Racial Disparities in Occult Hypoxemia and Clinically Based Mitigation Strategies to Apply in Advance of Technological Advancements

Christopher F Chesley, Meghan B Lane-Fall, Venkat Panchanadam, Michael O Harhay, Arshad A Wani, Mark E Mikkelsen, and Barry D Fuchs

BACKGROUND: Pulse oximetry is the mainstay of patient oxygen monitoring. Measurement error from pulse oximetry is more common for those with darker skin pigmentation, yet this topic remains understudied, and evidence-based clinical mitigation strategies do not currently exist. Our objectives were to measure the rate of occult hypoxemia, defined as arterial oxygen saturation (S_{aO_2}) < 88% when pulse oximeter oxygen saturation was between 92–96%, in a racially diverse critically ill population; to analyze degree, direction, and consistency of measurement error; and to develop a mitigation strategy that minimizes occult hypoxemia in advance of technological advancements. METHODS: We performed a multi-center retrospective cohort study of critically ill subjects. RESULTS: Among 105,467 paired observations from 7,693 subjects, we found occult hypoxemia was more common among minority subjects. The frequency of occult hypoxemia was 7.9% versus 2.9% between Black and white subjects, respectively, (P < .001). Pulse oximeter measurement errors were inconsistent throughout a patient encounter, with 67% of encounters having a range of intra-subject measurement errors > 4 percentage points. In 75% of encounters, the intra-subject errors were bidirectional. S_{aO_2} < 88% was less common at higher pulse oximeter oxygenation ranges (4.1% and 1.8% of observations among Black and white subjects at a pulse oximeter threshold of 94–98%). Although occult hypoxemia was further reduced at oxygenation saturation range 95–100%, the frequency of hyperoxemia (partial pressure of arterial oxygen > 110 mm Hg) became more common, occurring in 42.3% of Black and 46.0% of white observations. CONCLUSIONS: Measurement error in pulse oximetry is common for all racial groups, but occult hypoxemia occurred most commonly in Black subjects. The highly variable magnitude and direction of measurement error preclude an individualized mitigation approach. In advance of technological advancements, we recommend targeting a pulse oximetry saturation goal of 94–98% for all patients. *Key words: oximetry; racial bias; measurement error; disparities; hypoxia.* [Respir Care 2022;67(12):1499–1507. © 2022 Daedalus Enterprises]

Introduction

In emergency departments, clinical wards, ICUs, operating rooms, and procedural suites, pulse oximeters are indispensable to assessing oxygenation for the acutely and seriously ill. Pulse oximeters use the absorption of red- and near-infrared wavelength light (at 660 nm and 940 nm, respectively) in pulsatile arterial blood to estimate arterial oxygen saturation (S_{aO_2}).[1] Most approved oximeters have a reported accuracy specification between 1–2%, with blood co-oximetry serving as the accepted standard.[2] However, many factors encountered during routine medical care impact the reliability of pulse oximetry.[3,4]

For over 35 years, it has been known that pulse oximeters are comparatively inaccurate for individuals with dark skin pigmentation.[4-12] A recent study brought this issue to international attention[13] and invigorated clinicians, investigators, and regulatory agencies to take action. In the study by Sjoding and colleagues,[13] the investigators found that pulse oximeter readings frequently overestimated the S_{aO_2} in hospitalized subjects. Inaccuracies were disproportionately present in subjects of color such that occult hypoxemia, defined as S_{aO_2} <

88% among subjects with pulse oximeter saturation between 92–96%, was 3 times more likely in Black compared to white subjects.[13] Subsequent studies[14,15] have confirmed that occult hypoxemia is more common among patients of color and is associated with poor clinical outcomes. However, no previous study has used data related to pulse oximeter measurement error to develop an intervention designed to minimize disparities.

SEE THE RELATED EDITORIAL ON PAGE 1633

Our goals for this study were several-fold. First, given evidence that suggested certain pulse oximeters may be more accurate in patients of color,[7] we examined the frequency of occult hypoxemia across different pulse oximeters in a racially diverse multi-center cohort of acutely ill subjects. We then examined whether the observed bias was the result of dynamic changes in hospitalized subjects by conducting analyses in steady-state conditions. Second, once confirmed, we sought to develop a health system-wide approach to mitigate racial disparities in occult hypoxemia and to avoid the potential risk of occult hyperoxemia, defined as an arterial oxygen partial pressure > 110 mm Hg,[16] in advance of needed technological advancements. We conducted a series of analyses to understand magnitude, direction, and consistency of measurement bias to inform the development of a systematic approach to mitigate the development of occult hypoxemia, especially for patients with darker skin pigmentation, for whom Black race is often used as a proxy.

QUICK LOOK

Current knowledge

Pulse oximetry is the standard for oxygen monitoring of critically ill patients, but inaccuracies have been reported for dark skinned patients as a result of pulse oximeter measurement error. This is a serious limitation of commonly used respiratory monitoring devices that has been known for over 3 decades and is the source of disparities of occult hypoxemia disproportionately experienced among Black patients.

