The Effects of Motion Artifact and Low Perfusion on the Performance of a New Generation of Pulse Oximeters in Volunteers Undergoing Hypoxemia

Hartmut Gehring MD, Christoph Hornberger PhD, Holger Matz ME, Ewald Konecny PhD, and Peter Schmucker MD

INTRODUCTION: Motion artifact and low perfusion often lead to faulty or absent pulse oximetry readings in clinical practice. OBJECTIVE: Determine the impact of motion artifact and low perfusion on newly introduced pulse oximetry technologies during hypoxemic episodes in healthy volunteers. METHODS: Five different pulse oximeters from 4 manufacturers (the Datex Ohmeda 3900P; the Agilent; the Nellcor N-3000; the Nellcor N-395; and the Schiller OX-1, which is the European version of the Ivy SatGuard 2000 with Masimo SET) were compared with respect to their ability (separated or in combination) to provide accurate readings in the presence of motion artifact and low perfusion. Four of these oximeters represent the latest available oximetry technology, and one (the N-3000) represents a previous generation of oximeters. Oxygen saturation values ($S_{pO_2}$) and pulse rate from the oximeters were recorded during episodes of induced hypoxemia in 10 healthy volunteers. Standardized and repeatable motion artifacts were generated by a motion machine and by having the test subject perform tapping and scratching motions. Perfusion to the finger was reduced by an inflatable balloon impinging on the brachial artery. The pulse oximetry readings from the test oximeters were compared to readings from control pulse oximeters on the unper- turbated reference hand. The pulse rates from the test oximeters were compared to the electrocardiographically-measured heart rate. RESULTS: The frequency of faulty readings was increased by increasing motion interference and decreasing perfusion. The $S_{pO_2}$ deviation was within $\pm 3\%$ of the reference reading $> 95\%$ of the time for all instruments during the control desaturation period in the absence of motion and with normal perfusion. With the combination of motion and low perfusion, the $S_{pO_2}$ error was within $\pm 3\%$ less than 62% of the time for all oximeters tested. A significant difference in the frequency of large $S_{pO_2}$ errors was observed only in the direct comparison of the N-395 and N-3000. The N-395 exhibited less frequent $S_{pO_2}$ error exceeding 6% of $S_{pO_2}$ in the combination of the most challenging situations (motion and motion with reduced perfusion). In the same situation the Datex-Ohmeda 3900P and Nellcor N-3000 showed significantly higher pulse rate errors than the other devices (Datex-Ohmeda 3900P 53% of the time and N-3000 37% of the time). CONCLUSIONS: The established model of creating motion artifact and low perfusion is capable of simulating a hierarchy of severe clinical situations. With solely motion or solely reduced perfusion the percentage of errors exceeding $\pm 3\%$ of $S_{pO_2}$ increased by 20% and 10%, respectively, compared to the control period. Simultaneous presence of motion and reduced perfusion leads to a relative incidence of $> 35\%$ of errors $> 3\%$ of $S_{pO_2}$ for the various oximeters. In this situation the N-3000 and the Datex-Ohmeda 3900P exhibited differences between estimated pulse rate and electrocardiographically-measured heart rate $> 25$ beats/min $> 37\%$ of the time. Key words: pulse oximetry, artifact simulation, motion, reduced perfusion, desaturation, monitoring. [Respir Care 2002;47(1):48–60]
Introduction

Pulse oximetry is part of routine basic monitoring in anesthesia, along with electrocardiographic (ECG) and blood pressure monitoring. Pulse oximetry is commonly employed in nearly all areas where patients are at risk of hypoxemia.

In clinical practice, pulse oximetry has to cope with many factors that influence \( S_{pO_2} \) accuracy. In a randomized study of pulse oximetry performance, the pulse oximetry failure rate (temporarily and completely abandoned pulse oximetry monitoring) was 2.5% of all monitored patients, and it increased to 7.2% in patients with American Society of Anesthesiology status 4. The main reasons (59%) for temporary and complete pulse oximetry failure in that study were restricted peripheral perfusion and movement artifacts due to patient restlessness. Minor reasons (41%) for pulse oximetry failure were technical or practical problems (not further defined) and unknown.

In a retrospective study by Reich et al, the incidence of intraoperative loss of pulse oximetry readings for > 10 minutes was 9.2%. Predictors of pulse oximetry failure included American Society of Anesthesiology status 3, 4, and 5 patients undergoing vascular or cardiac surgery, and restricted finger perfusion caused by hypothermia or hypotension.

