Differential Pulse Oximeter Accuracy, Occult Hypoxemia Prevalence, and Clinical Outcomes by Patient Race/Ethnicity

August 2023

U.S. Department of Veterans Affairs
Veterans Health Administration
Health Services Research & Development Service

Recommended citation: Parr NJ, Beech EH, Young S. Differential Pulse Oximeter Accuracy, Occult Hypoxemia Prevalence, and Clinical Outcomes by Patient Race/Ethnicity: A Systematic Review. Washington, DC: Evidence Synthesis Program, Health Services Research and Development Service, Office of Research and Development, Department of Veterans Affairs. VA ESP Project #09-199; 2023.

AUTHORS

Author roles, affiliations, and contributions (using the CRediT taxonomy) are listed below.

Author	Role and Affiliation	Report Contribution
Nicholas J. Parr, PhD, MPH	Associate Director & Research Scientist, ESP Coordinating Center, Portland VA Health Care System Portland, OR	Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Writing – original draft, Writing – review & editing, Project administration, Supervision
Erin H. Beech, MA	Senior Research Associate, ESP Coordinating Center, Portland VA Health Care System Portland, OR	Methodology, Investigation, Data curation, Writing – review & editing, Project administration, Supervision
Sarah Young, MPH	Research Associate, ESP Coordinating Center, Portland VA Health Care System Portland, OR	Data curation

PREFACE

The VA Evidence Synthesis Program (ESP) was established in 2007 to conduct timely, rigorous, and independent systematic reviews to support VA clinicians, program leadership, and policymakers improve the health of Veterans. ESP reviews have been used to develop evidence-informed clinical policies, practice guidelines, and performance measures; to guide implementation of programs and services that improve Veterans' health and wellbeing; and to set the direction of research to close important evidence gaps. Four ESP Centers are located across the US. Centers are led by recognized experts in evidence synthesis, often with roles as practicing VA clinicians. The Coordinating Center, located in Portland, Oregon, manages program operations, ensures methodological consistency and quality of products, engages with stakeholders, and addresses urgent evidence synthesis needs.

Nominations of review topics are solicited several times each year and submitted via the ESP website. Topics are selected based on the availability of relevant evidence and the likelihood that a review on the topic would be feasible and have broad utility across the VA system. If selected, topics are refined with input from Operational Partners (below), ESP staff, and additional subject matter experts. Draft ESP reviews undergo external peer review to ensure they are methodologically sound, unbiased, and include all important evidence on the topic. Peer reviewers must disclose any relevant financial or non-financial conflicts of interest. In seeking broad expertise and perspectives during review development, conflicting viewpoints are common and often result in productive scientific discourse that improves the relevance and rigor of the review. The ESP works to balance divergent views and to manage or mitigate potential conflicts of interest.

ACKNOWLEDGMENTS

The authors are grateful to Becky Baltich Nelson, MLS, MS, for literature searching, Payten Sonnen for editorial and citation management support, external peer reviewers, and the following individuals for their contributions to this project:

Operational Partners

Operational partners are system-level stakeholders who help ensure relevance of the review topic to the VA, contribute to the development of and approve final project scope and timeframe for completion, provide feedback on the draft report, and provide consultation on strategies for dissemination of the report to the field and relevant groups.

John W. Nord, MD, MSc
Acting Deputy Chief Officer
Veterans Health Administration (VHA) Office of Specialty Care Services

Joseph Francis, MD, MPH
Executive Director of Analytics and Performance Integration
VHA

Mel L. Anderson, MD, MACP
Executive Director
VHA National Hospital Medicine Program

Thomas S. Valley, MD
Research Scientist and Staff Physician
Center for Clinical Management Research (VA Ann Arbor Healthcare System)

Disclosures

This report was prepared by the Evidence Synthesis Program Center located at the **VA Portland Health Care System,** directed by Mark Helfand, MD, MPH, MS, and funded by the Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development.

The findings and conclusions in this document are those of the author(s) who are responsible for its contents and do not necessarily represent the views of the Department of Veterans Affairs or the United States government. Therefore, no statement in this article should be construed as an official position of the Department of Veterans Affairs. The final research questions, methodology, and/or conclusions may not necessarily represent the views of contributing operational and content experts. No investigators have affiliations or financial involvement (*eg*, employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties) that conflict with material presented in the report.

Main Report

Evidence Synthesis Program

TABLE OF CONTENTS

Background 4
Methods 5
Registration and Review 5
Key Questions and Eligibility Criteria 5
Searching and Screening 5
Data Abstraction and Risk of Bias Assessment 5
Synthesis 6
Results 8
Literature Flow Diagram 8
Overview of Included Studies 9
Main Findings 9
Table. Pulse Oximeter Accuracy, Occult Hypoxemia Prevalence, and Clinical Outcomes by Patient Race/Ethnicity 10
Discussion 15
Conclusions 16
References 17
Appendix 21

ABBREVIATIONS TABLE

aHR	Adjusted hazard ratio
aOR	Adjusted odds ratio
A_{RMS}	Accuracy root mean square error
CI	Confidence interval
FDA	US Food and Drug Administration
IQR	Interquartile range
k	Number of studies
PI	Prediction interval
PR	Prevalence ratio
N	Number of patients
N_{Obs}	Number of paired observations
SaO_2	Arterial oxygen saturation
SD	Standard deviation
SpO_2	Peripheral oxygen saturation measured by pulse oximetry
VHA	Veterans Health Administration

BACKGROUND

Pulse oximeters are used in many clinical settings and provide a rapid and noninvasive means of measuring oxygen saturation. Despite this utility, pulse oximeters may over- or underestimate a patient's arterial oxygen saturation (SaO_2). Overestimating oxygen saturation is especially concerning when pulse oximeter readings of peripheral oxygen saturation (SpO_2) indicate a normal blood oxygen level while a patient is actually in a hypoxemic state—a situation known as *occult hypoxemia*. Potential clinical impacts of occult or undetected hypoxemia include delayed or inadequate treatment, premature treatment de-escalation or discharge, and ultimately, greater morbidity and mortality.[1,2]

Inaccurate pulse oximeter readings can occur for a variety of reasons, including severe anemia, excessive blood carbon monoxide levels, impaired circulation (hypoperfusion), and patient movement.[3,4] Because pulse oximeters rely on the transmission of light through the skin to estimate SaO_2, skin pigmentation level may also influence pulse oximeter accuracy. Differences in pulse oximeter accuracy by patient race/ethnicity have been observed in clinical settings for several decades, but the COVID-19 pandemic has heightened concern that pulse oximeters may routinely be less accurate in patients with darker skin pigmentation. In the United States, the Food and Drug Administration (FDA) regulates pulse oximeter accuracy and recommends that pulse oximeter performance be within 3% of CO-oximetry in arterial blood (as assessed with accuracy root mean square error, or A_{RMS}).[5] A recent synthesis of pulse oximeter accuracy studies published prior to mid-2021 by Shi et al[6] reported a somewhat larger pulse oximeter A_{RMS} among Black or African American ("Black") patients compared with White patients (2.3% versus 1.6%).

