OPEN ACCESS

Randomised crossover study on pulse oximeter readings from different sensors in very preterm infants

Christian Achim Maiwald [1,2], Christoph E Schwarz [2,3], Katrin Böckmann,[2] Laila Springer [2], Christian F Poets [2], Axel Franz [1,2]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/archdischild-2023-325961).

[1]Department of Pediatrics, Center for Pediatric Clinical Studies (CPCS), University Hospital Tübingen, Tübingen, Germany
[2]Department of Neonatology, Tübingen University Hospital, Tübingen, Germany
[3]Department of Neonatology, University of Heidelberg, Heidelberg, Germany

Correspondence to
Professor Christian F Poets, Department of Neonatology, University Hospital Tubingen Department of Pediatrics, Tübingen, 72076, Germany; christian-f.poets@med.uni-tuebingen.de

Received 13 June 2023
Accepted 4 December 2023

To cite: Maiwald CA, Schwarz CE, Böckmann K, *et al*. *Arch Dis Child Fetal Neonatal Ed* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/archdischild-2023-325961

ABSTRACT

Objective In extremely preterm infants, different target ranges for pulse oximeter saturation (SpO_2) may affect mortality and morbidity. Thus, the impact of technical changes potentially affecting measurements should be assessed. We studied SpO_2 readings from different sensors for systematic deviations.

Design Single-centre, randomised, triple crossover study.

Setting Tertiary neonatal intensive care unit.

Patients 24 infants, born at <32 weeks' gestation, with current weight <1500 g and without right-to-left shunt via a patent ductus arteriosus.

Interventions Simultaneous readings from three SpO_2 sensors (Red Diamond (RD), Photoplethysmography (PPG), Low Noise Cabled Sensors (LNCS)) were logged at 0.5 Hz over 6 hour/infant and compared with LNCS as control using analysis of variance. Sensor position was randomly allocated and rotated every 2 hours. Seven different batches each were used.

Outcomes Primary outcome was the difference in SpO_2 readings. Secondary outcomes were differences between sensors in the proportion of time within the SpO_2-target range (90–95 (100)%).

Results Mean gestational age at birth (±SD) was $27^{4/7}$ ($\pm 2^{3/7}$) weeks, postnatal age 20 (±20) days. 134 hours of recording were analysed. Mean SpO_2 (±SD) was 94.0% (±3.8; LNCS) versus 92.2% (±4.0; RD; p<0.0001) and 94.5% (±3.9; PPG; p<0.0001), respectively. Mean SpO_2 difference (95% CI) was −1.8% (−1.9 to −1.8; RD) and 0.5% (0.4 to 0.5; PPG). Proportion of time in target was significantly lower with RD sensors (84.8% vs 91.7%; p=0.0001) and similar with PPG sensors (91.1% vs 91.7%; p=0.63).

Conclusion There were systematic differences in SpO_2 readings between RD sensors versus LNCS. These findings may impact mortality and morbidity of preterm infants, particularly when aiming for higher SpO_2-target ranges (eg, 90–95%).

Trial registration number DRKS00027285.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ SpO_2 target ranges affect outcome of extremely preterm infants. Current recommendations on SpO_2 targets are based on one instrument brand and sensor type.

WHAT THIS STUDY ADDS

⇒ Some new generation sensors resulted in SpO_2 readings that were 2% lower than with the previous standard. This may lead to higher oxygen levels and thus potentially affect oxygen-related morbidity and mortality in extremely preterm infants.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Recommendations on SpO_2 targets should refer to a specific instrument brand and sensor type. There is a need for better standardisation of SpO_2 technology.

INTRODUCTION

Medical oxygen is one of the most common drugs administered in neonatal intensive care units (NICUs).[1] The majority of infants with a gestational age (GA) at birth <32 weeks require supplemental oxygen, and both too much and too little oxygen may impact on outcome. Therefore, considerable effort has been, and continues to be, employed for achieving an optimal oxygen supply strategy in these infants.

In a recent Cochrane report comparing two different SpO_2-target ranges (85–89% vs 91–95%, effective difference 2.8%) that was largely based on the results of the NeoPROM collaboration, the higher range was associated with lower rates of death and necrotising enterocolitis (NEC), but higher rates of retinopathy of prematurity (ROP) and bronchopulmonary dysplasia.[2] Therefore, the higher SpO_2 targets (eg, 90–94%, with alarm limits of 89% and 95%) are recommended by experts in the field[3] and also in consensus guidelines for the treatment of neonates.[4]

While the NeoPROM studies were performed exclusively with Masimo SET oximeters with 'Low Noise Cabled' Sensors (LNCS), the manufacturer currently recommends the use of 'Red Diamond' sensors (RD) because of reportedly improved accuracy when compared with arterial haemoglobin oxygen saturation by co-oximetry (SaO_2; ±3% points vs ±1.5% points in SpO_2 in LNCS vs RD sensors[5 6]). The Photoplethysmography (PPG) sensors have the same accuracy as LNCS, but would have the advantage of enabling wireless transmission.[7] In NeoPROM, a 2.8% difference in achieved SpO_2 changed the outcome, and therefore, any change in measurement technology (or components

