Skin Color and Ear Oximetry*

Andrew I. Ries, M.D., F.C.C.P.; Lela M. Prewitt; and Jeffery J. Johnson

Measurements of arterial blood oxygen saturation from two ear oximeters were compared with 655 simultaneously drawn arterial blood samples in 187 patients grouped by skin color quantified by the Munsell color system. Technical problems including warning lights and messages with the two ear oximeters were recorded. There were significantly more technical problems in patients with the darkest skin color associated with inability to obtain a reading or warning messages indicating poor tissue penetration of the signal (18 and 15 percent vs 1 percent). When readings could be obtained, the ear oximetry readings were found to be slightly less accurate in the darker patient groups. These findings suggest that dark skin color may affect the performance and accuracy of ear oximeters, including the newer type of pulse oximeters. (Chest 1989; 96:287-90)

Cutaneous oximetry is a valuable tool for the noninvasive estimation of arterial oxygen saturation. Several studies have established the overall accuracy (95 percent confidence limits) to be ±3 to 5 percent saturation under various clinical conditions.1-5 Several factors including carboxyhemoglobin,3,4 bilirubin,6 and exercise3,7,8 may affect accuracy. A few studies have examined the effect of skin color on cutaneous oximetry and reported variable results.1,3,9,10 However, the influence of skin color has not been examined systematically or with a replicable method of quantification. Therefore, we conducted this study to evaluate the effect of skin color determined by a standard, quantifiable method on the accuracy of two ear oximeters in pulmonary patients.

Materials and Methods

Patients

We studied prospectively 187 pulmonary patients referred to our laboratory for clinical exercise testing. A standard protocol was followed for either steady-state or incremental, maximal exercise testing for each patient. Steady-state tests involved exercise for at least four minutes at a level approximating the patient’s normal daily activity level. Maximal tests employed one minute work increments to a symptom-limited maximum. Patients exercised either walking on a treadmill or seated on a cycle ergometer.

Skin Color

The Munsell color system was used to assess skin color.11,18 This is a standard method for color notation which defines any shade of color on three scales: hue, chroma, and value. Hue represents color, or location on the wheel of primary colors. Chroma represents the saturation of color, or “the departure from a gray of the same lightness.” Value represents lightness from 0 for ideal black to 10 for ideal white. Value (lightness) has the same relative scale across all hues and chroma allowing for a consistent evaluation of skin lightness/darkness regardless of skin color.

The Munsell color charts consist of loose leaf neutral gray pages with matte color tiles. Each page contains multiple tiles with one hue designation and a range of chroma and value characteristics. Under uniform lighting conditions, an observer can compare the subject’s skin color to a single tile at a time, masking the adjacent tiles with a neutral gray mask. Each tile has a small hole at its base to allow close comparison with the skin. When the best matching tile is located, its hue, chroma, and value notations are recorded.

During skin color assessment, patients were seated in the same place in the laboratory to allow for standard lighting. Since we were primarily interested in skin lightness and darkness (value) rather than color (hue) per se, a single hue chart (SYR) was used. The SYR hue chart corresponded most closely to the range of skin colors observed. A technician of normal color vision selected the tile from the SYR hue chart which best matched the skin at the placement site for the ear probes. Value and chroma notations for this tile were then recorded.

Rest and Exercise Testing

Following skin color assessment, an indwelling radial artery catheter was inserted and the ear oximeters were secured in place. Both pinnae were vigorously massaged for 20 seconds with an alcohol pad. One oximeter (Ohmeda/Biox III) was placed on the patient’s right ear lobe and another oximeter (Hewlett-Packard 47901A) on the patient’s left ear anti-helix. Each probe was secured with the manufacturer’s ear probe retainer. A gauze bandage was wrapped around the patient’s head to further stabilize the oximeter probes. The oximeters were used simultaneously on each patient whenever possible. Both ear oximeters were calibrated prior to each study and were used according to the manufacturers’ instructions. The Ohmeda/Biox III was used in the “fast” response mode.

If a reading was not immediately obtainable from either oximeter, that probe was removed and the ear was rubbed again for an additional 20 seconds with an alcohol pad. The probe was then replaced. If, after repeated adjustment, the oximeter still failed to give a reading, the specific technical problem noted by the instrument was recorded.

During each exercise test, arterial blood samples were drawn (within 5 seconds) into heparinized 3 ml syringes, immediately placed in ice, and analyzed at the completion of the test. Readings from both ear oximeters were recorded simultaneously with the withdrawal of each arterial blood sample. If a warning light or other technical problem with either oximeter occurred at the time of the arterial blood sample, it was noted regardless of whether a reading could be obtained.

