BMJ Open

Protocol for a multisite, observational clinical study of the association between skin colour and pulse oximeter accuracy in children undergoing cardiac catheterisation (PACH study)

Halley Ruppel ,[1,2] Liming Huang,[3] Christopher J Petit,[4,5] Athar M Qureshi,[6,7] Christopher P Bonafide,[2,8,9] Elizabeth E Foglia,[10] Jeffery Feldman,[11,12] Meghan B Lane-Fall,[13] Jennifer M Lynch,[13,14,15] Anna Chien,[16] Michelle Dunn,[1] Michael L O'Byrne[2,9,17,18]

To cite: Ruppel H, Huang L, Petit CJ, *et al*. Protocol for a multisite, observational clinical study of the association between skin colour and pulse oximeter accuracy in children undergoing cardiac catheterisation (PACH study). *BMJ Open* 2025;**15**:e100391. doi:10.1136/bmjopen-2025-100391

► Prepublication history for this paper is available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2025-100391).

Received 08 February 2025
Accepted 12 March 2025

For numbered affiliations see end of article.

Correspondence to
Dr Halley Ruppel;
hruppel@nursing.upenn.edu

ABSTRACT

Introduction Prospective, real-world clinical studies of the association between skin color and pulse oximeter (SpO2) accuracy in children are needed to address the limitations of previous research. Such studies are essential for generating evidence for clinicians, regulators and industry. This is the protocol for a multisite study funded by the National Heart, Lung, and Blood Institute (R01HL171313; 1 January 2024–31 December 2028).

Methods and analysis In this pragmatic, observational study conducted in three large paediatric cardiac catheterisation centres in the USA, children undergoing cardiac catheterisation with directly measured arterial oxygen saturation will be prospectively enrolled. The outcome variable (SpO_2 bias) is the difference between contemporaneous paired measurements of pulse oximetry (SpO_2) and the standard reference comparator, arterial blood sample oxygen saturation (SaO_2), obtained during the catheterisation procedure. The independent variable is an objective measure of skin colour obtained via spectrophotometry. Our primary analysis is a multivariable regression model testing the relationship between skin colour and SpO_2 bias, after adjusting for covariates. We will also conduct a moderator analysis to identify factors that may affect the magnitude of the association. The target sample size is 584 participants.

Ethics and dissemination This study was approved by the University of Pennsylvania Institutional Review Board (#854895) under expedited review. Study risks are minimal. Parental permission, and child assent when applicable, are obtained prior to enrolment. In accordance with the NIH Public Access Policy, publications associated with the study will be made publicly available through PubMed Central. The analytic dataset will be contributed to a repository for future use. In collaboration with a children's hospital-based research family advisory council, interpretation and dissemination of the results for lay, clinical and scientific audiences will be considered.

Trial registration number Although not a clinical trial, this observational study is registered on ClinicalTrials.gov (identifier: NCT06529575) for public awareness.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This multisite real-world clinical study of pulse oximeter accuracy across a range of skin colours in children addresses limitations of prior studies, including larger sample size, diversity of the eligible patient pool and the use of different pulse oximeter devices across sites.
⇒ Conducting the study in the cardiac catheterisation lab reduces burden to paediatric patients, as much of the required data are already collected in the course of clinical care and are highly accurate because patients are sedated and closely monitored.
⇒ The patient population included in the study is generally haemodynamically stable and has a broad range of expected oxygen saturations due to the variety of congenital heart lesions treated in the cardiac catheterisation laboratory.
⇒ As a pragmatic real-world study, the study does not include standardised placement location of the pulse oximeter sensor across participants, which could be viewed as a limitation, and each participant will only have one arterial oxygen saturation measurement reflective of systemic saturation.

INTRODUCTION

The association between skin colour and pulse oximeter (SpO_2) accuracy has been studied for decades, with several studies demonstrating systematic overestimation of true arterial oxygen saturation (SaO_2) in adults with darker skin colours.[1–5] Less is known about SpO_2 accuracy across skin colour in children, which may be different than adults due to age-related differences in characteristics such as skin thickness, melanin concentration, tissue optical properties and peripheral perfusion. Several recent studies of SpO2 in children and infants demonstrated greater

overestimation of SaO_2 in hypoxaemic black patients compared with white patients using retrospective data from the electronic health record (EHR).[6–8] Retrospective EHR data-based studies lack direct measurement of skin colour, relying instead on race and ethnicity variables. Race and ethnicity, especially in the USA, may be associated with skin colour, but are flawed proxies. A few small prospective clinical studies in children and infants have not identified an association between directly measured skin colour and SpO_2 accuracy.[9 10]