What this paper contributes to our knowledge

Occult hypoxemia was common for Hispanic and Asian/Pacific Islander subjects in addition to Black subjects. In addition, this study demonstrates that the direction and magnitude of measurement error preclude a patient-based approach to minimize disparities (ie, applying a "correction factor" would not accurately target arterial oxygen saturation). Rather, a threshold-based approach for patients of all ethnic backgrounds, in which oxygen saturations are targeted between 94–98%, is the most informed approach to minimize harms from occult hypoxemia related to pulse oximeter measurement error.

Methods

Study Design

We conducted a multi-center, observational cohort study of critically ill subjects hospitalized at Penn Medicine, located in Philadelphia, Pennsylvania, between January 1, 2019–January 15, 2021. The study was reviewed and approved as a quality improvement project by the institutional review board of the University of Pennsylvania (number 844981).

Dr Chesley is affiliated with Pulmonary, Allergy, and Critical Care Division, Department of Medicine, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania; Palliative and Advanced Illness Research Center, Department of Medicine, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania; and Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, Pennsylvania. Dr Lane-Fall is affiliated with Palliative and Advanced Illness Research Center, Department of Medicine, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania; Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, Pennsylvania; Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania; and Department of Anesthesiology and Critical Care, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania. Dr Panchanadam is affiliated with Penn Value Improvement, Penn Medicine, University of Pennsylvania, Philadelphia, Pennsylvania. Dr Harhay is affiliated with Pulmonary, Allergy, and Critical Care Division, Department of Medicine, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania; Palliative and Advanced Illness Research Center, Department of Medicine, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania; Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, Pennsylvania; and Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania. Drs Wani and Fuchs are affiliated with Pulmonary, Allergy, and Critical Care Division, Department of Medicine, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania. Dr Mikkelsen is affiliated with Division of Pulmonary Sciences and Critical Care Medicine, Department of Medicine, University of Colorado, Anschutz Medical Campus, Aurora, Colorado.

The authors have disclosed no conflicts of interest.

Dr Chesley received funding from the American Thoracic Society Fellowship in Health Equity.

Drs Mikkelsen and Fuchs are co-senior authors.

Correspondence: Christopher F Chesley MD, 3400 Spruce Street, 839 West Gates Building, Philadelphia, PA 19104. E-mail: christopher.chesley@pennmedicine.upenn.edu.

DOI: 10.4187/respcare.09769

Study Population

We included patients admitted to ICUs at the Hospital of the University of Pennsylvania and Penn Presbyterian Medical Center.

Measurements of Oxygen Saturation

At the Hospital of the University of Pennsylvania, S_{pO_2} was measured using Covidien Nellcor oximeters and OxiMax technology disposable finger sensors (Medtronic, Mansfield, Massachusetts). At Penn Presbyterian Medical Center, S_{pO_2} was measured using Masimo pulse oximeters and LNCS disposable sensors (Masimo, Irvine California). S_{aO_2} was measured in both sites by multi-wavelength co-oximeters (ABL90 [Danaher, Washington, District of Columbia] at the Hospital of the University of Pennsylvania; GEM 4000 [Instrumentation Laboratory, Bedford, Massachusetts] at Penn Presbyterian Medical Center).

In a pilot study that preceded this work, we examined 40 pulse oximetry measurements at the time of arterial blood gas (ABG) at Penn Presbyterian Medical Center. ABGs were analyzed by multi-wavelength co-oximetry using a broad-spectrum spectrometer. We determined the presence of dyshemoglobins, including carboxyhemoglobinemia (COHb) and methemoglobinemia (MetHb), during the pilot. We did not observe evidence of dyshemoglobinemia, as none of the readings were out of normal range (average COHb was 0.99, and average MetHb was 0.83). As we observed bias similar in frequency to that observed by Sjoding et al[13] in the pilot, we did not assess for dyshemoglobinemia in our current study.

Data Collection

We retrospectively abstracted clinical information from the electronic health record. We limited our study to critically ill patients. Consistent with prior studies, we paired S_{pO_2} with S_{aO_2} measurements via ABG obtained within 10 min of each other.[13] We also extracted subject age, sex, race, ethnicity, and hemoglobin on the day of ABG. Subject race and ethnicity were self-reported or identified by clinical staff if clinical condition precluded subject self-report. Racial and ethnic study categories were based on National Institutes of Health recommendations[17] for reporting in clinical research and related guidance.

Outcomes

The primary outcome was frequency of occult hypoxemia, defined as $S_{aO_2} < 88\%$ when S_{pO_2} was within one of 3 specified normal ranges.[13] For our primary analyses, we examined frequency of occult hypoxemia when pulse oximeter oxygen saturation was between 92–96%.[13] We subsequently examined how frequently S_{aO_2} was $< 88\%$ at the higher goal S_{pO_2} ranges 94–98% and 95–100%. To understand how frequently occult hypoxemia occurs at the subject level, we reported frequency of occult hypoxemia at paired observational level and at subject level. An instance of occult hypoxemia identified at paired observational level for a given subject's hospitalization would result in the subject being categorized as having experienced occult hypoxemia. The secondary outcome was frequency of arterial hyperoxemia, defined as partial pressure of arterial oxygen > 110 mm Hg on ABG.[16]

Steady-State Subgroup Analysis

Because measurements displayed on pulse oximeters might not immediately reflect an abrupt change in arterial saturation given the delay associated with signal averaging algorithms, we identified a subset of samples that had a 60-min period of S_{pO_2} stability. Stability was defined as having all S_{pO_2} readings within 3 percentage points of each other during the 30 min before and after paired measurements. In sensitivity analyses, we further restricted oxygen saturation range to be within 1 or 2 percentage points in the 30 min preceding, and after, paired samples.

