"New-Generation” Pulse Oximeters in Extremely Low-Birth-Weight Infants

How Do They Perform in Clinical Practice?

Anne C. van der Eijk, MSc; Sandra Horsch, MD, PhD; Paul H. C. Eilers, PhD, BSc; Jenny Dankelman, PhD; Bert J. Smit, MD, PhD

ABSTRACT
The aim of this study was to evaluate the performance of “new-generation” pulse oximeters in extremely low-birth-weight (≤1000 g) infants. In a prospective crossover observational study, the performance of pulse oximeters of 3 brands (Masimo, Nellcor, and Philips) was evaluated by dual Spo2 measurement in ELBW infants. Disposable probes of either equal or different brands were placed around both feet of the patient simultaneously for approximately 4 hours. Probes were switched between feet every hour. Absolute differences in Spo2 values (ΔSpo2) and the bias between brands were studied. Nine ELBW infants were included (gestational age: mean ± SD = 26/7 ± 1/7 weeks). The median (range) ΔSpo2 was 2% (0%-26%). In 9% of the time, ΔSpo2 was 5% or more. The variance of the difference of the 3 pulse oximeter brands was not significantly different. No consequent bias between brands was found. Simultaneously obtained pulse oximeter measurements from the feet of ELBW infants differ from each other. Our results suggest that it is not the brand but the handling of the pulse oximeter in clinical practice, such as the place and positioning of the probe, that influences the performance of the pulse oximeter the most. Improvement in the accuracy of oxygen-monitoring techniques for ELBW infants is required.

Key Words: monitoring, oxygen, premature, preterm infants, pulse oximetry

In extremely low-birth-weight (ELBW) infants, oxygenation is monitored closely because of its narrow therapeutic range. Pulse oximetry is a patient-friendly and noninvasive method, there are some well-known disadvantages of this technique. The accuracy of pulse oximetry is easily influenced by movement artifacts, low blood perfusion, and ambient light. Because of these limitations, pulse oximetry is notorious for its high alarm rate in the NICU. Furthermore, the technique was not designed for detecting hyperoxia and cannot counteract for the presence of fetal hemoglobin. The latter 2 limitations are especially important in preterm infants because neonatal hyperoxia is associated with injury to amongst others the developing brains, lung, and retina.

In the last 2 decades, several brands have introduced “motion-resistant” pulse oximeters into the market. The performance of both these “new-generation” pulse oximetry devices and the conventional ones has been evaluated extensively in (preterm) neonates.
Although conclusions varied between studies, Giuliano et al\textsuperscript{23} concluded in a review that the clinical performance of new-generation pulse oximeters is better than conventional pulse oximeters. They did not find evidence for differences in performances between brands.

In the years that pulse oximetry technology has advanced, survival rates of ELBW infants have increased remarkably.\textsuperscript{24,25} Among survivors, the incidence of oxidative stress–related diseases is considerable.\textsuperscript{26} Hence, reliable monitoring of oxygenation is essential in this most vulnerable group of patients.

In our department, caregivers had diverse impressions of the performance of pulse oximeter brands when used with ELBW infants. Therefore, the aim of this study was to evaluate the performance of new-generation pulse oximeters in ELBW infants. To objectively evaluate the performance, a prospective crossover observational study was designed in which pulse oximeter measurements of either equal or different brands were compared by simultaneous dual measurement.

**PATIENTS AND METHODS**

The prospective crossover observational study was performed at a level III-c\textsuperscript{27} NICU. Approval was obtained from the ethical review board of the Erasmus Medical Center—Sophia Children’s Hospital, Rotterdam, the Netherlands. Informed written consent by the parents was obtained for all included infants prior to inclusion.

**Patients**

Infants were eligible for inclusion when the gestational age was less than 30 weeks, the birth weight was 1000 g or less, and the postnatal age of the infant was less than 14 days. Data recording started when it was possible to place a disposable pulse oximeter probe on both feet (for study purpose) and on the left hand (for clinical purpose). Infants were excluded when the status of the skin was poor (eg, hematoma or skin lesions), or when infants were, according to the responsible medical staff member, too unstable to handle.

