranges investigated by the published RCTs, targeting an SpO_2 range of 91–95% is clearly safer than 85–89%. On the other hand, the COT did not find such a difference at 18 months and does not support this suggestion. Importantly, none of the RCTs, and no other published studies, have shown that an intended SpO_2 of 91–95% is the safest range for clinical practice and that using any other target is a bad choice. Not using SpO_2 of 85–89% does not necessarily mean that infants have to be treated with a desired SpO_2 target of 91–95%.

Additionally, there were no differences in neurodevelopment, intracranial haemorrhage, patent ductus arteriosus or bronchopulmonary dysplasia between the two target groups. Necrotising enterocolitis was more frequent with SpO_2 85–89% in BOOST II, but not in the COT or SUPPORT trials. Finally, when the SUPPORT investigators reported longer-term results from their prespecified hypotheses (9), they found no significant differences in the composite outcome of death or neurodevelopmental impairment among extremely premature infants randomly assigned to a lower- or higher-target range of oxygen saturation at 18–22 months. However, as reported at 36 weeks (8), mortality was modestly increased in the lower-target group (85–89%) and severe ROP was greatly increased in the higher-target group (91–95%). COT (10), as mentioned, found no differences in the composite outcome of death or disability at 18 months. The primary outcome variable has yet to be reported by BOOST II.

COMMENTS ON THE DIFFERENT FINDINGS IN THE RCTS

The differences between the findings of these RCTs may be due to many factors, including important differences in design, implementation and data handling. These issues may also affect the ability to perform and interpret meta-analyses, as briefly mentioned later (15,26–28).

The COT did not perform an interim analysis and ascertained the primary outcome variable before analysing any of the trial data. Some of the findings reported by the BOOST II trial were after interim analysis, repeated testing and with data-driven investigator-initiated questions. The possible impact of performing interim analysis is addressed towards the end of this paper. Other differences in the reported data include that the mortality rates were lower in the COT (15.9%) than the SUPPORT (20.1%) and BOOST II trials (23.1%). These dissimilar rates may also partly explain the differences in mortality. Furthermore, a smaller proportion of the COT infants had median SpO_2 <85% and >95%, and the distributions of SpO_2 in the two treatment groups overlapped less. Additionally, there were some differences in clinical protocols. For example, alarm settings varied, and the lower saturation thresholds that would trigger an alarm were not prescribed in the UK protocol, which formed part of BOOST II. These, and other

differences, may explain why there was no excess mortality in the low-target group in the COT or excess ROP in the high-target group.

SPO_2 MONITORING

The following observations are based on extensive peer reviewed publications, on what is summarised above and on previously acquired knowledge (17,19,28–46). There have been more than 200 publications on the subject since the late 1990s, and as space is limited in the journal, a fuller reference list is available from the corresponding author.

All SpO_2 monitors have an inherent bias when they function according to their specification. However, the failure to provide accurate readings varies between monitors. Also, SpO_2 monitors offer different accuracy levels and precision and differ in the time they take to respond and in the number of false alarms, missed events and detection of true events.

Specificity and sensitivity rates of false-positives and false-negatives and therefore receiver operating curves are also different between different SpO_2 monitors. Finally, they also differ in their ability to provide accurate measurements through motion and low perfusion. The SpO_2 monitor chosen by the investigators for the important RCTs covered by this review (8–11) features signal extraction technology (SET) and is manufactured by Masimo (Masimo Corporation, Irvine, CA, USA). More than 100 publications have stated that it functions more accurately and precisely than any other SpO_2 monitor, and it is approved by the FDA to measure through motion and low perfusion.

Due to bias and in line with the Gaussian curve of normal distribution, the SpO_2 read-out is within ± one standard deviation (SD) of the true arterial saturation 68% of the time. For example, when the SpO_2 read-out is 90%, and the equipment bias is 1%, 68% of the time the true arterial saturation will be 89%, 90% or 91%. If, on the other hand, the monitor has a bias of 3%, then when the monitor reads an SpO_2 of 90%, the arterial saturation will be between 87% and 93% for 68% of the time. In two individuals with exactly the same SpO_2 read-out, the true arterial SaO_2 could therefore be 1–3% different from each other for 32% of the time. Therefore, a difference of 1–2% in the reading of the SpO_2 monitor may be inconsequential. This lends support to using wider target ranges and suggests that it does not make much sense to spend much time arguing if an SpO_2 of 89% is significantly worse than an SpO_2 of 91% for an individual baby.

The algorithms of SpO_2 study monitors were changed during the randomised studies, but both algorithms function within specifications approved by FDA and do not affect care when addressing the needs of an individual baby. The COT, which did not performed any interim analysis, found no significant differences in the SpO_2 readings when it compared the software of the two algorithms, with differences of 0.3% and 0.5% in the low- and high-target groups, respectively. In addition, in the COT, replacing the oximeter software was not associated with improved targeting of the SpO_2 values in either group.