Statistical Analysis

We conducted a series of analyses to understand magnitude, direction, and consistency of intra-subject measurement errors during a subject's hospitalization. These analyses were used to determine whether one could reliably use the initial error measured between paired and S_{aO_2} values to correct all subsequent S_{pO_2} measurements. We calculated the range and SD in errors measured from all paired samples available throughout a subject's hospitalization to quantitate the intra-subject variability in S_{pO_2} measurement error. We used Bland-Altman plots to visually compare distribution and SD of errors measured in all available sample pairs between Black and white subjects. Bland-Altman plots depict paired differences between S_{pO_2} and S_{aO_2} measurements as a function of mean value of the 2 paired oxygen saturation readings. In addition, for each subject encounter, we counted the numbers of paired observations that either overestimated or underestimated S_{aO_2} to examine directionality and consistency. For an alternative mitigation strategy, we also examined how frequently S_{aO_2} was $< 88\%$ at progressively higher-goal S_{pO_2} ranges, including 94–98%[18] and 95–100%. We also measured rate of hyperoxemia across these S_{pO_2} goal ranges.

We report counts and percentages. We used chi-square and Kruskal-Wallis tests when comparing categorical and continuous (respectively) variables between groups. We used $P \leq .05$ to signify statistical significance. We used multivariable logistic regression to adjust for age, sex, and hemoglobin on the day of ABG to examine the association between self-

Table 1. Subject Characteristics and Rate of Occult Hypoxemia by Race and Ethnicity

Subject Characteristics	White	Black	Latinx	Asian/Pacific Islander	Indigenous	Other	Unknown	Total
Total subjects	4,621 (60)	1,919 (25)	226 (3)	239 (3)	17 (0.2)	220 (3)	451 (6)	7,693 (100)
Total paired observations	59,975 (57)	25,148 (24)	4,279 (4)	3,836 (4)	216 (0.2)	3,933 (4)	8,080 (8)	10,5467 (100)
Number of paired observations per subject	6 (2–12)	4 (1–12)	8 (3–21)	7 (2–14)	8 (3–10)	7.0 (2.0–17.5)	6 (2–16)	6 (2–12)
Minimum, maximum number of paired observations per subject	1, 322	1, 497	1, 232	1, 268	1, 57	1, 278	1, 335	1, 497
Age, y	66.0 (56–74)	62.0 (52–71)	58.5 (44–68)	61.0 (48–70)	66.0 (51–69)	64.0 (54–72)	64.0 (52–73)	64.0 (54–73)
Female	1,730 (37.4)	976 (50.9)	84 (37.2)	96 (40.2)	9 (52.9)	87 (39.5)	176 (39.0)	3,158 (41.1)
Rate of occult hypoxemia at paired observational level*†	373/13,060 (2.9)	382/4849 (7.9)	43/901 (4.8)	35/746 (4.7)	1/66 (1.5)	37/898 (4.1)	63/1,875 (3.4)	934/22,395 (4.2)
Rate of occult hypoxemia at subject level*†	264/2,397 (11.0)	203/922 (22.0)	20/128 (15.6)	21/112 (18.8)	1/9 (11.1)	19/120 (15.8)	38/268 (14.2)	566/3,956 (14.3)

Data are presented as *n*, *n* (%), or median (interquartile range).
Occult hypoxemia is defined as arterial oxygen saturation < 88% when S_{pO_2} saturation was between 92–96%.
*An instance of occult hypoxemia identified at paired observational level for a given subject's hospitalization would result in the patient being categorized as having experienced occult hypoxemia.
†$P < .001$ for comparing rates across self-reported race and ethnicity.

Table 2. Rate of Occult Hypoxemia at Individual Oxygen Saturation Values, Stratified by Self-Identified Race

S_{pO_2}	White	Black
92	73/1,098 (6.6)	90/504 (17.9)
93	69/1,544 (4.5)	75/683 (11.0)
94	71/2,268 (3.1)	65/877 (7.4)
95	81/3,327 (2.4)	79/1,202 (6.6)
96	79/4,823 (1.6)	73/1,583 (4.6)

Data are presented as % or *n* (%).
Occult hypoxemia is defined as arterial oxygen saturation < 88% at specified values within the goal range for S_{pO_2} saturation. Only Black and white subjects are presented due to sample size.

reported race and occult hypoxemia.[7,19] We used Stata 15.1 (StataCorp, College Station, Texas) and Tableau Desktop Professional Edition Version 2020.1.4 (Tableau Software, Seattle, Washington) to perform analyses.

Results

We examined 105,467 paired observations from 7,693 unique subjects (Table 1). Since there were no differences in results when measurements were compared across hospitals, the data were combined and the results reported accordingly. Specifically, at 2 hospitals using different pulse oximeters and disposable sensors and different blood gas analyzers, we found that occult hypoxemia was more common among Black subjects compared with white subjects.