There are multiple possible reasons for differential pulse oximeter accuracy in patients of differing skin colours. The skin pigment melanin is present in higher concentrations in patients with dark skin colour.[11] Melanin both scatters and partially absorbs the light emitted and measured by pulse oximeters. In addition, many pulse oximeter calibration tests are thought to have been conducted on patients with light skin colour.[2 11 12] We use the term 'skin colour', rather than melanin or pigmentation, in this protocol because our evaluation includes multiple dimensions of colour.

Real-world clinical evaluations of pulse oximetry accuracy in children are essential to inform clinical care and to guide product development efforts and governmental regulation of medical devices. Large prospective paediatric studies that address the limitations of the previous research are needed. In January 2024, our team received funding from the United States National Institutes of Health National Heart, Lung, and Blood Institute to conduct a prospective multisite study of pulse oximeter accuracy in children (R01HL171313; 1 January 2024–31 December 2028) (Box 1. Here, we report our study protocol.

Preliminary studies

Our team conducted two preliminary single-site studies in preparation for the current research: (1) a retrospective study using EHR data[7] and (2) a prospective feasibility pilot study. The preliminary studies were funded by the Center for Pediatric Clinical Effectiveness at the Children's Hospital of Philadelphia Research Institute. We conducted the retrospective study to detect any signal of differential SpO_2 accuracy across patients of black and white race in a large sample of paediatric patients (n=774). The study was conducted in paediatric patients undergoing cardiac catheterisation with existing blood SaO_2 that could be matched to simultaneously recorded SpO_2 measurements. After adjusting for covariates, we found greater overestimation of SaO_2 by SpO_2 in black patients than white patients.[7] In other words, the pulse oximeter systematically falsely reported a higher oxygen saturation for black patients compared with white patients at the same SaO_2 value.

Because of the limitations of using race, rather than skin colour, we developed and pilot tested a protocol for a prospective study in the same population. We enrolled 29 patients between 2021 and 2023, assessing protocol acceptability and feasibility. Skin pigmentation was assessed using Munsell and Fitzpatrick charts.[13 14] Of 34 patients approached, 29 (85%) enrolled. We also identified optimal recruitment procedures. Participants' SaO_2 values ranged from 70–99%. The skin colour of 11 participants was rated Munsell value ≤4 (dark), one was rated Munsell value of 6 (medium) and 17 were rated Munsell value of 7–8 (light). The pilot study was not powered to detect significant differences in pulse oximeter accuracy across skin colour. Together, findings from the retrospective and prospective pilot studies informed the protocol and power analysis for the current study.

Box 1 Study timeline

IRB Approval: December 2023.
Recruitment initiated: August 2024.
Target enrolment anticipated: August 2027.

METHODS AND ANALYSIS

Objectives

The primary objective of the current study is to quantify the association between skin colour and SpO_2 accuracy in children. We hypothesise that in children with darker skin colours, SpO_2 will overestimate SaO_2 as compared with children with lighter skin colours. In adults, research indicates that the relationship between skin colour and SpO_2 accuracy is moderated by other factors.[1 15 16] We, therefore, also hypothesise that the relationship between skin colour and SpO_2 accuracy will be moderated by several factors, such as age, biological sex and peripheral perfusion.

Our secondary objective is to explore the mechanistic link between self-identified race and SpO_2 accuracy. We hypothesise that skin colour will mediate an observed relationship between self-identified race and SpO_2 accuracy, as this is the leading mechanistic hypothesis for prior studies' findings of an association between race and SpO_2 Bias,[6 7 17 18] but the link has not been clearly established.[19] If no mediation is found, it is possible that other factors not accounted for in these studies explain the observed relationship.

Design

This is a pragmatic, prospective, observational, multisite study conducted among paediatric patients undergoing cardiac catheterisation. This is a relatively haemodynamically stable population where most study data can be obtained from data already collected in the course of clinical care, minimising burden to study participants. We also anticipate variability in SpO_2 values in this population, as many paediatric patients undergoing cardiac catheterisation are not fully saturated at baseline.