**Methods**

Two, brand new, disposable new-generation pulse oximeter probes from 1 of the 3 brands used in the study were placed around either foot by the nurse taking care for the infant. The devices used to collect the data and the characteristics for each pulse oximeter brand are shown in Table 1. The researcher (A.C.v.dE.) told the nurse which probe should be placed on which foot. For example, probe I of brand A (IA) was placed on the left foot and probe II of brand A (IIA) on the right foot. To prevent interference of light, and thus erroneous measurements, all probes used during this study were covered by an extra opaque wrap, as suggested by Ralston et al\textsuperscript{28} and Fouzas et al.\textsuperscript{29}

After minimal 50 minutes of dual Sp\textsubscript{O}\textsubscript{2} monitoring, the probes were switched over to the other foot. Table 2 provides the study design and especially which probe was positioned on which foot. Thus, following the example, probe IA was now placed on the right foot and probe IIA on the left foot. After the second hour, one of the probes was removed and replaced by a pulse oximeter from a different brand; probe IA stayed on the right foot, probe IIA was removed from the infant, and probe I of brand B (IB) was positioned on the left foot. After the third hour, the pulse oximeter probes were switched again to the other foot. Thus, probe IB was now positioned on the right foot and the probe IA on the left foot. After the fourth hour, the recordings ended, and both the pulse oximeter probes were removed from the infant. To prevent that caregivers would be influenced by the 2 extra pulse oximeter measurements in their daily care, Sp\textsubscript{O}\textsubscript{2} values obtained by the 2 additional pulse oximeters were blinded for the nursing staff.

A multiparameter patient monitor was used for monitoring Sp\textsubscript{O}\textsubscript{2} with Masimo or Philips probes (MP50, modules: M1020B A01 and A03, MMS: M3001 A; Philips Medical Systems, Boeblingen, Germany). Dual Sp\textsubscript{O}\textsubscript{2} monitoring by Nellcor was performed by 2 stand-alone pulse oximeters (model: N600x). The averaging time of the software of both Masimo and Philips was 5 seconds. The Nellcor software could not be set to 5 seconds, but it required a range for the averaging time. The range closest to 5 seconds was chosen; subsequently, the averaging time was 2 to 4 seconds. A laptop was used to record Sp\textsubscript{O}\textsubscript{2} measurements (Software: “Tera Term,” open source, and “TrendFace”; Ixellence GmbH, Wildau, Germany).

**Data analysis**

To prevent artifacts in Sp\textsubscript{O}\textsubscript{2} values caused by (re)positioning of the probes and to make sure that the data used for analysis were exactly the same length, time periods of 45 minutes were analyzed. The data obtained in the first minute after the moment that both pulse oximeters provided an Sp\textsubscript{O}\textsubscript{2} value were excluded from analysis. The next 2nd up to 46th minutes of collected data were used for analysis. All Sp\textsubscript{O}\textsubscript{2} data were interpolated to 1 Hz and rounded to integers. The absolute difference between simultaneously obtained Sp\textsubscript{O}\textsubscript{2} values was defined as ΔSp\textsubscript{O}\textsubscript{2}. When one or both pulse oximeters did not provide an Sp\textsubscript{O}\textsubscript{2} value, these time periods were defined as “dropouts” and excluded from the calculation of ΔSp\textsubscript{O}\textsubscript{2}.

There was no gold standard to compare the pulse oximeter measurements with; ΔSp\textsubscript{O}\textsubscript{2} is always the difference between 2 sensors (either of equal or different
brands). So it is not possible to detect any bias. We investigated whether there is a difference in variance. The variance of a difference of the 2 pulse oximeters is the sum of the variance of the individual variances. The study setup was designed in such a manner that all possible combinations of brands were tested, thus a linear regression model, for the observed variance of the difference, could be developed. The independent variables are indicators of 2 brands that are compared, and the dependent variable is the observed variance.

RESULTS

Patients
Nine infants, all white, with a mean ± SD gestational age of 263/7 ± 14/7 weeks, birth weight of 825 ± 136 g, and postnatal age of 7 ± 4 days, were included. The patient characteristics are shown in Table 3. In total, 36 periods of 45 minutes (ie, 27 hours) of dual Spo2 measurements were analyzed. During the study, infants were either on invasive ventilation (9 hours) or continuous positive airway pressure (18 hours).