POINTS WITH POTENTIAL VALUE FOR CLINICAL CARE

Alarms

There are many alarms in the NICU over the course of a day, and this has been identified as a safety issue. Many of the alarms are not related to true clinical problems and are considered false alarms. There are also alarms that are true alarms but do not require any intervention. Obviously, the narrower the SpO_2 targeting range and alarm limits chosen, the greater the number of alarms. In addition, the high alarm is violated more frequently than the low alarm. For example, a study of alarm limits for pulse oximetry in very preterm infants at the Royal Women's Hospital in Melbourne, using an SpO_2 target of 85–94%, found that the lower alarm limit was set correctly 91% of time, but the upper limit only was set correctly 23% of the time. It is noteworthy that for 24% of the time the upper alarm limit was set at 100% (45). Due to legitimate concerns about hypoxia, many registered nurses will change the upper-limit alarm more frequently, but not the lower one, if it becomes difficult to maintain the infant within a narrow range. We also know that there is less response to high alarms in clinical care. All this is associated with an increased tolerance of SpO_2 ≥95%, which would increase the percentage of time with hyperoxia; as a study of 1000 samples found that at an SpO_2 > 94%, 60% of the samples had a partial pressure of oxygen in arterial blood (PaO_2) of more than 80 mmHg (19). *The alarm limits should be set no more than 1% or 2% above or below the chosen target range and should always be on.* The lower alarm should always be ≥85%. We consider that in infants breathing supplemental oxygen, the high alarm should be set at 95% to avoid PaO_2 of more than 80 mmHg, and the high alarm value chosen should always be operative while the infant is breathing a fraction of inspired oxygen (FiO_2) of more than 0.21. It cannot be overemphasised that it is clinically mandatory to choose and set alarms appropriately and to ensure the alarm system is always operative. Of course, when an infant previously receiving oxygen reaches room air, then the upper alarm limit needs to be deactivated.

One issue that may affect clinical practice in the NICU is that some monitors have a desaturation alarm limit as well as a low-limit alarm, and each can be set independently with different SpO_2 values. Some of these monitors have a soft volume and yellow light for the low-limit alarm. If the SpO_2 falls below the set desaturation limit, the alarm sound changes markedly, becoming strident, and the light turns red, signalling that there is a potentially serious situation. Interestingly, the high alarm only has a softer sound and a yellow light, with no option for a more strident alarm or red light.

Averaging time and sensitivity

Pulse oximeter saturation values are usually obtained by averaging preceding measurements. Some SpO_2 monitors

allow clinicians to modify the averaging time and sensitivity or both at the same time. This has an impact on the monitor's response and therefore on what the clinicians observe in the digital display in the same infant at the same time. With monitors that provide this function, the number of desaturations, and therefore alarms, is greater with the short, 2-sec averaging time. However, this does not reach significance when only desaturations of clinically significant duration are considered. On the other hand, a long averaging time of 16 sec reduces the detection of brief periodic desaturation events and of those of greater severity. It may also interpret a cluster of shorter events as a single, prolonged episode and, as a result, potentially overestimate the frequency of long events (40). Another study that recorded desaturations of SpO_2 <80% found that there were 339 when using an average time of 16 sec and 1958 when an average time of 3 sec was used. There was a significantly lower-pulse oximeter saturation nadir with the shorter averaging time, while the maximum duration was significantly longer with the 16 sec averaging time (44). These findings demonstrate that there are pros and cons with each of the settings, depending on the infant's condition. Careful attention should therefore be paid to averaging time and sensitivity in the monitors used for clinical care. The performances of monitors manufactured by different companies vary, and some have default settings that cannot be modified, and it is our responsibility to understand the monitors we use in clinical care. In the delivery room, when changes may occur rapidly during resuscitation, maximum sensitivity and an average time of 2 sec is recommended. However, this is not the case for a critically ill infant in the NICU, as these settings could alarm the clinician by alerting him or her to very brief and clinically insignificant episodes. In the NICU setting, an average time of eight to 12 sec with intermediate sensitivity may be adequate.

SPO_2 TARGETS

A narrow SpO_2 target of 85–89% should not be used clinically, as it may be associated with increased mortality. However, the COT trial findings do not support recommendations that targeting SpO_2 in the upper 80% range should be avoided. At the high end, there is no evidence that an SpO_2 target of 91–95% or 90–95% should be universally used in clinical practice. There are several reasons why an SpO_2 target of 91–95% is of concern. Firstly, it is a challenge to maintain infants in a tight or narrow SpO_2 target range (10,11,24,25,29,30), which is associated with more fluctuations in SpO_2 and periods of hypoxia and hyperoxia, with more time spent in hyperoxia. Moreover, this range will trigger more alarms, and it is likely that SpO_2 of more than 94% will occur more frequently if the low alarm is set at 89% and oxygen is increased when the alarm sounds. Secondly, this higher target was associated with significantly higher rates of severe ROP in two of the RCTs. These combined factors could lead to a resurgence of ROP. It could therefore be risky to develop guidelines that recommend SpO_2 targeting of 90–95% in clinical practice. In contrast, wider intermediate targets may allow for easier care and better compliance, and they have been associated with a decreased rate of severe ROP, without an increased morbidity or mortality (16–25).