Setting

Data are collected from patients undergoing cardiac catheterisation procedures at three major Children's

Hospitals in the USA that participate in the Congenital Catheterisation Research Collaborative. The three clinical sites represent some of the largest paediatric cardiac catheterisation programmes in the country and are located in highly diverse cities in the USA (Philadelphia, Houston and New York City), which will optimise the diversity of the sample.

Organisational structure and responsibilities

The study is overseen by an executive committee, comprising the study principal investigators (HR and MLO'B) and biostatistician (LH). The executive committee makes decisions related to study protocol and design. The study coordinating centre is located at the University of Pennsylvania and is responsible for obtaining ethics approval, establishing reliance agreements between site institutional review boards (IRBs), conducting training and maintaining active communication with the study sites. Data coordination also occurs at the University of Pennsylvania. Each of the three clinical sites has a site principal investigator and clinical research coordinator team. The study steering committee, which meets quarterly, consists of co-investigators with experience in dermatology, neonatology, paediatric cardiology, anaesthesiology and hospital medicine. The steering committee reviews progress and provides input on study protocol and design decisions. An advisory committee, consisting of a biomedical engineer, dermatologist, sociologist and demographer, and anaesthesiologist with expertise in biomedical optics, meets annually and as needed to provide subject matter expertise.

Inclusion and exclusion criteria

Children are included if they are <18 years old, ≥1 month old and are undergoing a cardiac catheterisation procedure that typically involves direct SaO_2 measurement by co-oximeter from a location reflective of systemic arterial saturation as part of the cardiac catheterisation procedure (eg, annual transplant evaluation, aortic valvuloplasty or diagnostic haemodynamic evaluation). Eligibility is determined by the site principal investigator (an interventional cardiologist), based on knowledge of the procedure and operator practice. Children are excluded if they have differential cyanosis across extremities, are receiving intravenous vasoconstrictor agents prior to the procedure or have a ventricular assist device. Although children may undergo multiple procedures during the study period, individuals will only be enrolled once.

Sample size

Our power calculation was based on our primary outcome, SpO_2 bias ($SpO_2 - SaO_2$).[20] We assumed a multivariable regression model with SpO_2 bias as a continuous outcome and skin colour as a single continuous predictor, adjusting for 10 covariates. Our sample size calculation is based on preliminary results from our pilot studies. Because we do not have reliable data on the distribution of skin colours in our patient population, we will conduct a formal interim analysis after 50% of the study sample has been recruited to adjust sample size estimations and revise recruitment strategies as needed.

In the preliminary results from our prospective pilot study, we found a difference in mean SpO_2 bias between patients of dark skin colour and light skin colour of 0.25 and an SD of 3.1—a small effect size of 0.08. To account for missing data and to ensure we have adequate distribution of skin colours in our sample, we are proposing to enrol a conservative sample size, sufficient to detect a small effect size based on Cohen's interpretation of effect sizes for regression analysis. A sample size ranging 113–211 would achieve 80% power to detect a small effect size (f^2) ranging 0.08–0.04, attributable to one independent variable using an F-test with a significance level (alpha) of 0.05 to test the main effect, adjusting for 10 covariates. To test for moderation, simulation studies have shown that the sample size required is approximately four times what is required for a main effect analysis.[21] Therefore, to ensure adequate sample for the moderation and mediation analyses, we seek to have complete data from at least 584 unique participants (type I error =0.05 and power =80%). Statistical versus clinical significance will be considered when interpreting results. Clinical significance may vary across clinical specialty, and we will partner with members of the interdisciplinary clinical team to understand clinical relevance in their field.

To achieve our target sample size, we anticipate recruiting for approximately 3 years across three sites, which conservatively assumes our study team will be able to approach >50% of eligible participants, accounting for study team percent effort and time off. Study sites each perform approximately 700–1000 cases annually. In our pilot study (conducted at a single site), we found that approximately 15–20% of cardiac catheterisation cases met our eligibility criteria. The current study expands eligibility criteria. We are also accounting for approximately 10–15% refusal rate and a diminishing eligible pool over time, given that a portion of catheterisation procedures are performed annually in the same patients.

Recruitment and consent

We are employing a pragmatic consecutive sampling approach. After screening the cardiac catheterisation laboratory schedule and medical records and verifying eligibility, trained study personnel approach the child and parent/legal guardian to verbally describe the study prior to the cardiac catheterisation procedure. To address workflow differences at the sites, patients may be approached at their preprocedure outpatient visit, on the day of the procedure, or on an inpatient unit if the patient is hospitalised. The study team member obtains written permission from the parent/legal guardian and assent from the child (when applicable). Families who speak a language other than English are approached with appropriate interpreter services.