Spo2 values
For each infant, the median (min-Q1-Q3-max) of the Spo2 was determined, using the Spo2 measurements of the research probe placed on the infant for the complete 4 × 45 minutes of recording (the 2nd till 46th minutes from each measurement period; see Table 3). In the 27 hours of dual Spo2 recording, 6 dropouts were observed in total in 4 different infants (4 times Philips, once Masimo, and once Nellcor). These dropouts lasted between 2 and 21 seconds (in total 0.5% of the recording time). Never more than 1 pulse oximeter dropped out at the same time.

Table 1. Pulse oximeter characteristics and specifications of the data recording

<table>
<thead>
<tr>
<th>Pulse oximeter characteristics</th>
<th>Experimental setup</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Brand</strong></td>
<td><strong>Software</strong></td>
</tr>
<tr>
<td>Masimo</td>
<td>LNOP NeoPt-L &lt;1 kg</td>
</tr>
<tr>
<td>Philips</td>
<td>FAST</td>
</tr>
<tr>
<td>Nellcor OxiMax</td>
<td>Softcare SC-PR &lt;1.5 kg</td>
</tr>
</tbody>
</table>

Abbreviation: Spo2, oxygen saturation measured by a pulse oximeter.

Table 2. Study designa

<table>
<thead>
<tr>
<th>Number of subjects</th>
<th>Dual Spo2 measurement by</th>
<th>Pulse oximeter probe</th>
<th>Left foot</th>
<th>Right foot</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Masimo SET-Masimo SET</td>
<td>Masimo probe 1</td>
<td>Masimo probe 2</td>
<td>∼1 h</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Masimo SET-Nellcor OxiMax</td>
<td>Masimo probe 2</td>
<td>Masimo probe 1</td>
<td>∼1 h</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Nellcor OxiMax-Nellcor OxiMax</td>
<td>Nellcor probe 1</td>
<td>Nellcor probe 1</td>
<td>∼1 h</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Nellcor OxiMax-Philips FAST</td>
<td>Nellcor probe 2</td>
<td>Nellcor probe 2</td>
<td>∼1 h</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Philips FAST-Philips FAST</td>
<td>Philips probe 1</td>
<td>Philips probe 2</td>
<td>∼1 h</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Philips FAST-Masimo SET</td>
<td>Masimo probe 1</td>
<td>Masimo probe 1</td>
<td>∼1 h</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: Spo2, oxygen saturation measured by a pulse oximeter.

aTo compare each of the 3 brands with each other, 3 different tests were performed. Each test was performed with 3 subjects (column 1). Two pulse oximeter probes of either equal or different brands were placed on the feet of the subject for approximately 4 hours (column 2). During these 4 hours, pulse oximeter probes were switched between feet every hour (column 3).
Desaturations of less than 70% showed by at least 1 pulse oximeter were observed in 34 moments in 5 different infants. In 11 (32%) of the 34 occasions, the second pulse oximeter showed a desaturation of less than 70% as well. Hence, in 23 (67%) of the 34 occasions, only 1 of the pulse oximeters showed a desaturation of less than 70%. In 6 (26%) of 23 times, the desaturation lasted for more than 20 seconds.

Figures 1 to 4 show examples of typical situations that occurred during the recording and corresponding frequency histograms of the Spo2 measurements. In Figure 1, the Spo2 values of the right foot are continuously higher than Spo2 values of the left foot. After switching the probes to the other foot, the measurements from the right foot are still higher than the left foot. Figure 2 shows that ΔSpo2 varies over time in
After changing the probes to the other foot, $\Delta$SpO$_2$ is smaller. Figure 3 shows 2 examples of a desaturation of less than 80% by only one of the pulse oximeters. Figure 4 shows 2 desaturations of less than 50% by 2 pulse oximeters at the same time.