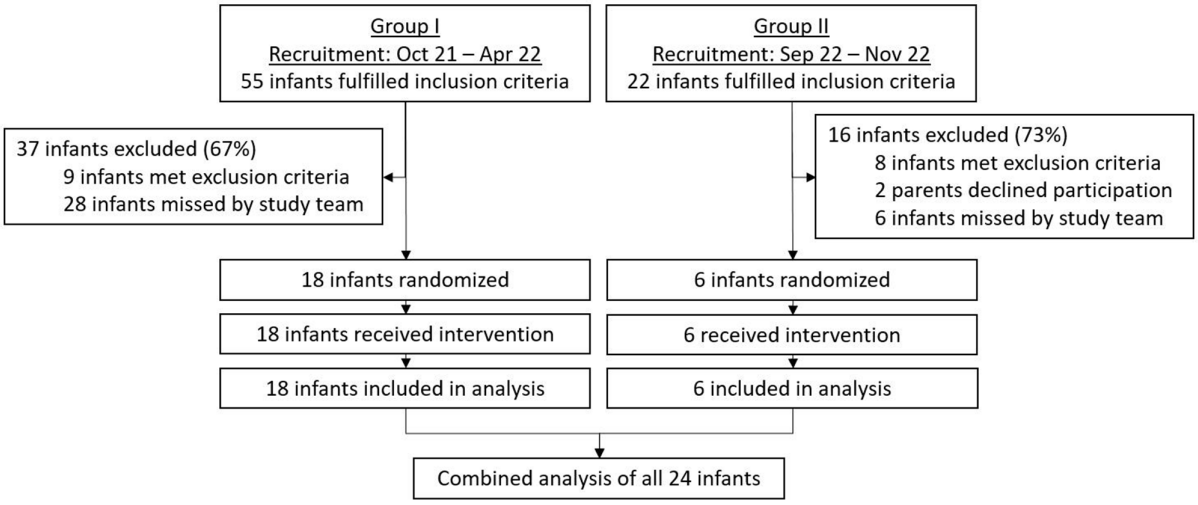

Figure 1 Study flow diagram.

thereof) should be carefully assessed for their potential impact on achieved SpO_2 in this very vulnerable population of extremely preterm infants. Consequently, we performed a head-to-head comparison between SpO_2 readings from two new sensor types (RD; PPG) against our local standard, the LNCS.

MATERIAL AND METHODS

Study design

This is a single-centre, randomised, triple cross-over, prospective observational study of CE-marked medical devices applied according to their intended use.

Patients

Infants born at <32 weeks GA and with an excluded bi-directional or right-to-left shunt through a patent ductus arteriosus on echocardiography were screened during their postnatal hospitalisation; those receiving less than 12 feeds per day (to align study-driven changes in sensor site with clinically indicated disturbance, ie, feeding and nursing) or on palliative care were excluded. Due to six possible randomisation clusters, we initially planned to examine 18 infants (group 1; three infants per cluster) and then added another six infants to exclude sensor batch-related differences (group 2; study flow diagram (figure 1)). The study protocol required group 1 to include at least nine infants each with a current GA <28 weeks and receiving supplemental oxygen (FiO_2>0.21).

Setting

This study took place in the tertiary NICU at the Department of Neonatology, University Hospital Tübingen, Germany.

Equipment

'Radical 7' oximeters, 2012 version (MCU: 1064; Tech-card: 7e23 (RD and LNCS) and 7f10 (PPG); processor: V.1.5.5.8i) were used. Docking stations were RDS-1 (ASCII1 IAP Flexport 5143) and trends were downloaded using the Masimo Instrument Configuration Tool (V.1.2.5.1, 2020). Sensor types were LNCS as the local standard (Masimo internal Order No: 1862 and for 2 infants <800 g: 1901); for comparison, we used RD (Order No: 4003) and PPG (Order No: 4585). In group 1, we used a single batch per sensor type in all infants; in group 2 every recording was performed with different batches for all sensor types (online supplemental table 1) to exclude biased results due to production errors. All devices and sensors were produced by Masimo, Irvine, California, USA.

Procedures

Parents of eligible infants were approached and written informed parental consent was obtained. The three different SpO_2 sensors were simultaneously attached to three IV-access-free limbs. Limbs were numbered clockwise in supine position, starting on the right hand. Sensor types were randomly allocated to sensor sites (see: Randomisation).

Sensors were placed and repositioned every 2 hours, exclusively during care periods or meals. Data from all three sensors were simultaneously recorded at a sampling rate of 0.5 Hz for a total duration of 6 hours (ie, each sensor type and position for at least 2 hours each. The expected 10.800 measurements per patient were considered to be sufficient to demonstrate any clinically relevant difference. The 2-hour period was chosen to meet nursing practices and to avoid sensor changes independent of care rounds. Averaging time was set to 2–4 s.