As presented in Table 1, at the level of paired observations, frequency of occult hypoxemia was more common among Black subjects compared to white subjects, at 7.9% and 2.9%, respectively, ($P < .001$). At the subject level, this meant that 22% of Black subjects and 11% of white subjects experienced occult hypoxemia during their hospitalization. Among other races and ethnicities, the frequency of occult hypoxemia ranged between 1.5–4.8% (Table 1). Occult hypoxemia was more common at lower S_{pO_2} saturations and for Black subjects (Table 1–2). For example, at S_{pO_2} 92%, S_{aO_2} was < 88% in 17.9% and 6.6% of observations from Black and white subjects, respectively. As presented in Table 3, the association between self-reported race and occult hypoxemia persisted after adjustment for age, sex, and hemoglobin.

Steady State Analyses

We examined 25,678 pulse oximetry measurements, from 5,262 subjects, where paired ABG sample was obtained while S_{pO_2} was in "steady-state." As shown in Figure 1, the frequency of occult hypoxemia and magnitude of difference between white and Black subjects (6.2% vs 1.2%) were similar to the primary analyses,

Table 3. Relationship Between Self-Identified Race and Ethnicity and Occult Hypoxemia, With and Without Adjustment for Potential Confounders

	White	Black	Latinx	Asian/Pacific Islander	Indigenous	Other
OR	Reference	2.91	1.70	1.67	0.52	1.46
95% CI		(2.51–3.37)	(1.23–2.35)	(1.17–2.39)	(0.07–3.78)	(1.04–2.06)
P		< .001	.001	.004	.52	.031
Adjusted	Reference	2.84	1.69	1.64	0.51	1.41
OR *95% CI		(2.44–3.30)	(1.22–2.34)	(1.15–2.34)	(0.07–3.72)	(0.99–1.99)
P		< .001	.002	.007	.51	.053

*Adjusted for age, sex, and hemoglobin on the day of arterial blood gas.
OR = Odds ratio

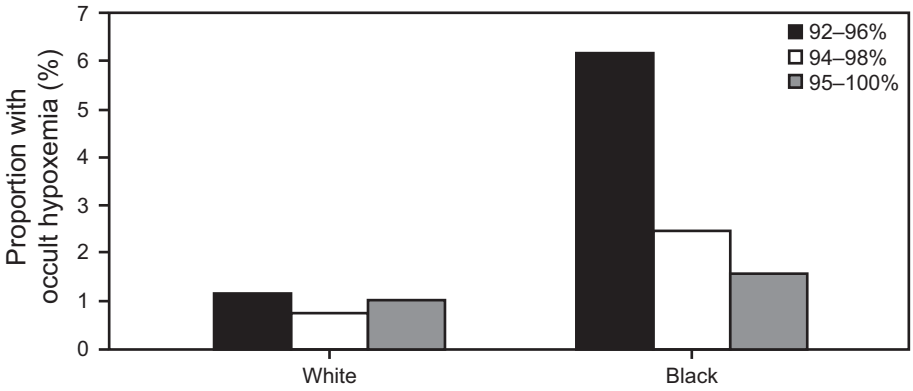

Fig. 1. Rates of occult hypoxemia, defined as arterial oxygen saturation < 88% across 3 pulse oximeter oxygen saturation ranges (92–96%, 94–98%, and 95–100%), in the steady-state subgroup. Subjects in the steady-state subgroup had pulse oximeter saturation that changed ≤ 3% for 30 min before and after paired collection of samples. Due to sample sizes, only Black and white subjects are included.

Table 4. Maximum Difference in Measurement Error Between Paired Arterial and Pulse Oximeter Oxygen Saturation Levels, at the Subject Encounter Level, Stratified by Self-Reported Race

Study Group	Difference in Measurement Error			
	0	≤ 4	≤ 7	≤ 10
Black	7 (1)	236 (24)	408 (41)	529 (54)
White	61 (2)	1,036 (36)	1,691 (59)	1,994 (70)
Total	68 (2)	1,272 (33)	2,099 (54)	2,523 (65)

Data are presented as *n* (%).
In total, intra-subject measurement error was ≤ 4 percentage points in 33% of subject encounters, with the remaining 67% of encounters having intra-subject measurement error > 4 percentage points.

which included all paired observations when S_{pO_2} was in 92–96% range.

Mitigation Strategies

Comparing multiple paired samples from individual subjects, we found large differences in magnitude of errors in S_{pO_2} measurements. As shown in Table 4, errors varied by < 4% points in saturation in only 33% of all subjects. Consistent with the above findings, the intra-subject SD of measurement errors was large and was greater in Black compared to white subjects (Fig. 2). Furthermore, as shown in Figure 3, directionality of intra-subject measurement errors was inconsistent such that in > 3 of 4 subjects errors were bidirectional. The findings from these 3 analyses were confirmed when performed using the steady-state cohort.