The conditions that are most challenging for pulse oximeters are patient motion and low perfusion, which can be present in critical situations such as surgery, during patient transport, in emergency departments, and in neonatal intensive care units. We believe that a clinically relevant performance test of pulse oximeters should include a standardized simulation of motion and restricted perfusion during changes in arterial oxygen saturation. So under those conditions we tested instruments from the most recently introduced generation of pulse oximeters from 4 manufacturers. Our goal was to address the question, “Do the new oximeters improve the accuracy of tracking \( S_{pO_2} \) and pulse rate during induced patient hypoxemia in the presence of motion artifact and low perfusion?”

The new-generation instruments we tested were: the 3900P, software version 2.000/04.000 (Datex Ohmeda, Helsinki, Finland); the Agilent with CMS monitor software Rev B.0 (Philips Medical Systems, Böblingen, Germany); Nellcor N-395 with Oximeter XL (software version 1.6.1.0) (Tyco Healthcare Group, Pleasanton, California); the Ivy SatGuard 2000 with Masimo Signal Extraction Technology (SET) (Model 2000, Ivy 2356-00-04 REV 01 DSP 2.2.0.0, MS1:HW REV: A ID:2 SMC: 1.0.2.2, sold by the Swiss company Schiller as model OX-1) (Ivy Biomedical Systems, Branford, Connecticut; Masimo, Irvine, California). The Nellcor N-3000 (software version 3.03 001) was also tested as a representative of the previous generation of oximeters.

Methods

Ten healthy volunteers (8 men and 2 women, 30–37 years of age) participated in the study. The study was approved by the Ethics Committee of the Medical University of Lübeck. Prior to the study, participants were checked by an independent cardiologist and had no evidence of vascular, cardiopulmonary, or other systemic disease. After written informed consent was obtained, all volunteers participated in a training desaturation procedure one month before the motion study, to get used to the test situation.

To assess the effect of motion artifact and low perfusion on the pulse oximeters’ ability to accurately detect changes in arterial oxygen saturation, we tested the oximeters during a controlled desaturation procedure in which both hands were at heart level. The reference oxygen saturation was recorded on the right hand (reference hand), using 2 Nellcor N-3000 pulse oximeters. The left hand (test hand) was subjected to motion and low perfusion. Heart rate was recorded with an ECG, using electrodes placed on the chest.

Controlled Desaturation Procedure

Participants breathed through a tight-fitting face mask connected to a Trajan 808 (Dräger Medizintechnik, Lübeck, Germany) with fresh gas supply of 14–17 L/min, using a valveless high-flow system. The expiration side of the Y-piece was 1 m in length, providing uncontrolled inspiratory gas mixture. Fraction of inspired oxygen was varied by adjusting oxygen and nitrogen with the system’s flow meters. Inspiratory and expiratory oxygen and carbon dioxide concentration and breathing pattern were continuously monitored by a Datex Ultima V (Datex Medical Instruments, Helsinki, Finland). The bidirectional, pressure-based flow sensor (D-lite, Datex Medical Instruments, Helsinki, Finland) with side stream system was introduced into the air flow between the Y-piece and the mask.
Motion

We used a mechanical motion generator to move the test hand with repeatable computer-generated irregular patterns, to produce a standardized motion artifact (Figure 1). The maximum motion amplitude was 38 mm, which is a range used in previous studies and which represents a substantial challenge to pulse oximeters. The minimum period for a 38 mm movement was about 0.1 second. Irregular patterns were thought to be more realistic than simple repetitive sine wave motions. A similar motion simulator was used by Barker and Shah in a previous study. Additionally, in order to simulate more clinically relevant motion patterns, we asked each volunteer to perform a tapping motion for 120 seconds, followed 60 seconds later by a scratching motion for 120 seconds. A training video showing the desired motion patterns was presented to the volunteers before the study. The voluntary motion patterns were synchronized using a nonperiodic acoustic signal provided to the volunteers to define start and stop points for each tapping and scratching motion.

Low Perfusion

We produced varying levels of finger perfusion by compressing the brachial artery of the left upper arm. To control arterial perfusion we used a specially designed clamp that encircles the arm, without causing venous congestion, while holding a balloon in a fixed position above the artery. A similar system was established in a previous study. The perfusion was controlled by inflating the balloon under continuous pressure measurement. The balloon pressure was increased until the modulation depth of the photoplethysmographic signal was decreased to less than 2%. The perfusion on the reference hand was monitored with a Nellcor N-395 pulse oximeter. In this way the perfusion of the fingers was held at a nearly constant level, lower than the perfusion of the reference hand.