FiO_2 was manually or automatically controlled (if infant participated in our multicentre FiO_2 controller trial[8]) to achieve SpO_2 values within the target range of 90–95% according to the SpO_2 readings of the LNCS.

Randomisation

Six different algorithms for changing the three sensors, each with different starting positions (see online supplemental table 2 for randomisation clusters), were randomly assigned with appropriate allocation concealment using consecutively numbered sealed opaque envelopes.

Blinding

Since the different sensors have different patient cable/sensor interfaces, blinding was not feasible.

Efforts to reduce bias and to assess potentially influencing variables

- Deviations based on limb allocation
  - Echocardiography: All infants had routine echocardiography at maximum 48 hours before start of recording to exclude right-to-left ductal shunting
  - Two-hourly, clockwise rotation of sensor positions.

Table 1 Outcome measurements

	LNCS		RD			PPG		
SpO_2—all readings (%)								
Mean (±SD)								
All infants	94.0	(±3.8)	92.2	(±4.0)	p<0.0001	94.5	(±3.9)	p<0.0001
Group 1	94.6	(±3.6)	92.5	(±3.7)	–	95.1	(±3.6)	–
Group 2	92.6	(±4.0)	91.3	(±4.4)	–	92.7	(±4.3)	–
Mean difference to LNCS (95% CI of mean)								
All infants	–	–	−1.84	(−1.85 to −1.83)	–	0.46	(0.45 to 0.47)	–
Group 1	–	–	−2.04	(−2.05 to −2.03)	–	0.58	(0.57 to 0.59)	–
Group 2	–	–	−1.27	(−1.30 to −1.25)	–	0.11	(0.09 to 0.14)	–
% time in SpO_2 target range								
Mean (±SD) with SpO_2 90–95/100% (n=18)	87.4	(±13.9)	81.3	(±18	p<0.001	83.9	(±18.8)	p=0.63
% time above target range (SpO_2>95%)								
Mean (±SD) for FiO_2>0.21 (n=8)	9.5	(±7.0)	4.9	(±3.5)	p=0.42	17.6	(±11.0)	p=0.09
% time below target range								
Mean (±SD) for SpO_2 80–89%; all infants	8.8	(±7.2)	17.8	(±14.3)	p<0.001	8.5	(±8.1)	p=0.62
Mean (±SD) for SpO_2: <80%; all infants	0.6	(±0.7)	1.2	(±1.4)	p=0.001	0.7	(±1.0)	p>0.99

LNCS, Low Noise Cabled Sensors; PPG, Photoplethysmography; RD, Red Diamond.

  - Randomised assignment of starting position with adequate allocation concealment
- Deviations based on signal quality
  - Bedside nurses were advised to check (and if necessary correct) sensor position in the event of persistently poor signal quality ('low signal IQ'-alarm) but not if there were discrepancies between readings
  - Exclusion of data lines with invalid values (ie, if at any given time point any of the three SpO_2 or pulse rate readings showed either 'no value' or 'zero' or an exception code such as 'sensor OFF' or 'low signal IQ' and all data recorded during care periods (to exclude any impact of motion artefacts).
  - Comparison of pulse rate readings in the analysed data to check the validity of the recordings
- Influence of batches
  - After recruitment of 18 infants with RD and PPG sensors from a single batch (group1), we repeated measurements in six additional infants (group 2) using a different batch for all sensor types in each infant to rule out that the observation made was based on a single batch and possibly biased by production errors.

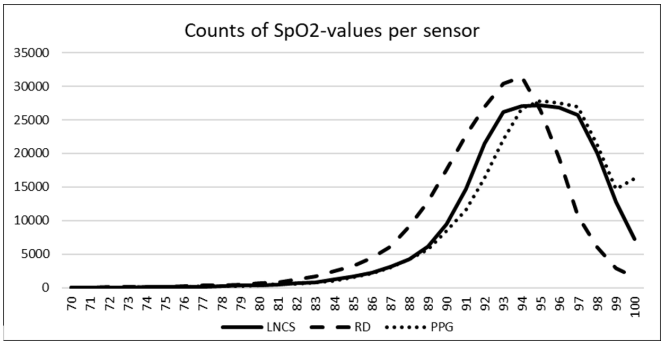

Figure 2 Counts of SpO_2 values per sensor in all infants. LNCS, Low Noise Cabled Sensors; PPG, Photoplethysmography; RD, Red Diamond

Outcome variables

Primary outcome was the SpO_2 difference (95% CI) between RD or PPG sensors compared with LNCS as control. Therefore, mean values (±SD) were compiled for every infant over all sensor positions and compared between sensor types. Secondary outcomes were the proportion of time in SpO_2 target (90–95% for infants in FiO_2>0.21 and 90–100% for infants in FiO_2=0.21) and the proportion of time above target (only for infants with FiO_2>0.21). Infants with an FiO_2 of both, 0.21 and >0.21, were excluded because FiO_2 was not logged. Proportion of time with SpO_2 below target was calculated for all sensors in all infants. FiO_2 was controlled throughout the study according to LNCS readings.