The COVID-19 pandemic has also drawn attention to whether racial or ethnic minority patients are at greater risk of occult hypoxemia due to pulse oximeter inaccuracy. A widely discussed retrospective study[7] published in late 2020 analyzed nearly 50,000 paired SpO_2–SaO_2 measurements, finding that the prevalence of occult hypoxemia was over 3 times greater among Black patients compared with White patients (11.4% versus 3.4%). In late 2022, the FDA convened a public meeting of its Anesthesiology and Respiratory Therapy Devices Panel in response to ongoing concerns about racial biases in pulse oximeter accuracy and unrecognized hypoxemia. Guidance on the use and interpretation of pulse oximeters was subsequently updated and emphasizes that pulse oximeter readings are *estimates* of oxygen saturation that provide "more utility for trends over time instead of absolute thresholds" at the individual patient level.[8]

A number of studies of racial or ethnic disparities in occult hypoxemia prevalence have been recently published. Most utilize large health system databases and also report data on pulse oximeter accuracy, which goes some way to addressing the sparsity of accuracy data in racial and ethnic minority patients in the pre-COVID-19 literature reviewed by Shi et al. The aim of the present review was to provide an up-to-date synthesis of evidence on racial and ethnic disparities in the accuracy of pulse oximeters, the prevalence of occult hypoxemia, and clinical outcomes associated with occult hypoxemia. This review was developed in response to a request from the Veterans Health Administration (VHA) National Hospital Medicine Program and Office of Specialty Care Services.

METHODS

REGISTRATION AND REVIEW

A preregistered protocol for this review can be found on the PROSPERO register of systematic reviews (CRD42023402152). A draft version of this report was reviewed by external peer reviewers; their comments and author responses are located in the Appendix.

KEY QUESTIONS AND ELIGIBILITY CRITERIA

The following key questions were the focus of this review:

Key Question 1	Does pulse oximeter accuracy differ by patient race/ethnicity?
Key Question 2	Does the prevalence or risk of occult hypoxemia differ by patient race/ethnicity?
Key Question 2b	If present, are racial/ethnic disparities in occult hypoxemia associated with differences in treatment delivery or harms?

Eligible studies must have been conducted among adults in inpatient or outpatient healthcare settings. Accuracy studies were required to report paired SpO_2–SaO_2 readings measured within 10 minutes of one another. Studies reporting occult hypoxemia prevalence were required to define occult hypoxemia as, at minimum, arterial oxygen saturation ≤ 88% despite a pulse oximeter reading > 88%. Stricter criteria (*eg*, pulse oximeter reading > 92%) were permitted. Studies that induced hypoxemia in a controlled setting were ineligible. We considered evidence on clinical outcomes (treatment delivery or harms) of occult hypoxemia only when studies reported within-group comparisons by occult hypoxemia status. For example, we included studies that examined whether there was a relationship between occult hypoxemia and in-hospital mortality by comparing patients of the same race or ethnicity with and without occult hypoxemia (then examined whether this relationship differed across race/ethnicity groups), but we did not include studies that investigated whether in-hospital mortality differed by patient race or ethnicity in general.

SEARCHING AND SCREENING

A research librarian searched Ovid MEDLINE, CINAHL, Scopus, the Cochrane Database of Systematic Reviews, and AHRQ and HSR&D databases for relevant studies published through February 2023 (see Appendix for complete search strategies). Additional citations were identified by hand-searching reference lists. English-language titles, abstracts, and full-text articles were independently reviewed by 2 investigators, and disagreements were resolved by consensus.

DATA ABSTRACTION AND RISK OF BIAS ASSESSMENT

Participant characteristics, study methodological details, and outcome data were abstracted from all included studies. The internal validity (risk of bias) of included studies was rated using tools appropriate to each study type: the ROBIS[9] tool for systematic reviews, the QUADAS-2[10] tool for studies providing accuracy data, and the QUIPS[11] tool for studies of occult hypoxemia prevalence or occult hypoxemia as a risk factor for downstream clinical outcomes. Potential biases in prevalence estimates were captured in the *prognostic factor measurement* domain of the QUIPS tool; we supplemented criteria for this domain with biases specific to prevalence assessment.[12] Risk of bias rating and data abstraction were first completed by 1 investigator then checked by another, and disagreements were resolved by consensus. See Appendix for complete risk of bias ratings.

SYNTHESIS

When 3 or more sufficiently comparable studies were available, study findings were pooled using random-effects meta-analysis. We identified a well-conducted systematic review and meta-analysis of pulse oximeter accuracy by Shi et al,[6] which synthesized accuracy data available through mid-2021. To expedite the present review, we took the approach of updating this review with the most recent evidence available. This involved pooling mean bias and precision data from studies published since the end search date of the review (replicating the methods used by Shi et al), then synthesizing the pooled estimates from new evidence with those reported in the existing review. Each of these steps is described in detail below. To harmonize risk of bias ratings, we also re-assessed risk of bias of studies included in the existing systematic review using the approach described above.

For newly identified studies, we first adjusted reported standard deviations (SDs) for each race/ethnicity group in each study using the formula shown below.[13] The purpose of this adjustment, which was also implemented by Shi et al, was to produce more conservative estimates of precision that account for repeated observations of the same patient. The number of paired oximetry measurements per patient was calculated by dividing the total number of oximetry measurements by the number of patients (*eg*, in a group of 50 patients in which 500 paired measurements were collected, the approximate number of measurements per patient would be 500/50 = 10 measurements).

$$Adjusted\ SD = \sqrt{Reported\ SD^2 \times \left(\frac{Total\ Measurements - 1}{Total\ Measurements - Measurements\ per\ Patient}\right)}$$

We also replicated the use of hierarchical random-effects models, cluster-robust confidence intervals, and degrees of freedom calculated using the Satterwaithe approximation. These approaches better account for dependency among estimates; that is, when estimates from the same study are more similar than they would be if they were from a different study and patient sample. The final stage of analysis of mean bias and precision data was to synthesize pooled estimates reported by Shi et al with those from new studies. We implemented the recommendation of Tang et al[14] for updating a meta-analysis with new results based on comparable methods, which was to use a fixed-effect model to synthesize the final pooled estimates (*ie*, a pooled estimate corresponding to all evidence prior to September 2021 from Shi et al, and a pooled estimate of evidence from September 2021 through February 2023).

To synthesize occult hypoxemia prevalence estimates, we used meta-analytic generalized random-effects logistic models or hierarchical approximations when dependent prevalence estimates were available. Prevalence estimates were transformed using the standard logit transformation for analysis, and back-transformed for interpretation and reporting. To facilitate comparison of occult hypoxemia prevalences between race/ethnicity groups, we also calculated and synthesized prevalence ratios (PRs). Conventional random-effects models were used to pool PRs[i] as well as adjusted odds ratios (aORs) of occult hypoxemia by race/ethnicity group. Models incorporated the Knapp-Hartung adjustment to

[i] We chose standard random-effects models over more complex options (*eg*, log-binomial models) for prevalence ratios because: 1) overall, occult hypoxemia was fairly rare, and therefore logarithms of prevalence ratios and odds ratios would be expected to have similar analytic properties; 2) no studies reported zero occult hypoxemia events, meaning we were not required to use continuity corrections to carry out standard random-effects modeling (analysis in "one step" using log-binomial or similar models would be preferable in such a scenario); and 3) the standard approach was more compatible with methods to account for dependency in reported estimates.

better account for uncertainty in heterogeneity estimation,[15,16] and when necessary, we used methods for handling dependent data comparable to those described above.

For all analyses, heterogeneity was estimated using (restricted) maximum-likelihood estimation and is presented as 95% prediction intervals (PIs). Prediction intervals describe the likeliest range of true effects (*eg*, true occult hypoxemia prevalence) across studies and provide an estimate of the magnitude and direction of effects that would be found in future studies similar to those included in a synthesis.[17] All meta-analyses were conducted using the *metafor*[18] package for R (R Foundation for Statistical Computing, Vienna, Austria). When fewer than 3 comparable studies were available for a given outcome—or studies were judged to be too disparate in methodological or participant characteristics—we described evidence narratively.