When we evaluated paired samples from progressively higher pulse oximeter oxygenation ranges, the frequency of S_{aO_2} being < 88% decreased, whereas the frequency of hyperoxemia increased (Fig. 4). At S_{pO_2} range 94–98%, S_{aO_2} was < 88% in 4.1% and 1.8% of observations from Black and white subjects, respectively. This change was associated with a modest increase in frequency of occult hyperoxemia. Although frequency of S_{aO_2} being < 88% was further reduced at S_{pO_2} range 95–100%, the rate of occult hyperoxemia rose to 42.3% and 46.0% in Black and white subjects, respectively. Notably, disparity between races persisted even at the highest S_{pO_2} range, with S_{aO_2} being < 88% in 2.5% of observations from Black subjects compared to 1.4% of observations from white subjects ($P < .001$).

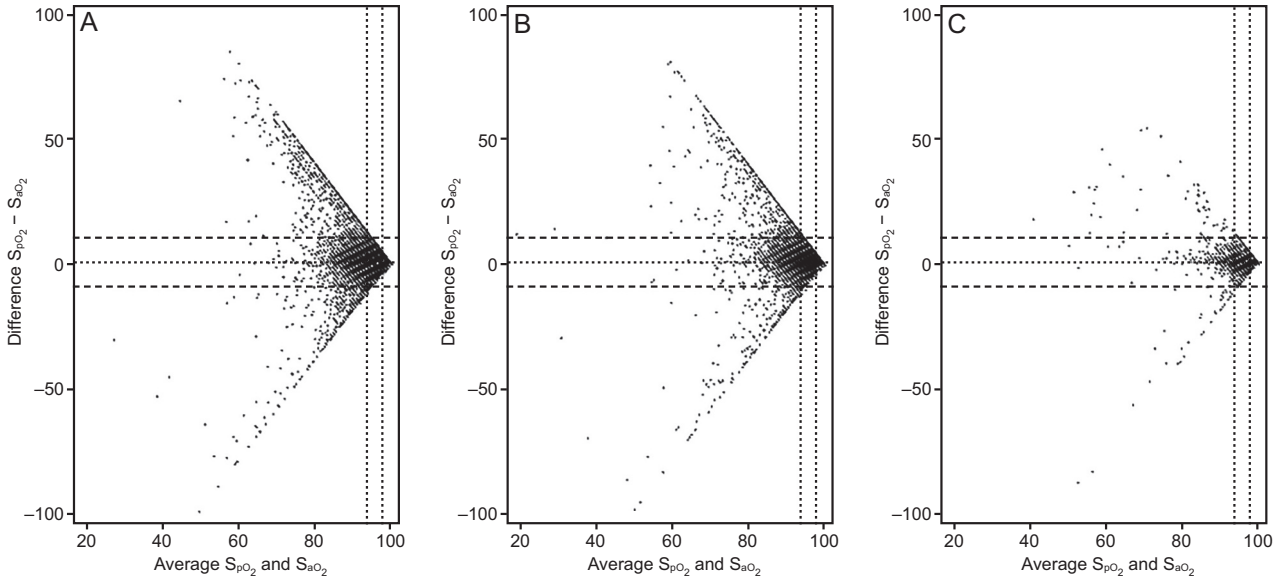

Fig. 2. Bland-Altman plot, stratified by self-reported race. A: White, B: Black, and C: Hispanic. The y axis presents difference between pulse oximeter and arterial oxygen saturations among paired observations. The x axis depicts the mean oxygen saturation for each paired measurement. Dotted lines on the y axis refer to the lower and upper bounds of the 95% confidence interval for the mean difference between S_{pO_2} and S_{aO_2}. Dotted lines on the x axis indicate the 94% and 98% oxygen saturation threshold.

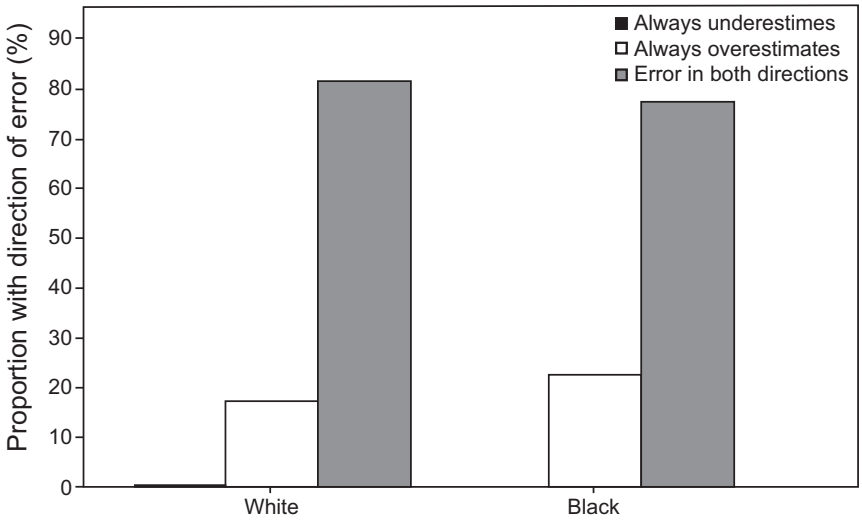

Fig. 3. Directionality of measurement error between pulse oximetry and arterial oxygen saturation (S_{aO_2}) measures, stratified by self-reported race. Always underestimates refers to proportion of subjects who had pulse oximeter saturation that was always $< S_{aO_2}$. Always overestimates refers to proportion of subjects who had pulse oximeter saturation that was always $> S_{aO_2}$. Error in both directions refers to proportion of subjects who had errors in both directions. Due to sample sizes, only Black and white subjects are included.