In many clinical centres, there is no accurate recording and analysis of SpO_2 and significant under recognition of alarms. Nursing notes only account for 25–30% of true desaturation episodes (46), and neonatal medical notes reflect even less, at around 7%. Even if this is far from accurate, there are almost no recordings of higher than expected or intended SpO_2 in many NICUs around the world. It is therefore much more accurate to use normograms with SET monitors (24,25).

In clinical practice, adherence to planned SpO_2 targets varies between shifts and centres, mainly due to the SpO_2 targets chosen, differences in education, staff commitment to this issue, workload and accuracy and precision of monitors used. In addition, the education programmes in most nursing and medical schools emphasise the problems of hypoxia and the need to correct it, much more than the complications of hyperoxia. This means that actual versus intended SpO_2 levels are very variable (18,24,25,29,30). When the SpO_2 target is at the higher range, SpO_2 values occur much more frequently above than below target. One study of random observations, after clinical practice was changed to aiming for SpO_2 in the 88–93% range, showed that the target was achieved 72% of the time; SpO_2 was <85% 10% of the time and, it was 96–100% 18% of the time (18). Another study reported on 58 000 h recorded on 153 infants, with a target SpO_2 of 85–89%. The actual SpO_2 was 85–95% for 65% of the time, more than 95% for 14% of the time and <85% for 21% of the time. When the target SpO_2 was 91–95%, it was actually 85–95% *for* 54% of the time, more than 95% for 36% of the time and <85% for only 10% of the time (47). It is well known by clinicians that some infants breathing oxygen have more oscillation in SpO_2 than others, and it is impossible to maintain the SpO_2 within a narrow target like 90–93% for a prolonged period of time in most of them, even when the nurse to patient ratio is one-to-one, which very rarely happens in NICUs. Therefore, a narrow high target is difficult to achieve and can lead to increased tolerance of an SpO_2 level of more than 95%. To improve the achievement of planned targets, it is necessary to use wider targets and ensure that the whole care team functions with the same goals. In this way, measures can be taken when defined targets are not respected or ignored and when an individual baby is off target and the primary nurse is carrying out other important chores. All these measures are just as important during transport and any surgical or other procedures, to try to prevent hyperoxia and hypoxia (48). Finally, in addition to teamwork among all neonatal healthcare providers, including parents in the objective of avoiding hyperoxia may improve safety.

Finally, there are other important measures that will prevent hyperoxia. Pre-oxygenation is a dangerous and ineffective practice that should be eradicated in newborns of any gestational age, and the justification for this has been

summarised elsewhere (48). Hyperoxia should not be induced in any neonate during and after endotracheal tube suction or invasive procedures, including surgery, and it should not be allowed in the periods after extubation. Lastly, practices that can lead to severe and prolonged hyperoxaemia, like nitrogen washout without any proven benefit in important outcomes and the hyperoxic challenge test, should also be eliminated from clinical practice.

RECOMMENDATIONS FOR CARE PROVIDERS REGARDING SPO_2 RANGES, ALARM LIMITS AND VARYING AVERAGING TIME

Based on all the available evidence, the most prudent approach would be not to tolerate possible hypoxia and not to accept SpO_2 values associated with potential hyperoxia in infants breathing supplemental oxygen. Some would like to keep a preterm baby at a steady SpO_2 of 91–92% most of the time, but we learnt a long time ago that things don't work out like that way in clinical practice. It may be sensible to consider a lower SpO_2 limit of >85% and a higher limit of ≤95% when a preterm infant is breathing FiO_2 of more than 0.21 (49). This is an intended range, which is very different to saying that an SpO_2 of 85% or 95% is normal. Rather than using narrower intended targets, like 85–89%, 90–93% or 91–95%, it may be safer to use wider intermediate SpO_2 targets such as 86–93%, 87–94%, 88–94% or something similar. However, these intermediate target ranges may be not very different from one another, due to the inherent bias of SpO_2 monitoring and the significant difference between monitors we described briefly above. The reason why we mention several ranges is that they are, or were, used in different clinical centres in the real world. For clinical practice, with the current state of knowledge, we would recommend using intention to treat or SpO_2 target of 87–94% with a monitor as the one used in the RCTs.