Median (range) $\Delta$SpO$_2$ was 2 (0%-26%), and $\Delta$SpO$_2$ varied over time. In 9% of the time, $\Delta$SpO$_2$ was 5% or more (drop-out time excluded). In 31 of the 36 time periods, the standard deviation of the difference was between 1.6% and 3.8%. In the remaining 5 time periods, the standard deviation of the difference was higher (range: 6.0%-9.4%). These relatively large differences were related to deep desaturations in only one pulse oximeter (eg, Figure 3), and/or because there was a short time difference between the pulse oximeter measurements. These situations are present in clinical practice but are not representative for normal operation. Their impact on the linear regression model was large. Therefore, it was decided to exclude these exceptional situations in the calculations, resulting in variances of 2.0, 2.9, and 3.3 for the 3 brands, with standard errors (SEs) around 0.7. The SE of the differences between the pulse oximeter measurements was 1.0. Dividing the differences of the means by the SE of the differences results in 0.8, 0.5, and 1.3. Even in case of a normal distribution, the results would have to be larger than 1.96 to indicate a significant difference ($P < .05$). Thus, no significant differences between the variances of the difference of the 3 brands were found. It is more convenient to express variances as standard deviations, by taking square roots. We obtained the values 1.4, 1.7, and 1.8 in units of $\Delta$SpO$_2$.

**DISCUSSION**

Simultaneously obtained SpO$_2$ measurements with new-generation “motion-resistant” pulse oximeters from both feet of 9 ELBW infants differ from each other (median [range] $\Delta$SpO$_2$ was 2.0 [0%-26%]), both for equal and different brands. The differences found between the pulse oximeter measurements varied over time among infants for all 3 tested brands. These differences may have an effect on the outcome of extremely preterm infants because the oxygenation of these vulnerable patients has a narrow therapeutic range. The variances of the fluctuations of the 3 pulse oximeter brands were not significantly different. Because it is unlikely that the SpO$_2$ values in both feet of these stable infants are really different, our results suggest the handling of the pulse oximeter in clinical practice influences the performance of the pulse oximeter the most.

When studying pulse oximetry in clinical practice, there is a lack of continuous knowledge about the real SaO$_2$ values of the patient. Because the actual SaO$_2$ value was not known, there was no criterion standard. Consequently, it is impossible to tell whether both, one, or no pulse oximeters provided SpO$_2$ values that were...
Figure 3. (A and B) Two sets of 10 minutes of dual Spo2 measurement in 1 patient. Both figures show short desaturations less than 80% by only one of the pulse oximeters. (A) Spo2 values obtained by equal brands. (B) Spo2 values obtained by different brands. (C and D) The frequency histograms for proportion of time spent at each saturation value for the complete 45 minutes of recording, including the 10 minutes of measurement shown in (A) and (B).

Figure 4. (A and B) Two sets of 10 minutes of dual Spo2 measurement in 2 different patients. Both figures show large desaturations registered by both pulse oximeters. (A) Spo2 values obtained by equal brands. (B) Spo2 values obtained by different brands. (C and D) The frequency histograms for proportion of time spent at each saturation value for the complete 45 minutes of recording, including the 10 minutes of measurement shown in (A) and (B).
matching the actual SaO2 values. Thus, on the basis of the setup of this study, it is impossible to judge which of the 3 brands performed best in terms of correct measurement or other characteristics (eg, easy to handle, time to first measurement). However, finding the best performing pulse oximeter was not the goal of the study. An earlier performed review by Giuliano et al23 already mentioned that they did not find evidence for differences in performances between pulse oximeter brands. Our results confirm their conclusion by showing that the variances of the difference of the 3 pulse oximeter brands did not differ significantly. We aimed to evaluate the performance of new-generation pulse oximeters in ELBW infants in clinical practice. Knowledge about the usefulness of pulse oximeters in ELBW infants helps us to improve strategies for ventilation and patient care in general.

A limitation of this study was the different averaging time between pulse oximeter brands. It has been shown that the averaging time of pulse oximeter software influences the saturations values on the monitor.21,22 A shorter averaging time makes the pulse oximeter more sensitive for changes in SpO2 but increases the chance of a false alarm due to artifacts. However, the averaging time was comparable (between 2 and 5 seconds) for all pulse oximeters. In addition, this limitation reflects a problem in clinical practice that caregivers confront with.

Variation in pulse oximetry readings is inherent to pulse oximetry and has been mentioned before.2,6,7,15−21 However, especially in ELBW infants, these (large) inconsistent deviations are, regarding the risks for hypoxia and hyperoxia, undesirable and should be minimized. The causes for the deviations as seen in this study are in literature often indicated as multifactorial and could find their origin in the patient, the software of the pulse oximeter, and/or the probe position and placement. The difficulty in clinical practice is that the actual cause of the deviations cannot always be identified.

