

Dark Skin Decreases the Accuracy of Pulse Oximeters at Low Oxygen Saturation: The Effects of Oximeter Probe Type and Gender

John R. Feiner, MD

John W. Severinghaus, MD

Philip E. Bickler, MD, PhD

INTRODUCTION: Pulse oximetry may overestimate arterial oxyhemoglobin saturation (SaO_2) at low SaO_2 levels in individuals with darkly pigmented skin, but other factors, such as gender and oximeter probe type, remain less studied.

METHODS: We studied the relationship between skin pigment and oximeter accuracy in 36 subjects (19 males, 17 females) of a range of skin tones. Clip-on type sensors and adhesive/disposable finger probes for the Masimo Radical, Nellcor N-595, and Nonin 9700 were studied. Semisupine subjects breathed air-nitrogen- CO_2 mixtures via a mouthpiece to rapidly achieve 2- to 3-min stable plateaus of SaO_2 . Comparisons of SaO_2 measured by pulse oximetry (SpO_2) with SaO_2 (by Radiometer OSM-3) were used in a multivariate model to assess the source of errors.

RESULTS: The mean bias ($SpO_2 - SaO_2$) for the 70%–80% saturation range was 2.61% for the Masimo Radical with clip-on sensor, -1.58% for the Radical with disposable sensor, 2.59% for the Nellcor clip, 3.6% for the Nellcor disposable, -0.60% for the Nonin clip, and 2.43% for the Nonin disposable. Dark skin increased bias at low SaO_2 ; greater bias was seen with adhesive/disposable sensors than with the clip-on types. Up to 10% differences in saturation estimates were found among different instruments in dark-skinned subjects at low SaO_2 .

CONCLUSIONS: Multivariate analysis indicated that SaO_2 level, sensor type, skin color, and gender were predictive of errors in SpO_2 estimates at low SaO_2 levels. The data suggest that clinically important bias should be considered when monitoring patients with saturations below 80%, especially those with darkly pigmented skin; but further study is needed to confirm these observations in the relevant populations.

(Anesth Analg 2007;105:S18–23)

Pulse oximetry estimates arterial hemoglobin oxygen saturation (SaO_2) from the ratio of the pulsatile to the total transmitted red light divided by the same ratio for infrared light transilluminating a finger, ear, or other tissue. The SaO_2 estimated by pulse oximetry (SpO_2) should therefore be independent of skin pigmentation and many other variables, such as hemoglobin concentration, nail polish, dirt, and jaundice. Several large controlled studies, including one comparing 380 African American and Caucasian subjects

reported no significant pigment-related errors in pulse oximeter measurements at normal SaO_2 (1,2).

We recently examined the performance of three pulse oximeters (Nonin Onyx, Nellcor N-595, and Novametric 513) and found a positive bias in SpO_2 in darkly pigmented subjects as large as 7% in the SaO_2 range of 70%–79% (3). That study, because of its relatively small size (21 subjects) and study design (subjects with either very light or very darkly pigmented skin, no intermediates), did not have the power to determine the relationship between differences in skin pigment, whether gender affected errors, and if oximeter probe type (e.g., clip-on versus adhesive/disposable) contributed to the errors.

A significant issue for pulse oximeter accuracy is finger size and geometry. In 20 yr of testing pulse oximeters, it is our impression that women, especially those with smaller fingers, tend to exhibit greater bias and variability in oximeter performance, especially at low SaO_2 , although this has not been systematically examined.

On the basis of the above considerations, we designed this study to 1) determine the effect of a range of human skin pigmentation on pulse oximeter accuracy at a range of SaO_2 from 70% to 100%; 2)

From the Department of Anesthesia and Perioperative Care, University of California at San Francisco, San Francisco, California.

Accepted for publication July 27, 2007.

Dr. Philip Bickler, Section Editor for the ISIAMOV 2007 supplement, was revised from all editorial decisions related to this manuscript.

Supported by a research fund generated from the clinical testing of pulse oximeters.

No manufacturers were involved in the selection of subjects, design, or conduct of experiments, data analysis, or manuscript preparation.

Address correspondence and reprint requests to John R. Feiner, MD, Clinical Sciences 555, Box 0648, University of California Medical Center, 513 Parnassus Avenue, San Francisco, CA 94143-0648. Address e-mail to feinerj@anesthesia.ucsf.edu.

