Racial and Ethnic Bias in Pulse Oximetry and Clinical Outcomes
Valeria S. M. Valbuena, MD, MSc; Raina M. Merchant, MD, MSHP; Catherine L. Hough, MD, MSc

Assessment of blood oxygen saturation is an important measure of health on which many diagnostic and treatment decisions are based. Blood oxygen saturation is most commonly assessed via pulse oximetry, with increasing use across the home, clinic, and hospital settings during the COVID-19 pandemic. However, inaccuracies in pulse oximetry measurement have come under scientific scrutiny over the past 2 years. Professional organizations, lawmakers, and the public have actively engaged with the issue. A known design flaw of the pulse oximeter is that patients with darker skin (compared with lighter skin) are more likely to experience occult hypoxemia—defined conceptually as substantial arterial hypoxemia (SaO2) detected on blood gas but not noted on simultaneous pulse oximetry (SpO2). Less is known about whether underdiagnosis of hypoxemia for historically marginalized racial and ethnic groups has consequences.

In their article “Racial and Ethnic Discrepancy in Pulse Oximetry and Delayed Identification of Treatment Eligibility Among Patients With COVID-19,” Dr Fawzy and colleagues evaluate this measurement bias in pulse oximetry among a population of hospitalized patients diagnosed with COVID-19. The authors show measurement bias was associated with differences in delay in recognizing patients’ eligibility for COVID-19 treatment, with racial and ethnic minoritized groups being more affected. Although less visible, this is similar to the ways in which race- and ethnicity-based cutoffs in pulmonary function tests and estimated glomerular filtration fraction lead to underestimation of disease severity among patients of racial and ethnic minoritized groups, which in turn delays treatment of severe lung disease and limits access to timely kidney replacement therapy, including transplantation.

The study by Dr Fawzy and colleagues was a retrospective analysis using a large database from 5 well-resourced hospitals. As in other investigations of pulse oximetry, the authors evaluated clinical SaO2 and SpO2 measurements within 10 minutes of each other, a design that allows for head-to-head comparison of pulse oximetry estimations of arterial oxygenation with the gold standard from blood gas analysis. Using race and ethnicity as a proxy for skin color, the authors chose to aggregate the data for Black–Hispanic patients and Black patients into a single study group.

Dr Fawzy and colleagues reported higher rates of occult hypoxemia (SaO2 <88% when SpO2 was 92%–96%) for Asian, Black, and Hispanic patients compared with non-Hispanic White patients, as well as delayed or unrecognized eligibility for the US Centers for Disease Control and Prevention guideline-concordant COVID-19 treatments (eg, steroids, remdesivir) for Black and Hispanic patients. Using underdetection of a treatment threshold to assess the consequences of undetected hypoxemia was a novel and key feature of this study. Remdesivir has US Food and Drug Administration approved use for patients whose SpO2 level is less than 94%, and dexamethasone is only recommended for patients with COVID-19 who are receiving supplemental oxygen. These findings suggest that underdetection of even mild hypoxemia can delay guideline-appropriate care. Increased mortality and/or differential mortality in patients of racial and ethnic minoritized groups has been demonstrated in studies of acute respiratory distress syndrome, a condition defined by arterial hypoxemia and for which treatment decisions frequently are based on pulse oximetry. Racial and ethnic discrepancy in the accuracy of commonly used pulse oximetry might play an important role in the outcome disparities described in the pulmonary literature.

A recent investigation by Wong and colleagues similarly demonstrated consequences of underdetection of hypoxemia among a population of patients with respiratory disease. In this multicenter, retrospective, cross-sectional study using data from 3 publicly available databases that included 215 hospitals and 382 intensive care units, the authors associated higher odds of hospital mortality, elevated lactate, and high sequential organ failure assessment scores with a higher prevalence of occult hypoxemia among Asian, Black, and Hispanic patients compared with White patients. The outcome of interest (occult hypoxemia) was defined differently than in Dr Fawzy and colleagues’ study (ie, SaO2 < 88% when SpO2 > 88%), yet both studies showed that underdetected arterial hypoxemia was harmful and more frequently encountered in patients of racial and ethnic minoritized groups.

Based on prior investigations and the most recent studies on pulse oximetry, 2 questions stand: (1) Why has this racial and ethnic discrepancy in the accuracy of commonly used pulse oximetry not been addressed? (2) How should clinicians respond given the increased awareness of the problem and some of its ramifications?