In theory, SaO2 levels in both feet of a healthy infant are the same because both feet are in the postductal region of the body.30−33 Therefore, the origin of differences in simultaneously obtained SpO2 values must be found in the pulse oximetry measurement. It is conceivable that there are differences between extremities of the infant such as movement and low blood perfusion. The differences between extremities could be the reason for the deviating SpO2 values in Figure 1. The well-known “motion artifacts” in pulse oximetry are mainly due to irregular venous blood flow during motion. Pulse oximeters are based on the assumption that there is only pulsating blood flow in the arterial compartment; the irregular flow of the venous blood influences the SpO2 measurement. A low blood perfusion may influence the SpO2 measurement due to a poor signal quality caused by a lack of clear pulsations.

Poor positioning of the probes can also cause differences between pulse oximeter measurements, which may be the case in Figure 2. False readings due to inappropriate probe placement were first mentioned in the 1990s.34,35 Although manufacturers since then have been working on the improvement of the software to recognize sensor displacement, appropriate probe placement is still a requirement to obtain reliable pulse oximeter readings. In this study, the probes were brand new and positioned carefully to obtain high performance. In daily practice, it is possible that deviations of pulse oximeters may even be larger because of poor positioning of the probes and because of using probes longer than recommended.36 To prevent movement of the probe, the probe should be positioned firmly and close to the skin, preferably with an extra cover to prevent influence of ambient light.20,26 However, the weak skin of ELBW infants often hampers the firm positioning of probes.

A challenge for pulse oximeter manufacturers is to deal with moments of movements, low blood perfusion, and poor sensor placement, leading to unreliable measurement. On the one hand, manufacturers want to provide an SpO2 value as often as possible and, on the other hand, there are moments where the measurement is not reliable enough to provide SpO2 values. Because the software used by the manufacturers differs, there are moments that pulse oximeters of one brand do not provide SpO2 values whereas others do, and vice versa.

The challenges in the development of pulse oximetry have led to a relative wide accuracy range. Technical specifications for pulse oximeters used for neonates suggest an accuracy of “±2 or 3 digits (1 SD)” between 70% and 100% saturation. The accuracy under 70% saturation is often unspecified. The “±2 or 3 digits (1 SD)” means that the standard deviation is equal to 2 or 3 digits. In other words, in case of a standard deviation of “2 digits,” an SpO2 value of 90% is with 0.68 certainty within 90 ± 2 digits (ie, 88%-92%) and with 0.95 certainty within 90 ± 4 digits (ie, 86%-94%). The latter range is of the same magnitude as the latest advised allowed range for the SpO2 in ELBW infants (ie, 85%-93%).37 In case of a standard deviation of 3%, the accuracy range of pulse oximeters is even wider. This relative wide accuracy range is, regarding the risk for developing hypoxia and hyperoxia in ELBW infants, undesirable and should be improved.

Besides the possible improvements that should be made in pulse oximetry (probes) by the manufacturers, caregivers themselves can play an important role in optimizing pulse oximetry. Caregivers
References


32. Meier-Stauss P, Bucher H, Hürlimann R, König V, Huch R. Pulse oximetry used for documenting oxygen saturation and are the users of the technique and have to interpret pulse oximeter measurements. Although the basics of pulse oximetry are described extensively in literature, it has been shown that knowledge about the working principles, the limitations, and the correct method for interpretation of pulse oximetry among caregivers still need improvement. Therefore, the working principles of pulse oximetry and the interpretation of the Spo2 values should be educated extensively and frequently among caregivers.

To conclude, in clinical practice, simultaneously obtained pulse oximeter measurements from both feet of ELBW infants differ from each other, both for equal and different brands. No systematic deviations between pulse oximeter measurements were found in all 3 tested brands. Our results suggest that it is not the brand but the handling of the pulse oximeter in clinical practice, such as the place and positioning of the probe, that influences the performance of the pulse oximeter the most. Although the introduction of pulse oximetry into the NICU has led to a major improvement in patient monitoring, the limitations of pulse oximetry emphasize the need for the improvement and/or the development of new techniques for both reliable and user-friendly oxygenation measurements in ELBW infants.