The averaging time for clinical care varies with monitors, and in many of them, this cannot be changed. There is always a trade-off when choosing a shorter or longer averaging time, if the monitor allows for that. The Masimo SET, which is the monitor used in all the NICU RCTs discussed in this review, needs to be set at eight to 12 sec and definitely at no more than 20 sec, to avoid seeing very brief episodes that are probably meaningless in routine care and at the same time ensure that situations that could require a clinical response are seen quickly. However, in the delivery room or during acute resuscitation, using a 2-sec averaging time could be a better option for rapidly identifying clinical situations and treatment responses. The sensitivity also varies. Most of the time, it should be on default, but high sensitivity is better in the delivery room or during resuscitations.

For infants breathing supplemental oxygen, we suggest setting the low alarm at 1% below the lower SpO_2 chosen in the clinically intended range, but never <85%, and the high alarm at 1% above the higher SpO_2 chosen, but never more than 95%. For example, if the intended range is 87–94%, the alarms would be set at 86% and 95%. Finally, the delay time should not be longer than 20 sec, to ensure that significant events are not missed.

BRIEF COMMENTS ON THE RCTS

We are all proud of scientific research successes, but sometimes the findings are just too good to be true (50). It is not disputed that RCTs are the gold standard, but just like gold, there are RCTs of different carats and many RCTs are <50% pure. These latter studies cannot be considered as unshakable evidence.

In 2005, Dr Ioannidis started releasing papers that challenged the foundations of medical research (15,27,50–53), stating that 'A pervasive theme of ancient Greek literature is that you need to pursue the truth, no matter what the truth might be'. His studies have shown that many published research findings are false or completely wrong, and he has pointed out that 'The probability that a research claim is true depends on many factors, and for many current scientific fields, claimed research findings may often be simply accurate measures of the prevailing bias'. For example, Ioannidis states that a research finding is less likely to be true when more teams are involved in a scientific field and are chasing statistical significance. If time is of the essence, and to beat the competition, each team may prioritise and disseminate its most impressive 'positive' results. In addition, randomised trials that are stopped early, after fewer than 500 events, may result in large overestimations of the treatment effect (54,55). In the BOOST II trial, for example, there were 153 such events. There also maybe a problem with meta-analyses of trials that were stopped early (56). The use of an interim analysis, employed by some RCTs, carries a statistical risk that, by chance, the observed effect might not represent the true effect. This may further distort the picture and may indeed lead to even smaller probabilities of the research findings being true, as proven by Ioannidis in his published articles. Ioannidis also sustains that there is an intellectual conflict of interest that pressures researchers to present the findings that are most likely to secure funding or give them more prestige. Of significant concern is his statement that much of what biomedical researchers conclude in published studies is misleading and that researchers are frequently chasing findings that will advance their careers, rather than good science. He has also stated that 'Even highly regarded researchers at prestigious institutions sometimes churn out attention-grabbing findings rather than findings likely to be right'. Ioannidis has also commented on the use of the peer-review process to suppress opposing views, and sadly, this is a well-known issue in the real world.

In addition to all of the above, it has been identified that good care should not be based solely on the findings of RCTs (15,26–28,50–53,57–59). Acknowledging that evidence may still be misinterpreted or distorted by recalcitrant proponents of entrenched practices and other biases, the authors state that rational, quantitative evidence may not necessarily be the only, or even the main, factor driving healthcare decisions. Besides, many papers report

significant results, and only a few published papers report neutral results, negative results or adverse events. It has been estimated that about 90% of papers report positive results and that this increases down the hierarchy of the sciences (*59)*. If one searches for a low p value, one can always find significant results. Results of medical research should not be reported as significant or nonsignificant, but they should be interpreted in the context of the type of study and other available evidence. Bias or confounding factors should always be considered for findings with low p values. To stop medical research being discredited by chance findings, we need more studies that avoid all the well-known errors that can increase the chance of a significant low p value but do not really reflect what happens in the real world. (28). Also, repeated testing and interim analysis may result in overestimations of the treatment effect. This may lead to a small probability of some of the research findings actually being true. Of course, we are not addressing at all the issues of scientific ethics and fraud, which are completely different matters and of great concern to science and psychology (60).

The correctable weaknesses that still persist in the design, conduct, and analysis of biomedical and public health research studies, produce misleading results and waste valuable resources. Statistical precision is often used in a misleading way, and the arbitrary choice of analyses might affect the reported findings (15,26,27,57). Unfortunately, consensus building groups and bedside care providers remember the novel findings of some studies, despite significant flaws, when taking care of patients.