Three of the test instruments provide a pulse strength index that is related to the perfusion or percent modulation of the photoplethysmographic signal. In the case of the Agilent monitor it is called perfusion index and is the percentage modulation of the transmission signal corrected for the influence of saturation. Normal values are between 0.1% and 20%. The Nellcor N-395 provides a pulse amplitude indicator, which is a number between 0 and 200. The Datex Ohmeda 3900P provides a pulsation index ranging from 0.00 to 9.99. For all instruments, a higher value means larger amplitude pulses. To make the values comparable, the Nellcor pulse amplitude value was divided by 10 and the Datex-Ohmeda pulsation index was multiplied by 2. All values are thus comparable to the Agilent percentage value, ranging from 0 to 20%. This variable was used to monitor our method of reducing perfusion.

Protocol

Figure 2 describes the procedure. The subject’s oxygen saturation started at 100% and was decreased to 90%, where it was held constant for about 1 minute. Then it was decreased further to about 80%, where it again was held constant. This was followed by an increase to 90% (constant period again), and then an increase to 100% oxygen saturation, where it was held constant for at least 2 minutes. Motion started at the 100% plateau. In Figure 2, motion periods and periods when the balloon was inflated are indicated by bars. The maximum pressure was adjusted individually for each volunteer. At the end of the study, we added another desaturation period, which involved low perfusion but no motion artifact.

Data Acquisition

In order to report the displayed oxygen saturation and the alarm status of the instruments, we continuously recorded the following variables:

- \( S_{\text{O}_2} \) of every instrument simultaneously (control pulse oximeters and pulse oximeters being tested)
- Pulse rate
- ECG-measured heart rate, recorded with the Agilient CMS monitor

---

Fig. 1. The motion machine (top) and the tapping (bottom left) and scratching (bottom right) motions. The left part of the table supporting the arm and hand is pivoted and the rod on the right side moves the table up and down. The arm pressure cuff for the reduction of perfusion is indicated at the left side.
An index related to the perfusion or percent modulation (Agilent perfusion index, Nellcor N-395 pulse amplitude, Datex-Ohmeda pulsation index).

- All alarm variables and warnings

**Settings**

The alarm limit for low saturation was set to 92% for all instruments. The pulse limits were set to 40 (low pulse rate) and 150 (high pulse rate).

The Agilent response time was set to medium, which is about 10 seconds.

The Datex-Ohmeda averaging time was set to 6 seconds.

The Ivy 2000 with Masimo SET was set to 10 seconds averaging, and its sensitivity was set to normal.

The response time of the Nellcor pulse oximeters is not accessible to users. It is nominally about 7 seconds.

**Sensors**

For the sensors on the test hand we used adhesive sensors so as to avoid the additional contribution to motion artifact that can be expected from the more massive reusable sensors. The sensors we used were:

- OxyTip adult sensor (Datex-Ohmeda)
- Nellcor Oxisensor II D25 (Tyco Healthcare Group) for Agilent, Nellcor N-3000, and Nellcor N-395
- Masimo SET LNOP Adt (Masimo) for the Ivy 2000 with Masimo SET

All sensors were protected with opaque covers to prevent light from one sensor reaching another sensor and interfering with performance. The pulse oximeters on the test hand were positioned on randomly-selected fingers, differently for each volunteer.

**Reference**

Two Nellcor N-3000 pulse oximeters equipped with DS-100A Durasensors on fingers of the nonmoving and normally perfused right hand represented the reference. These 2 reference oximeters were used to control the desaturation procedure and served as the standard for comparing the pulse oximetry readings from the moving test hand. The mean values of the pulse oximetry readings from those 2 devices had been compared against arterial blood samples in previous desaturation procedures, and the mean value was an accurate measure of the arterial oxygen saturation.

**Statistics**

The continuously recorded data were analyzed with respect to the number of deviations of pulse oximetry readings and pulse rate, and with respect to alarms. The analysis was divided among various periods, each period representing a different level of signal interference:

- Period 1: no motion, normal perfusion
- Period 2: no motion, low perfusion
- Period 3: motion, normal perfusion
- Period 4: motion, low perfusion

In periods 3 and 4, all motion patterns (machine motion, tapping, and scratching) were merged.

*Pulse oximetry error* was defined as the $S_{pO_2}$ of the test hand minus the $S_{pO_2}$ of the reference hand. The $S_{pO_2}$ of the test hand was the reading of the pulse oximeter under test.
The $S_{pO2}$ of the reference hand was the mean value of the $S_{pO2}$ readings of the 2 Nellcor N-3000s on the reference hand.