Patient and public involvement statement

We sought feedback from our institution's research family advisory council on study procedures and consent language prior to initiating recruitment. Council members indicated that families may be surprised to learn that pulse oximeter accuracy is a concern, considering its routine use and suggested that the study team be prepared to answer further questions about why the concern exists. Council members identified language, such as skin colour and melanin, that might be triggering for some families in the healthcare context and encouraged clear emphasis that the problem is with the technology, not with an individual's skin colour, and that there are ongoing efforts nationally to ensure the technology is universally accurate, regardless of skin colour. A diverse sample is essential to achieving study objectives, and given known challenges with diversity in clinical research,[22] we will continue to engage the family advisory council as needed to address recruitment procedures. We will also return to the council to discuss interpretation and dissemination of findings for a lay audience.

Data sources

Primary outcome measure

Our outcome variable (SpO_2 bias) is the difference between contemporaneous paired measurements of SpO_2 and the standard reference comparator, SaO_2 (ie, SpO_2 measurement error).[20 23] Prior studies have used the outcome of rates of 'occult hypoxaemia' (eg, SpO_2>92% when SaO_2<88%) in patients of black and white races.[6 17 18] We will calculate occult hypoxaemia rates (by skin color), but they are not our primary outcome as these rates are affected by the prevalence of hypoxaemia (SaO_2<88%) in the sample.[24 25]

SpO_2 is monitored continuously during catheterisation procedures, typically with excellent reliability, as the patient is under sedation or general anaesthesia, with minimal to no movement artefact. The SpO_2 wave form is closely monitored by the anaesthesiologist and poor readings are typically immediately corrected. The pulse oximeter is also often under the sterile procedural drapes, minimising sensor exposure to ambient light. SpO_2 values are captured in the EHR at a frequency of at least every 1 min at the three clinical sites.

SaO_2 measurements are routinely obtained during the types of cardiac catheterisation procedures included in the study. Many catheterisation procedures only have one SaO_2 measurement that will meet our criteria of being sampled from a location reflective of systemic arterial saturation and analysed by co-oximetry. SaO_2 obtained via co-oximeter is considered the gold standard comparator for SpO_2.[23] At the clinical sites, when a blood sample is drawn from the arterial line during the procedure, it is immediately passed to a nurse or technician, who runs the sample on a co-oximeter located in the catheterisation lab (study sites use the Avoximeter 1000E or Avoximeter 4000 co-oximeters). The SaO_2 result is immediately entered into the medical record and verbally reported back to the operator to ensure the value is consistent with their understanding of the patient's physiology. As such, the time stamp and value for the SaO_2 measurement in the medical record have a high degree of accuracy, as compared with blood gases obtained in other clinical settings, where there may be a delay between obtaining and running the specimen. For analysis, we extract from the medical record one paired SpO_2 and SaO_2 measurement per patient (matched by time, to the minute). We also obtain SpO_2 values in the 3 min preceding and following the index time, to assess for stability of the measurement.

Independent variable

For our primary analysis, our independent variable of interest is objective measurement of skin colour. The OpenOximetry project at the University of California San Francisco is generating standard-setting protocols for skin colour measurement for pulse oximeter studies,[26 27] which have also been adapted into the recent U.S. Food and Drug Administration (FDA) draft recommendations for pulse oximeter clinical performance testing.[28] To harmonise our results with other studies, we are using a similar skin colour measurement protocol, adapted for paediatrics. Our primary independent variable is individual typology angle (ITA), a summary metric of standard colour values, L* (black/white) and b* values (yellow/blue), for which a standard classification for skin colour exists (very light to dark).[29 30] Skin colour measurements are obtained via Konica Minolta Spectrophotometer CM-700d and taken in five locations (forehead, palm of hand, back of hand, dorsal distal phalanx below the nail bed and palmar distal phalanx) using the 3 mm aperture. Three measurements are taken in each location without lifting the spectrophotometer. Valid and reliable measurements on the fingers are challenging in some smaller children (ie, spectrophotometer target mask opening cannot be completely covered), in which case measurements from the palm and back of hand may be used in analysis. Study team members are trained to use the spectrophotometer by a manufacturer representative or another trained study team member.