Copyright © 2007 International Anesthesia Research Society

DOI: 10.1213/01.ane.0000285988.35174.d9

Table 1. Demographics

Age	29 ± 5 (19–44)
Gender (M/F)	19/17
Ethnicity	
African American	17
Caucasian	12
Hispanic	3
Indian	1
Filipino	2
Vietnamese	1

determine whether gender affects pulse oximeter accuracy; and 3) determine whether probe type (adhesive disposable versus clip-on) affects the accuracy of SaO_2 estimates/ SpO_2 values.

METHODS

The University of California at San Francisco Committee on Human Research approved the study, and informed consent was obtained from all subjects. Thirty-six healthy, nonsmoking subjects were studied. None of the subjects had lung disease, obesity, or cardiovascular problems. They ranged in age between 19 and 44 yr. We specifically sought subjects with a range of skin pigment for this study. Each subject's skin was categorized as light (Caucasian), dark (African American), or intermediate (others, Table 1), and confirmed in subject photographs by an observer blinded to saturation data/oximeter performance.

Subjects were studied while reclining in a semisupine position (approximately 30° head up) and deliberately hyperventilating air-nitrogen- CO_2 mixtures via a mouthpiece from a partial rebreathing circuit with 10- to 20-L/min fresh gas inflow. CO_2 was added to the breathing circuit to maintain end-tidal P_{CO_2} in low normal range. A nose clip prevented breathing of room air. A 22-gauge radial artery catheter was placed to facilitate arterial blood sampling for measurements of SaO_2 . Six oximeters were mounted on each subject's fingers: a Nellcor N-595 with the OxiMax A disposable adhesive finger probe (Nellcor Inc., Pleasanton, CA), a clip-type Nellcor finger probe, a Masimo Radical (Masimo Inc., Irvine, CA) clip type and a Masimo Radical adhesive disposable type probe, and a Nonin (Nonin Inc., Plymouth, MN) 9700 clip type and a Nonin 9700 disposable adhesive probe. SpO_2 from the analog output connectors of the oximeters and the end-tidal gas values were recorded by a computer running LabVIEW 7.1 (National Instruments, Austin, TX). Oximeter probes were not mounted on the thumbs or little fingers.

A series of 11 stable target SaO_2 plateaus between 60% and 100% were achieved by an operator who adjusted the inspired air-nitrogen- CO_2 mixture breath by breath in response to the estimated SaO_2 derived from an oxyhemoglobin dissociation curve determined for each individual subject from mass spectrometer end-tidal gas analysis. Input parameters for

the computer oxygen dissociation curve included arterial pH estimated from end-tidal partial pressure of CO_2 , base excess adjusted if needed after each arterial blood sample was analyzed, and alveolar-arterial oxygen difference as needed, especially at low SaO_2 , to attempt to match the predicted with the measured SaO_2 . At each level, arterial blood was sampled after a plateau of 30–60 s had been achieved, followed by a second sample at the same plateau 30 s later. To promote good finger blood flow, each hand was wrapped in a warming pad. Five-second averages of SpO_2 values were taken at the end of the stable plateau and 30 s earlier, corresponding to the two arterial blood samples. SpO_2 values were eliminated for obvious signal dropout, or failure to reach an appropriate stable plateau. Functional SaO_2 ($HbO_2/[hemoglobin + HbO_2]$) was determined by multiwavelength oximetry (Radiometer OSM-3, Copenhagen, Denmark). Fractional SaO_2 was calculated from the functional SaO_2 and the measured levels of carboxyhemoglobin and methemoglobin. Quality control standards were run each day.

Bias was computed as SpO_2 minus SaO_2 from the reading of each oximeter and the corresponding blood sample value and is reported as mean ± sd. The relationship of bias to SaO_2 was analyzed by linear regression. Bias was also analyzed within decadal subgroups of SaO_2 (60%–70%, 70%–80%, 80%–90%, and 90%–100%). The effect of skin pigment group (dark, light, and intermediate), SaO_2 range or gender on bias was determined by Kruskal-Wallis or Wilcoxon's signed rank test, since variances were unequal. Multiple comparisons within skin pigment groups were made by the Tukey-Kramer HSD technique. Root mean square (RMS) error (square root of the sum of $[SpO_2 - SaO_2]^2$ divided by number of samples) was calculated as a measure of error magnitude. To analyze the effect of gender and skin pigment, RMS error was calculated for individual subjects within decadal SaO_2 ranges. The effect of skin pigment, SaO_2 interval and gender on RMS error was determined by Kruskal-Wallis or Wilcoxon's signed rank test. Multivariate (mixed-effects) models were used to analyze the effect of skin pigment, gender, age, hemoglobin, oximeter and probe combination, and SaO_2 on bias and RMS error. $P < 0.05$ was considered statistically significant. Statistical analysis was performed with JMP 5.1 (SAS Institute, Cary, NC).