For the purposes of synthesizing and reporting findings, the consensus reached among investigators was that the following categorizations best balanced data availability with variation in race/ethnicity groups reported across studies: Patients described as Black or African American ("Black"), patients described as White or Caucasian ("White"), and patients described as Asian, Latino or Hispanic, Native American or Indigenous, or other or mixed race/ethnicity. Where possible, we disaggregated findings for the latter group by specific race or ethnicity identities. Studies also varied in whether prevalence of occult hypoxemia was reported by patients or by paired observations. We synthesized each type of prevalence estimate separately. Finally, for consistency with existing literature on pulse oximeter accuracy, we also report A_{RMS} values calculated from pooled mean bias and precision results. It is important to note, however, that interpretation of A_{RMS} is not straightforward: A value between 2–3%, for example, equates to roughly 66% of SpO_2 readings being within 2–3% of corresponding SaO_2 readings and 95% being within 4–6% of SaO_2 readings.[8]

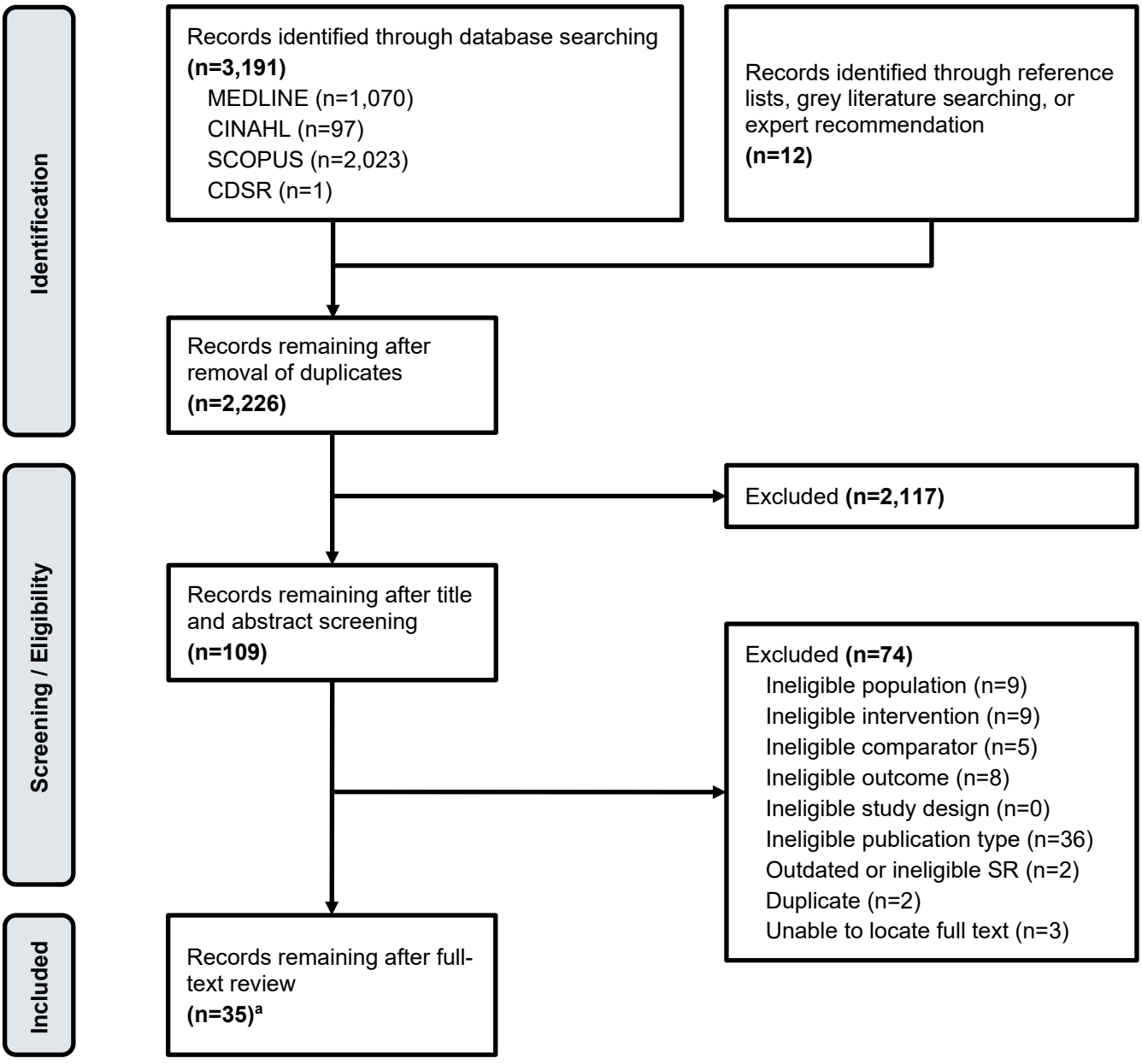
Strength of Evidence

After synthesizing available evidence, we rated the strength of evidence for each outcome based on the methodology and risk of bias of available studies, the consistency and certainty of findings, and the directness of outcomes (whether reported outcomes are relevant to patients and providers). We used the following general algorithm: *high strength* evidence consisted of multiple, large studies with consistent and precise findings at low risk of bias, and clinically relevant outcomes; *moderate strength* evidence consisted of multiple studies with consistent and precise findings at low to moderate risk of bias, and clinically relevant outcomes; *low strength* evidence consisted of a single study, or multiple small studies, with moderate to high risk of bias, inconsistent or imprecise findings, and/or outcomes with limited clinical relevance; and *insufficient* evidence consisted of a single study with moderate or high risk of bias, or no available studies. Conclusions using *likely* (*eg*, "Pulse oximeters likely overestimate Black patients' blood oxygen saturation level") are based on moderate strength evidence, while those using *may* are based on low strength evidence.

RESULTS

LITERATURE FLOW DIAGRAM

The literature flow diagram summarizes the results of the study selection process. A full list of excluded studies is provided in the Appendix.



Notes. [a]34 primary studies and 1 systematic review.

Abbreviations. CDSR=Cochrane Database of Systematic Reviews; CINAHL=Cumulative Index to Nursing and Allied Health Literature; SR=systematic review.

OVERVIEW OF INCLUDED STUDIES

Our search identified 109 potentially relevant articles after deduplication and title and abstract screening. Of these, 34 primary studies and 1 existing systematic review and meta-analysis[6] on pulse oximeter accuracy met eligibility criteria. Eighteen studies[19-36] were included in the identified systematic review, which together provided 21,269 paired observations from 3,176 patients. These studies were published between 1985 and 2021 and varied considerably in setting and patient characteristics (details of all included studies are provided in the Appendix). We found 2 additional accuracy studies published in the 1980s[37,38] but did not include these in formal syntheses because they used pulse oximeters that are no longer commercially available.

We located 4 recent observational studies[3,4,39,40] that reported pulse oximeter accuracy data in sufficient detail for meta-analyses; together these studies contribute 241,680 new paired observations in 102,841 patients. Four other accuracy studies[36,41-43] did not provide adequate outcome data or information about race/ethnicity groups to be included in formal syntheses. The 4 studies contributing accuracy data noted above also reported occult hypoxemia prevalence; in total, we included 11 observational studies[1-4,7,35,39,40,44-46] that reported occult hypoxemia prevalence and 4 studies[1,2,39,46] that examined the association between occult hypoxemia and clinical outcomes by patient race/ethnicity. Most newly identified studies used data from patients receiving acute care in academic or community hospitals or health systems. One study was limited to patients undergoing anesthesia[4] and 2 studies included surgical inpatients.[2,40] Patient race/ethnicity was generally self-reported.