Putting all these matters aside, one of the most confusing aspects of the data presented by the SUPPORT and BOOST II trials is that there was no real clinical rationale provided to explain why mortality was higher in the 85–89% saturation groups. Other than flaws in design and analysis, which can occur in RCTs according to Ioannidis and others, the evidence from those trials suggests that the increased deaths were not caused by tissue hypoxia. For example, all the infants who died were reported in the BOOST II appendix, in Tables S 3.1, with the original oximeters, and S 3.2, with the revised oximeters. The vast majority died due to severe respiratory distress syndrome, severe bronchopulmonary dysplasia, grade 3–4 intraventricular haemorrhage, septicaemia and 'other' causes. We can identify absolutely no differences in those tables between the low and high-SpO_2 groups or between the original and revised oximeters. We, and many others, are uncertain about why the association of SpO_2 85–89% with increased mortality occurred in two of the trials and not in the COT trial. In the current neonatal era, there are often multiple factors associated with neonatal deaths in the NICU, and we can only speculate that this discrepancy was due to one or more of those or to chance. We can only wonder what would have happened if, by chance, the 91–95% SpO_2 group had been associated with increased mortality. We are hopeful that some of these considerations are helpful to the reader and to NICU care providers.

SUMMARY OF FINDINGS

The best oxygen profiles to reduce ROP hazards, while enhancing preterm infant health and development, are still unknown. While respectfully acknowledging the existence of uncertainty, we feel that any practice guidelines or recommendations on any topic should be developed avoiding extremes and biases that could potentially have ill effects on many newborn infants. This paper provides the arguments that are needed to avoid such extremes and biases during SpO_2 monitoring. We hope that our contribution will stimulate further debate on the subject and reduce the current clinical confusion, so that doctors, nurses and respiratory therapists providing bedside care for babies make better choices that avoid both hypoxaemia and hyperoxaemia in daily practice. Rational, quantitative evidence may not necessarily be the only, or even the main, factor driving healthcare decisions.

Over the past year, many clinicians, mostly in the United States, have expressed concerns about legal action if they allow babies to have 'low' SpO_2 values in the mid- to high 80s. It has been said that it is good to practise clinical medicine based on the best interpretation of the available evidence and not to practise in a defensive manner. No one should ever practise out of fear of the legal system, and we are certain that this also includes SpO_2 targeting in very preterm infants. If a very preterm infant that had SpO_2 readings of 86 or 87% at times was to die, this could not be considered the cause of the infant's death. Furthermore, there is absolutely no evidence to prove that choosing an SpO_2 in the mid- to high 80s as the low limit is associated with, or causes, mortality. Of course, this is different to implementing an intention to treat today with a low SpO_2 limit of 85% and a high one of 89%, as these limits have been associated with a moderate increase in mortality in two RCTs. Similarly, if a very preterm newborn infant develops severe ROP, having had SpO_2 readings of 95% or 97% at times, this could not be identified as the cause of ROP. Again, this is not the same as choosing an intention to treat with an SpO_2 target of 91–95%, which has been associated with a significant increase in ROP in two RCTs. Saturations within the range of 85–95% largely exclude hyperoxia in preterm infants <29 weeks (61).

We need to acknowledge the potential weaknesses of the RCTs summarised in this paper, which include the fact that repeated testing and interim analysis may result in overestimations of the treatment effect and lead to a small probability of some of the research findings being indeed true. We also need to recognise, and admit to, the real and persisting need for further improvements in what is currently known and used in neonatal care for neonatal oxygenation and for this complex issue of SpO_2 monitoring and targeting. For example, our clinical understanding of oxygenation status needs to be improved, and hopefully, this will happen in the not too distant future, with non-invasive monitors that can accurately measure regional tissue oxygenation, oxygen content, oxygen delivery and tissue oxygen needs. Improvements in SpO_2 targeting and monitoring require better education and understanding by

bedside care providers of the issues at stake and the differences in available monitors. In addition, temporal quantification of oxygen saturation ranges and better documentation of alarms are essential (24,25,44–46). Furthermore, improving oxygen saturation targeting by automated methods in preterm infants receiving invasive and non-invasive ventilation, by continuous positive airway pressure or nasal cannula, will improve the efforts to reduce hyperoxia in the NICU.

Our view is that if a centre has never used an SpO_2 target range of 85–89% before, it should not use it now and that if it has, then it should give serious consideration to modifying that practice. The same consideration should be given to an intention to treat with a target of 91–95% to avoid dangerous hyperoxaemia, a resurgence of ROP, altered brain development and other morbidities associated with hyperoxaemia.

Based on current knowledge, we must aim to avoid both hypoxaemia and hyperoxaemia in daily clinical practice. With regard to hyperoxaemia, pure oxygen should not be given if the baby's colour does not look good or seems to be cyanotic; instead, oxygenation should be assessed using an SpO_2 monitor. Using skin colour to detect hyperoxia is of little or no use, as several studies have shown that it has wide interobserver variability and a low correlation with SpO_2 (62–66). A recent meta-analysis and review, that included 11 diagnostic studies with 5787 patients, showed that neither single nor combined symptoms and signs satisfactorily predicted hypoxaemia in young children (62). Other previous studies showed that pulse oximetry is the best indicator of hypoxaemia in children (63,64), and this has been corroborated by large clinical studies on screening for critical congenital heart disease (65,66). However, improved access to pulse oximetry is still needed in many delivery rooms in the developed world and in many delivery rooms and neonatal units in developing countries. If SpO_2 does not show hypoxaemia, then oxygen should not be administered. If it does show hypoxaemia, then FiO_2 should be used at a dose that corrects the hypoxaemia, while avoiding hyperoxaemia. To achieve this in small preterm infants, it is safer to widen the target, using SpO_2 targets like 86–93%, 87–94% or 88–94% or similar. It is imperative to control the low limit but also to increase compliance with the upper limit and avoid a flip-flop phenomenon that leads to acute and significant changes and fluctuations in SpO_2. Hopefully, this will improve with the advent of automated control of oxygenation (30,67).