For the $S_{pO2}$ error, categories were defined as shown in Table 1. The neutral range was defined as $\pm 3\%$ of the $S_{pO2}$ value, since that is the range that many pulse oximetry manufacturers define as the range of accuracy in the presence of motion (without motion it is often given as $\pm 2\%$ of the $S_{pO2}$ value). These categories are subdivided into false low (index L) and false high (index H) $S_{pO2}$ readings.

From other studies it is known that the incidence of large positive or negative pulse oximetry errors ($\Delta S_{pO2}$) increases as an effect of poor signal quality, as occurs during motion and low perfusion. A broadening of the error distribution is observed, whereas the mean value of $\Delta S_{pO2}$ is less affected. For that reason, when specifying the accuracy of their instruments, most manufacturers provide larger standard deviations under motion conditions. Comparing pulse oximeters in different situations requires a test that is sensitive to the error distribution, so we applied the Wilcoxon matched-pairs signed-rank test to the absolute values of the errors ($|\Delta S_{pO2}|$). In addition, to test the incidence of $S_{pO2}$ errors exceeding 6% and 10%, respectively, a sign test was used. For this purpose new variables $Y_{PO1}$ and $Y_{PO2}$ were generated:

$$Y_{POi} = \begin{cases} 1 & \text{if } \Delta S_{pO2} > c \\ 0 & \text{else} \end{cases}$$

for $c = 6\% S_{pO2}$ and $c = 10\% S_{pO2}$.

The whole protocol described above was performed for each volunteer. $S_{pO2}$ readings from the pulse oximeters

---

Table 1. Definition of Error Categories*

<table>
<thead>
<tr>
<th>False Cumulative</th>
<th>False Low</th>
<th>Subcategory</th>
<th>False High</th>
<th>Subcategory</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>\Delta S_{pO2}</td>
<td>\leq 3$</td>
<td>$-3 \leq \Delta S_{pO2} \leq 0$</td>
<td>$1_L$</td>
</tr>
<tr>
<td>$</td>
<td>\Delta S_{pO2}</td>
<td>&gt; 6$</td>
<td>$-6 \leq \Delta S_{pO2} &lt; -3$</td>
<td>$2_L$</td>
</tr>
<tr>
<td>$</td>
<td>\Delta S_{pO2}</td>
<td>&gt; 10$</td>
<td>$-10 \leq \Delta S_{pO2} &lt; -6$</td>
<td>$3_L$</td>
</tr>
<tr>
<td>$</td>
<td>\Delta S_{pO2}</td>
<td>&gt; 10$</td>
<td>$\Delta S_{pO2} &lt; -10$</td>
<td>$4_L$</td>
</tr>
</tbody>
</table>

*The first two columns define the absolute error categories. In the other columns these categories are subdivided into false low (index L) and false high (index H) $S_{pO2}$ values.

$|\Delta S_{pO2}|$ = absolute value of the error of the oxygen saturation value measured via pulse oximetry.

$S_{pO2}$ = oxygen saturation value measured via pulse oximetry.

See text for explanation of subcategories.

---

**Fig. 3.** Oxygen saturation measured via pulse oximetry ($S_{pO2}$) versus time for 4 test oximeters and the reference $S_{pO2}$ during a desaturation in Period 1 (no motion and normal perfusion). The $S_{pO2}$ differences between the oximeters are within the defined range of $\pm 3\%$.**
were sampled at 1 Hz. However, pulse oximeters do not create new values every second, so adjacent data points are not independent, because of the different averaging algorithms employed in the various oximeters. In order to avoid an unreasonable level of data generation, data were resampled at 0.1 Hz. This new data set was treated as independent data used to generate the error distribution, which served as the basis for the statistical tests (in Figs. 3 and 4 the original 1 Hz sampling rate is displayed). These tests were applied separately for each volunteer, since it is known that pulse oximeters show subject-specific performance because of interindividual variability.

For every pair of pulse oximeters, the $S_pO_2$ errors of those 2 instruments were compared directly at the selected time points. With 5 pulse oximeters there are 10 possible pairs. As only 4 pulse oximeters were attached to the hand of each volunteer (we did not use the thumb), for each pair of instruments the data of 6 volunteers were available.

Let us now consider one particular period and one pair of pulse oximeters. In order to account for possibly subject-dependent variances, we applied 2 tests (Wilcoxon signed-rank test and sign test), to the data of each volunteer separately. This resulted in $Z$ values $Z_1$ through $Z_6$ (ie, transformed p values, one for each volunteer). A combined $Z$ value for the 6 volunteers was computed using the formula:

$$Z_{\text{sum}} = \sum Z_i / \sqrt{n}$$  \hspace{1cm} (2)

Under the null hypothesis of no difference between the 2 instruments, every single $Z_i$ as well as $Z$ has a standard...
normal distribution. This leads to a p value of $p = \text{NCDF}(Z)$, where NCDF denotes the standard normal distribution function.