Discussion

In this retrospective observational study, we found that occult hypoxemia was disproportionately more common among minority groups. Whereas occult hypoxemia was most common among Black subjects, it was also more common among Latinx, Asian, and Pacific Islander subjects compared to white subjects. Though attenuated, we found these disparities were present among subjects who had reached a steady state of oxygen stability by pulse oximeter. Our findings support the notion that patients with darker skin pigmentation are vulnerable to pulse oximeter mismeasurement. Lastly, we

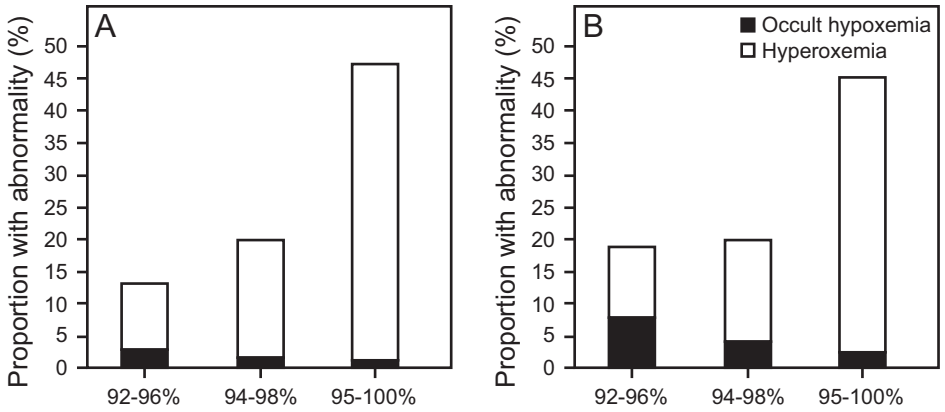

Fig. 4. Rates of occult hypoxemia, defined as an arterial oxygen saturation < 88%, and hyperoxemia, defined as partial pressure of arterial oxygen > 110 mm Hg on arterial blood gas, across 3 pulse oximeter oxygenation saturation ranges (92–96%, 94–98%, and 95–100%). A: White subjects, and B: Black subjects. Due to sample sizes, only Black and white subjects are included.

found that frequency of occult hypoxemia increased as measured S_{pO_2} decreased, suggesting that targeting a higher S_{pO_2} may reduce the occurrence of clinically unrecognized hypoxemia.

Our study adds to a growing body of literature that confirms that pulse oximeters frequently overestimate S_{aO_2}, especially for individuals with dark skin pigmentation.[4-7,13-15,19] Given the widespread use of pulse oximetry, technological advancements are urgently needed. However, until such time that technological advancements have been achieved and are accessible to centers who care for patients of color who are at greatest risk for occult hypoxemia, health systems must consider practice changes designed to mitigate these disparities. The prevailing mechanism used to explain the tendency of pulse oximeters to overestimate true S_{aO_2} is that darkly pigmented skin absorbs more light in red and infrared wavelengths, leading these devices to erroneously interpret increased light absorbance as greater oxygen saturation. From this framework, errors in oxygen saturation estimated by pulse oximeter would be anticipated to be positive (ie, greater than the true arterial saturation) and if consistently positive could perhaps be mitigated by assuming a lower S_{aO_2} than that displayed by pulse oximeter. In addition, one would expect that in a given patient the degree of pigmentation would impact all samples similarly and cause a consistent error. However, our study demonstrates that this is not the case. Rather, among all races, and especially at lower oxygen saturations, we found estimated bias in oxygen saturation by pulse oximeters was inconsistent and often erroneous in both positive and negative direction. As a result, beyond melanin, other factors may play a role, including poor perfusion, motion degradation, and poorly fitting or functioning pulse oximeter devices.[19] Whereas measurement error related to dark skin pigmentation may be a major driver of disparities, future mechanistic studies designed to examine the aforementioned factors, and the potential interaction between such factors and skin pigment, are needed to inform and aid necessary technological advances. Further, we encourage future studies to consider how disparities in the quality of oxygen monitoring, from clinician selection of appropriately fitting oxygen sensors to clinician monitoring of pulse oximeter waveform quality, may be influenced by patient race.

The inconsistencies in magnitude and direction of errors in pulse oximetry preclude an individualized approach to setting oxygenation targets. Our study reveals that the measurement error over the course of a hospitalization was not a fixed one, and therefore, an assessment of error observed in a single paired sample cannot be applied as a correction factor for subsequent S_{pO_2} measurements.

In contrast, we found that raising the oxygen saturation range from 92–96% to 94–98% was associated with a lower frequency of occult hypoxemia with only a modest increase in hyperoxemia. Although further elevation of oxygen saturation range to 95–100% further reduced the frequency of occult hypoxemia, this incremental benefit appeared to be outweighed by the increase in hyperoxemia. Of note, raising oxygenation range to 94–98% mitigated risk of occult hypoxemia in Black subjects but did not eliminate it.

