

The Pulse Oximeter Is Amazing, but Not Perfect

Philip Bickler, M.D., Ph.D., Kevin K. Tremper, Ph.D., M.D.

It is amazing to think that a device developed for aviation research during World War II and implemented into anesthesia care in the mid-1980s is now likely the most common medical device on Earth—with the possible exception of the thermometer. In the 1940s, Glenn Allen Millikan developed an ear oximeter to estimate hemoglobin saturation for determination of when high-altitude World War II fighter and bomber pilots required supplemental oxygen.¹ The Millikan oximeter required cumbersome calibration and so remained a research device for the next 30 yr. It was the ingenious observation of Takou Aoyagi in 1974 that the arterial signal could be obtained without calibration, if you assume that the only pulsatile absorber in the tissue is the pulsing arterial blood. Aoyagi's pulse oximeter involved a relatively simple device of light-emitting diodes and photodiode detectors and used the pulsatile absorbance signal to estimate arterial hemoglobin saturation. Once an empiric calibration for this approach was established for human subjects in controlled laboratory hypoxia conditions, no further calibration of pulse oximeters is needed. For the most part, the device today is the same as the original pulse oximeters of the 1980s.

Before the pulse oximeter, clinicians relied on observed cyanosis to detect hypoxemia. Julius Comroe at the University of California, San Francisco (San Francisco, California), conducted volunteer studies in the 1940s using a Millikan oximeter to assess the accuracy of clinicians' ability to detect clinical cyanosis.² He demonstrated that these observations were unreliable until saturation was less than 80% and still highly variable between subjects and observers. This study was conducted in a well-lit setting in White subjects, recognizing that darker skin would make the detection of anoxemia more difficult, hence, the need



“How does nonpulsatile skin pigmentation affect the pulse oximeter accuracy...?”

for and rapid adoption of pulse oximetry as a standard of care. In this issue of *ANESTHESIOLOGY*, Burnett *et al.* analyze the accuracy of pulse oximeters in a large retrospective cohort of anesthetized patients with varying degrees of skin pigmentation.³ It was noted in previous studies of volunteer subjects and intensive care unit patients that pulse oximeter readings were erroneously higher at lower saturations in patients with darker skin.⁴ This current operating room study analyzed 11 yr of data from 46,000 patients under anesthesia. Their surrogate marker for skin pigmentation was self-reported race in the categories of White, Black, Hispanic, Asian, and Other. They estimated arterial oxygen saturation (SaO_2) by calculating saturation from blood gas data. The traditional way of assessing the accuracy of two methods of measuring a variable (*e.g.*, oxygen saturation measured by pulse oximetry [SpO_2] *vs.* SaO_2) is by a bias analysis—that is, the mean difference between the two measures and the SD of those differences. The bias being the average difference is the systematic error, and the SD of differences (or precision, as it is sometimes called) the random error.

In addition to determining the bias and precision by skin pigmentation groups, they also chose a clinical measure of the incidence of unrecognized hypoxemia defined as a saturation SaO_2 less than 88% when the SpO_2 reading was greater than 92 to 96%. In this analysis, they found that the incidence of occult hypoxemia differed with skin pigmentation (*e.g.*, White, 1.1%; Hispanic, 1.8%; and Black, 2.1%). The good news is this is a low incidence; the better news is that for the group with SpO_2 greater than 96%, incidence was rare, and there were no differences among racial/ethnic groups. So, the clinical bottom line is to keep the SpO_2 greater than 96%. If that cannot be achieved by increasing fraction of inspired oxygen or modifying positive

Image: Adobe Stock.

Accepted for publication February 11, 2022. Published online first on March 18, 2022. This editorial accompanies the article on p. 688.

Philip Bickler, M.D., Ph.D.: Department of Anesthesiology, University of California, San Francisco, San Francisco, California.

Kevin K. Tremper, Ph.D., M.D.: Department of Anesthesiology, University of Michigan, Ann Arbor, Michigan.

Copyright © 2022, the American Society of Anesthesiologists. All Rights Reserved. *Anesthesiology* 2022; 136:670–1. DOI: 10.1097/ALN.0000000000004171

end-expiratory pressure, an invasive blood sample may be considered.