Note that we have 40 different p values corresponding to all combinations of 4 periods and 10 pairs of pulse oximeters. In order to account for multiple testing, we made a Bonferroni adjustment and replaced each individual p value with the value of $p = \frac{0.05}{40}$. Then any p value of $\leq 0.05$ is significant at that level and can be interpreted.

The times when some pulse oximeters displayed a zero reading or no reading are listed separately.

**Pulse rate error** was defined as the pulse rate measured by pulse oximetry minus the heart rate measured via ECG and recorded using an Agilent CMS monitor. For the pulse rate error, 2 error categories were defined: pulse error $\leq 10$ beats/min and pulse error $\leq 25$ beats/min. We defined the 10 beats/min range as the neutral range. This represents the range covering more than 90% of all readings when no motion is present. The relatively large deviations between oximetry-measured pulse rate and ECG-measured heart rate are caused in part by the differing averaging times of the pulse oximeters and the ECG instrument. The pulse oximeters use the averaging times described above (approximately 10 s), whereas the ECG monitor computed a beat-to-beat heart rate or computed heart rate using 2 neighboring RR intervals. For that reason, the natural heart rate variability was the cause of at least some of the discrepancy between the 2 measurements. The statistical analysis was performed in the same way as described for the $S_{pO_2}$ error.

**Selection of Periods**

The motion period was defined as the period in which the hand was moving, plus the 20 seconds after the end of the motion, to take into account the pulse oximeters’ response times.

Despite setting the oximeters’ response times to similar values, there were small differences in dynamic behavior between the various pulse oximeters. We therefore only used periods in which the changes in the $S_{pO_2}$ values of the reference pulse oximeters were $< 0.3\%$ of $S_{pO_2}$ per second. This was done by forming the first derivative of the $S_{pO_2}$ signal and neglecting all periods wherein the $|AS_{pO_2}/s|$ was $< 0.3$, and, in addition, the periods 5 seconds before and after these intervals. That included all periods with desaturation and resaturation, with maximum $S_{pO_2}$ alterations of $20\%$ of $S_{pO_2}$ per minute.

**Results**

Figure 3 shows an $S_{pO_2}$ recording during a desaturation in Period 1 (no motion and normal perfusion). The differences between the oximeters are within the defined range of $\pm 3\%$.

**Change of Perfusion**

Figure 4B shows the change of perfusion caused by inflating the balloon positioned at the brachial artery of the test arm, while figure 4A demonstrates the corresponding desaturation values of the pulse oximeter during this period. The SpO2 data from the various pulse oximeters of the test hand and reference hand are in a good agreement.

**$S_{pO_2}$ Error**

Figure 5 shows the percentage of the total time the $S_{pO_2}$ error was in categories 4L (low) through 4H (high) during Periods 2–4. The increasing challenge for the pulse oximeters is visible in the increase of column heights for the higher error categories. In the period with motion and low
perfusion, all instruments tested showed high \( S_pO_2 \) errors. The selected Periods 2 (low perfusion) through 4 (motion and low perfusion) fulfill the requirement of representing different levels of interference.

The false low readings are more pronounced than the false high readings because the magnitude of false high readings is limited by the maximum possible \( S_pO_2 \) value of 100%. As the true \( S_pO_2 \) values were always in the range of \( 70–100\% \), the maximum positive deviation was 30% at the lowest saturations tested, and less for larger saturations. For false negative values there is a wider range of possibilities.

Table 2 shows the results of Wilcoxon matched-pairs signed-rank test on the absolute errors of the \( S_pO_2 \) values and the results of sign test.

| Pair            | Wilcoxon Test | Sign Test | \( |\Delta S_pO_2| > 6 \) | \( |\Delta S_pO_2| > 10 \) |
|-----------------|---------------|-----------|----------------|----------------|
|                 | Period 1 | Period 2 | Period 3 | Period 4 | Period 1–3 | Period 4 | Period 1–4 |
| Datex–Agilent   | NS        | NS       | NS        | NS        | NS        | NS       | NS         |
| Datex–N-395     | NS        | NS       | NS        | NS        | NS        | NS       | NS         |
| Datex–N-3000    | Datex*    | NS       | NS        | NS        | NS        | NS       | NS         |
| Ivy–Datex       | NS        | NS       | Ivy*      | NS        | NS        | NS       | NS         |
| Ivy–Agilent     | Ivy*      | NS       | NS        | NS        | NS        | NS       | NS         |
| Ivy–N-395       | N-395*    | NS       | NS        | NS        | NS        | NS       | NS         |
| Ivy–N-3000      | Ivy*      | NS       | Ivy*      | NS        | NS        | NS       | NS         |
| N-395–Agilent   | N-395*    | NS       | NS        | NS        | NS        | NS       | NS         |
| N-395–N-3000    | N-395*    | NS       | NS        | NS        | NS        | NS       | NS         |
| N-3000–Agilent  | Agilent*  | NS       | Agilent*  | NS        | NS        | NS       | NS         |

\( \Delta S_pO_2 \) = absolute value of the error of the oxygen saturation value measured via pulse oximetry.