Based on our analyses, until more accurate pulse oximeters are widely available, we recommend that targeting a pulse oximetry saturation range of 94–98% be the preferred mitigation strategy to reduce the prevalence of occult hypoxemia in patients who are now being targeted to an oxygen saturation goal of 92–96% and have adopted such a strategy within our health system. This recommendation aligns with established guideline recommendations[18] and may reduce prevalence of occult hypoxemia independent of race, with only a modest increase in risk of exposure to hyperoxemia. Of note, these recommendations do not apply to patients currently targeted to a goal oxygen saturation 88–92% (eg, those with chronic lung disease) since

systematically raising F_{IO_2} in these patients may cause acute hypercapnia. Based on our results, we discourage clinicians from the practice of relying on episodic ABG results performed as a mitigation strategy, which could potentially lead to differential rates of ABG sampling based on race, with unclear benefit.

Although recent trials[20] suggest that a modest increase in the oxygen saturation goal to the level proposed herein will not prolong duration of mechanical ventilation, we acknowledge that further study is necessary to directly assess this and other potential consequences of adopting the strategy we suggest. Likewise, we acknowledge that raising the goal oxygen saturation threshold modestly increases the risk of hyperoxemia. Although recent studies[20-22] suggest risks associated with hyperoxia may have been overstated, further study is warranted.

Our study further highlights the urgent need for more accurate pulse oximeter technology. Whereas approval of these devices relies on internal validation of these devices by medical device companies themselves,[2] disparities related to pulse oximeter measurement error are unlikely to be fully mitigated until such time as better technologies are available. In addition, conditions during validation studies tend to be idealized simulations of clinical stability that fail to account for factors that can lead to fluctuations in patient oxygen saturations.[2,3,19,23-25]

The racial disparities we observed in this study are likely a result of inadequate requirements for accounting for skin pigment variation in the device approval process. The FDA requires validation studies for new pulse oximeters to enroll a minimum of 10 subjects and at least 2 of these subjects be patients with dark skin pigmentation. If sample sizes are larger than 13, at least 15% of enrolled patients must have dark skin pigmentation.[26] Whereas enrolling patients with dark skin color is a necessary step to address disparities, it is insufficient if race and ethnicity considerations are not incorporated into a rigorous analytical plan designed to assess for accuracy and performance in both idealized and dynamic clinical conditions. We call upon medical device manufacturers and federal regulatory bodies to develop new standards for pulse oximeter measurement devices that explicitly take patient racial and ethnic background into account during device development, recruitment, data analysis, and results reporting to avoid disparities such as these in the future. A statement that acknowledges this limitation of pulse oximeters is not enough.[27]

Our study has limitations. As a retrospective observational study, we are unable to confirm that our recommended strategy will mitigate disparities. Empirically derived, confirmatory studies designed to test its effectiveness are needed. In addition, we acknowledge as a limitation that we used self-reported race as a proxy for skin color and did not assess important patient characteristics, including patient skin pigment, presence of vascular disease and other comorbid conditions, body mass index, tissue perfusion, and other underlying disease that might be prevalent in different proportions among subjects of different races. Additionally, because fluctuations in pulse oximetry readings may be due to periodic breathing in patients who are hypoxic, it is possible that at least some of the variability in measurement error was due to this phenomenon rather than to the pulse oximeter sensor. Prospective studies, designed to account for these potential confounders in addition to subject motion, ambient light, and fingernail polish,[3] are warranted. Although we performed our study in 2 hospitals with different patient populations using different pulse oximeters, our study was done in a single health system and may not generalize to centers that use different oximeters. However, given the consistency of results across multiple studies, including one[19] that included reusable finger clips, disposable adhesive finger, and disposable adhesive forehead sensors, this is unlikely. While skin tone differences are likely to account for disproportional errors in pulse oximetry, further study is required. Further, confirmatory studies powered to permit robust analyses across all racial and ethnic minority groups including Indigenous, Asian, Native Hawaiian and other Pacific Islanders, and Hispanic peoples, are warranted. Finally, we could not assess the clinical impact of occult hypoxemic events. However, recent work[14,15] revealing an association between unrecognized hypoxemic episodes and poor clinical outcomes substantiated our intuition and supports the notion that strategies to mitigate risk of occult hypoxemia are necessary in advance of technological advances.

Conclusions

In this retrospective observational study conducted across 2 hospitals, we found that occult hypoxemia was more common among all subjects of color and Black subjects in particular. Our data provide evidence for hospitals to implement an interim mitigation strategy of raising oxygen saturation goals to 94–98% for all patients, as this should have a favorable impact on reducing hypoxemic events without significantly increasing exposure to hyperoxemia. As this strategy does not eliminate disparity in occult hypoxemia experienced by Black patients, additional studies and innovations in pulse oximetry technology are urgently needed to eliminate risk across all patients.

REFERENCES

1. Sinex JE. Pulse oximetry: principles and limitations. Am J Emerg Med 1999;17(1):59-67.
2. Batchelder PB, Raley DM. Maximizing the laboratory setting for testing devices and understanding statistical output in pulse oximetry. Anesth Analg 2007;105(6 Suppl):S85-S94.
3. Louie A, Feiner JR, Bickler PE, Rhodes L, Bernstein M, Lucero J. Four types of pulse oximeters accurately detect hypoxia during low perfusion and motion. Anesthesiology 2018;128(3):520-530.