Statistical analysis

Time stamp, SpO_2 and pulse rate readings were downloaded as CSV files and compiled using Microsoft Office Excel 2019 (V.1808). Analysis was descriptive using mean (±SD) and Friedman test performed if the mean difference was >0.1 in any comparison, using Prism V.9.4.1 (GraphPad, Boston, USA). p<0.05 was considered statistically significant. Bland-Altman plots for visualisation of differences were created for individual values of SpO_2 and pulse rate in both groups and sensor comparisons (RD vs LNCS and PPG vs LNCS).

RESULTS

Patients

Twenty-four infants (12 female) were recruited between 10/2021 and 11/2022.

In group 1, we recruited 10 girls and 8 boys; 8 infants had a GA<28 weeks. Mean GA (±SD) at birth was $28^{0/7}$ ($\pm 2^{3/7}$) weeks and mean birth weight (±SD) 925 (±345) g. Mean postnatal age (±SD) was 18 (±21) days.

In group 2, we recruited two girls and four boys with a mean GA at birth (±SD) of $26^{4/7}$ ($\pm 2^{1/7}$) and a mean birth weight (±SD) of 714 (±241) g. Mean postnatal age (±SD) was 26 (±18) days.

For a more detailed description of weight and GA distributions, see online supplemental table 3: demographic data.

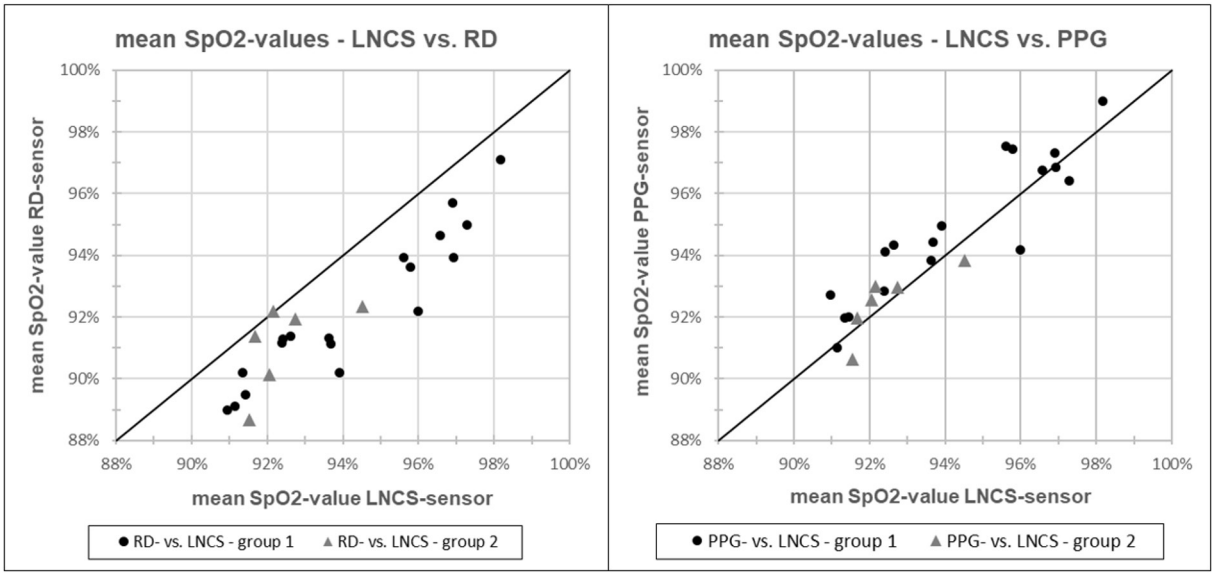

Figure 3 Comparison of mean SpO_2-values of all 24 infants. LNCS, Low Noise Cabled Sensors; PPG, Photoplethysmography; RD, Red Diamond.

Data

147.2 hours of data were recorded (group 1: 110.2 hours; group 2: 37.0 hours). After exclusion of invalid data, we analysed 241.595 data points (group 1: 178.426; group 2: 63.169), corresponding to 134.2 hours (91.2% of recorded data) and a mean duration (±SD) of 5.6 hours (±0.5) per patient.

Between sensor comparisons

For all measurements, mean pulse rates were identical for LNCS, RD and PPG sensors. These and between-sensor differences in pulse rate for individual measurements are represented in the online supplemental table 4 and figure 1.