Other variables

Skin complexion

We are obtaining visual (subjective) assessment of skin colour using the Monk Skin Tone Scale (MST). The MST is intended to assess overall complexion and will not be used in modelling.[31] MST will be used to describe the study sample and correlation between the MST and ITA will be explored. Study team members were trained using a module developed for the study, incorporating freely available MST-annotated images.[32]

Race and ethnicity

Participant race and ethnicity are captured using a self-report form with standard US federal reporting categories (completed by the parent/guardian or the participant) as

well as from the medical record. Race documented in the medical record may be inaccurate more than 50% of the time,[33] in part because race/ethnicity may be assigned by hospital staff without input from patients or their caregivers.

Other demographic data

Age, sex and body mass index are extracted from the medical record.

Clinical data

The following clinical data are not reliably recorded in the medical record and are therefore obtained from direct consultation with a clinician in the catheterisation procedure: location of the pulse oximeter sensor, pulse oximeter manufacturer/device (sites use either Masimo, Nellcor or Nonin), perfusion index if available and clinical assessment of peripheral perfusion (capillary refill and tactile extremity temperature). Studies demonstrate that capillary refill time is a clinically meaningful and prognostic assessment[34–36] and has been used in prior studies of pulse oximeter accuracy in children to assess perfusion.[15] We are not standardising the sensor location across participants due to differences in optimal placement for clinical care that may vary by patient age, condition, positioning, room setup and comorbidities. However, we distributed an informational module about the study to the anaesthesia teams for awareness.

Other clinical data are obtained from the procedure and anaesthesia medical record. These data include: diagnoses, procedure type, and inotrope, blood product and airway use during the procedure. We also obtain the following data with a timestamp as close to the index SaO_2 time as available: heart rate, temperature, blood pressure, blood gas results (including haemoglobin) and supplemental oxygen use.

Study procedures

After obtaining consent and assent, the study team member obtains the child's self-reported race and ethnicity from the participant or parent/legal guardian. The study team member documents skin colour using the MST scale and measures objective skin colour using the spectrophotometer. Additional data are documented during the patient's catheterisation procedure and extracted from the medical record, as described above.

Statistical analysis

Demographic and clinical characteristics will be summarised by standard descriptive summaries (eg, means and SD for continuous variables such as age and percentages and counts for categorical variables such as sex). Descriptive statistics for SpO_2, SaO_2, SpO_2 bias, skin colour (ITA) and MST will also be generated. Comparisons between ITA categories for demographic and clinical characteristics, as well as the key variables used in the primary and secondary analyses, will be performed using either the t-test or Wilcoxon signed-rank test for continuous variables and χ^2 or Fisher's exact test for categorical variables, as appropriate.

Primary analysis

Our primary analysis is a multivariable regression model. The model will test the relationship between skin colour (ITA) and SpO_2 bias, after adjusting for covariates. Our three hypothesised moderators (sex, age and perfusion) will be tested separately, by adding terms for each moderator and an interaction term with skin colour. Other potential covariates include: blood pressure, temperature, haemoglobin, haemoglobinopathies, medication use, blood products, probe location and study site.[7 15]

In addition to reporting the overall association between skin colour and SpO_2 bias, we will stratify the predicted values from the multivariable model to assess the relationship between skin colour and SpO_2 bias by SpO_2 category (eg, $SpO_2 \geq 92\%$, SpO_2 84–91% and $SpO_2 < 84\%$), as possible given the distribution of SpO_2 readings in the study population. This stratification will identify whether the magnitude of the relationship between skin colour, and SpO_2 bias is different at different levels of SpO_2, without including SpO_2 as a covariate in the model. Including SpO_2 as a covariate in the model is not recommended because SpO_2 is used to calculate the outcome variable (SpO_2 bias). As with our previous work,[7] we will stratify by SpO_2 category, rather than SaO_2 category, because clinicians can use this information to consider the potential magnitude of discrepancy between a given SpO_2 value and the true SaO_2 value in clinical practice.

Secondary analysis/additional analyses

We also seek to establish the relationship between self-identified race and SpO_2 Bias and assess the mediation effect of skin colour on this relationship. As in the primary analysis, the outcome variable will be SpO_2 bias (ie, SpO_2–SaO_2). The independent variable, race, is self-identified. We will use the ITA value calculated from the spectrophotometer measurements as the mediator. We will quantify the direct, indirect and total effects in terms of the linear regression coefficients from models regressing bias on race, regressing skin colour on race and finally regressing bias on skin colour and race. Regression models will adjust for covariates.