Several comments and questions related to the conclusions of the work from Burnett *et al.* are appropriate. Pulse oximeter errors and bias can be caused by motion (shivering), interfering substances (carboxyhemoglobin, methemoglobin), mistiming of blood draws and oximeter readings, optical interference, probe misplacement, and low perfusion. In the case of low perfusion, reductions in the pulse oximeter signal are compounded by light absorption by melanin. In the study by Burnett *et al.*, none of these factors were controlled. In addition, Black patients were older on average (62 vs. 52 yr) and had a greater incidence of kidney failure and diabetes than White patients. Could these disparities have contributed to lower perfusion, and amplified the calibration error? Further study is needed.

During the early years of pulse oximetry, it is likely that the in-human testing of pulse oximeters involved almost all White patients. It was not until 2005 that the limitations of this approach were identified.⁴ It was noted that darker skin pigmentation caused a positive bias—that is, the pulse oximeter reading was higher than the actual saturation measured on a blood sample. The error was greater in saturation ranges that were less than 80%, but some bias existed at higher ranges as well. A recent controlled laboratory study, with multiple types of good-quality oximeters for sale in 2017 to 2020, found that oximeters *still* read 1 to 2% too high in patients with darker skin who were near the critical 90% hypoxia threshold.⁵

How does nonpulsatile skin pigmentation affect the pulse oximeter accuracy, particularly in producing a positive bias that might cause hypoxemia to be missed? The pulse oximeter measures the ratio of the pulse's added absorbance in red and infrared light transmitted through the tissue. The ratio (R) is then empirically calibrated with human volunteer data to produce an “ R calibration curve,” (*e.g.*, $R = 3.4$, $SpO_2 = 100\%$; $R = 1.0$, $SpO_2 = 85\%$). If the pigment acts as a variable light filter for the transmitted light frequency, the peak frequency could slightly change and produce a slightly changed R and slightly altered SpO_2 value. The red light-emitting diodes used in pulse oximeters do not produce a single wavelength of red light, but rather a bell curve distribution of wavelengths, and the shorter wavelengths of this distribution are more heavily absorbed by melanin than the longer. In effect, the R curve for patients with darker skin pigmentation needs to be different than that for White patients.

Another unanswered important clinical question is regarding patients who live in the low saturation ranges. The current study provides reassurance when SpO_2 is greater than 96%, but what about children with cyanotic heart disease? These patients live in the most error-prone range of SpO_2 . During general anesthesia, it is rare to have patients with sustained saturations in the low 90s or 80s, but it is part of the treatment plan for patients with cyanotic heart disease. This population needs further study.

Overall, the work of Burnett *et al.* provides clinical context to errors in a device that we depend on every day. It gives us new targets for saturation ranges to be safe for all patients. The pulse oximeter is incredibly useful and reliable for medical monitoring, and it works on a tremendous range of patients. Even in the very low ranges of saturation where there are few to no calibration data, its directional trends are very useful. We should give thanks to the ingenuity of Takuo Aoyagi every day when we are reassured by that *beep, beep, beep*.

Competing Interests

The authors are not supported by, nor maintain any financial interest in, any commercial activity that may be associated with the topic of this article.

Correspondence

Address correspondence to Dr. Tremper: ktremper@umich.edu

References

1. Tremper KK, Barker SJ: Pulse oximetry. *ANESTHESIOLOGY* 1989; 70:98–108
2. Comroe JH Jr, Botelho S: The unreliability of cyanosis in the recognition of arterial anoxemia. *Am J Med Sci* 1947; 124:1–6
3. Burnett GW, Stannard B, Wax DB, Lin H-M, Pyram-Vincent C, DeMaria S Jr, Levin MA: Self-reported race/ethnicity and intraoperative occult hypoxemia: A retrospective cohort study. *ANESTHESIOLOGY* 2022; 136:688–96
4. Bickler PE, Feiner JR, Severinghaus JW: Effects of skin pigmentation on pulse oximeter accuracy at low saturation. *ANESTHESIOLOGY* 2005; 102:715–9
5. Okunlola OE, Lipnick MS, Batchelder PB, Bernstein M, Feiner JR, Bickler PE: Pulse oximeter performance, racial inequity, and the work ahead. *Respir Care* 2022; 67:252–7