MAIN FINDINGS

Pulse Oximeter Accuracy

Pooled mean bias and precision findings by race/ethnicity groups are shown in the Table below. In Black patients, pulse oximeters appear to overestimate blood oxygen saturation by an average of 1.5% compared with CO-oximetry in arterial blood (pooled mean bias = 1.54, 95% CI [0.99, 2.10]), based on 37,562 paired observations in 14,626 patients. Mean bias among Black patients was considerably larger than among White patients (0.62, 95% CI [-0.08, 1.32]; N_{Obs} = 154,286) or patients identifying as Asian, Latino or Hispanic, Native American or Indigenous, or other or mixed race/ethnicity (0.31, 95% CI [0.09, 0.54]; N_{Obs} = 71,101). Precision of pulse oximeter readings was comparable across race/ethnicity groups, with pooled SDs ranging from 1.61 to 1.98.[ii] In recent studies that contribute most available oximetry data, mean bias was larger in all race/ethnicity groups compared with earlier evidence. The most substantial increase from older studies was among Asian, Latino or Hispanic, Native American or Indigenous, or other or mixed race/ethnicity patients (from 0.31% to 1.41%). Precision estimates were comparable across groups but considerably larger than in earlier studies (pooled SDs from 4.30 to 5.23), indicating greater variability in bias between patients.

The calculated A_{RMS} for all accuracy data was 1.64 for Black patients, 0.78 for White patients, and 0.75 for Asian, Latino or Hispanic, Native American or Indigenous, or other or mixed race/ethnicity patients. A_{RMS} values in newly identified studies were much larger: 2.57, 1.83, and 2.03, respectively.

[ii] Shi et al also pooled studies that reported accuracy by patient pigmentation level (high pigmentation, medium pigmentation, or low pigmentation). Pooled mean bias was largest among patients with high pigmentation (1.11, 95% CI [0.29, 1.93]). Pooled SDs were similar across pigmentation levels and comparable to findings in race/ethnicity groups (1.47 to 1.52).

Table. Pulse Oximeter Accuracy, Occult Hypoxemia Prevalence, and Clinical Outcomes by Patient Race/Ethnicity

Patients Described as Black or African American				
Accuracy	**Patients**	**Paired Obs.**	**Mean Bias [95% CI]**	**Precision [95% CI]**
Shi 2022 Review (*k* = 9)	459	5,753	1.52 [0.95, 2.09]	1.68 [1.32, 2.14]
Recent Studies (*k* = 4)	14,167	31,809	1.96 [-0.57, 4.50]	5.23 [1.50, 18.28]
Updated Estimate (*k* = 13)	14,626	37,562	1.54 [0.99, 2.10]	1.75 [1.38, 2.22]
Occult Hypoxemia Prevalence			**% [95% CI]**	**[95% PI]**
Patients (*k* = 5)	34,869		11.4 [4.6, 25.5]	[1.2, 58.1]
Paired Observations (*k* = 11)		76,177	6.9 [3.9, 11.9]	[0.9, 36.9]
Clinical Outcomes				
Eligibility Recognition (*k* = 1)	928		aHR: 0.71 [0.63, 0.80] (ref: White)	
Treatment Delay (*k* = 1)	681		Median (IQR): 7.0 h (1.9–20.8)	
Length of Stay (*k* = 1)	26,032		Difference: -3.0 d (*p* = .00)	
In-hospital Mortality (*k* = 1)	26,032		Difference: +5.9% (*p* < .001)	
Patients Described as Asian, Latino or Hispanic, Native American or Indigenous, or Other				
Accuracy	**Patients**	**Paired Obs.**	**Mean Bias [95% CI]**	**Precision [95% CI]**
Shi 2022 Review (*k* = 3)	522	2,646	0.31 [0.09, 0.53]	1.55 [0.53, 4.53]
Recent Studies (*k* = 3)	23,001	68,455	1.41 [-2.65, 5.47]	4.30 [0.63, 29.17]
Updated Estimate (*k* = 6)	23,523	71,101	0.31 [0.09, 0.54]	1.98 [0.78, 5.04]
Occult Hypoxemia Prevalence			**% [95% CI]**	**[95% PI]**
Patients (*k* = 5)	25,130		9.7 [3.2, 26.1]	[0.6, 66.6]
Paired Observations (*k* = 6)		89,243	5.3 [2.0, 13.0]	[0.4, 42.8]
Clinical Outcomes				
Eligibility Recognition (*k* = 1)	Asian: 25; Hispanic: 445		aHR: 0.97 [0.62, 1.50]; 0.77 [0.66, 0.89] (ref: White)	
Treatment Delay (*k* = 1)	Asian: 21; Hispanic: 323		Median (IQR): 7.7 h (3.5–13.6); 5.0 h (1.2–15.8)	
Length of Stay (*k* = 1)	Asian: 1,919; Hispanic: 2,397		Difference: +0.5 d (*p* = .02); +0.8 d (*p* < .01)	
In-hospital Mortality (*k* = 1)	Asian: 1,919; Hispanic: 2,397		Difference: +5.6% (*p* = .13); +4.6% (*p* = .06)	
Patients Described as White or Caucasian				
Accuracy	**Patients**	**Paired Obs.**	**Mean Bias [95% CI]**	**Precision [95% CI]**
Shi 2022 Review (*k* = 13)	2,195	12,870	0.55 [-0.21, 1.31]	1.55 [1.32, 1.83]
Recent Studies (*k* = 4)	65,673	141,416	1.01 [-0.78, 2.79]	4.63 [1.89, 11.37]
Updated Estimate (*k* = 17)	67,868	154,286	0.62 [-0.08, 1.32]	1.61 [1.37, 1.89]
Occult Hypoxemia Prevalence			**% [95% CI]**	**[95% PI]**
Patients (*k* = 5)	109,286		6.5 [2.8, 14.2]	[0.8, 36.8]
Paired Observations (*k* = 11)		337,976	3.3 [1.8, 5.9]	[0.4, 22.0]
Clinical Outcomes				
Treatment Delay (*k* = 1)	427		Median (IQR): 5.3 h (1.4–15.2)	
Length of Stay (*k* = 1)	57,623		Difference: -0.5 d (*p* < .01)	
In-hospital Mortality (*k* = 1)	57,623		Difference: +11.1% (*p* < .001)	

Abbreviations. aHR=adjusted hazard ratio; CI=confidence interval; d=day; h=hour; IQR=interquartile range; k=number of studies; PI=prediction interval.

Although available studies of pulse oximeter accuracy generally use observational designs, we judged newly identified studies to be at moderate or low overall risk of biased findings. The most common concerns were unclear detail about the pulse oximeter or CO-oximeter devices used, lack of reporting on blinding (*ie*, whether pulse oximeter results were read without knowledge of SaO_2), and limitations arising from use of health record data, including unclear detail about patient characteristics. Shi et al found that excluding studies with high risk of bias did not substantively alter conclusions. No newly identified studies included in syntheses were rated at high risk of bias.