Interim analysis

We cannot anticipate the distribution of skin colours in our patient population, as this information is not collected in clinical care. In the proposed study, we have tried to ensure that our sample will include a variety of skin colours by partnering with study sites that serve diverse patient populations. We will conduct a formal interim analysis after 50% of the study sample has been recruited to assess enrolment across subgroups, re-estimate sample size, and adjust recruitment strategies as needed. Our team is partnering with our institution's health equity research consortium to consider strategies for representativeness in our study sample.

DATA MANAGEMENT AND QUALITY ASSURANCE

Quality assurance procedures include a weekly review of data entry completeness and consistency, conducted by the study coordinating centre at the University of Pennsylvania. Study data are recorded in a REDCap database[37 38] hosted at the University of Pennsylvania, with features such as branching logic and range limits to facilitate accurate data entry. Clinical sites will maintain master logs linking participant identifiers to coded study identification numbers. Consent forms and source documents will be stored securely at clinical sites. Prior to initiating data collection, site study teams undergo training on data entry into the REDCap database. At least yearly (or as needed), the coordinating centre will conduct monitoring visits for each data collection site, to review study procedures, source documents and the regulatory binder. The goals of the monitoring visits are to ensure the study is conducted in accordance with the protocol, support the site in resolving issues and provide further education as needed.

ETHICS AND DISSEMINATION

This study was approved by the University of Pennsylvania IRB (#854895) under expedited review. The University of Pennsylvania IRB is serving as the IRB of record and reliance agreements were established with participating clinical sites. Study risks are not greater than minimal. The primary risk is breach of privacy and confidentiality. Participants could also experience momentary embarrassment or discomfort from the skin color assessment. The risks of the study are outweighed by the benefit of the knowledge to be gained for the paediatric healthcare community. Parental permission, and child assent when applicable, are obtained prior to any study procedures.

To ensure the results of the study contribute to the knowledge base, we commit to timely publication of the primary outcomes paper, regardless of the significance of the results. Papers, abstracts and products resulting from the study will be approved by the study executive committee. In accordance with the US National Institutes of Health Public Access Policy, publications associated with the study will be made publicly available through PubMed Central. In addition, the final analytic dataset from the study will be contributed to a repository for use by future researchers (demographic information, clinical data and skin color measurement data) in a manner that protects participant identities. Our team will collaborate with our institution's research family advisory council and health equity research consortium in the interpretation and dissemination of the results for both lay and clinical audiences. We will identify additional dissemination strategies as appropriate to communicate results beyond the scientific community.

Author affiliations
[1]Department of Family and Community Health, University of Pennsylvania School of Nursing, Philadelphia, Pennsylvania, USA
[2]Clinical Futures, a Center of Emphasis in the CHOP Research Institute, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA
[3]Department of Biobehavioral Health Sciences, University of Pennsylvania School of Nursing, Philadelphia, Pennsylvania, USA
[4]Division of Cardiology, Department of Pediatrics, Columbia University Vagelos College of Physicians and Surgeons, New York, New York, USA
[5]NewYork-Presbyterian Children's Heart Center, New York, New York, USA
[6]The Lillie Frank Abercrombie Division of Cardiology, Texas Children's Hospital, Houston, Texas, USA
[7]Department of Pediatrics, Baylor College of Medicine, Houston, Texas, USA
[8]Division of General Pediatrics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA
[9]Department of Pediatrics, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA
[10]Division of Neonatology, Department of Pediatrics, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA
[11]Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA
[12]University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA
[13]Department of Anesthesiology and Critical Care, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA
[14]Division of Cardiothoracic Anesthesiology, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA
[15]Department of Bioengineering, University of Pennsylvania School of Engineering and Applied Science, Philadelphia, Pennsylvania, USA
[16]Department of Dermatology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA
[17]Division of Cardiology, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA
[18]Leonard Davis Institute and Center for Cardiovascular Outcomes, Quality, and Evaluative Research, The University of Pennsylvania, Philadelphia, Pennsylvania, USA

Acknowledgements We thank Courtney Boen, PhD (Brown University) for her role as a study advisor. We also thank Subhash Aryal, PhD (Johns Hopkins University) for conceiving of the original statistical approach. We thank The Children's Hospital of Philadelphia (CHOP) Research Institute Family Advisory Council and Chén Kenyon, MD, MSHP and Jordan Wood, MPH from the CHOP Research Institute Health Equity Research Methods Initiative. We also thank the clinical site research teams at each of our data collection sites for their dedicated effort to the success of this study.

