Validity of pulse oximetry during maximal exercise in normoxia, hypoxia, and hyperoxia

YOSHIKI YAMAYA, HARM J. BOGAARD, PETER D. WAGNER, KYUICHI NIIZEKI, AND SUSAN R. HOPKINS

Department of Medicine, University of California, San Diego, La Jolla, California 92093

Received 27 April 2001; accepted in final form 29 August 2001

The purpose of this study was to evaluate new pulse oximeter technologies during exercise in a variety of study populations. We tested two devices: the Ivy 2000 (Masimo, Irvine, CA) with the LNOP-Adt finger sensor and two Nellcor N-395 (Oxismart XL, Mallinckrodt, St. Louis, MO) pulse oximeters equipped with either a RS-10 forehead sensor or a D-25 digit sensor. These devices were used in normal subjects, athletes, and patients with either chronic obstructive pulmonary disease or chronic heart failure. We hypothesized that the combination of motion-tolerant pulse oximeter and RS-10 forehead sensor, because of the previously mentioned problems with motion artifact and digital perfusion during exercise, would provide the most valid noninvasive measure of SaO2. We chose to compare athletes and patients with normal subjects because monitoring SaO2 in these two groups is likely to be problematic for the athletes because of poor signal detection caused by the diversion of large amounts of oxygen saturation when signals are weak (e.g., low perfusion) or corrupted by motion artifact (1, 2, 15).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

DURING EXERCISE TESTING, it is often desirable to monitor arterial oxygen saturation of hemoglobin (SaO2) especially in the context of pulmonary or cardiovascular disease or when subjects are breathing hypoxic gas mixtures as part of research protocols. Pulse oximeters, because they are noninvasive and obviate the need for arterial catheterization, are often used in this context. In addition, many researchers investigating pulmonary gas exchange during exercise utilize pulse oximetry either to prescreen research subjects before more invasive studies or in place of direct measurement of SaO2 using cooximetry (3, 7, 8, 11).

Pulse oximeters use a light source and photodiode light detector to measure the amount of light passing through an arteriolar bed. SaO2 can be estimated noninvasively because the light-absorbing characteristics of hemoglobin differ between oxyhemoglobin and deoxyhemoglobin. Although well accepted for use in resting subjects, using pulse oximetry during exercise for accurate measurement of SaO2 has been problematic for several reasons. First, depending on the sensor site, sensors are subjected to varying degrees of motion resulting in signal corruption and thus inaccurate estimations of saturation (9). Furthermore, sensors placed on the digits are even more susceptible to this problem during cycle exercise because gripping the handlebars results in weakening or even complete loss of signals (17, 19). Recently developed pulse oximeters offer potential advantages because they utilize advanced signal-processing methodologies in an attempt to provide continuous and accurate measurements of oxygen saturation when signals are weak (e.g., low perfusion) or corrupted by motion artifact (1, 2, 15).

The purpose of this study was to evaluate new pulse oximeter technologies during exercise in a variety of study populations. We tested two devices: the Ivy 2000 (Masimo, Irvine, CA) with the LNOP-Adt finger sensor and two Nellcor N-395 (Oxismart XL, Mallinckrodt, St. Louis, MO) pulse oximeters equipped with either a RS-10 forehead sensor or a D-25 digit sensor. These devices were used in normal subjects, athletes, and patients with either chronic obstructive pulmonary disease or chronic heart failure. We hypothesized that the combination of motion-tolerant pulse oximeter and RS-10 forehead sensor, because of the previously mentioned problems with motion artifact and digital perfusion during exercise, would provide the most valid noninvasive measure of SaO2. We chose to compare athletes and patients with normal subjects because monitoring SaO2 in these two groups is likely to be problematic for the athletes because of poor signal detection caused by the diversion of large amounts of oxygen saturation when signals are weak (e.g., low perfusion) or corrupted by motion artifact (1, 2, 15).