RESULTS

Demographics

Seventeen female and 19 male subjects participated. Seventeen subjects were dark, 7 intermediate, and 12 light skinned. No differences in the distribution of male and female subjects were found in the different skin pigment categories. Subjects averaged 29 ± 5 yr (19–44 yr) (Table 1). Dark-skinned subjects were slightly older ($P < 0.01$), but there was no age difference between light- and intermediate-skinned subjects.

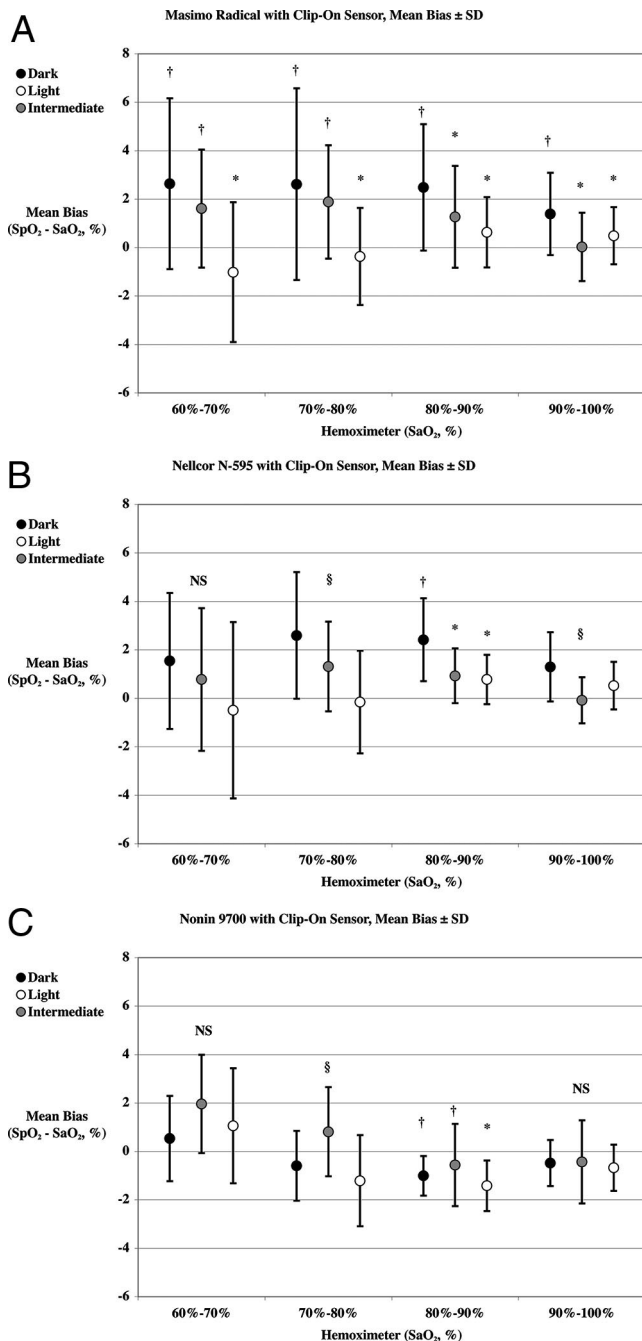


Figure 1. Bias (mean \pm SD) for the three oximeters with clip-on finger probes in decadal ranges of oxyhemoglobin saturation. Bias is calculated as SpO_2 (oximeter-measured value of oxyhemoglobin saturation) minus SaO_2 (oxyhemoglobin saturation measured by a Radiometer OSM-3 multiwavelength oximeter). Light-skinned subjects are indicated by open circles, dark-skinned subjects by closed circles, and intermediate skin pigments with gray circles. Oximeter are: A. Masimo Radical, B. Nellcor N-595, and C. Nonin 9700. Statistical differences are indicated: *different from †; or §, all values statistically different from each other.