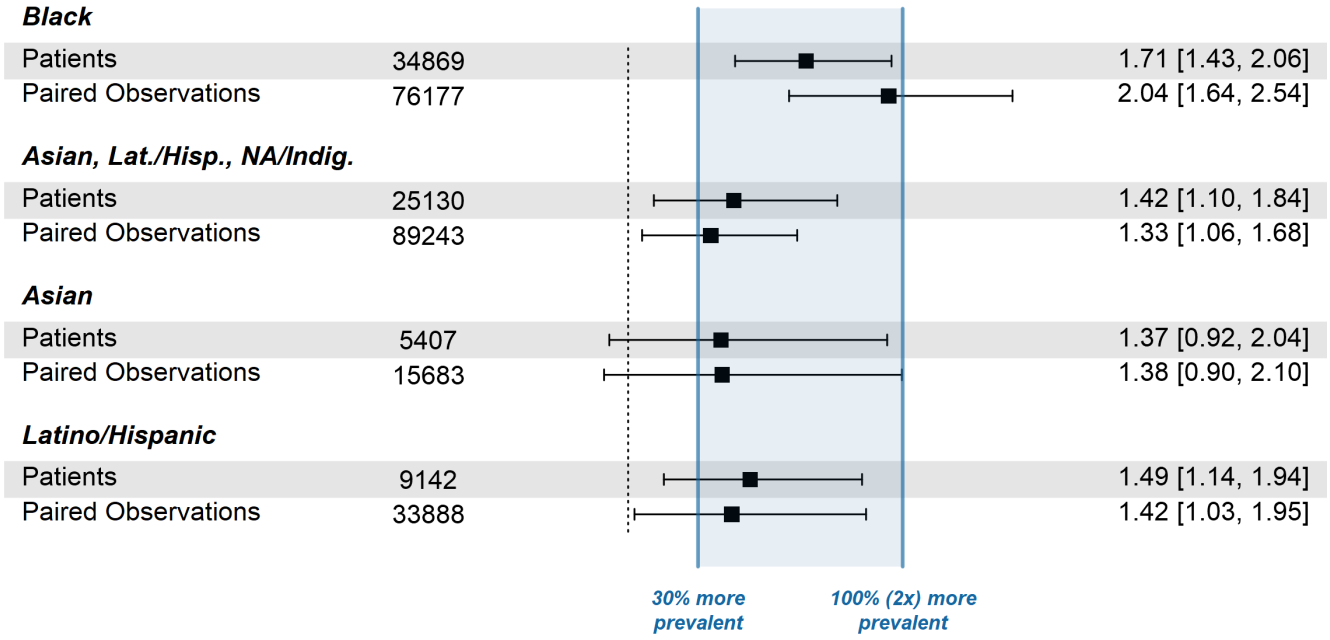
Evidence on differential pulse oximeter accuracy by patient race/ethnicity from studies published prior to mid-2021 was rated as low or very low strength (certainty) by Shi et al. Despite pooled mean bias and SD values being larger in recent studies, the substantial increase in available data and consistent magnitude of findings led us to increase the strength of evidence supporting the finding that pulse oximeters overestimate SaO_2 in Black patients to moderate. Although there were comparable increases in the amount of data available for other race/ethnicity groups, there was also much greater variability in paired readings. Given this, we did not upgrade the strength of evidence supporting the conclusion that pulse oximeters *do not* overestimate SaO_2 to a clinically important degree among Asian, Latino or Hispanic, Native American or Indigenous, or other or mixed race/ethnicity patients (from low certainty). The same conclusion was upgraded to low strength or certainty (from very low certainty) for White patients because of the substantial increase in the number of observations contributing to pooled estimates.

Occult Hypoxemia Prevalence

Occult hypoxemia prevalence by race/ethnicity groups is shown in the Table above. The pooled prevalence of occult hypoxemia was highest among Black patients (11.4%, 95% CI [4.6, 25.5]; N = 34,869) and Asian, Latino or Hispanic, Native American or Indigenous, or other or mixed race/ethnicity patients (9.7%, 95% CI [3.2, 26.1]; N = 25,130), compared with White patients (6.5%, 95% CI [2.8, 14.2]; N = 109,286). Corresponding prevalence ratios (PRs) are shown in the Figure below. Compared with White patients, the prevalence of occult hypoxemia was 71% greater among Black patients (pooled PR = 1.71, 95% CI [1.43, 2.06]) and 42% greater among Asian, Latino or Hispanic, Native American or Indigenous, or other or mixed race/ethnicity patients (pooled PR = 1.42, 95% CI [1.10, 1.84]). For Black patients, an even larger disparity was apparent at the observation level (pooled PR = 2.04, 95% CI [1.64, 2.54]).

Four studies[2,4,35,44] reported odds of occult hypoxemia adjusted for potential confounders (in most cases, patient demographics, comorbidities, and therapeutic variables such as use of vasopressors). When results were pooled, Black patients had twice the odds of experiencing occult hypoxemia compared with White patients (aOR = 1.99, 95% CI [1.15, 3.41]; N = 8,410). Reported odds ratios from all studies were similar in magnitude and consistent in direction. Three[2,4,44] of the 4 studies also included Asian, Latino or Hispanic, Native American or Indigenous, or other or mixed race/ethnicity patients, whose odds of hypoxemia were not significantly greater than those of White participants (aOR = 1.36, 95% CI [0.82, 2.25]; N = 20,536). Most odds ratios in this population were consistent in direction but differed in magnitude. Prevalence ratios and odds ratios followed similar patterns in each group.

Figure. Occult Hypoxemia Prevalence Ratios by Patient Race/Ethnicity



Notes. All prevalence ratios use White patients as the reference group, and the dashed line corresponds to a prevalence ratio of 1.0 (no difference in prevalence compared with White patients). *N* indicates the number of patients in the race/ethnicity group shown and does not include the number of White patients comprising the reference group. Patients reported as *Other race/ethnicity* are included in the group *Asian, Lat./Hisp., NA/Indig* (Asian, Latino or Hispanic, Native American or Indigenous).

Most studies providing data on occult hypoxemia prevalence were rated at moderate risk of biased findings for the same reasons as accuracy studies (given the use of accuracy data to define the occult hypoxemia outcome). One study[35] reporting odds of occult hypoxemia adjusted only for patient sex and measured SpO_2, compared with the other available studies which controlled for a more comprehensive set of potential confounders. Considering both prevalence and associational findings together, occult hypoxemia is likely more common among Black patients than among White patients (moderate strength of evidence). Asian, Latino or Hispanic, Native American or Indigenous, or other or mixed race/ethnicity patients may also experience occult hypoxemia more frequently than White patients, but the inconsistency of findings across studies and sparse data for some subgroups led us to rate evidence supporting this conclusion as low strength.

One recent study[40] used approximately 30,000 paired SpO_2–SaO_2 observations from patients treated in VHA surgical and general (non-intensive) care settings between 2013 and 2019. The study found larger bias and worse precision in Black VHA patients compared with White patients, and reported prevalences of occult hypoxemia (defined as SaO_2 less than 88% despite SpO_2 of 92% or greater) that were among the highest of any included study: 19.6% for Black patients, 16.2% for Hispanic or Latino patients, and 15.6% for White patients. The difference in likelihood of occult hypoxemia between Black patients and White patients remained after controlling for patient sex, age, and comorbidities ($p < .001$). Investigators also analyzed whether SpO_2 and SaO_2 values were consistent across pairs of patient measurements taken the same day. Black patients had a higher likelihood of occult hypoxemia

at a second oximetry reading than White patients, even when patients in both groups did not show occult hypoxemia on their first oximetry reading taken the same day (*ie*, Black patients' probability of occult hypoxemia was more variable among measurement instances than that of White patients). Data used in the study were drawn from the Corporate Data Warehouse, and in our appraisal, reasonable efforts were made to generate a dataset representative of the general VHA patient population and to limit selection biases that could reduce the generalizability of study findings.

Clinical Outcomes

We found few studies that examined the association between occult hypoxemia (or differential pulse oximeter accuracy) and clinical outcomes by patient race/ethnicity, and evidence was considered insufficient to make firm conclusions about this relationship. Nonetheless, there does appear to be some signal that Black patients with undetected hypoxemia could experience poorer treatment delivery and clinical outcomes than White patients with undetected hypoxemia.