The purpose of this study was to evaluate new pulse oximeter technologies during exercise in a variety of study populations. We tested two devices: the Ivy 2000 (Masimo, Irvine, CA) with the LNOP-Adt finger sensor and two Nellcor N-395 (Oxismart XL, Mallinckrodt, St. Louis, MO) pulse oximeters equipped with either a RS-10 forehead sensor or a D-25 digit sensor. These devices were used in normal subjects, athletes, and patients with either chronic obstructive pulmonary disease or chronic heart failure. We hypothesized that the combination of motion-tolerant pulse oximeter and RS-10 forehead sensor, because of the previously mentioned problems with motion artifact and digital perfusion during exercise, would provide the most valid noninvasive measure of SaO2. We chose to compare athletes and patients with normal subjects because monitoring SaO2 in these two groups is likely to be problematic for the athletes because of poor signal detection caused by the diversion of large amounts of oxygen saturation when signals are weak (e.g., low perfusion) or corrupted by motion artifact (1, 2, 15).

The purpose of this study was to evaluate new pulse oximeter technologies during exercise in a variety of study populations. We tested two devices: the Ivy 2000 (Masimo, Irvine, CA) with the LNOP-Adt finger sensor and two Nellcor N-395 (Oxismart XL, Mallinckrodt, St. Louis, MO) pulse oximeters equipped with either a RS-10 forehead sensor or a D-25 digit sensor. These devices were used in normal subjects, athletes, and patients with either chronic obstructive pulmonary disease or chronic heart failure. We hypothesized that the combination of motion-tolerant pulse oximeter and RS-10 forehead sensor, because of the previously mentioned problems with motion artifact and digital perfusion during exercise, would provide the most valid noninvasive measure of SaO2. We chose to compare athletes and patients with normal subjects because monitoring SaO2 in these two groups is likely to be problematic for the athletes because of poor signal detection caused by the diversion of large amounts of oxygen saturation when signals are weak (e.g., low perfusion) or corrupted by motion artifact (1, 2, 15).
Table 1. Subject descriptive characteristics

<table>
<thead>
<tr>
<th></th>
<th>Group 1 (n = 9)</th>
<th>Group 2 (n = 6)</th>
<th>Group 3 (n = 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yr</td>
<td>33.5 ± 6.8</td>
<td>28.8 ± 5.8</td>
<td>56.7 ± 16.6*</td>
</tr>
<tr>
<td>Height, cm</td>
<td>171.0 ± 8.7†‡</td>
<td>173.7 ± 3.9†</td>
<td>179.7 ± 9.6†</td>
</tr>
<tr>
<td>Weight, kg</td>
<td>68.9 ± 8.2</td>
<td>70.3 ± 3.9</td>
<td>95.8 ± 21.2*</td>
</tr>
<tr>
<td>Normoxic VO2max, l/min</td>
<td>3.13 ± 0.56</td>
<td>4.12 ± 0.34†</td>
<td>1.83 ± 0.75*</td>
</tr>
<tr>
<td>Normoxic VO2max, ml·kg⁻¹·min⁻¹</td>
<td>45.6 ± 6.2</td>
<td>58.7 ± 4.0†</td>
<td>18.8 ± 5.1*</td>
</tr>
<tr>
<td>Hypoxic (FIO2 = 0.12) VO2max, l/min</td>
<td>2.59 ± 0.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypoxic VO2max, ml·kg⁻¹·min⁻¹</td>
<td>37.2 ± 6.4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Values are means ± SE. VO2max, maximal oxygen consumption; FIO2, inspired oxygen fraction. *Significantly different from group 1 and 2 (P < 0.05). †Significantly different from group 1 (P < 0.05). ‡Significantly different from group 3 (P < 0.05).

The order of the two exercise tests was balanced in group 1, and subjects rested for ~1 h between tests.