Mean SpO_2 values were significantly lower with RD sensors (92.2% vs 94.0%; p<0.0001) and significantly higher with PPG sensors (94.5% vs 94.0%; p<0.0001) compared with LNCS. Mean differences (95% CI) between simultaneous SpO_2 values were −1.84% (–1.85% to −1.83%) for RD versus LNCS and 0.46% (0.45% to 0.47%) for PPG versus LNCS (table 1, online supplemental file 1). The graphical illustration of counts for all SpO_2 values also showed a deviation towards lower values for the RD sensor compared with the LNCS and PPG sensor (figure 2). Additionally, all infants had a lower mean SpO_2 with RD sensors compared with LNCS, while mean SpO_2 was similar for PPG sensors versus LNCS (figure 3). In periods with SpO_2 between 90% and 95% as measured by LNCS, the mean SpO_2 was 93.0% (±1.5) for LNCS, 91.5% (±2.6) for RD and 93.7% (±2.7) for PPG.

Proportion of time in SpO_2 target (90–95% for 8 infants with FiO_2 continuously >0.21 and 90–100% for 10 infants with FiO_2 continuously =0.21)

Compared with LNCS (which had been used to control FiO_2), mean proportion of time with SpO_2 in target was significantly lower with RD, but similar with PPG sensors (table 1).

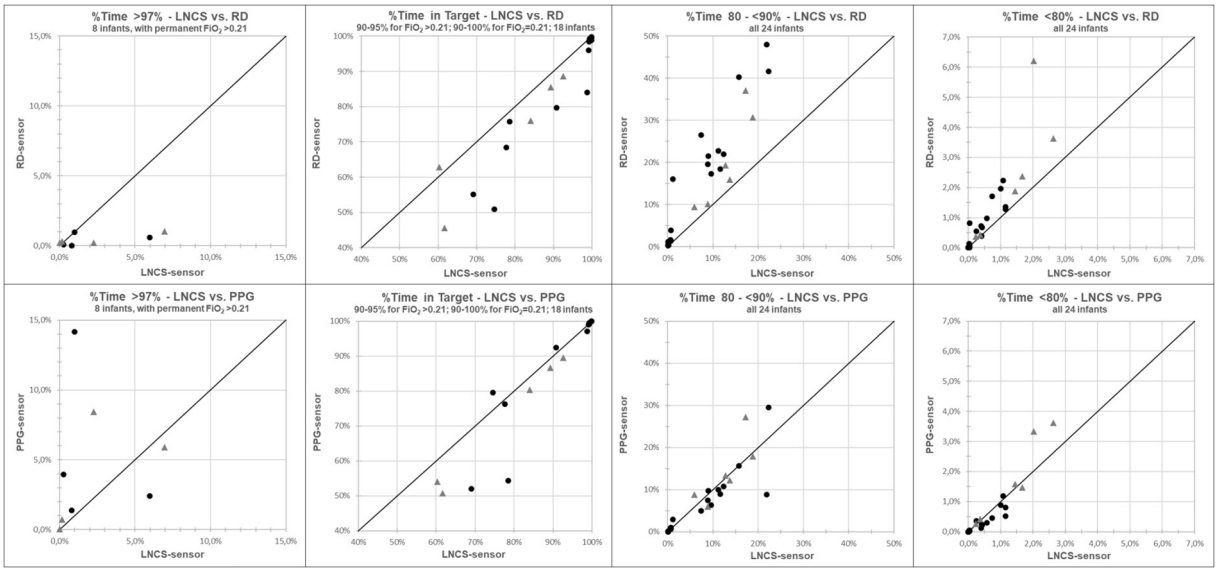

Figure 4 Distributions of proportions of time in- and outside of SpO_2-target range. LNCS, Low Noise Cabled Sensors; PPG, Photoplethysmography; RD, Red Diamond.

Only one infant at an FiO_2 of 0.24–0.28 spent a higher proportion of time in target with RD compared with LNCS (figure 4). This infant had a high proportion of time above the target range with LNCS and a mean difference in SpO_2 of −1.88% between RD and LNCS.

Proportion of time above target (eight infants with FiO_2 continuously>0.21)

The mean proportion of time spent above the target range was not statistically significantly different across sensors (table 1).

Proportion of time below target range (all 24 infants)

The mean proportion of time with SpO_2 80–89% and with SpO_2<80% was increased for RD sensors compared with that for LNCS and similar for PPG sensors compared with LNCS (table 1 and figure 4).

DISCUSSION

To our knowledge, this is the first study systematically comparing SpO_2 readings obtained with different sensor types from the same manufacturer in the vulnerable population of extremely preterm infants most in need of tight oxygen targeting. Previous studies compared instruments from different manufacturers[9–14] or SpO_2 with SaO_2 to verify, for example, the impact of skin colour or fetal haemoglobin.

Whereas most neonatologists will be familiar with the fact that simultaneous pulse oximetry readings from different limbs are not identical for substantial proportions of time, even if identical technology and equipment is used, our finding of a systematic deviation between LNCS and RD sensors is disturbing.