*From the Division of Pulmonary and Critical Care Medicine, Department of Medicine, University of California, San Diego. This research was supported in part by NIH Grant No. R808827 from the Division of Research Resources for the Clinical Research Center.

Manuscript received May 19; revision accepted December 9.
Reprint requests: Dr. Ries, UCSD Medical Center HH72E, 225 Dickinson Street, San Diego 92103
Arterial blood samples were analyzed in duplicate for pH, Po₂, and PcO₂ using blood gas analyzers. Direct measurements of percentage of oxyhemoglobin (SaO₂), carboxyhemoglobin (COHb), and methemoglobin (MetHb) were made on a four-wavelength spectrophotometer calibrated once daily using a reference solution with viscosity adjusted to approximate that of whole blood. Accuracy and precision were validated once daily with commercially prepared bovine hemoglobin solutions. Each sample was checked on a second CO-oximeter and reanalyzed if the measured hemoglobin levels were more than 0.5 g/dl apart. The average imprecision (observed minus mean) of repeated SaO₂ measurements in our laboratory using blood tonometry to an SaO₂ of approximately 95 percent is 0.5 ± 0.3 percent (mean ± SD).

Data Analysis

For analysis, patients were placed into one of four groups based on the Munsell color system assessment of skin value as follows: group 1, value of 8 (very light); group 2, value of 7 (light); group 3, value of 6 (average); and group 4, value less than or equal to 5 (moderately dark to very dark). These value scores corresponded to rows of lightness (value) on the SYR hue chart.

The frequency and type of technical problems recorded for each oximeter were tabulated by patient group and analyzed by chi square (with Yates correction for small cell sizes). Measurements of SaO₂ were analyzed separately for each ear oximeter by two-way analysis of variance for group assignment (1 to 4) vs measurement technique (CO-oximeter or ear oximeter). The frequency of arterial blood samples with an elevated COHb level (>4 percent) was not significantly different between groups (chi square). Therefore, data were analyzed without correction of the ear oximeter SaO₂ measurement for COHb. Overall, elevated COHb measurements were observed in 15 percent of samples (4 percent of patients) for the Ohmeda/Biox oximeter and 21 percent of samples (6 percent of patients) for the Hewlett-Packard oximeter.

Results

Of the 187 patients studied, the Ohmeda/Biox III oximeter was used for 476 rest and exercise samples in 136 patients and the Hewlett-Packard oximeter was used for 524 rest and exercise samples in 154 patients. The instruments were used simultaneously for 345 rest and exercise samples in 103 patients.

Technical problems with ear oximetry readings were noted in 10 out of 136 patients (7 percent) tested with the Ohmeda/Biox oximeter and five out of 154 patients (3 percent) with the Hewlett-Packard oximeter.

For the Ohmeda/Biox oximeter, technical problems included an activated warning message of “probe off,” “probe low,” or “low perfusion” observed at some point during testing in ten patients (seven percent). The frequency of all technical problems was not significantly different between groups; however, the frequency of the “probe low” warning was significantly higher in group 4 (p<0.05 by chi square). The “probe low” signal was observed in four of 22 (18 percent) patients in group 4, but in only one of 114 (one percent) patients in groups 1 to 3. According to the instrument documentation, this message corresponds to an insufficient amount of light penetrating the tissue sample.

For the Hewlett-Packard oximeter, technical problems were limited to activation of the “off ear” lamp at some point during testing in five patients. These included three of 20 (15 percent) patients in group 4 vs two of 134 (1 percent) patients in groups 1 to 3 (p<0.01 by chi square). In only one of these patients did the observed warning message coincide with inability to obtain a saturation reading. According to the instrument documentation, this warning generally indicates too little hemoglobin in the field of view.

Results of the two-way analysis of variance of SaO₂ measurements by patient group are presented in Table 1. For both oximeters, there were significant differences between patient groups and between measurement techniques (ear oximeter and CO-oximeter). The Ohmeda/Biox oximeter tended to overestimate directly measured SaO₂, whereas the Hewlett-Packard oximeter tended to underestimate it. The interactive effect of patient group X measurement technique was found to reach borderline significance for the Ohmeda/Biox (p = 0.09) and to be statistically significant for the Hewlett-Packard (p<0.05). These interactive effects were due to greater differences between ear oximeter and co-oximeter SaO₂ measurements in groups 3 and 4 (the darker two patient groups).