Results of these preliminary tests were used to select workloads that represented 30, 60, 90, and 100% of VO2 max in blood to working muscles during high-intensity exercise and for the patients because of a potential circulatory compromise caused by chronic disease.

METHODS

This study was approved by the Human Subjects Committee of the University of California, San Diego. Nineteen subjects were recruited by advertisement and, after giving written informed consent, agreed to further study. Subjects belonged to one of the following three groups. Group 1 was healthy active nonsmoking adults (n = 9), group 2 was healthy nonsmoking competitive cyclists (n = 6), and group 3 (n = 4) was patients with either chronic heart failure or chronic obstructive pulmonary disease.

Preliminary screening. A screening history and physical examination was performed, and female subjects were screened for pregnancy. Maximal oxygen consumption (VO2 max) was determined on an electronically braked cycle ergometer (Excaliber, Quinton Instruments, Gronigen, Netherlands). After a 5-min warm up at 25–100 W, groups 1 and 2 rode a progressive exercise test (25–30 W/min) until they were unable to continue. Group 3 protocol was similar except that the work rate increment was 10–20 W/min. Heart rate was monitored by cardiac monitor (Lifepak 6, Physio-control, Redmond, WA). Subjects breathed through a nonrebreathing valve (2700, Hans-Rudolph, Kansas City, MO). Expired gas was sampled continuously from a heated mixing chamber, and oxygen and carbon dioxide concentrations were measured (mass spectrometer-1100, Perkin-Elmer, Pomona, CA). Expired gas flow was measured by using a pneumotach (no. 3 Fleisch) and differential pressure transducer (Validyne, DP45-14, Northridge, CA), and electrical signals from the mass spectrometer and pneumotach were logged at 100 Hz by using a 12-bit analog-to-digital converter. Ventilation (Ve) oxygen consumption (VO2) and carbon dioxide production (VCO2) were calculated by using a commercially available software package (Consentius Technologies, Salt Lake, UT). VO2 max was calculated as the average of the four highest consecutive 15-s measurements of VO2.

All subjects fulfilled at least two of the following four criteria for VO2 max: 1) heart rate ≥ age predicted maximum; 2) respiratory exchange ratio > 1.10; 3) no further increase or a decrease in VO2 with increasing workload; and 4) no further increase in heart rate despite an increase in workload. Group 1 also underwent a similar protocol, breathing 12% oxygen.

$\text{VO2max, l/min}$
groups 1 and 2, and 25, 50, 75, 90, and 100% of \( \text{Vo}_{2\text{max}} \) in group 3.

Subject preparation. Under continuous electrocardiogram (ECG) monitoring (LifePak 6), a 20-gauge arterial cannula was placed in the radial artery of the nondominant hand by using a sterile technique. We followed the instruction as supplied by the manufacturer, attaching the Icy 2000 with Masimo’s LNOP-Adt sensor and the Nellcor N-395 with the D-25 digit sensor (N-395/D-25) to the digits of opposite hands. A light-opaque shield was lightly taped over the digit to exclude interfering ambient light. Digit placement (digit 2, 3, and 4) and hand selection (right or left) were randomized between subjects. The Nellcor N-395 with a RS-10 forehead sensor (N-395/RS-10) was positioned over the left eyebrow on the forehead and secured with a terrycloth sweatband.

Data collection protocol. Each subject was seated on the cycle ergometer and breathed through the mouthpiece for 10 min before the start of resting measurements. Data collection consisted of duplicate 3-ml samples of arterial blood, which were collected at rest and in the last 2 min of each exercise level (30, 60, 90, and, in some subjects, 100% of \( \text{Vo}_{2\text{max}} \) in groups 1 and 2 and 25, 50, 75, and, in some subjects, 90–100% of \( \text{Vo}_{2\text{max}} \) in group 3). Group 1 performed the exercise protocol in normoxia and while breathing hypoxic gas (inspired oxygen fraction = 0.12), group 2 exercised in normoxia only, and group 3 exercised in both normoxia and 100% oxygen. \( \text{SaO}_2 \) by pulse oximetry was obtained by interfacing the three devices with a portable computer and recording data over a 30-s period simultaneously from all three devices and synchronously with arterial blood sampling. Data in which poor signal detection was evident (heart rate deviated 10 or more beats/min from that measured by ECG) were identified so that analyses could be conducted with and without these data points.