Both new sensors (PPG and RD) showed statistically significant differences in mean SpO_2 compared with LNCS, but for the PPG sensor (differing from LNCS technology only in wireless transmission), this mean difference in SpO_2 was smaller, less reproducible (figure 3) and there was no difference in the proportion of time outside the SpO_2-target range, indicating that subsequent clinical practice of FiO_2 control would not be different after changing sensors from LNCS to PPG. These findings agree with the expectation that the wireless transmission should have no effect on the SpO_2 readings. In contrast, the difference between RD sensors and LNCS was of clinical importance and found in every infant. The relevant difference in proportion of time outside the target range may indicate that using RD sensors for FiO_2 control would have resulted in relevantly higher oxygen exposure.

Since pulse detection is essential for pulse oximetry, the exact concordance of mean pulse rates between all sensor types confirms that care was taken to avoid any systematic bias in sensor application and that data collection and processing were of high quality. Whereas pulse rate measurements directly rely on the detection of an alternating signal, SpO_2 measurements are more complex as they rely on the relative extinction of light of at least two wavelengths within this alternating signal over a non-alternating background to approximate arterial oxygen saturation, which is more sensitive to external perturbations. This is supported by the observation that the coefficient of variation (ie, the SD divided by the mean) for SpO_2 measurements is much higher than for pulse rate measurements. According to the manufacturer, LNCS yield an SD of ±3% and RD sensors of 1.5% within 70–100% SaO_2. This means that at an SaO_2 of 90%, 95% of SpO_2 readings will be between 84% and 96% for LNCS and between 87% and 93% for RD sensors.

Comparing this imprecision in SpO_2 readings, given the narrow target ranges of 90–95% currently recommended for extremely preterm infants, is worrying, as is the systematic mean difference of almost 2% between readings from LNCS and RD sensors, independent of mean SpO_2 and across all batches tested.

This is particularly true because the Cochrane analysis of the NeoPROM studies reported significant and clinically relevant differences concerning the risk of death, NEC or ROP with an effective difference in SpO_2 of only 2.8%.[2] We believe that the difference in mean SpO_2 between RD sensor versus LNCS, although likely imperceptible during routine neonatal care, might be clinically relevant. Patients who are within the SpO_2 target range based on LNCS readings are below target for substantial proportions of time based on RD sensor readings, likely resulting in systematically higher FiO_2 settings with the use of RD sensors, which in turn may impact on clinical outcome. Therefore, a switch from LNCS to RD sensors may potentially have the same clinical consequences as changing the SpO_2-target ranges from 90–95% to 92–97%, which may have only a debatable impact on the proportion of time with PaO_2 values >80 mm Hg (eg, in the studies by Bachman *et al*,[15] Wackernagel *et al*,[16] Christie *et al*[17]), but the clinical impact on oxygen-related morbidity and mortality has not yet been explored.

One limitation of our study is that our data do not allow to assess the accuracy of SpO_2 readings with the different sensor types in comparison to SaO_2. However, because current recommendations on SpO_2 targeting are based on measurements with LNCS, we aimed to verify the agreement of newly introduced sensors with the previous 'standard'.

CONCLUSION

Our study results show a systematic difference in SpO_2 readings between RD sensors and LNCS. Particularly for NICUs that aim for the upper NeoPROM target range (91–95%, centre value 93%), this may result in an unintendedly high oxygen exposure when replacing LNCS by RD sensors without adjusting the SpO_2 target range (ie, a median value of 93% with RD technology might represent a value of 95% with the LNCS). This may impact clinical outcomes in extremely preterm infants and should lead to caution when implementing changes in SpO_2 technology in an NICU, irrespective of the manufacturer and also when transferring an SpO_2 target range from one to another oximeter technology. Independent international standardisation of pulse oximetry technology would be desirable.

Acknowledgements We gratefully acknowledge the families for their willingness to participate in this study, as well as the contribution from Masimo Corporation, Irvine, for providing the required LNCS, RD and PPG sensors for this study.

Contributors CAM, AF, LS and CFP were involved in the study design. CAM, CES and KB collected the data and performed together with AF echocardiographic examinations (all echocardiographic examinations were validated by AF). CAM analysed the data, drafted the first version of the manuscript and is responsible for the overall content as the guarantor. AF, CFP, LS, KB and CES reviewed the manuscript with respect to clinical interpretation of the data and made important contributions. All authors have reviewed and approved the final version of the manuscript.

Funding Masimo Corporation, Irvine, provided the required LNCS, RD and PPG sensors for this study. No other specific grants were received from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests CFP received advisory board-honoraria from Masimo, Irvine, California in 09/2020. All other authors have indicated they have no conflicts of interests relevant to this article to disclose. AF and CFP declare that Masimo generously supported SpO_2 measurements in a previous and an ongoing clinical trial. In this study, Masimo provided also the required LNCS, RD and PPG sensors. However, Masimo had no impact on the design of this study, analysis of the data and writing of this manuscript.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by Ethics Committee of University Hospital Tuebingen, reference: 366/2021BO2. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

ORCID iDs
Christian Achim Maiwald http://orcid.org/0000-0002-2240-6488
Christoph E Schwarz http://orcid.org/0000-0003-0975-1575
Laila Springer http://orcid.org/0000-0001-6284-9673
Christian F Poets http://orcid.org/0000-0002-1072-0066
Axel Franz http://orcid.org/0000-0001-5066-0857