NS = no significant difference.

* = significant difference (\( p \leq 0.05 \)). The more error-prone pulse oximeter is named.

Pulse oximeters have proven reliable for detecting the threat of hypoxemia and, together with capnography, they represent the fundamental monitoring for adverse anesthesia-related events.\(^6\) The increasing challenges posed by clinical and ambulatory use of pulse oximeters have been met by their manufacturers; new algorithms for signal evaluation have been implemented into the devices on a regular basis. Our goal was to evaluate the performance of a new generation of pulse oximeters under conditions where interference of signal registration might be expected: move-

---

**Discussion**

Pulse oximeters have proven reliable for detecting the threat of hypoxemia and, together with capnography, they represent the fundamental monitoring for adverse anesthesia-related events.\(^6\) The increasing challenges posed by clinical and ambulatory use of pulse oximeters have been met by their manufacturers; new algorithms for signal evaluation have been implemented into the devices on a regular basis. Our goal was to evaluate the performance of a new generation of pulse oximeters under conditions where interference of signal registration might be expected: move-
ment of the sensor-carrying finger and decreased perfusion in the measurement area. The arm motion was standardized, reproducible, and designed to produce challenging conditions for the pulse oximeters, and the motion was applied during controlled desaturation procedures in healthy volunteers, because detection of hypoxemia is the important task for pulse oximeters monitoring in the clinic. In developing this method, similarity to clinical conditions had to be guaranteed. Patients in the recovery room are often restless, and perfusion in the finger can be restricted by vasoconstriction due to peripheral hypothermia, hypovolemia, or catecholamine release.\textsuperscript{1,2} Similar situations occur during weaning from mechanical ventilation in intensive care or during sedation for bronchoscopy.

We used motion created by a motion machine, as has been done by Barker et al.\textsuperscript{3,7} which has the advantage of standardized reproducibility for each volunteer. For an improved modeling of clinical reality, we also introduced tapping and scratching movements, which are more challenging for the pulse oximetry sensors. We also tested oximeter performance under low perfusion conditions. Instead of using a software simulator,\textsuperscript{8} we chose to change the arterial perfusion, and we created a control mechanism for arterial perfusion. The pulse oximeters tested indicated a change in the pulse amplitude, whereas the $S_{pO_2}$ reading was still working properly. Unlike the procedure used by Falconer and Robinson,\textsuperscript{4} we used not a rigid clamp but an inflatable balloon, to avoid tissue damage. All volunteers found the procedure to be agreeable and painless.

$S_{pO_2}$ Error

For the pulse oximeters tested in the present study, the $S_{pO_2}$ error was within $\pm 3\%$ of the reference reading $> 90\%$ of the time during low perfusion and $> 80\%$ during motion (see Fig. 5). During the combination of motion and reduced perfusion the $S_{pO_2}$ error was within $\pm 3\%$ less than 62\% of the time for all the tested pulse oximeters. In the present study no dramatic differences between the oximeters were observed. Although there were significant
Fig. 7. Comparison of oximeter-measured pulse rate to electrocardiographically measured heart rate. Relative incidence of pulse errors (in beats/min) in (A) Period 1 (no motion and normal perfusion) and (B) Period 4 (motion and low perfusion). C: Percentage of readings within the limits pulse error $\leq 10$ beats/min and pulse error $\leq 25$ beats/min in Period 4 (motion and low perfusion).
differences between oximeters within certain volunteers, those differences cancelled out when all the results were combined. The differences observed in Period 1 were statistically significant, but were not meaningful, because the $\text{SpO}_2$ errors remained in the neutral range in most cases ($\Delta\text{SpO}_2 \leq 3\%$). The pulse oximeters showed different responses to increasing motion levels. Some display zero readings when they failed to compute a saturation, whereas some others (e.g., the Nellcor N-3000) freeze the $\text{SpO}_2$ display and warn the user with a flashing display. In this study the displayed value, although it might be frozen, was included in the analysis. The percentage of zero values was largest for the Agilent device, followed by the Ivy 2000 with Masimo SET. The other instruments did not display zero values.