Blood-gas measurements. Arterial samples were maintained on ice until analyzed for hemoglobin concentration and \( \text{SaO}_2 \) using an IL 682 cooximeter (Instrumentation Laboratories, Lexington, MA). Blood gas analyses were completed within 30 min of data collection. \( \text{SaO}_2 \) was measured as

\[
\text{SaO}_2 = 100 \times \frac{\text{F}_{O_2}\text{Hb}}{100 - (\text{F}_{CO_2}\text{Hb} + \text{F}_{met}\text{Hb})}
\]

where \( \text{F}_{O_2}\text{Hb} \) is the oxyhemoglobin fraction, \( \text{F}_{CO_2}\text{Hb} \) is the carboxyhemoglobin fraction, and \( \text{F}_{met}\text{Hb} \) is the methemoglobin fraction.

Statistical analyses. We compared measured \( \text{SaO}_2 \) by cooximetry to each of the pulse oximeters by using linear regression (Statview 5.0, SAS, Cary, NC). Prediction limits of \( \pm 95\% \) for the prediction of \( \text{SaO}_2 \) by pulse oximeter as a function of measured \( \text{SaO}_2 \) were generated for each device. The \( \pm 95\% \) prediction limits can be thought of as the \( \pm 2 \) SD limits of the least squares regression line at any point along the regression line. To examine the effect of poor signal detection, heart rate by ECG was compared with heart rate by pulse oximetry, and separate regression analyses were performed after eliminating all \( \text{SaO}_2 \) data in which the simultaneously obtained heart rate deviated by \( \pm 10 \) beats/min from that measured by ECG. In addition, data obtained while breathing 100% oxygen are on the flat part of the oxyhemoglobin equilibrium curve. These data were not included as part of the regression analyses and were analyzed separately. Bias (or mean error; calculated as (\( \text{SaO}_2 \) pulse-oximetry – \( \text{SaO}_2 \) cooximetry)/\( n - 1 \)) and precision (SD of \( \text{SaO}_2 \) pulse-oximetry – \( \text{SaO}_2 \) cooximetry; the smaller the SD, the greater the precision) were calculated for each device. As for the regression analyses, these parameters were calculated for all data points and also after eliminating data points in which there were poor pulse rate signal detections. Significance was accepted at \( P < 0.05 \), two-tailed. Data are presented as means ± SD.

RESULTS

Subject descriptive data are given in Table 1. As expected, athletes had a greater \( \text{Vo}_{2\text{max}} \) than healthy normal subjects (group 1; \( P < 0.005 \)) or patients (group 2; \( P < 0.0001 \)). Patients were significantly older and heavier than either the athletes (\( P < 0.001 \)) or the healthy normal subjects (\( P < 0.001 \)).

\( \text{SaO}_2 \) measured by pulse oximetry during exercise in normoxia and hypoxia is compared with results obtained by direct arterial blood measurements in Fig. 1 and Table 2. Figure 1 shows the correlation between \( \text{SaO}_2 \) measured by cooximetry and by pulse oximetry for each device for all subjects. All three devices showed significant correlations between cooximetry and pulse oximetry values. However, there were considerable differences between devices. The N-395/RS-10 forehead \( \text{SaO}_2 \) (Fig. 2A) was very highly correlated with \( \text{SaO}_2 \) measured by using cooximetry (\( R^2 = 0.90, P < 0.0001 \)), and values obtained were centered around the line of identity (slope = 1.009, intercept = -0.52). Poor pulse-rate or arterial signal detection was evident in 13 (8.1%) measurements of \( \text{SaO}_2 \) and, when these data were eliminated, \( R^2 \) increased to 0.94, although slope and intercept of the relationship did not change appreciably (slope = 0.998, intercept = 0.65).