Bias

With the exception of the Masimo Radical with the adhesive finger probe, all instruments and probe combinations showed positive bias in intermediate- and dark-skinned subjects at low SaO_2 (60%–70% and 70%–80% saturation decades, Figs. 1 and 2). The

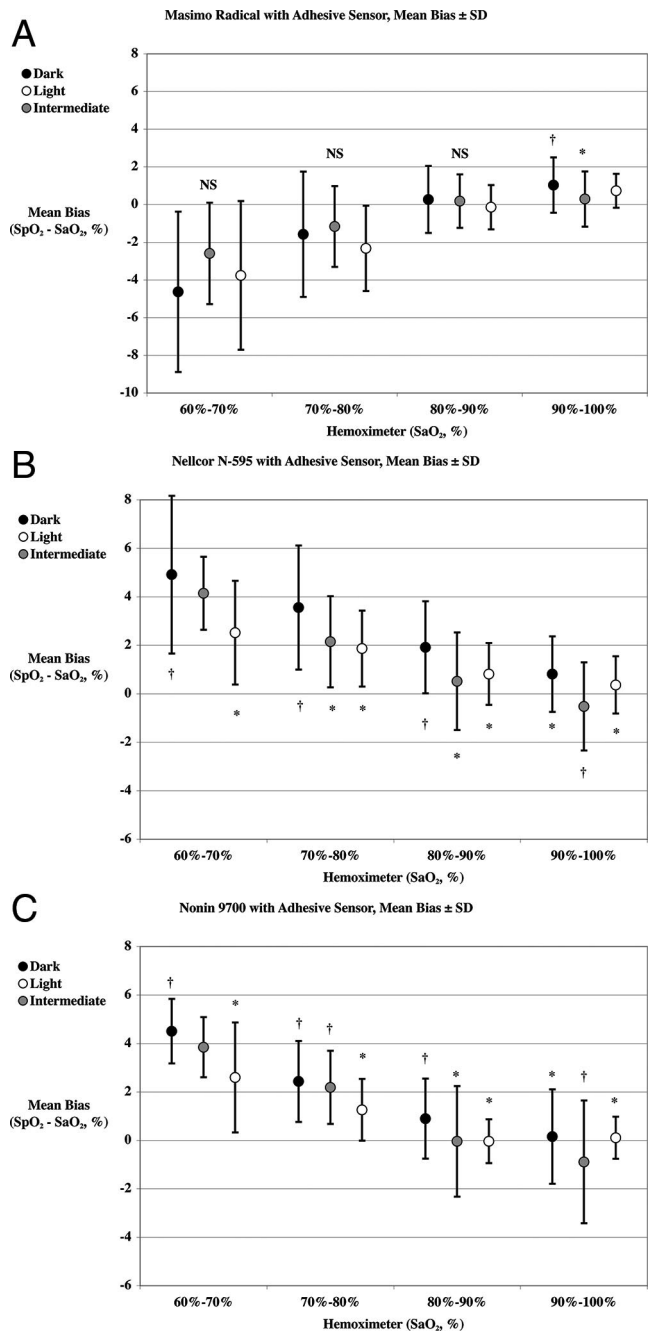


Figure 2. Bias (mean \pm SD) for three oximeters with adhesive/disposable finger probes. Bias is calculated as SpO_2 (oximeter-measured value of oxyhemoglobin saturation) minus SaO_2 (oxyhemoglobin saturation measured by a Radiometer OSM-3 multiwavelength oximeter). Light-skinned subjects are indicated by open circles, dark-skinned subjects by closed circles, and intermediate skin pigments with gray circles. Oximeters are as follows: A. Masimo Radical, B. Nellcor N-595, and C. Nonin 9700. Statistical differences are indicated: *different from †; or §, all values statistically different from each other.

greatest degree of bias was found with the adhesive probes; the Nellcor and Nonin adhesive probes showed bias of 4.5%–4.9% at 60%–70% saturation and 2.4%–3.6% bias at 70%–80% saturation with dark-skinned subjects. At the extreme, the Nellcor adhesive and the Masimo with clip-on probe read on average

Table 2. *P* Values for Multivariate Analysis of Bias ($\text{SpO}_2 - \text{Sao}_2$)