Contributors HR, MLO'B, CPB and EEF conceived of the study. JF, MBL-F, AC and JML provided subject matter expertise and protocol refinement. HR, MLO'B, AMQ and CJP initiated the study design and MD supported implementation. LH provided statistical oversight and will conduct the primary statistical analysis. HR drafted the manuscript; all authors offered critical revisions and approved the final manuscript. HR is the guarantor.

Funding This research reported in this article was supported by the National Heart, Lung, and Blood Institutional of the National Institutes of Health under award number R01HL171313. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The research was also supported by The Children's Hospital of Philadelphia Center for Pediatric Clinical Effectiveness pilot grant (2021-2022).

Disclaimer This funding sources had no role in the design of this study and will not have any role during its execution, analyses, interpretation of the data or decision to submit results.

Competing interests EEF discloses a previous relationship as a paid consultant for Medtronic, which ended in 2022. MBL-F is a paid consultant for Janssen Pharmaceuticals (work unrelated to the content of this manuscript). She is on the Boards of Directors for the Anesthesia Patient Safety Foundation and the Foundation for Anaesthesia Education and Research. MLO'B is a member of the Data Safety and Monitoring Board for iECURE, a gene therapy company. AMQ is a consultant for Medtronic, W.L. Gore and Associates and B. Braun. AMQ is also a member of the Data Safety and Monitoring Board for Medtronic. JF is a consultant for Micropore, Inc., Becton-Dickinson and GE Healthcare.

Patient and public involvement Patients and/or the public were involved in the design, conduct, reporting or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; peer reviewed for ethical and funding approval prior to submission.