Historical neglect of patients of racial and ethnic minoritized groups and a diminished concern for their health outcomes may explain in part why this phenomenon—differential pulse oximetry accuracy—has been recognized for more than 30 years and has not been corrected. Using White patients as the standard in biomedical design has led to both differential care and innovation inertia for optimizing the way devices and algorithms work for patients of racial and ethnic minoritized groups. The economics associated with ignoring the issue cannot be excluded from this discussion. Hospitals and practitioners continue to buy and use these devices despite their inaccuracy for non-White patients. The observation that designing a new de-
vice and exchanging millions of machines in hospitals and clinics across the country may be deemed unpopular could suggest that racial equity in patient care is not something these institutions are willing to pay for—or at least not enough of a priority to insist on devices that work equally. Similarly, changing this injustice paradigm presents a challenge for clinicians, in part because of the passivity of medical education—students of medicine are taught how the system works, not how to change it.1,2

There is no single easy solution to the problem of pulse oximetry’s racial and ethnic bias. Although an easy bedside adjustment for the measurement disparity is an attractive idea, there is no proven approach because of the noise and variable bias of the currently available devices. Interestingly, the design flaws of the pulse oximeter have been corrected in some devices. These improved devices use more wavelengths of light and perform equally for different skin tones; they have been used in limited settings but have not been widely produced or distributed.1

The next generation of oxygen monitors should be designed and tested to work equally on all patients. However, this is an unlikely scenario without market pressure. Hospital systems, medical practitioners, and regulatory entities need to press for regulatory scrutiny and design equity by limiting purchasing choices to devices with equal performance for patients of all skin tones. In the meantime, a clinical strategy for avoiding undertreating hypoxia in patients with darker skin is, unfortunately, lowering the threshold for suspected disease and obtaining more arterial blood gases. Ultimately, pulse oximetry is a (flawed) measure of blood arterial saturation, which is connected to tissue hypoxia through oxygen delivery and uptake. There are multiple clinical trends that can be used to globally assess tissue hypoxia in patients with an array of illnesses. Increasing invasive testing, such as arterial blood gases, puts patients of racial and ethnic minority groups at higher risk of rare but possible complications, as well as physical pain, but these need to be balanced against the risks of hypoxemia, which are not uncommon.

Although differences in pulse oximetry may seem to be most meaningful near treatment thresholds, certain patient populations (ie, Black patients undergoing extracorporeal membrane oxygenation for acute respiratory distress syndrome) have high rates of occult hypoxemia,4 indicating a need for greater awareness when treating these groups. Moreover, with children, even mild occult hypoxemia can have substantial long-term effects.5 Why tolerate a device that consistently works less well in patients with darker skin tones?

Additionally, the association of underdetection of occult hypoxemia with treatment delay cannot be dissociated from the multilevel injustices and persistent disparities underlying the worse COVID-19 outcomes experienced by patients of racial and ethnic minority groups. These larger and systemic issues exist outside of device design issues demonstrated by these investigations.6 Although the device measurement error is real and based purely on optics, the decision to do nothing about a faulty device is a human one, and one that can and should be corrected.

ARTICLE INFORMATION

Author Affiliations: Department of Surgery, University of Michigan, Ann Arbor (Valbuena); Center for Healthcare Outcomes and Policy, University of Michigan, Ann Arbor (Valbuena); National Clinician Scholars Program, University of Michigan, Ann Arbor (Valbuena); Department for Emergency Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Merchant); Center for Digital Health, Penn Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Merchant); Division of Pulmonary Medicine and Critical Care Medicine, Department of Internal Medicine, Oregon Health Sciences University, Portland (Hough).

Corresponding Author: Valeria S. M. Valbuena, MD, MSc, Department of Surgery, University of Michigan, 1500 E Medical Center Dr, 2110 Taubman Center, SPC 5346, Ann Arbor, MI 48109 (valeria@umich.edu).

Published Online: May 31, 2022.
doi:10.1001/jamainternmed.2022.1903

Conflict of Interest Disclosures: Dr Valbuena is supported by the Institute for Healthcare Policy and Innovation Clinician Scholars Program and the US National Institutes of Health (NIH: No 5T32HS000053-29) outside the submitted work. Dr Merchant reported support from NIH (Nos. ROIHL14184-401 and K25 HL157621), NIH (No. RO1MH127686), and NIDA (No. IR1DA050761) and serving as a member of Accelerate Health Equity, all outside the submitted work. Dr Hough reported grants from NIH (Nos. K24HL141526, RO1NR016459, U01AT0099741, RO1HL132232) and from the American Lung Association; and compensation for DSMB membership for Quantum Leap, all outside the submitted work. No other disclosures were reported.

REFERENCES