Oximeter	Gender	Skin pigment	Sao ₂	Skin × Sao ₂	Gender × Sao ₂
Masimo Radical Clip-On	0.3	0.02	<0.0001	<0.0001	<0.0001
Nellcor N-595 Clip-On	0.8	0.0009	0.0005	<0.0001	<0.0001
Nonin 9700 Clip-On	0.4	0.06	<0.0001	<0.0001	<0.0001
Masimo Radical Adhesive	0.4	0.7	<0.0001	0.0007	<0.0001
Nellcor N-595 Adhesive	0.001	0.05	<0.0001	<0.0001	0.02
Nonin 9700 Adhesive	0.4	0.2	<0.0001	<0.0001	<0.0001

Sao₂ represents decadal range: 70%–80%, 80%–90%, and 90%–100%.

Table 3. *P* Values for Multivariate Analysis of Root Mean Square Error

Oximeter	Gender	Skin pigment	Sao ₂	Skin × Sao ₂	Gender × Sao ₂
Masimo Radical Clip-On	0.09	0.01	<0.0001	0.01	0.8
Nellcor N-595 Clip-On	0.13	0.001	<0.001	0.049	0.9
Nonin 9700 Clip-On	0.03	0.07	0.02	0.7	0.003
Masimo Radical Adhesive	0.004	0.2	<0.0001	0.4	0.2
Nellcor N-595 Adhesive	0.01	0.049	<0.0001	0.3	0.3
Nonin 9700 Adhesive	0.7	0.02	<0.0001	0.3	0.051

Sao₂ represents decadal range: 70%–80%, 80%–90%, and 90%–100%.

nearly 10 points differently in dark-skinned subjects at 60%–70% saturation and seven points at 70%–80% saturation.

Pulse oximeter bias was significantly influenced by Sao₂ range for all oximeters and probe combinations (Figs. 1 and 2). The relationship between bias and Sao₂ was found to be similar statistically, whether analyzed by decadal Sao₂ range (Figs. 1 and 2) or by linear regression analysis (not shown). The pattern of bias varied according to oximeter configurations, with all except the Masimo Radical and adhesive probe showing increasing positive bias at low Sao₂. The analysis revealed essentially the same results when performed for fractional Sao₂ in place of functional Sao₂ (not shown).

Gender is a statistically significant determinant of pulse oximeter bias, with the magnitude of the gender bias differences varying with oximeter/probe type and Sao₂ range. With five of the six oximeter/probe combinations, females had greater bias in saturation estimates over the saturation range from 60% to 100%.

Multivariate Analysis: Effects of Skin Pigment, Gender, and Sao₂ on Bias

Multivariate analysis was used to examine relationships between skin pigmentation classification, gender, age, hemoglobin, Sao₂ and probe/oximeter combination and the outcome variable bias. The *P* values for these analyses are presented in Table 2 for each oximeter/probe combination separately.

Sao₂ was a highly statistically significant predictor of bias in all analyses. Although the results in Table 2 are shown for modeling Sao₂ as a continuous variable, the results were essentially the same when the analysis was performed with decadal ranges of Sao₂.

The effect of skin pigment on bias is also presented in Table 2. Skin pigment significantly predicted bias for the oximeters listed. Similar conclusions were reached when performing the analysis by the ethnicities shown in

Table 1. However, the most robust statistical model included gender, skin pigment, and Sao₂, with the interaction terms of skin pigment with Sao₂ (Skin × Sao₂) and gender with Sao₂ (Gender × Sao₂) included. The interaction for gender with skin pigment was not statistically significant. Age was not statistically significant.

Hemoglobin and gender were too significantly correlated to be separated in the multivariate analysis, 11.8 ± 1.2 (women) vs 13.9 ± 0.8 (men), $P < 0.0001$. Hemoglobin was statistically significant in place of gender, but with both variables in the model, neither was clearly better.

RMS Error

Sao₂ range was a highly significant determinant of RMS error for all the groups shown in the tables except for male subjects using the Nonin 9700 and a permanent probe. Univariate analysis of the effect of skin pigment showed that dark-skinned subjects tended to have higher RMS errors, although this was not statistically significant within every Sao₂ range. RMS error for dark-skinned subjects was more than 3.0% in several of the gender/Sao₂ combinations.