One multicenter study[46] in 215 US hospitals and 315 intensive care units reported length of stay and in-hospital mortality by whether patients experienced occult or "hidden" hypoxemia (defined as SaO_2 less than 88% despite SpO_2 of 88% or greater in observations separated by no longer than 10 minutes). For Black patients ($N = 26{,}032$), occult hypoxemia was associated with significantly shorter length of stay compared with Black patients without occult hypoxemia (-3.0 days, $p < .01$). Length of stay for White patients ($N = 57{,}623$) with occult hypoxemia was also significantly shorter than White patients without occult hypoxemia, but by only 0.5 days on average ($p < .01$). Length of stay was significantly longer for Asian ($N = 1{,}919$) and Hispanic ($N = 2{,}397$) patients with occult hypoxemia, but by less than 1 day on average compared with patients without occult hypoxemia. In-hospital mortality was more common in patients experiencing occult hypoxemia regardless of race or ethnicity, with the largest difference among White patients (11.1% greater than White patients without occult hypoxemia, $p < .001$). Length of stay comparisons were unadjusted, while mortality comparisons were adjusted for patient age, sex, and Sequential Organ Failure Assessment (SOFA) score.

A second study,[2] which used health record data from 26,603 patients in intensive care or undergoing surgery at 3 US academic medical centers, reported that occult hypoxemia was associated with fewer hospital-free days and greater in-hospital mortality after adjusting for patient age, sex, comorbidities, setting (intensive care unit or surgery), and acuity. Neither outcome significantly differed by patient race or ethnicity, though Black, Asian, or American Indian patients together made up a relatively small proportion of the patient sample (2,110 versus 24,493 White patients). Only simultaneously collected SpO_2–SaO_2 readings were used in the study, and occult hypoxemia was defined as SaO_2 less than 88% despite SpO_2 of 92% or greater.

A smaller study[1] in patients evaluated in the emergency department or hospitalized for COVID-19 in the Johns Hopkins Health System ($N = 1{,}903$) examined whether patients predicted to have an SaO_2 of 94% or less prior to a measured SpO_2 of 94% or less (*ie*, an unrecognized hypoxemic state) experienced delayed recognition of treatment eligibility or delayed treatment initiation. Failure to recognize eligibility or delayed recognition of eligibility was significantly more likely among both Black patients (adjusted hazard ratio [aHR] = 0.71, 95% CI [0.63, 0.80]) and Hispanic or Latino patients (aHR = 0.77, 95% CI [0.66, 0.89]) with unrecognized hypoxemia compared with White patients with unrecognized hypoxemia. Analyses were adjusted for patient demographics, comorbidities, acuity, and laboratory values (*eg*, hemoglobin). In the subset of patients eventually recognized as treatment eligible, the median delay to eligibility recognition was about 2 hours longer

for Black patients and Asian patients compared with Hispanic or Latino and White patients (see Table above; $p = .01$ for Black patients versus White patients).

A final observational study[39] used a causal inference methodology to assess whether differential bias in pulse oximeter measurements led to poorer COVID-19 treatment and health services outcomes for Black patients in a large California integrated health system. Compared with White patients, overestimation of SaO_2 among Black patients (N = 1,699) was associated with a significantly lower likelihood of hospital admission (-3.1%, 95% CI [-3.4, -2.8]), initiation of treatment with dexamethasone (-3.1%, 95% CI [-3.4, -2.7]), delivery of supplemental oxygen (-4.5%, 95% CI [-4.9, -4.2]), and post-discharge return to the hospital (-1.2%, 95% CI [-1.9, -0.5]). Overestimation of SaO_2 was predicted to result in a 37-minute delay in dexamethasone initiation (95% CI [20.1, 54.3]) and a 279-minute delay in supplemental oxygen initiation (95% CI [181.0, 376.0]). Results were adjusted for patient age, sex, common comorbidities, and homelessness and insured statuses.

DISCUSSION

Based on a large body of evidence, pulse oximeters likely overestimate Black patients' blood oxygen saturation level. Most available oximetry data is from recent studies conducted in contemporary hospital and health system settings and using modern pulse oximeters. In these studies, pulse oximeters also appear to have the largest bias and greatest imprecision among Black patients, though some degree of bias and considerable imprecision is evident regardless of patient race/ethnicity. Occult hypoxemia is likely more common among Black patients compared with White patients. Asian, Latino or Hispanic, Native American or Indigenous, or other or mixed race/ethnicity patients may also experience occult hypoxemia more frequently than White patients, but evidence in these groups is less certain. Studies examining differences in clinical outcomes associated with occult hypoxemia or biases in pulse oximeter accuracy are methodologically inconsistent but provide suggestive evidence that Black patients with undetected hypoxemia could experience poorer treatment delivery and clinical outcomes than White patients with undetected hypoxemia.

Re-assessing oxygen saturation in arterial blood more routinely—particularly for patients who show signs or symptoms of arterial hypoxemia—has been proposed as one step to reduce the risk of occult hypoxemia and subsequent harms arising from biases in pulse oximeter readings.[40] Another approach suggested to mitigate this risk is to raise the oxygen saturation range from 92–96% to 94–98%, though this may increase risk of hyperoxemia.[44] Applying a skin-tone-based correction factor to readings from currently available pulse oximeters has also been discussed,[47] but such adjustments have a controversial history[48-50] and may have limited efficacy because the accuracy and reliability of pulse oximetry readings is influenced by a number of factors that cannot be accounted for in a single correction factor.[47,49] A broader recommendation has been to revise guidelines for pulse oximeter validation studies, in particular to require enrolling more patients with darker skin pigmentation, testing oximeters under real-world health care conditions, and incorporating perfusion into validation requirements.[44,51]

As noted earlier, the systematic review by Shi et al[6] identified some studies that reported pulse oximeter accuracy by skin pigmentation level rather than race or ethnicity. We focused on the latter because most available accuracy data is from recent studies that report their results by race/ethnicity (likely due to the use of patient health record data). A concern with the use of race/ethnicity in accuracy studies is that it may introduce spurious variation across studies and lead to unexpected or clinically counterintuitive findings, given that individuals with a wide range of skin pigmentation levels may identify with the same race/ethnicity.

For instance, groups made up of patients identified as Black through health record data may, in one hypothetical study, be composed of individuals who on average have lighter skin pigmentation compared with a Black patient group in a different study. Variation in sample composition could be caused by many factors, including geographic or other contextual differences between the studies, or may simply be due to chance. In the former study, pulse oximeter bias in the Black patient group may be smaller and closer to levels observed among White patients, while in the latter study, bias in the Black patient group may be more substantial. Underlying differences in sample composition by skin pigmentation level, therefore, resulted in inconsistent evidence about the same race/ethnicity group. Moreover, findings from any single study (*eg*, the first study, in which bias was similar across groups) could be inappropriately generalized because they do not account for the composition of patient groups relative to skin pigmentation level. Given these considerations, use of objective skin pigmentation

metrics when determining the accuracy of pulse oximeters has been proposed.[51] Extending this recommendation to prospective research on disparities in occult hypoxemia is likely warranted.

Finally, it has been acknowledged that addressing biases in pulse oximeter readings and in the care that follows from those readings requires fundamental advancements in the technologies used for routine oxygen saturation monitoring. Fawzy et al[1] state:

> "Although increased awareness of the limitations of pulse oximetry may mitigate some of the adverse effects…the race and ethnicity-based discrepancy of pulse oximetry exposes a fundamental flaw in the acquisition rather than interpretation of data, although all the aforementioned biases are associated with systematic underdiagnosis of disease or withholding of therapies for racial and ethnic minority groups."