Table 2. Bias and precision of \( \text{SaO}_2 \) measurement by subject group and pulse oximeter

<table>
<thead>
<tr>
<th>Group 1 (Normal Subjects)</th>
<th>Group 2 (Athletes)</th>
<th>Group 3 (Patients)</th>
<th>Average Across Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-395/RS-10 forehead</td>
<td>0.4 ± 3.0^a,b</td>
<td>-0.3 ± 1.5^b,c</td>
<td>1.1 ± 1.1^a,b</td>
</tr>
<tr>
<td>(n = 90)</td>
<td></td>
<td>(n = 41)</td>
<td>(n = 30)</td>
</tr>
<tr>
<td>N-395/D-25 finger</td>
<td>-1.4 ± 7.9</td>
<td>-3.1 ± 7.1</td>
<td>-2.1 ± 5.2</td>
</tr>
<tr>
<td>(n = 90)</td>
<td></td>
<td>(n = 38)</td>
<td>(n = 29)</td>
</tr>
<tr>
<td>Ivy 2000 finger</td>
<td>-2.2 ± 4.9</td>
<td>-0.4 ± 4.1^b</td>
<td>-3.8 ± 6.9^c</td>
</tr>
<tr>
<td>(n = 89)</td>
<td></td>
<td>(n = 41)</td>
<td>(n = 30)</td>
</tr>
<tr>
<td>Average across devices</td>
<td>-1.1 ± 5.7</td>
<td>-1.3 ± 5.0</td>
<td>-1.6 ± 5.4</td>
</tr>
<tr>
<td>(n = 269)</td>
<td></td>
<td>(n = 120)</td>
<td>(n = 89)</td>
</tr>
</tbody>
</table>

Values are means of bias ± precision (SD); \( n \), number of observations. \(^a\)Significantly different from Ivy 2000 finger (\( P < 0.05 \)).

\(^b\)Significantly different from N-395/D-25 finger (\( P < 0.05 \)). \(^c\)Significantly different from groups 1 and 2 (\( P < 0.05 \)). \(^d\)Significantly different from group 1 (\( P < 0.05 \)). \(^e\)Significantly different from group 3 (\( P < 0.05 \)).
Correlation for the Ivy 2000 finger sensor was not as close (Fig. 2B; $R^2 = 0.78$, $P < 0.0001$), and there was a tendency for the device to underestimate $\text{SaO}_2$, particularly under hypoxic conditions, (slope $= 1.23$, intercept $= -24.2$). Many of the measurements suffered from poor pulse-rate signal detection (76 measurements; 48.4%), and eliminating these measurements increased $R^2$ to 0.87. However, even with this modification, slope of the relationship was unchanged (1.20) and intercept was still markedly negative ($-19.9$). Thus adequate pulse-rate signal detection does not prevent the problem of underestimating $\text{SaO}_2$ in this device.

The N-395/D-25 finger sensor utilized the same pulse oximeter tested for the N-395/RS-10 but provided data that were much less closely correlated with $\text{SaO}_2$ measured by cooximetry (Fig. 2C; $R^2 = 0.52$, $P < 0.0001$). Although the N-395/D-25 finger sensor underestimated $\text{SaO}_2$ at all levels, measurements were not greater during hypoxia than during normoxia (slope $= 1.004$, intercept $= -2.32$). Poor pulse-rate signal detection was present in 43 (26.9%) measurements, and removal of these data from the regression improved the strength of the relationship substantially ($R^2 = 0.87$) but did not significantly alter slope and intercept of the relationship (1.04 and $-3.30$, respectively).