Multivariate Analysis: Effects of Skin Pigment, Gender, and Sao₂ on RMS Error

Table 3 shows the results of multivariate analysis of how RMS error is influenced by gender, skin pigment, and Sao₂. Sao₂ range was a consistent predictor of RMS error, with error increasing at low Sao₂. Skin pigment and gender, or both, were also statistically significant factors with various oximeter–probe combinations. The interaction term for skin pigment with Sao₂ and gender with Sao₂ were also significant for some oximeter/probe combinations. Two oximeters, the Nonin 9700 with clip-on probe and the Masimo radical with the adhesive probe, did not show an effect of skin pigment, although both had a significant

gender effect. Age was not a statistically significant factor on RMS error magnitude.

DISCUSSION

Confirming a previous study (3), but expanding the oximeters and probes studied, we found that pulse oximeters generally overestimate SaO_2 in hypoxic subjects (SaO_2 values below 80%.) This bias was generally the greatest in dark-skinned subjects, intermediate for intermediate skin tones, and least for lightly pigmented individuals, although SpO_2 was underestimated by one type of oximeter, the Masimo Radical with the adhesive/disposable probe.

Theoretically, the ratio of the pulsatile to the total transmitted red light divided by the same ratio for infrared light should be dependent only on arterial saturation, making pulse oximetry independent of skin color. In practice, venous and tissue pulsation by mechanical force from nearby arteries causes deviations from this ideal. The chromatic characteristics of skin color arise from the interactions of light (primarily absorption and scattering) with the epidermis and the dermis. Because deoxyhemoglobin and melanin are the primary light absorbers in skin at the wavelength used for hemoglobin absorbance, the effective light-path for red light through the finger will vary with skin pigmentation. The magnitude of this effect will vary with skin pigment, tissue perfusion, and with finger geometry, and apparently with oximeter design. Therefore, pulse oximeter performance must be partly based on empirically determined correction factors obtained by *in vivo* comparison of oximeter readings with measured SaO_2 from arterial blood samples of volunteer subjects.

In our 20 yr of testing pulse oximeter accuracy, and probably in other testing laboratories, the majority of subjects have been light skinned. Most pulse oximeters have probably been calibrated using light-skinned individuals, with the assumption that skin pigment does not matter. In addition, several studies have shown that skin pigment does not produce errors in SpO_2 estimates at high SaO_2 ranges (1,2). The current data show that skin pigment introduces a positive bias at low SaO_2 in the Nonin, Masimo, and Nellcor instruments. Our previous study (3) reported similar findings for the Nellcor N-595 clip-on sensors, and for Nonin Onyx and Novamatrix 513.

An important finding in the current study is that there is probably a continuous quantitative relationship between skin pigment and oximeter bias. This is seen in Figures 1 and 2 in which, for every oximeter and probe combination, light skin produced the smallest bias across the SaO_2 range, intermediate skin tones an intermediate degree of bias, and dark skin the greatest bias. Our earlier study (3) maximized power by studying only two extremes of skin pigment: very dark and very light. The current study enrolled subjects with a full range of skin pigment. The fact that the

SpO_2 bias changes in an ordered relationship with skin pigment strongly indicates that our findings are due to skin pigment itself, rather than from some other effect. To further pursue this question, we performed all the univariate and multivariate analyses of bias with skin color quantified with a numerical scale (Munsell number) measured by a commercial color laboratory from photographs of the subjects' hands. This analysis, done by treating skin Munsell number as continuous variable, was not significantly more robust statistically than the simpler categorization of dark, light, and intermediate skin pigment groups, and we have omitted the data here.

As clearly seen in Tables 2 and 3 and Figures 1 and 2, SaO_2 is a highly significant determinant of pulse oximeter bias. The most common pattern with early model oximeters was negative bias at lower SaO_2 levels. Newer oximeters exhibit both positive and negative bias patterns for different oximeter/probes combinations. Because of the fundamental influence of SaO_2 on bias, it is essential to account for this effect when analyzing additional influences such as skin pigment and gender. Therefore, all data are presented in the context of SaO_2 range, and analyzed with respect to SaO_2 range or continuous SaO_2 value.