REFERENCES

1 Cherian S, Morris I, Evans J, *et al*. Oxygen therapy in preterm infants. *Paediatr Respir Rev* 2014;15:135–41.
2 Askie LM, Darlow BA, Davis PG, *et al*. Effects of targeting lower versus higher arterial oxygen saturations on death or disability in preterm infants. *Cochrane Database Syst Rev* 2017;4:CD011190.
3 Saugstad OD. Oxygenation of the immature infant: a commentary and recommendations for oxygen saturation targets and alarm limits. *Neonatology* 2018;114:69–75.
4 Sweet DG, Carnielli V, Greisen G, *et al*. European consensus guidelines on the management of respiratory distress syndrome - 2019 update. *Neonatology* 2019;115:432–50.
5 Masimo International. Product information - newborn sensors; copyright 2020 downloaded on May, 08. 2023. Available: https://www.masimo.com/siteassets/us/documents/pdf/plm-10199b-product-information-newborn-sensors-us.pdf
6 Masimo International. Brochure - RD SET sensors; copyright 2020 downloaded on May, 08. 2023. Available: https://www.masimo.com/globalassets/us/documents/pdf/plm-10041f-brochure-rd-set-sensors-us.pdf
7 Masimo International. Product information - radius PPG; copyright 2020 downloaded on May, 16. 2023. Available: https://www.masimo.com/siteassets/us/documents/pdf/plm-11706e-product-information-radius-ppg-us.pdf
8 Maiwald CA, Niemarkt HJ, Poets CF, *et al*. Effects of closed-loop automatic control of the Inspiratory fraction of oxygen (Fio(2)-C) on outcome of extremely preterm infants - study protocol of a randomized controlled parallel group multicenter trial for safety and efficacy. *BMC Pediatr* 2019;19:363.
9 Thilo EH, Andersen D, Wasserstein ML, *et al*. Saturation by pulse oximetry: comparison of the results obtained by instruments of different brands. *J Pediatr* 1993;122:620–6.
10 Dawson JA, Saraswat A, Simionato L, *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.
11 Katayama S, Shima J, Tonai K, *et al*. Accuracy of two pulse-oximetry measurements for Intellivent-ASV in mechanically ventilated patients: a prospective observational study. *Sci Rep* 2021;11:9001.
12 Hummler HD, Engelmann A, Pohlandt F, *et al*. Decreased accuracy of pulse oximetry measurements during low perfusion caused by sepsis: is the perfusion index of any value *Intensive Care Med* 2006;32:1428–31.
13 Hummler HD, Engelmann A, Pohlandt F, *et al*. Accuracy of pulse oximetry readings in an animal model of low perfusion caused by emerging pneumonia and sepsis. *Intensive Care Med* 2004;30:709–13.
14 Hummler HD, Pohlandt F, Franz AR. Pulse oximetry during low perfusion caused by emerging pneumonia and sepsis in rabbits. *Crit Care Med* 2002;30:2501–8.
15 Bachman TE, Newth CJL, Iyer NP, *et al*. Hypoxemic and hyperoxemic likelihood in pulse oximetry ranges: NICU observational study. *Arch Dis Child Fetal Neonatal Ed* 2019;104:F274–9.
16 Wackernagel D, Blennow M, Hellström A. Accuracy of pulse oximetry in preterm and term infants is insufficient to determine arterial oxygen saturation and tension. *Acta Paediatr* 2020;109:2251–7.
17 Christie FG, Kelly R, Boardman JP, *et al*. Measuring oxygenation in newborn infants with targeted oxygen ranges (MONITOR): a randomised crossover pilot study. *Arch Dis Child Fetal Neonatal Ed* 2023;108:638–42.

Online supplemental material

Supplementary Table 1: Lot-numbers used in group 2

	LNCS	PPG	RD
group2 - infant 1	22CXP	20KGJ	22E53
group2 - infant 2	22CMY	A20EG73	21GWF
group2 - infant 3	19MSK	22FSO	21CTK
group2 - infant 4	22FAK	22DMQ	22FUA
group2 - infant 5	22EUC	A21CJ31	21M7Z
group2 - infant 6	22CCV	22H6J	21JZC

Supplementary Table 2: Randomization-Clusters

Phase	**I (2 hours)**			**II (2 hours)**			**III (2 hours)**		
Sensor position	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
Variant 1	**LNCS**	**RD**	**PPG**	**PPG**	**LNCS**	**RD**	**RD**	**PPG**	**LNCS**
Variant 2	**PPG**	**LNCS**	**RD**	**RD**	**PPG**	**LNCS**	**LNCS**	**RD**	**PPG**
Variant 3	**RD**	**PPG**	**LNCS**	**LNCS**	**RD**	**PPG**	**PPG**	**LNCS**	**RD**
Variant 4	**LNCS**	**PPG**	**RD**	**RD**	**LNCS**	**PPG**	**PPG**	**RD**	**LNCS**
Variant 5	**RD**	**LNCS**	**PPG**	**PPG**	**RD**	**LNCS**	**LNCS**	**PPG**	**RD**
Variant 6	**PPG**	**RD**	**LNCS**	**LNCS**	**PPG**	**RD**	**RD**	**LNCS**	**PPG**