ACKNOWLEDGEMENTS

The authors are very grateful to Annette Whibley for her extensive effort and commitment to review and improve the English language of this manuscript.

CONFLICT OF INTEREST

Augusto Sola is the Vice President of Medical affairs in Neonatology at the Masimo Corporation, Irvine, California, USA, No author or individual received any guidance, honorarium, grant or other form of payment from any source to produce the manuscript.

References

1. Sola A, Rogido MR, Deulofeut R. Oxygen as a neonatal health hazard: call for détente in clinical practice. *Acta Paediatr* 2007; 96: 801–12.
2. Comroe JH Jr, Forster RE, Dubois AB, Briscoe WA, Carlsen E, editors. *Textbook the lung: clinical physiology and pulmonary function tests*. Chicago, IL: Year Book Medical Publishers, 1955.
3. Comroe JH Jr, Bahnson ER, Coates EO Jr. Mental changes occurring in chronically anoxemic patients during oxygen therapy. *J Am Med Assoc* 1950; 143: 1044–8.
4. Comroe JH Jr, Fowler WS. Lung function studies. VI. Detection of uneven alveolar ventilation during a single breath of oxygen. *Am J Med* 1951; 10: 408–13.
5. Comroe JH. Retrospectroscope: insights into medical discovery. Menlo Park, CA: Von Gehr Press, 1977. ISBN 0-9601470-1-2.
6. Comroe JH. The location and function of the chemoreceptors of the aorta. *Am J Physiol* 1939; 127: 176–91.
7. Saugstad O. Take a breath–but do not add oxygen (if not needed). *Acta Paediatr* 2007; 96: 798–800.
8. SUPPORT Study Group of the Eunice Kennedy Shriver NICHD Neonatal Research Network. Target ranges of oxygen saturation in extremely preterm infants. *N Engl J Med* 2010; 362: 1959–69.
9. Vaucher YE, Peralta-Carcelen M, Finer NN, Carlo WA, Gantz MG, Walsh MC, et al. Neurodevelopmental outcomes in the early CPAP and pulse oximetry trial. *N Engl J Med* 2012; 367: 2495–504.
10. Schmidt B, Whyte RK, Asztalos EV, Moddemann D, Poets C, Rabi Y, et al. Effects of targeting higher vs lower arterial oxygen saturations on death or disability in extremely preterm infants: a randomized clinical trial. *JAMA* 2013; 309: 2111–20.
11. BOOST II United Kingdom Collaborative Group, BOOST II Australia Collaborative Group, BOOST II New Zealand Collaborative Group, Stenson BJ, Tarnow-Mordi WO, Darlow BA, et al. Oxygen saturation and outcomes in preterm infants. *N Engl J Med* 2013; 368: 2094–104.
12. Bancalari E, Claure N. Oxygenation targets and outcomes in premature infants. *JAMA* 2013; 309: 2161–2.
13. Saugstad OD, Aune D. Optimal oxygenation of extremely low birth weight infants: a meta-analysis and systematic review of the oxygen saturation target studies. *Neonatology* 2014; 105: 55–63.
14. Alderliesten T, Lemmers PM, Smarius JJ, 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.
15. Ioannidis JPA. Why most published research findings are false. *PLoS Med* 2005; 2: e124.
16. Hartnett ME, Lane RH. Effects of oxygen on the development and severity of retinopathy of prematurity. *J AAPOS* 2013; 17: 229–34.
17. Chow LC, Wright KW, Sola A; CSMC Oxygen Administration Study Group. Can changes in clinical practice decrease the incidence of severe retinopathy of prematurity in very low birth weight infants? *Pediatrics* 2003; 111: 339–45.
18. Deulofeut R, Critz A, Adams-Chapman I, Sola A. Avoiding hyperoxia in infants <1250 g is associated with improved short- and long-term outcomes. *J Perinatol* 2006; 26: 700–5.