A pulse oximeter’s response time is the period required for the device’s algorithm to orient itself reliably to alterations in oxygen saturation. To obtain a similar response from all oximeters tested, we set the response times to almost identical values (within $\pm 2\,\text{s}$). Figure 3 shows the response to a desaturation period. To impose identical requirements of the oximeters, we excluded periods with $\text{SpO}_2$ alterations exceeding $20\%$ per minute. This lies within the required limits for rapid desaturation periods, as described by Barker and Shah.\(^5\) Because of their non-zero response times, pulse oximeters tend to display artifact-influenced values some seconds after the artifact has stopped. We therefore included the 20 seconds after the end of the artifact into the artifact period. During profound artifact conditions, when there is no reliable signal detection, pulse oximeters tend to “freeze” the last apparently reliable value. This process leads to false readings, especially in cases where hypoxemia may be threatening. The falsifying effect is strongest when a rapid change in oxygen saturation is provoked during the disrupted period. In this study we observed this freezing effect, and the $\text{SpO}_2$ errors created by those periods are included in the results.

### Pulse Error

Measurement of pulse rate and its representation in visual and acoustic form is essential for supervising patients intraoperatively during anesthesia or in a pre-clinical situation of an emergency case. It has not yet been established from which source the signal should ideally be taken.\(^6\) During anesthesia it is clinical practice to use the pulse rate display in the following way, because:

1. It underlines the credibility of the $\text{SpO}_2$ measurements. When the ECG-measured heart rate agrees with the pulse rate of the pulse oximetry it is concluded clinically that the pulse curve produced by the pulse oximeter corresponds to the patient’s true pulse.
2. It can replace supervision of cardiac function when an ECG is unavailable or disrupted.
3. From the acoustic variation of the peeping tone dependent on changes in oxygen saturation, the clinician gains additional information on the oxygenation state of the patient.

Figures 7A and 7B show the pulse rate error distributions for Periods 1 and 4, respectively. Consistent with the pulse oximeters’ evaluation algorithms, minor differences between ECG-measured heart rate and oximeter-measured pulse rate are normal, as can be seen from the Gaussian distribution curves in Figure 7A. The distributions of all the tested pulse oximeters are almost identical, and the differences are not statistically significant. Figure 7C shows

### Table 3. Results of Wilcoxon Matched-Pairs Signed-Rank Test on the Absolute Pulse Errors and the Results of Sign Test

<table>
<thead>
<tr>
<th>Pair</th>
<th>Wilcoxon Test</th>
<th>Sign Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Period 3</td>
<td>Period 4</td>
</tr>
<tr>
<td>Datex–Agilent</td>
<td>Datex*</td>
<td>Datex*</td>
</tr>
<tr>
<td>Datex–N-395</td>
<td>Datex*</td>
<td>Datex*</td>
</tr>
<tr>
<td>Datex–N-3000</td>
<td>Datex*</td>
<td>Datex*</td>
</tr>
<tr>
<td>Ivy–Datex</td>
<td>Datex*</td>
<td>Datex*</td>
</tr>
<tr>
<td>Ivy–Agilent</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Ivy–N-395</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>N-395–Agilent</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>N-395–N-3000</td>
<td>NS</td>
<td>N-3000</td>
</tr>
<tr>
<td>N-3000–Agilent</td>
<td>NS</td>
<td>N-3000</td>
</tr>
</tbody>
</table>

$|\text{pulse error}|$ = absolute value of the pulse error.

NS = no significant difference.

* = significant difference ($p < 0.05$). The more error-prone pulse oximeter is named.

In periods 1 and 2 all comparisons were NS.
the cumulative percentages in the 2 error categories (pulse error \(\leq 10\%\) and pulse error \(\leq 25\%\)) for Period 4. With the Agilent, Ivy 2000 with Masimo SET, and Nellcor N-395 the pulse error was > 25 beats/min < 10\% of the time. With the N-3000 the pulse error was > 25 beats/min 37\% of the time. With the Datex-Ohmeda the pulse error was > 25 beats/min 53\% of the time. The paired comparison of the oximeters gives a significantly higher pulse error for the Datex-Ohmeda device than all other devices in all error categories (see Table 3). The Nellcor N-3000 showed significantly higher pulse errors than the Agilent, Ivy 2000 with Masimo SET, or Nellcor N-395. There are no significant differences between the Agilent, Ivy 2000 with Masimo SET, and Nellcor N-395 in any of the error categories.