Gender was a statistically significant univariate predictor of pulse oximeter bias. This observation may relate to smaller finger size and a correspondingly smaller pulsatile signal detected by the sensor, a variant of the problem with light attenuation due to dark skin pigmentation. The female subjects had lower levels of hemoglobin, which was statistically significant in place of gender. Because lower hemoglobin was so highly associated with female gender, it was not possible to statistically separate the contributions of gender and low hemoglobin to oximeter error or bias. Anemia was previously reported to increase bias in pulse oximetry (4). We also cannot eliminate the influence of other confounding variables that we did not measure.

The univariate analyses presented in the tables and figures have important limitations. For example, the possibility of confounding influence among skin pigment, gender, SaO_2 , and age might affect the results. Further, we recognize that while specific combinations of gender, skin pigment, and SaO_2 values produce statistically significant differences, the overall statistical significance of the variables themselves is of greater relevance. Multivariate analysis (Table 2) confirmed this highly significant influence of SaO_2 on bias. The most robust statistical model included the interaction of SaO_2 with both gender and skin pigment. Also, interaction terms were more important than the influence of both variables themselves. This interaction is clearly seen in the data: bias is not greatly affected by gender or skin pigment at higher SaO_2 , but increases with oxyhemoglobin desaturation. Age was not a statistically significant variable in the multivariate analysis, although the age range of our

healthy subjects was smaller than might be encountered clinically.

This study and a previous study (3) have found that five types of oximeters show varying degrees of bias at low Sao_2 , suggesting that this bias is a general feature of oximeter design. However, the clip-on probes generally exhibit less bias than the adhesive probes, suggesting that design of sensors or software can affect oximeter accuracy. Because we tested only three types of oximeters in this study and three in a previous study (3), our results may not apply to pulse oximeters made by other manufacturers.

Variability, as measured by RMS error, also increased because of skin pigment. Unlike bias, where positive and negative values may average to give a low mean value, RMS error will also be high with higher variability of bias. RMS error increases at lower Sao_2 , which was the most important variable. Gender was a small, but statistically significant, factor for several oximeter-probe combinations. Skin pigment was not a statistically significant variable for two oximeters: the Nonin 9700 with clip-on probe and the Masimo radical with adhesive probe. However, similar to the effect of skin pigment on bias, the RMS error was higher in darkly pigmented individuals. RMS error for Sao_2 of 70%–100% is used for Food and Drug Administration device approval. Above 3.0 is not acceptable, yet there were several instances of oximeter-probe combinations with values over 3.0. Nearly all values over 3.0 are seen for darkly pigmented subjects, and none for lightly pigmented.

The magnitude of the pulse oximeter error due to dark skin pigmentation is relatively small at Sao_2 values more than 80%, and probably of no general

clinical significance. However, in individuals with darkly pigmented skin, bias of up to 8% was observed at lower Sao_2 levels, which may be quite significant under some circumstances. For example, in congenital heart disease, many patients have stable low Sao_2 values, and accuracy in an outpatient setting or during surgery may be desired. At high altitude, pulse oximeters are frequently relied on for accurate readings in both research and clinical settings, and the bias at lower Sao_2 may be important. Furthermore, studies examining ethnic differences in treatment responses based on Sao_2 as determined by pulse oximetry (e.g., treatment of lung diseases) need to account for differences in oximeter readings between light and darkly pigmented subjects. This may be relevant for the United States Food and Drug Administration or to other regulatory bodies to consider this in designing pulse oximeter accuracy standards.

We conclude that skin pigmentation, gender, and type of oximeter probe all affect pulse oximeter bias and error at low Sao_2 with a bias of approximately +5% at 75% saturation seen in three manufacturers' instruments. These deviations may be clinically relevant in some situations.

REFERENCES

1. Adler JN, Hughes LA, Vivilecchia R, Camargo CA Jr. Effect of skin pigmentation on pulse oximetry accuracy in the emergency department. *Acad Emerg Med* 1998;5:965–70
2. Bothma PA, Joynt GM, Lipman J, Hon H, Mathala B, Scribante J, Kromberg J. Accuracy of pulse oximetry in pigmented patients. *S Afr Med J* 1996;86:594–6
3. Bickler PE, Feiner JR, Severinghaus JW. Effects of skin pigmentation on pulse oximeter accuracy at low saturation. *Anesthesiology* 2005;102:715–19
4. Severinghaus JW, Koh SO. Effect of anemia on pulse oximeter accuracy at low saturation. *J Clin Monit* 1990;6:85–8