Bias and precision of the $\text{SaO}_2$ measurement from the three devices compared with the cooximeter are given in Table 2 and Fig. 2. Averaged over all subjects, the N-395/RS-10 forehead device had significantly lower bias and greater precision than the two finger probes (precision $= 2.5$ for N-395/RS-10 vs. 5.2 and 7.3 for IVY 2000 and N-395/D-25, respectively). Eliminating data points with poor pulse-rate signal detection had a minimal effect on these values (precision $= 2.0$ for N-395/RS-10 vs. 4.3 and 8.4 for IVY 2000 and N-395/D-25, respectively). The N-395/D-25 and N-395/RS-10 performed similarly in all three subject groups, although there were minor differences between athletes and patients for the N-395/RS-10. However, the IVY 2000 was significantly worse in group 3 compared with athletes and normal subjects.

Bias and precision $\text{SaO}_2$ data from patients exercising in 100% oxygen are presented in Table 3. In this data set, $\text{SaO}_2$ approaches 100% [all partial pressure of oxygen (PaO2) values were $>500$ Torr]. This approach essentially eliminates any variation in the independent variable (i.e., $\text{SaO}_2$ measured by cooximetry). Therefore, any deviation of $\text{SaO}_2$ measured by the three devices is related to issues such as perfusion and/or signal detection. Under these conditions, the N-395/RS-10 forehead sensor had significantly less bias and greater precision than the other two devices ($P < 0.05$).

Bias and precision of the heart rate data, which can be used as an index of the adequacy of pulse-rate signal detection for the different devices, are presented in Table 4. Averaged across all subject groups, there were significant differences between devices. The N-395/RS-10 had significantly less bias and greater precision of heart rate measurements compared with the other two devices ($P < 0.001$), and the N-395/D-25 demonstrated significantly less bias compared with the IVY 2000 ($P < 0.001$). There were no significant differences between subject groups for the N-395/RS-10. Both the N-395/D-25 and IVY 2000 significantly underestimated the heart rate compared with the ECG measure of heart rate. There were no significant differences in bias between subject groups for the N-395/D-25, but the Ivy
Table 3. Bias and precision by pulse oximeter in hyperoxia

<table>
<thead>
<tr>
<th>Device</th>
<th>Bias ± Precision (n)</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-395/RS-10 forehead</td>
<td>−0.24 ± 0.57 (27)</td>
<td>27</td>
</tr>
<tr>
<td>N-395/D-25 finger</td>
<td>−4.48 ± 6.29 (26)</td>
<td>26</td>
</tr>
<tr>
<td>Ivy 2000 finger</td>
<td>−5.20 ± 11.11 (24)</td>
<td>24</td>
</tr>
</tbody>
</table>

Values are means of bias ± precision (SD) for the two devices they evaluated during exercise in normoxia and hypoxia (one pulse oximeter equipped with an ear probe and the HP 47201A ear oximeter, which analyzes light from eight wavelengths).

Exercise induces potential problems related to the accuracy of pulse oximetry. For example, motion artifact can interfere with arterial signal detection, and new generation pulse oximeters such as the Masimo Ivy 2000 and the Nellcor Oxismart incorporate advanced signal processing to deal with nonstandard signal detection (1, 2, 15). In addition, poor perfusion states, such as those that occur with cool skin temperature, may induce significant bias (18). However, as can be appreciated from Figs. 1 and 2, significant measurement errors occurred even when adequate pulse-rate signal detection was evident. Conversely, especially with the N-395/RS-10, many data points with poor pulse-rate signal detection still provided reliable data. Thus this criterion alone cannot be used to judge the quality of the data obtained by using these devices. An estimate of the extent that a particular device underestimates SaO2 may be obtained by the administration of several breaths of hyperoxic gas mixtures (inspired oxygen fraction = 0.5–1.0) in the final few seconds of an exercise test. This will elevate SaO2 to ~100%, and any systematic bias by the device will become apparent. As can be seen in Table 3, the two devices using finger probes substantially underestimated SaO2 in patients breathing 100% oxygen by an average of −5%, which was similar to their performance compared with cooximetry during normoxia.