19. Castillo A, Sola A, Baquero H, Neira F, Alvis R, Deulofeut R, et al. Pulse oxygen saturation levels and arterial oxygen tension values in newborns receiving oxygen therapy in the neonatal intensive care unit: is 85% to 93% an acceptable range? *Pediatrics* 2008; 121: 882–9.
20. Bashambu MT, Bhola M, Walsh M. Evidence for oxygen use in preterm infants. *Acta Paediatr* 2012; 101(Suppl. 464): 29–33.
21. Wright KN, Sami A, Thompson L, Ramanathan R, Roy J, Farzavandi S. A physiologic reduced oxygen protocol decreases the incidence of threshold retinopathy of prematurity. *Trans Am Ophthalmol Soc* 2006; 104: 78–84.
22. Kapadia VS, Chalak LF, Sparks JE, Allen JR, Savani RC, Wyckoff MH. Resuscitation of preterm neonates with limited versus high oxygen strategy. *Pediatrics* 2013; 132: e1488–96.
23. Askie LM, Henderson-Smart DJ, Ko H. Restricted versus liberal oxygen exposure for preventing morbidity and mortality in preterm or low birth weight infants. *Cochrane Database Syst Rev* 2009; 21: CD001077.
24. Lim K, Wheeler KI, Gale TJ, Jackson HD, Kihlstrand JF, Sand C, et al. Oxygen saturation targeting in preterm infants receiving continuous positive airway pressure. *J Pediatr* 2014; 164: 730–736.e1.
25. Bizzarro MJ, Li FY, Katz K, Shabanova V, Ehrenkranz RA, Bhandari V. Temporal quantification of oxygen saturation ranges: an effort to reduce hyperoxia in the neonatal intensive care unit. *J Perinatol* 2014; 34: 33–8.
26. Enserink M. Scientific ethics. Final report on Stapel also blames field as a whole. *Science* 2012; 338: 1270–1.
27. Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014; 383: 166–75.
28. Sterne JAC. Smith GD Sifting the evidence—what's wrong with significance tests? *BMJ* 2001; 322: 226–31.
29. Hagadorn JI, Furey AM, Nghiem TH, Schmid CH, Phelps DL, Pillers DA, et al. Achieved versus intended pulse oximeter saturation in infants born less than 28 weeks' gestation: the AVIOx study. *Pediatrics* 2006; 118: 1574–82.
30. Bancalari E, Claure N. Control of oxygenation during mechanical ventilation in the premature infant. *Clin Perinatol* 2012; 39: 563–72.
31. Barker SJ, Shah NK. The effects of motion on the performance of pulse oximeters in volunteers. *Anesthesiology* 1997; 86: 101–8.
32. Bohnhorst B, Peter CS, Poets CF. Pulse oximeters' reliability in detecting hypoxemia and bradycardia: comparison between a conventional and two new generation oximeters. *Crit Care Med* 2000; 28: 1565–8.
33. Malviya S, Reynolds PI, Voepel-Lewis T, Siewert M, Watson D, Tait AR, et al. False alarms and sensitivity of conventional pulse oximetry versus the Masimo SET technology in the pediatric postanesthesia care unit. *Anesth Analg* 2000; 90: 1336–40.
34. Hay WW Jr, Rodden DJ, Collins SM, Melara DL, Hale KA, Fashaw LM. Reliability of conventional and new pulse oximetry in neonatal patients. *J Perinatol* 2002; 22: 360–6.
35. Brouillette RT, Lavergne J, Leimanis A, Nixon GM, Ladan S, McGregor CD. Differences in pulse oximetry technology can affect detection of sleep-disorderd breathing in children. *Anesth Analg* 2002; 94(Suppl. 1): S47–53.
36. Sahni R, Gupta A, Ohira-Kist K, Rosen TS. Motion resistant pulse oximetry in neonates. *Arch Dis Child Fetal Neonatal Ed* 2003; 88: F505–8.
37. Kawagishi T, Kanaya N, Nakayama M, Kurosawa S, Namiki A. A comparison of the failure times of pulse oximeters during blood pressure cuff-induced hypoperfusion in volunteers. *Anesth Analg* 2004; 99: 793–6.
38. Sola A, Chow L, Rogido MR. Pulse oximetry in neonatal care in 2005. A comprehensive state of the art review. *An Pediatr (Barc)* 2005; 62: 266–80.
39. Workie FA, Rais-Bahrami K, Short BL. Clinical use of new-generation pulse oximeters in the neonatal intensive care unit. *Am J Perinatol* 2005; 22: 357–60.
40. Ahmed SJ, Rich W, Finer NN. The effect of averaging time on oximetry values in the premature infant. *Pediatrics* 2010; 125: e115–21.
41. Baquero H, Alviz R, Castillo A, Neira F, Sola A. Avoiding hyperoxemia during neonatal resuscitation: time to response of different SpO2 monitors. *Acta Paediatr* 2011; 100: 515–8.
42. Shah N, Ragaswamy HB, Govindugari K, Estanol L. Performance of three new-generation pulse oximeters during motion and low perfusion in volunteers. *J Clin Anesth* 2012; 24: 385–91.
43. Dawson JA, Saraswat A, Simionato L, Thio M, Kamlin CO, Owen LS, et al. Comparison of heart rate and oxygen saturation measurements from Masimo and Nellcor pulse oximeters in newly born term infants. *Acta Paediatr* 2013; 102: 955–60.
44. Vagedes J, Poets CF, Dietz K. Averaging time, desaturation level, duration and extent. *Arch Dis Child Fetal Neonatal Ed* 2013; 98: F265–6.
45. Clucas L, Doyle LW, Dawson J, Donath S, Davis PG. Compliance with alarm limits for pulse oximetry in very preterm infants. *Pediatrics* 2007; 119: 1056–60.
46. Brockmann PE, Wiechers C, Pantalitschka T, Diebold J, Vagedes J, Poets CF. Under-recognition of alarms in a neonatal intensive care unit. *Arch Dis Child Fetal Neonatal Ed* 2013; 98: F524–7.
47. Gandhi B, Rich W, Finer N. Achieving targeted pulse oximetry values in preterm infants in the delivery room. *J Pediatr* 2013; 163: 412–5.
48. Sola A. Oxygen in neonatal anesthesia: friend or foe? *Curr Opin Anaesthesiol* 2008; 21: 332–9.
49. Sola A, Zuluaga C. Oxygen saturation targets and retinopathy of prematurity. *J AAPOS* 2013; 17: 650–2.
50. Ioannidis JPA. Research accomplishments that are too good to be true. *Intensive Care Med* 2014; 40: 99–101.
51. Lenzer J, Hoffman JR, Furberg CD, Ioannidis JP. Ensuring the integrity of clinical practice guidelines: a tool for protecting patients. *BMJ* 2013; 347: f5535.
52. Ioannidis JP. The importance of potential studies that have not existed and registration of observational data sets. *JAMA* 2012; 308: 575–6.
53. Prasad V, Ioannidis JP. Evidence-based de-implementation for contradicted, unproven, and aspiring healthcare practices. *Implement Sci* 2014; 9: 1–5.
54. Bassler D, Briel M, Montori VM, Lane M, Glasziou P, Zhou Q, et al. Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. *JAMA* 2010; 303: 1180–7.
55. Guyatt GH, Briel M, Glasziou P, Bassler D, Montori VM. Problems of stopping trials early. *BMJ* 2012; 344: e3863.
56. Bassler D, Montori VM, Briel M, Glasziou P, Walter SD, Ramsay T, et al. Reflections on meta-analyses involving trials stopped early for benefit: is there a problem and if so, what is it? *Stat Methods Med Res* 2013; 22: 159–68.
57. Sola A, Dieppa FD, Rogido MR. An evident view of evidence-based practice in perinatal medicine: absence of evidence is not evidence of absence. *J Pediatr (Rio J)* 2007; 83: 395–414.