ORCID iD
Halley Ruppel http://orcid.org/0000-0001-8326-5175

REFERENCES

1 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:S18–23.
2 Bickler PE, Feiner JR, Severinghaus JW. Effects of skin pigmentation on pulse oximeter accuracy at low saturation. *Anesthesiology* 2005;102:715–9.
3 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 Med* 2022;20:267.
4 Jubran A, Tobin MJ. Reliability of pulse oximetry in titrating supplemental oxygen therapy in ventilator-dependent patients. *Chest* 1990;97:1420–5.
5 Ries AL, Prewitt LM, Johnson JJ. Skin color and ear oximetry. *Chest* 1989;96:287–90.
6 Andrist E, Nuppnau M, Barbaro RP, *et al*. Association of Race With Pulse Oximetry Accuracy in Hospitalized Children. *JAMA Netw Open* 2022;5:e224584.
7 Ruppel H, Makeneni S, Faerber JA, *et al*. Evaluating the Accuracy of Pulse Oximetry in Children According to Race. *JAMA Pediatr* 2023;177:540–3.
8 Vesoulis Z, Tims A, Lodhi H, *et al*. Racial discrepancy in pulse oximeter accuracy in preterm infants. *J Perinatol* 2022;42:79–85.
9 Foglia EE, Whyte RK, Chaudhary A, *et al*. The Effect of Skin Pigmentation on the Accuracy of Pulse Oximetry in Infants with Hypoxemia. *J Pediatr* 2017;182:375–7.
10 Harris BU, Char DS, Feinstein JA, *et al*. Accuracy of Pulse Oximeters Intended for Hypoxemic Pediatric Patients. *Pediatr Crit Care Med* 2016;17:315–20.
11 Keller MD, Harrison-Smith B, Patil C, *et al*. Skin colour affects the accuracy of medical oxygen sensors. *Nature New Biol* 2022;610:449–51.
12 Sjoding MW, Ansari S, Valley TS. Origins of Racial and Ethnic Bias in Pulmonary Technologies. *Annu Rev Med* 2023;74:401–12.
13 Gitelson S. Determination of the skin color with the use of perforated Munsell color standards. “Visocolorimetry” of the skin. *Isr J Med Sci* 1965;1:994–1003.
14 Fitzpatrick TB. The validity and practicality of sun-reactive skin types I through VI. *Arch Dermatol* 1988;124:869–71.
15 Ross PA, Newth CJL, Khemani RG. Accuracy of pulse oximetry in children. *Pediatrics* 2014;133:22–9.
16 Gudelunas MK, Lipnick M, Hendrickson C, *et al*. Low Perfusion and Missed Diagnosis of Hypoxemia by Pulse Oximetry in Darkly Pigmented Skin: A Prospective Study. *Anesth Analg* 2024;138:552–61.
17 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.
18 Sjoding MW, Dickson RP, Iwashyna TJ, *et al*. Racial Bias in Pulse Oximetry Measurement. *N Engl J Med* 2020;383:2477–8.
19 Sjoding MW, Iwashyna TJ, Valley TS. Change the Framework for Pulse Oximeter Regulation to Ensure Clinicians Can Give Patients the Oxygen They Need. *Am J Respir Crit Care Med* 2023;207:661–4.
20 Batchelder PB, Raley DM. Maximizing the laboratory setting for testing devices and understanding statistical output in pulse oximetry. *Anesth Analg* 2007;105:S85–94.
21 Leon AC, Heo M. Sample Sizes Required to Detect Interactions between Two Binary Fixed-Effects in a Mixed-Effects Linear Regression Model. *Comput Stat Data Anal* 2009;53:603–8.
22 Rubin E, Editors. Striving for Diversity in Research Studies. *N Engl J Med* 2021;385:1429–30.
23 US Food and Drug Administration. Pulse oximeters—premarket notification submissions [510(k)s]: guidance for industry and Food and Drug Administration staff. 2013. Available: https://www.fda.gov/media/72470/download [Accessed 7 Feb 2025].
24 Pennello G. In Real-World Studies, Comparisons of Predictive Values Are Confounded by Prevalence. *Chest* 2022;162:e103.
25 United States Food and Drug Administration. FDA executive summary: review of pulse oximeters and factors that can impact their accuracy. 2022. Available: https://www.fda.gov/media/162709/download
26 Verkruysse W, Jaffe MB, Lipnick M, *et al*. Challenges of Subjective Skin Color Scales: The Case for the Use of Objective Pigmentation Measurement Methods in Regulatory Pulse Oximetry Studies. *Anesth Analg* 2024;139:e15–7.
27 Okunlola OE, Lipnick MS, Batchelder PB, *et al*. Pulse Oximeter Performance, Racial Inequity, and the Work Ahead. *Respir Care* 2022;67:252–7.
28 Food and Drug Administration. Pulse oximeters for medical purposes - non-clinical and clinical performance testing, labeling, and premarket submission recommendations: draft guidance for industry and Food and Drug Administration staff. 2025. Available: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/pulse-oximeters-medical-purposes-non-clinical-and-clinical-performance-testing-labeling-and
29 Ly BCK, Dyer EB, Feig JL, *et al*. Research Techniques Made Simple: Cutaneous Colorimetry: A Reliable Technique for Objective Skin Color Measurement. *J Invest Dermatol* 2020;140:3–12.
30 Chardon A, Cretois I, Hourseau C. Skin colour typology and suntanning pathways. *Int J Cosmet Sci* 1991;13:191–208.
31 Monk E. Monk skin tone scale. 2019. Available: https://skintone.google [Accessed 11 Mar 2025].
32 Schumann C, Olanubi GO, Wright A, *et al*. Consensus and subjectivity of skin tone annotation for ML fairness. *arXiv* 2024. Available: https://doi.org/10.48550/arXiv.2305.09073
33 Freed GL, Bogan B, Nicholson A, *et al*. Error Rates in Race and Ethnicity Designation Across Large Pediatric Health Systems. *JAMA Netw Open* 2024;7:e2431073.
34 Lima A, Jansen TC, van Bommel J, *et al*. The prognostic value of the subjective assessment of peripheral perfusion in critically ill patients. *Crit Care Med* 2009;37:934–8.
35 Jacquet-Lagrèze M, Wiart C, Schweizer R, *et al*. Capillary refill time for the management of acute circulatory failure: a survey among pediatric and adult intensivists. *BMC Emerg Med* 2022;22:131.
36 Fleming S, Gill P, Jones C, *et al*. The Diagnostic Value of Capillary Refill Time for Detecting Serious Illness in Children: A Systematic Review and Meta-Analysis. *PLoS One* 2015;10:e0138155.
37 Harris PA, Taylor R, Thielke R, *et al*. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.
38 Harris PA, Taylor R, Minor BL, *et al*. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform* 2019;95:103208.