The difference between ear oximeter and CO-oximeter measurements of SaO₂ was calculated for each sample for determination of descriptive statistics by patient group of the error in ear oximeter measurements. These data are presented in Figure 1 (Ohmeda/Biox) and Figure 2 (Hewlett-Packard) as “box and whisker” plots which indicate the means, standard deviations, quartiles, and outliers for each group. These results indicate that the fact that the distribution of differences in saturation in the darkest group (group

| Table 1 — Ear Oximeter and CO-oximeter SaO₂ Measurements for Patients Grouped by Skin Color* |
|---------------------------------------------------|---|---|---|---|
| Patient Group† | 1 | 2 | 3 | 4 |
| Ohmeda/Biox III Oximeter‡ | | | | |
| No. samples | 57 | 228 | 100 | 67 |
| SaO₂, % | 93.2±4.4 | 93.0±4.3 | 93.9±4.1 | 94.9±3.8 |
| CO-oximeter | | | | |
| SaO₂, % | 92.6±3.6 | 92.6±3.7 | 92.5±3.3 | 93.7±3.5 |
| Hewlett-Packard 47201A Oximeter‡ | | | | |
| No. samples | 97 | 265 | 95 | 64 |
| SaO₂, % | 91.0±6.1 | 92.5±3.8 | 92.0±3.6 | 93.0±2.6 |
| CO-oximeter | | | | |
| SaO₂, % | 91.3±5.0 | 92.5±3.5 | 92.6±3.1 | 93.9±3.0 |

*Results expressed as mean ± 1.0 standard deviation.
†Patients grouped by skin color (see text): Group 1, lightest; group 4, darkest.
‡F values for ANOVA:
<table>
<thead>
<tr>
<th>Effect</th>
<th>Patient Group</th>
<th>SaO₂ Technique</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ohmeda/Biox:</td>
<td>&lt;0.05</td>
<td>&lt;0.001</td>
<td>=0.09</td>
</tr>
<tr>
<td>Hewlett-Packard:</td>
<td>&lt;0.01</td>
<td>=0.001</td>
<td>&lt;0.05</td>
</tr>
</tbody>
</table>
this warning message was activated.

Similarly, for the Hewlett-Packard instrument, the "off ear" technical problem was observed in 15 percent of group 4 patients vs only 1 percent of patients in the other three groups. This warning message is supposed to correspond to insufficient arterIALIZATION of the ear. A numerical reading was obtained in two of the three group 4 patients with this technical problem.

When numerical readings were obtained, both oximeters demonstrated slightly less accuracy in the patients with darker skin color. The differences in means between ear oximeter and CO-oximeter SaO₂ measurements were slightly larger by approximately 0.5 percent saturation in the two darker groups. However, the clinical significance of these differences is probably marginal given the error in the absolute accuracy of ear oximetry measurements (95 percent confidence limits of ±3 to 5 percent). The box and whiskers plots in Figures 1 and 2 demonstrate that the distribution of differences did not differ appreciably among patient groups, indicating that there were

4) is similar to those in the other three groups is not due to an asymmetric or skewed distribution reflecting large differences in a few samples.

**DISCUSSION**

The results of this study indicate potential problems with the performance of cutaneous oximeters in some individuals with dark skin color. For both the newer "pulse-type" two wave-length Ohmeda/Biox III and the older eight wave-length Hewlett-Packard 47201A ear oximeter, technical problems were observed significantly more frequently in the darkest patients.

For the Ohmeda/Biox III instrument, technical problems were observed in 18 percent of the darkest (group 4) patients vs only 5 percent of all other patients. All of the problems in group 4 patients were "probe low" messages which indicate an insufficient amount of light penetrating the tissue. A numerical reading was obtained in only one of the group 4 patients while
not a greater number of "outliers" among the darker colored patients.

The reasons for the higher problem rate and reduced accuracy in darker individuals cannot be determined from these data. The newer "pulse-type" oximeters are supposed to correct for differences in baseline light absorption due to factors such as skin color by measuring only the change in light absorption during an arterial pulsation. However, our data would indicate that these instruments may not work properly in a significant proportion of darker patients.

One limitation of this study is the relatively small number of patients with darker skin color. This reflects the typical distribution of patients referred to our laboratory. In fact, even the darkest patient group (group 4) contained patients with a range of skin color from moderate to very dark. Thus, additional data from a larger sample of darker patients would be useful in better characterizing the effect of dark skin color on cutaneous oximetry performance.

These observations have important clinical implications in areas where cutaneous oximetry may be substituted for invasive arterial blood sampling to assess arterial oxygenation. They also highlight the importance of a thorough understanding of the accuracy and limitations of this new technology under clinical conditions in real patients.

In summary, we conclude that dark skin color may influence the performance and accuracy of cutaneous oximetry. In a significant proportion of the darker patients (approximately 20 percent), measurements were not obtained due to technical problems. When measurements could be obtained, however, the accuracy was only slightly less than that for individuals with lighter skin color.

REFERENCES
12. Munsell color charts for skin-hair-eye colors. Baltimore: Munsell Color Co