Supplementary Table 3: Demographic data

Pat-ID	Gender	Weight in g	PMA in weeks	Age in days	Weight in g	PMA in weeks	Mean FiO2	FiO2-Controlling in infants with FiO2 > 0.21
		on the day of birth		on the day of examination				
1	female	945	29 1/7	6	1090	30	0.21	-
2	female	1085	28 3/7	10	1190	29 6/7	0.21	-
3	female	600	26	40	1370	31 5/7	0.34 - 0.71	FiO2-C
4	male	760	27 1/7	13	920	29	0.21 - 0.27	-
5	male	730	27 6/7	12	990	29 4/7	0.21 - 0.25	-
6	male	540	28	10	630	29 3/7	0.21	-
7	female	1270	28 5/7	10	1370	30 1/7	0.21 - 0.24	-
8	male	1150	28 5/7	9	1260	30	0.21	-
9	female	820	27 6/7	9	890	29 1/7	0.21	-
10	female	1500	31 5/7	2	1450	32	0.21	-
11	female	1170	31 5/7	2	1210	32	0.21	-
12	male	670	24 5/7	6	640	25 4/7	0.22 - 0.28	RMC
13	female	1425	31	8	1478	32 1/7	0.21	-
14	female	1440	31	8	1494	32 1/7	0.21	-
15	male	860	26 6/7	13	1060	28 5/7	0.22 - 0.30	FiO2-C
16	female	440	23	54	880	30 5/7	0.21 - 0.25	-
17	male	960	27 6/7	24	1234	31 2/7	0.24 - 0.26	RMC
18	male	290	24 1/7	85	1270	36 2/7	0.21 - 0.24	-
Group 1:	10x female 8x male	925	28 0/7	18	1135	30 4/7	9x FiO2 = 0.21 5x FiO2 ≥ 0.21 4x FiO2 > 0.21	2x RMC 2x FiO2-C
		±345	±2 3/7	±21	±259	±2 1/7		
"+1"	female	990	30	14	1290	32	0.21	-
"+2"	male	550	24 2/7	51	1120	31 4/7	0.24 - 0.28	RMC
"+3"	female	755	25 1/7	12	830	26 6/7	0.24 - 0.28	FiO2-C
"+4"	male	550	24 1/7	51	990	31 3/7	0.30 - 0.40	FiO2-C
"+5"	male	1050	27 5/7	7	1000	28 5/7	0.21 - 0.23	-
"+6"	male	390	27 6/7	19	860	30 4/7	0.30 - 0.35	FiO2-C
Group 2:	2x female 4x male	714	26 4/7	26	1015	30 1/7	1x FiO2 = 0.21 1x FiO2 ≥ 0.21 4x FiO2 > 0.21	1x RMC 3x FiO2-C
		±241	±2 1/7	±18	±156	±1 6/7		
Total (mean ±SD)	12x female 12x male	873	27 4/7	20	1105	30 3/7	10x FiO2 = 0.21 6x FiO2 ≥ 0.21 8x FiO2 > 0.21	3x RMC 5x FiO2-C
		±335	±2 3/7	±20	±243	±2 0/7		

Supplementary Table 4: Comparison of pulse rates

	LNCS		**RD**			**PPG**		
Pulse-Rate (bpm)					***vs. LNCS***			***vs. LNCS***
Mean (±SD)								
- all infants:	***162.2***	***(±12.0)***	***162.2***	***(±11.9)***	***-***	***162.3***	***(±11.9)***	***-***
-- group 1:	*160.9*	*(±11.4)*	*160.9*	*(±11.3)*	*-*	*160.9*	*(±11.3)*	*-*
-- group 2:	*166.1*	*(±12.9)*	*166.1*	*(±12.9)*	*-*	*166.1*	*(±12.9)*	*-*
Mean difference to LNCS (95%-CI of mean)								
- all infants:	***-***	***-***	***-0.003***	***(-0.017 to 0.011)***	***-***	***-0.007***	***(-0.006 to 0.021)***	***-***

Supplementary Figure 1: Bland-Altman-Plots of pulse rate-readings and SpO_2-values (dots representing individual simultaneous SpO_2-readings; horizontal lines indicate mean deviation ±SD; X-axes indicating mean of two corresponding measurements; Y-axes indicating the difference between the two measurements), showing no systematic deviation in the comparisons of pulse rate between sensor types (Panels 1-4), and systematically lower SpO_2-values for RD vs LNCS (Panel 5+6) and slightly higher values for PPG vs. LNCS (Panel 7+8)