4. Jubran A, Tobin MJ. Reliability of pulse oximetry in titrating supplemental oxygen therapy in ventilator-dependent patients. Chest 1990;97 (6):1420-1425.
5. Emery JR. Skin pigmentation as an influence on the accuracy of pulse oximetry. J Perinatol 1987;7(4):329-330.
6. Bickler PE, Feiner JR, Severinghaus JW. Effects of skin pigmentation on pulse oximeter accuracy at low saturation. Anesthesiology 2005;102 (4):715-719.
7. Feiner JR, Severinghaus JW, Bickler PE. Dark skin decreases the accuracy of pulse oximeters at low oxygen saturation: the effects of oximeter probe type and gender. Anesth Analg 2007;105(6 Suppl): S18-S23.
8. Bothma PA, Joynt GM, Lipman J, Hon H, Mathala B, Scribante J, et al. Accuracy of pulse oximetry in pigmented patients. S Afr Med J 1996;86(5 Suppl):594-596.
9. 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(10):965-970.
10. Zeballos RJ, Weisman IM. Reliability of noninvasive oximetry in black subjects during exercise and hypoxia. Am Rev Respir Dis 1991;144(6):1240-1244.
11. Ries AL, Prewitt LM, Johnson JJ. Skin color and ear oximetry. Chest 1989;96(2):287-290.
12. Foglia EE, Whyte RK, Chaudhary A, Mott A, Chen J, Propert KJ, Schmidt B. The effect of skin pigmentation on the accuracy of pulse oximetry in infants with hypoxemia. J Pediatr 2017;182:375-377. e372.
13. Sjoding MW, Dickson RP, Iwashyna TJ, Gay SE, Valley TS. Racial bias in pulse oximetry measurement. N Engl J Med 2020;383 (25):2477-2478.
14. Henry NR, Hanson AC, Schulte PJ, Warner NS, Manento MN, Weister TJ, Warner MA. Disparities in hypoxemia detection by pulse oximetry across self-identified racial groups and associations with clinical outcomes. Crit Care Med 2022;50(2):204-211.
15. Wong A-KI, Charpignon M, Kim H, Josef C, de Hond AAH, Fojas JJ, et al. Analysis of discrepancies between pulse oximetry and arterial oxygen saturation measurements by race and ethnicity and association with organ dysfunction and mortality. JAMA Netw Open 2021;4(11): e2131674.
16. Mikkelsen ME, Anderson B, Christie JD, Hopkins RO, Lanken PN. Can we optimize long-term outcomes in acute respiratory distress syndrome by targeting normoxemia? Ann Am Thorac Soc 2014;11(4):613-618.
17. Flanagin A, Frey T, Christiansen SL; for the AMA Manual of Style Committee. Updated guidance on the reporting of race and ethnicity in medical and science journals. JAMA 2021;326(7):621-627.
18. O'Driscoll BR, Howard LS, Earis J, Mak V; British Thoracic Society Emergency Oxygen Guideline Group, BTS Emergency Oxygen Guideline Development Group. BTS guideline for oxygen use in adults in health care and emergency settings. Thorax 2017;72(Suppl 1):ii1-ii90.
19. 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(2):252-257.
20. Schjorring OL, Klitgaard TL, Perner A, Wetterslev J, Lange T, Siegemund M, et al; HOT-ICU Investigators. Lower or higher oxygenation targets for acute hypoxemic respiratory failure. N Engl J Med 2021;384(14):1301-1311.
21. Barrot L, Asfar P, Mauny F, Winiszewski H, Montini F, Badie J, et al; LOCO2 Investigators and REVA Research Network. Liberal or conservative oxygen therapy for acute respiratory distress syndrome. N Engl J Med 2020;382(11):999-1008.
22. Mackle D, Bellomo R, Bailey M, Beasley R, Deane A, Eastwood G, et al; ICU-ROX Investigators and the Australian and New Zealand Intensive Care Society Clinical Trials Group. Conservative oxygen therapy during mechanical ventilation in the ICU. N Engl J Med 2020;382(11):989-998.
23. Hinkelbein J, Genzwuerker HV, Sogl R, Fiedler F. Effect of nail polish on oxygen saturation determined by pulse oximetry in critically ill patients. Resuscitation 2007;72(1):82-91.
24. Ralston AC, Webb RK, Runciman WB. Potential errors in pulse oximetry. I. pulse oximeter evaluation. Anaesthesia 1991;46(3):202-206.
25. Webb RK, Ralston AC, Runciman WB. Potential errors in pulse oximetry. II. Effects of changes in saturation and signal quality. Anaesthesia 1991;46(3):207-212.
26. Food and Drug Administration - Center for Devices and Radiological Health. Pulse Oximeters - Premarket Notification Submissions [510(k) s] Guidance for Industry and Food and Drug Administration Staff. March 2013.
27. Food and Drug Administration. Pulse Oximeter Accuracy and Limitations: FDA Safety Communication. February 19, 2021.