The simulated interference of Period 4 clearly represents a challenge for pulse oximetry. From a clinical viewpoint, potentially dangerous situations can arise. With pediatric patients in particular, the gap is small between adequate oxygenation and hypoxemia with the threat of bradycardia, in which case reliable detection of heart rate and \(S_{\text{PO}_2}\) are of great value.\(^9\)

Previous studies reported better pulse oximetry performance with the Masimo signal extraction technology.\(^3,7,9\)

In high-noise settings such as in neonatal intensive care units\(^10\) and moving patients in the recovery room,\(^11\) a lower alarm frequency was reported. Furthermore, improved detection of hypoxemia and bradycardia was also observed.\(^9\)

Our results herein confirm that the range of performance in pulse rate accuracy with the Ivy 2000 with Masimo SET is also achieved by the Nellcor N-395 and the Agilent CMS monitor.

**Description of Algorithms**

Previous studies have compared the performance of different pulse oximeters in the clinical environment\(^11\) and in specially designed volunteer studies.\(^3,4,12\) In recent years, new algorithms in pulse oximeters have improved their performance. The basic principles of the algorithm in the Ivy 2000 with Masimo SET technology have been described by Barker and Shah.\(^3\)

The algorithm implemented in the Agilent device, according to Agilent, uses patented digital techniques in the frequency domain and an analysis of the resulting Fourier spectrum to remove noise from the signal. A fast Fourier transform generates Fourier spectral peaks, representing the pulse, the pulse harmonics, noise artifact, and any noise and artifact harmonics. Peaks from noise and artifact are eliminated by failure to pass a number of qualifying tests, leaving only the actual pulse and its harmonics. The tests base elimination on unrealistic \(S_{\text{PO}_2}\) values, pulse rate values, perfusion values, and statistical analysis of those values to eliminate nonsignal harmonics. Recently the latest Agilent technology (software version B.0) was tested against the Ivy 2000 with Masimo SET and the Nellcor N-3000 in a software simulation of critical clinical situations.\(^12\)

The Datex-Ohmeda 3900P uses a feature called TruTrak data sampling, which detects and calculates \(S_{\text{PO}_2}\) 30 times per second. By calculating all along the waveform, it is not prone to motion-induced information gaps seen with oximeters that calculate \(S_{\text{PO}_2}\) after detecting the peak and trough of a clean pulse. TruTrak data sampling determines how well the red and IR waveforms match, and, if they are too dissimilar, the data period is identified and isolated from the calculations. \(S_{\text{PO}_2}\) data are compared with previous data points and weighted according to the degree of consistency. Inconsistent values are assigned lower weights and thus have less impact on the \(S_{\text{PO}_2}\) averaging calculation.

The Nellcor N-395 oximeter introduces multiple algorithms for calculating pulse rate and \(S_{\text{PO}_2}\), along with arbitration procedures for choosing among the outputs of those algorithms. A basic pattern-matching algorithm is used to identify individual pulses to drive one of the pulse rate methods as well as the audible beep tone. A second algorithm, based on an adaptive comb filter, is used to track pulse rate in the presence of severe noise. This method functions by “locking on” to the rhythmic portion of the optical signal that is near the previously determined pulse rate. The first of 2 saturation algorithms uses a digital filter locked to the adaptive comb filter pulse rate to enhance the signal-to-noise ratio of the IR and red signals that are used to compute \(S_{\text{PO}_2}\). A second \(S_{\text{PO}_2}\) algorithm uses least squares fitting to find the best estimate of \(S_{\text{PO}_2}\) derived from continuously sampled IR and red data. The Nellcor N-395 is a member of the fourth generation of Nellcor pulse oximeters, whereas the N-3000 belongs to the third generation.

**Conclusions**

The pulse oximeters tested in this study reliably measured \(S_{\text{PO}_2}\), when movement was introduced as an artifact and when perfusion was reduced in the sensor area. Only during the clinically extreme condition of combined low perfusion and motion artifact was the reliable detection of \(S_{\text{PO}_2} < 62\%\) for the various devices. During combined reduced perfusion and motion artifact, 3 of the tested pulse oximeters also proved reliable for showing pulse rate, whereas the Nellcor N-3000 and the Datex Ohmeda 3900P showed clear differences from the ECG-measured heart rate—a fact that must be considered in the clinic.

**ACKNOWLEDGEMENTS**

We thank Agilent Technologies/Philips Medical Systems, Datex-Ohmeda, and Nellcor/Tyco Healthcare Group for supplying the instruments tested and the costs for the volunteers involved. We also thank...
Professor Lutz Dümbgen, Institute of Mathematics of the Medical University of Lübeck, for his advice on statistics.

REFERENCES