Improving pulse oximeter technology is an active research area. A small validation study[52] published earlier this year, for example, tested an investigational noninvasive oximeter that uses green rather than the conventional red light, is designed to target superficial skin layers to increase sensitivity to tissue hypoxia, and implements patient-specific skin tone calibration (rather than a pre-programmed correction factor). The study enrolled equal proportions of patients with fair skin, brown skin, and dark skin. Oxygen saturation readings from the novel oximeter were more highly correlated with blood-based oximetry ($r = 0.76$) than pulse oximeter readings ($r = 0.47$), and the novel oximeter was also able to accurately assess oxygen levels in cases in which the pulse oximeter failed, including a patient with very high skin pigmentation.

Limitations

Incorporating an existing review of pulse oximeter accuracy meant that we could not exclude studies with limited relevance to modern clinical practice among Veterans (*ie*, studies that used older pulse oximeter technologies or pediatric samples). Observations from these studies make up a small proportion of available accuracy data, so it is unlikely this limitation impacts the validity of the review's findings. We also used sequential data abstraction and risk of bias assessment rather than a fully independent (blinded) process.

CONCLUSIONS

Pulse oximeters likely overestimate Black patients' blood oxygen saturation level, increasing the risk for unrecognized or occult hypoxemia. Occult hypoxemia occurs to some degree in all races/ethnicities but is likely more common among Black patients compared with White patients. Findings of this review underscore that clinicians should be aware of the risk of occult hypoxemia in patients with darker skin pigmentation. Moreover, while pulse oximeter readings are on average fairly similar to arterial oxygen saturation levels, evidence from hospital and health system settings relevant to the VA suggests that the amount of bias could vary substantially from patient to patient regardless of their race/ethnicity. This finding implies that incorporating conventional race or ethnicity-based correction factors into pulse oximeters would not eliminate disparities in occult hypoxemia risk. Although proposed changes to clinical practice to accommodate bias and imprecision in pulse oximeters may help to mitigate harms in the near-term, advancements in noninvasive oximeter technology are needed. As the largest integrated health system in the United States, the VHA is uniquely positioned to cultivate innovations in oximeter technology.

REFERENCES

1. Fawzy A, Wu TD, Wang K, et al. Racial and ethnic discrepancy in pulse oximetry and delayed identification of treatment eligibility among patients with COVID-19. *JAMA Internal Medicine*. 2022;182(7):730-738. doi:https://dx.doi.org/10.1001/jamainternmed.2022.1906
2. Henry NR, Hanson AC, Schulte PJ, et al. Disparities in hypoxemia detection by pulse oximetry across self-identified racial groups and associations with clinical outcomes. *Critical Care Medicine*. 2022;50(2):204-211. doi:https://dx.doi.org/10.1097/CCM.0000000000005394
3. Bangash MN, Hodson J, Evison F, et al. Impact of ethnicity on the accuracy of measurements of oxygen saturations: A retrospective observational cohort study. *EClinicalMedicine*. 2022;48:101428. doi:https://dx.doi.org/10.1016/j.eclinm.2022.101428
4. Burnett GW, Stannard B, Wax DB, et al. Self-reported race/ethnicity and intraoperative occult hypoxemia: A retrospective cohort study. *Anesthesiology*. 2022;136(5):688-696. doi:https://dx.doi.org/10.1097/ALN.0000000000004153
5. US Food and Drug Administration. *Pulse oximeters − Premarket notification submissions: Guidance for industry and Food and Drug Administration staff*. 2013. https://www.fda.gov/media/72470/download
6. Shi C, Goodall M, Dumville J, et al. The accuracy of pulse oximetry in measuring oxygen saturation by levels of skin pigmentation: A systematic review and meta-analysis. *BMC Medicine*. 2022;20(1):267. doi:https://dx.doi.org/10.1186/s12916-022-02452-8
7. Sjoding MW, Dickson RP, Iwashyna TJ, Gay SE, Valley TS. Racial bias in pulse oximetry measurement. *New England Journal of Medicine*. 2020;383(25):2477-2478. doi:https://dx.doi.org/10.1056/NEJMc2029240
8. U.S. Food and Drug Administration. *Pulse oximeter accuracy and limitations: FDA safety communication*. 2022. https://www.fda.gov/medical-devices/safety-communications/pulse-oximeter-accuracy-and-limitations-fda-safety-communication
9. Whiting P, Savovic J, Higgins JP, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol*. 2016;69:225-34. doi:https://dx.doi.org/10.1016/j.jclinepi.2015.06.005
10. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*. 2011;155(8):529-536. doi:https://dx.doi.org/10.7326/0003-4819-155-8-201110180-00009
11. Hayden JA, van der Windt DA, Cartwright JL, Pierre Côté D, Bombardier C. Assessing bias in studies of prognostic factors. *Annals of Internal Medicine*. 2013;158(4):280-286. doi:https://dx.doi.org/10.7326/0003-4819-158-4-201302190-00009
12. Hoy D, Brooks P, Woolf A, et al. Assessing risk of bias in prevalence studies: Modification of an existing tool and evidence of interrater agreement. *J Clin Epidemiol*. 2012;65(9):934-9. doi:https://dx.doi.org/10.1016/j.jclinepi.2011.11.014
13. Tipton E, Shuster J. A framework for the meta-analysis of Bland–Altman studies based on a limits of agreement approach. *Statistics in Medicine*. 2017;36(23):3621-3635. doi:https://doi.org/10.1002/sim.7352
14. Tang LL, Caudy M, Taxman F. A statistical method for synthesizing meta-analyses. *Computational and Mathematical Methods in Medicine*. 2013;doi:https://dx.doi.org/10.1155/2013/732989
15. Knapp G, Hartung J. Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine*. 2003;22(17):2693-2710. doi:https://doi.org/10.1002/sim.1482

16. Rover C, Knapp G, Friede T. Hartung-Knapp-Sidik-Jonkman approach and its modification for random-effects meta-analysis with few studies. *BMC Med Res Methodol*. 2015;15:99. doi:https://dx.doi.org/10.1186/s12874-015-0091-1
17. Parr NJ, Schweer-Collins ML, Darlington TM, Tanner-Smith EE. Meta-analytic approaches for examining complexity and heterogeneity in studies of adolescent development. *J Adolesc*. 2019;77:168-178. doi:https://dx.doi.org/10.1016/j.adolescence.2019.10.009
18. *metafor: Meta-analysis package for R*. Version 4.2-0. The Comprehensive R Archive Network; 2023. https://cran.r-project.org/web/packages/metafor/index.html
19. Ries AL, Farrow JT, Clausen JL. Accuracy of two ear oximeters at rest and during exercise in pulmonary patients. *American Review of Respiratory Disease*. 1985;132(3):685-9. doi:https://doi.org/10.1164/arrd.1985.132.3.685
20. Gabrielczyk MR, Buist RJ. Pulse oximetry and postoperative hypothermia: An evaluation of the Nellcor N-100 in a cardiac surgical intensive care unit. *Anaesthesia*. 1988;43(5):402-404. doi:https://doi.org/10.1111/j.1365-2044.1988.tb09025.x
21. Ries AL, Prewitt LM, Johnson JJ. Skin color and ear oximetry. *Chest*. 1989;96(2):287-90. doi:https://doi.org/10.1378/chest.96.2.287
22. Escourrou PJ, Delaperche MF, Visseaux A. Reliability of pulse oximetry during exercise in pulmonary patients. *Chest*. 1990;97(3):635-8. doi:https://dx.doi.org/10.1378/chest.97.3.635
23. Jubran A, Tobin MJ. Reliability of pulse oximetry in titrating supplemental oxygen therapy in ventilator-dependent patients. *Chest*. 1990;97(6):1420-1425. doi:https://doi.org/10.1378/chest.97.6.1420
24. Lee KH, Hui KP, Tan WC, Lim TK. Factors influencing pulse oximetry as compared to functional arterial saturation in multi-ethnic Singapore. *Singapore Medical Journal*. 1993;34(5):385-7.
25. Bothma PA, Joynt GM, Lipman J, et al. Accuracy of pulse oximetry in pigmented patients. *South African Medical Journal*. 1996;86(5 Suppl):594-6.
26. McGovern JP, Sasse SA, Stansbury DW, Causing LA, Light RW. Comparison of oxygen saturation by pulse oximetry and co-oximetry during exercise testing in patients with COPD. *Chest*. 1996;109(5):1151-1155. doi:https://dx.doi.org/10.1378/chest.109.5.1151
27. Adler JN, Hughes LA, Vivilecchia R, Camargo CA, Jr. Effect of skin pigmentation on pulse oximetry accuracy in the emergency department. *Academic Emergency Medicine*. 1998;5(10):965-70. doi:https://doi.org/10.1111/j.1553-2712.1998.tb02772.x
28. Abrams GA, Sanders MK, Fallon MB. Utility of pulse oximetry in the detection of arterial hypoxemia in liver transplant candidates. *Liver Transplantation*. 2002;8(4):391-6. doi:https://doi.org/10.1053/jlts.2002.32252
29. 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. doi:https://dx.doi.org/10.1016/j.resuscitation.2006.06.024
30. Hinkelbein J, Koehler H, Genzwuerker HV, Fiedler F. Artificial acrylic finger nails may alter pulse oximetry measurement. *Resuscitation*. 2007;74(1):75-82. doi:https://dx.doi.org/10.1016/j.resuscitation.2006.11.018
31. Muñoz X, Torres F, Sampol G, Rios J, Martí S, Escrich E. Accuracy and reliability of pulse oximetry at different arterial carbon dioxide pressure levels. *Eur Respir J*. 2008;32(4):1053-1059. doi:https://dx.doi.org/10.1183/09031936.00126507
32. Ebmeier SJ, Barker M, Bacon M, et al. A two centre observational study of simultaneous pulse oximetry and arterial oxygen saturation recordings in intensive care unit patients. *Anaesthesia and Intensive Care*. 2018;46(3):297-303. doi:https://doi.org/10.1177/0310057x1804600307