We wish to emphasize that our conclusions are strongly influenced by the setting in which the device was used. Clearly, pulse oximeters offer several advantages in the clinical setting. These devices are noninvasive, easy to use, and do not require significant analysis time or maintenance of other equipment to obtain data. In addition, the nature of the bias of the devices we tested is such that, when significant errors are present, the tendency is toward underestimating the true SaO2. Consequently, pulse oximetry offers a conservative estimate of SaO2 during clinical exercise testing and thus is likely suitable for safety monitoring.

Table 4. Bias and precision of heart rate measurement by subject group and pulse oximeter

<table>
<thead>
<tr>
<th>Group</th>
<th>Group 2 (Athletes)</th>
<th>Group 3 (Patients)</th>
<th>Average Across Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-395/RS-10 forehead</td>
<td>−2.9 ± 13.0†‡</td>
<td>1.2 ± 1.9†‡</td>
<td>−0.8 ± 9.8†‡</td>
</tr>
<tr>
<td>(Normal Subjects)</td>
<td>(n = 90)</td>
<td>(n = 56)</td>
<td>(n = 176)</td>
</tr>
<tr>
<td>N-395/D-25 finger</td>
<td>−13.9 ± 33.6‡</td>
<td>−23.4 ± 37.0‡</td>
<td>−16.9 ± 35.0‡</td>
</tr>
<tr>
<td>(n = 90)</td>
<td>(n = 56)</td>
<td>(n = 176)</td>
<td></td>
</tr>
<tr>
<td>Ivy 2000 finger</td>
<td>−20.5 ± 33.6‡</td>
<td>−49.9 ± 41†‡§</td>
<td>−29.5 ± 35.4†‡</td>
</tr>
<tr>
<td>(n = 90)</td>
<td>(n = 30)</td>
<td>(n = 56)</td>
<td>(n = 176)</td>
</tr>
<tr>
<td>Average across devices</td>
<td>−12.5 ± 29.3‡</td>
<td>−23.8 ± 38.4‡</td>
<td>−16.6 ± 30.3‡</td>
</tr>
<tr>
<td>(n = 270)</td>
<td>(n = 90)</td>
<td>(n = 168)</td>
<td></td>
</tr>
</tbody>
</table>

Values are means of bias ± precision (SD). n, Number of observations. *Significantly different from Ivy 2000 finger (P < 0.05). †Significantly different from N-395/D-25 finger (P < 0.05). ‡Significantly different from group 1 (P < 0.01). §Significantly different from group 3 (P < 0.01).
PULSE OXIMETRY DURING EXERCISE

In addition to concerns related to accuracy during exercise, which we have briefly outlined, there are other issues to consider, such as sensor location. Finger sensors may be easier to place and to maintain in position in some patients. The forehead sensor offers a potential advantage in that it avoids the digits and the severe effects caused by gripping and motion. Also, the forehead site may be better than the earlobe, as signals are usually greater and the sensor can be easily secured and held in place with the addition of a headband. However, the signal obtained from forehead sensors is susceptible to contamination from venous blood and consequently will be biased toward low readings if central venous pressure is raised (12). This was not observed in the present study, even in patients with heart failure. However, it may be more of a problem with certain types of exercise, such as rowing, in which a Valsalva maneuver is performed at the start of the stroke. However, this may in part be prevented by the use of a compressive headband, such as the one used in the present study (4). The forehead sensor location may also interfere with secure placement of some types of headsets that support respiratory mouth pieces and vice versa. Finally, in one of our subjects in one instance, poor arterial signal acquisition occurred when the subject altered his facial expression during maximal exercise. This was accompanied by a loss of the pulse-rate signal and a markedly erroneous data point. As can be seen from Fig. 1, a similarly erroneous data point also occurred with no obvious loss of pulse-rate signal, and it is not possible, post facto, to determine the source of this error. Potential users of these devices should also be aware that significant differences in the time to detect the onset of hypoxia have been reported with different sensor locations (6). Because we tested healthy subjects and patients who were ambulatory and able to complete an exercise test, we cannot extend our findings to clinical situations other than exercise testing. It is possible that some devices using an ear sensor may offer similar performance to the forehead sensor, as these might be less susceptible to motion artifact and poor perfusion than finger sensors. However, because we did not test any ear sensors, we cannot comment on their performance.