58. Mol A. Proving or improving: on health care research as a form of self-reflection. *Qual Health Res* 2006; 16: 405–14.
59. Fanelli D. "Positive" results increase down the Hierarchy of the Sciences. *PLoS One* 2010; 5: e10068.
60. Enserink M. Scientific ethics. Fraud-detection tool could shake up psychology. *Science* 2012; 337: 21–2.
61. Quine D, Stenson BJ. Arterial oxygen tension (PaO2) values in infants <29 weeks of gestation at currently targeted saturations. *Arch Dis Child Fetal Neonatal Ed* 2009; 94: F51–3.
62. Zhang L, Mendoza-Sassi R, Santos JC, Lau J. Accuracy of symptoms and signs in predicting hypoxaemia among young children with acute respiratory infection: a meta-analysis. *Int J Tuberc Lung Dis* 2011; 15: 317–25.
63. Niermeyer S, Yang P, Shanmina, Drolkar, Zhuang J, Moore LG. Arterial oxygen saturation in Tibetan and Han infants born in Lhasa, Tibet. *N Engl J Med* 1995; 333: 1248–52.
64. Laman M, Ripa P, Vince J, Tefuarani N. Can clinical signs predict hypoxaemia in Papua New Guinean children with moderate and severe pneumonia? *Ann Trop Paediatr* 2005; 25: 23–7.
65. de-Wahl Granelli A, Wennergren M, Sandberg K, Mellander M, Bejlum C, Inganäs L, et al. Impact of pulse oximetry screening on the detection of duct dependent congenital heart disease: a Swedish prospective screening study in 39,821 newborns. *BMJ* 2009; 338: a3037.
66. Ewer AK, Middleton LJ, Furmston AT, Bhoyar A, Daniels JP, Thangaratinam S, et al. Pulse oximetry screening for congenital heart defects in newborn infants (PulseOx): a test accuracy study. *Lancet* 2011; 378: 785–94.
67. Zapata J, Gomez JJ, Araque Campo R, Matiz Rubio A, Sola A. A randomised controlled trial of an automated oxygen delivery algorithm for preterm neonates receiving supplemental oxygen without mechanical ventilation. *Acta Paediatr* 2014; 103, doi: 10.1111/apa.12684 (accepted).