33. Pilcher J, Ploen L, McKinstry S, et al. A multicentre prospective observational study comparing arterial blood gas values to those obtained by pulse oximeters used in adult patients attending Australian and New Zealand hospitals. *BMC Pulmonary Medicine*. 2020;20(1):7. doi:https://dx.doi.org/10.1186/s12890-019-1007-3
34. Harskamp RE, Bekker L, Himmelreich JCL, et al. Performance of popular pulse oximeters compared with simultaneous arterial oxygen saturation or clinical-grade pulse oximetry: A cross-sectional validation study in intensive care patients. *BMJ Open Respiratory Research*. 2021;8(1)doi:https://dx.doi.org/10.1136/bmjresp-2021-000939
35. Valbuena VSM, Barbaro RP, Claar D, et al. Racial bias in pulse oximetry measurement among patients about to undergo extracorporeal membrane oxygenation in 2019-2020: A retrospective cohort study. *Chest*. 2022;161(4):971-978. doi:https://dx.doi.org/10.1016/j.chest.2021.09.025
36. Wiles MD, El-Nayal A, Elton G, et al. Effect of patient ethnicity on the accuracy of peripheral pulse oximetry in patients with COVID-19 pneumonitis requiring mechanical ventilation. *Anaesthesia*. 2022;77(4):489-491. doi:https://dx.doi.org/10.1111/anae.15656
37. Cecil WT, Thorpe KJ, Fibuch EE, Tuohy GF. A clinical evaluation of the accuracy of the Nellcor N-100 and Ohmeda 3700 pulse oximeters. *J Clin Monit*. 1988;4(1):31-6. doi:https://dx.doi.org/10.1007/bf01618105
38. Wang YT, Poh SC. Noninvasive oximetry in pigmented patients. *Annals of the Academy of Medicine, Singapore*. 1985;14(3):427-9.
39. Sudat SEK, Wesson P, Rhoads KF, et al. Racial disparities in pulse oximeter device inaccuracy and estimated clinical impact on COVID-19 treatment course. *American Journal of Epidemiology*. 2023;192(5):703-715. doi:https://dx.doi.org/10.1093/aje/kwac164
40. Valbuena VSM, Seelye S, Sjoding MW, et al. Racial bias and reproducibility in pulse oximetry among medical and surgical inpatients in general care in the Veterans Health Administration 2013-19: Multicenter, retrospective cohort study. *BMJ*. 2022;378:e069775. doi:https://dx.doi.org/10.1136/bmj-2021-069775
41. Blanchet M-A, Mercier G, Delobel A, et al. Accuracy of multiple pulse oximeter brands in stable critically ill patients - Oxygap study. *Respiratory Care*. 2023;doi:https://dx.doi.org/10.4187/respcare.10582
42. Crooks CJ, West J, Morling JR, et al. Differential pulse oximetry readings between ethnic groups and delayed transfer to intensive care units. *QJM*. 2023;116(1):63-67. doi:https://dx.doi.org/10.1093/qjmed/hcac218
43. Nguyen LS, Helias M, Raia L, et al. Impact of COVID-19 on the association between pulse oximetry and arterial oxygenation in patients with acute respiratory distress syndrome. *Scientific Reports*. 2022;12(1):1462. doi:https://dx.doi.org/10.1038/s41598-021-02634-z
44. Chesley CF, Lane-Fall MB, Panchanadam V, et al. Racial disparities in occult hypoxemia and clinically based mitigation strategies to apply in advance of technological advancements. *Respiratory Care*. 2022;67(12):1499-1507. doi:https://dx.doi.org/10.4187/respcare.09769
45. Seitz KP, Wang L, Casey JD, et al. Pulse oximetry and race in critically ill adults. *Critical Care Explorations*. 2022;4(9)doi:https://dx.doi.org/10.1097/CCE.0000000000000758
46. Wong A-KI, Charpignon M, Kim H, 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 Network Open*. 2021;4(11):e2131674-e2131674. doi:https://dx.doi.org/10.1001/jamanetworkopen.2021.31674
47. Jamali H, Castillo LT, Morgan CC, et al. Racial disparity in oxygen saturation measurements by pulse oximetry: Evidence and implications. *Annals of the American Thoracic Society*. 2022;19(12):1951-1964. doi:https://dx.doi.org/10.1513/AnnalsATS.202203-270CME

48. Braun L. Race correction and spirometry: Why history matters. *Chest*. 2021;159(4):1670-1675. doi:https://doi.org/10.1016/j.chest.2020.10.046
49. Fawzy A, Valbuena VSM, Chesley CF, Wu TD, Iwashyna TJ. Dynamic errors in pulse oximetry preclude use of correction factor. *Annals of the American Thoracic Society*. 2022;20(2):338-339. doi:https://doi.org/10.1513/AnnalsATS.202210-872LE
50. Vyas DA, Eisenstein LG, Jones DS. Hidden in plain sight — reconsidering the use of race correction in clinical algorithms. *New England Journal of Medicine*. 2020;383(9):874-882. doi:https://dx.doi.org/10.1056/NEJMms2004740
51. Okunlola OE, Lipnick MS, Batchelder PB, Bernstein M, Feiner JR, Bickler PE. Pulse oximeter performance, racial inequity, and the work ahead. *Respiratory Care*. 2022;67(2):252-257. doi:https://dx.doi.org/10.4187/respcare.09795
52. Sanjay GG, Vinoop D, Georgios A. Innovative technology to eliminate the racial bias in non-invasive, point-of-care (POC) haemoglobin and pulse oximetry measurements. *BMJ Innovations*. 2023;9(2):73. doi:https://dx.doi.org/10.1136/bmjinnov-2022-001018