When precise measurement of oxygen transport is important, such as in answering research questions, measurement of $\mathrm{SaO}_2$ alone is probably not adequate, particularly during normoxic exercise, in which relatively small changes in $\mathrm{SaO}_2$ are associated with large differences in $\mathrm{PaO}_2$. For example, although the mean bias for the N-395/RS-10 for all patient groups was 0.3% $\mathrm{SaO}_2$, the precision was ±2.5% $\mathrm{SaO}_2$. With the use of a standard oxyhemoglobin equilibrium curve at a pH of 7.4, a $\mathrm{PaO}_2$ of 40 Torr and a temperature of 37°C, ±2.5% $\mathrm{SaO}_2$ encompasses a variation of almost 40 Torr in $\mathrm{PaO}_2$, assuming a normal resting $\mathrm{PaO}_2$ of 90 Torr (13). Additionally, during exercise in which $\mathrm{SaO}_2$ is also affected by temperature and pH, typical changes observed in maximal exercise, such as a reduction in pH from 7.4 to 7.2 and an increase in the temperature from 37 to 39.5°C, result in a ~4% decrement in $\mathrm{SaO}_2$ in the absence of any change in $\mathrm{PaO}_2$. Recently, Dempsey and Wagner (5) have offered a standardized definition of mild exercise-induced arterial hypoxemia as a fall in $\mathrm{SaO}_2$ to below 95% from a normal resting value of 98%. This value is very close to the precision of the most accurate device tested (N-395/RS-10) and for the two finger-probe devices, both of which had an average bias of −2.0%. Many data points could be expected to fit this definition on the basis of nonrandom measurement error alone. Ideally, if circumstances dictated the use of pulse oximetry, it would be desirable to validate the pulse oximetry device against direct measures on arterial blood in a representative sample of the study population.

In conclusion, the N-395/RS-10 forehead sensor offered greater validity of $\mathrm{SaO}_2$ measurements under all conditions and in all subject groups than the other two digital oximeters tested. Bias was negligible, however, and precision was ±2.5%. Both finger sensors showed significant negative bias and underestimated $\mathrm{SaO}_2$ during exercise. They also showed low precision (>5%). Eliminating data points with poor pulse-rate signal acquisition, as evidenced by an error in heart rate measurement, did not substantially improve the performance of these devices.

We would like to thank our subjects for their enthusiastic participation and Harrieth Wagner, Nick Busan, and Jeff Struthers for technical assistance.

This study was supported by National Center for Research Resources Grant MO1 RR-00827, National Heart, Lung, and Blood Institute Grant HL-17731, and Mallinckrodt.

REFERENCES

emia in elite endurance athletes at sea level. Eur J Appl Physiol
12. Sami HM, Kleinman BS, and Lonchyna VA. Central ve-
nous pulsations associated with a falsely low oxygen saturation measured by pulse oximetry. J Clin Monit 7: 309–312,
15. Sprague D, Richardson MS, Baish JW, and Kemp JS. A new system to record reliable pulse oximetry data from the
16. Stewart KG and Rowbottom SJ. Inaccuracy of pulse oximetry in patients with severe tricuspid regurgitation. Anaesthesia 46:
19. Webb RK, Ralston AC, and Runciman WB. Potential errors in pulse oximetry. II. Effects of changes in saturation and signal