Pulse oximeter performance and skin pigment: comparison of 34 oximeters using current and emerging regulatory frameworks

Caroline Hughes[1,2,3], Danni Chen MS[1,2], Tyler Law MD[1,2,4], Philip Bickler MD PhD[1,2], John Feiner MD[1,2], Leonid Shmuylovich MD PhD[5], Ella Behnke[1,2], Lily Ortiz MS[1,2,4], Gregory Leeb MBBS[1,2,4], Isabella Auchus MD[2], Fekir Negussie MPH[4], Ronald Bisegerwa MBChB MMed[4,6], René Vargas Zamora MS[1,2], Elizabeth Igaga MBChB MMed[4,6], Kelvin Moore, Jr. MD[1,4,7], Olubunmi Okunlola MD[8], Ellis Monk PhD[9], Jana Lyn Fernandez[1,4], Odinakachukwu Ehie MD[2,4], Bernadette Wilks MBBS[2,4], Koyinsola Oyefeso[1,4], Deleree Schornack[1,2], Cornelius Sendagire MBChB MMed[4,6,10], Michael S. Lipnick MD[1, 2, 4]

[1] University of California San Francisco Hypoxia Research Laboratory

[2] Department of Anesthesia and Perioperative Care, University of California San Francisco

[3] Saint Louis University School of Medicine

[4] University of California San Francisco Center for Health Equity in Surgery and Anesthesia

[5] Division of Dermatology, Department of Medicine, Washington University in Saint Louis

[6] Department of Anesthesia, College of Health Sciences, Makerere University

[7]Department of Surgery, Northwestern University Feinberg School of Medicine

[8] Department of Anesthesia, New York University Langone Hospital Brooklyn

[9] Department of Sociology, Harvard University

[10] Critical Care, D'Or Institute for Research and Education, Rio de Janeiro, Brazil

Corresponding Author:

Michael Lipnick

michael.lipnick@ucsf.edu

1001 Potrero Avenue

San Francisco, CA 94110

Acknowledgements

We are grateful to members of the Open Oximetry Collaborative Community who participated in the project's open forum discussions on protocols and approaches to data analysis, including James Ramsay, Jenna Lester, Alex Pogorzelski, Margaret Akey, Daryl Dorsey, Bob Kopotic and Sandy Weininger (ISO/IEC Conveners of the Oximeters Medical Device Standards), and many others.

Funding Statement: This study was conducted as part of the Open Oximetry Project funded by the Gordon and Betty Moore Foundation, Patrick J McGovern Foundation, PATH/UNITAID, and Robert Wood Johnson Foundation. Dr Ellis Monk's time utilized for data analysis, reviewing and editing was funded by grant number: DP2MH132941.

Conflicts of Interest: The UCSF Hypoxia Research Laboratory receives funding from multiple industry sponsors to test the sponsors' devices for the purposes of product

development and regulatory performance testing. This paper does not include data collected for sponsors. All data were collected from devices procured by the Hypoxia Research Laboratory for the purposes of independent research. No company provided any direct funding for this study, participated in study design, or was involved in analyzing data or writing the manuscript. None of the authors own stock or equity interests in any pulse oximeter companies.

Abbreviations and Acronyms:

A_{RMS} - Accuracy root mean square error

CI - Confidence interval

UCI - Upper confidence interval

DP - Dorsal distal phalanx

FDA - US Food and Drug Administration

ISO - International Organization for Standardization

ITA - Individual Typology Angle

KM - Konica Minolta

LME - Linear mixed effects (model)

LR - Linear regression (model)

MST - Monk Skin Tone (Scale)

NIH - National Institutes of Health

SaO_2 - Functional arterial hemoglobin oxygen saturation

SpO_2 - Pulse oximeter indirect measure of arterial hemoglobin oxygen saturation

UCSF - University of California San Francisco

Abstract

Background

The International Organization for Standardization (ISO) and US Food and Drug Administration (FDA) are updating regulations for pulse oximeters to reduce performance disparities linked to skin pigment. We tested common oximeters with current and anticipated regulatory frameworks. We hypothesized that not all oximeters show more positive bias in darkly vs lightly pigmented participants and that few oximeters would 'pass' the anticipated FDA regulations.

Methods

We used a controlled desaturation protocol to test 34 oximeters across arterial oxygen saturations (SaO_2) 70–100% in healthy adults. Based on what FDA and ISO had shared at the time of study design, we studied cohort sizes of $\geq$ 24 with $\geq$ 25% of participants being darkly pigmented. We used the subjective Monk Skin Tone (MST) scale and the objective individual typology angle (ITA) derived from a spectrophotometer to characterize skin pigment. The root mean square error (A_{RMS}), bias (mean of SpO_2 - SaO_2 error), and skin pigment differential bias were calculated. Monte Carlo simulation explored potential impacts of participant selection on device passing.

Results

For cohorts of 24 participants, 28/34 oximeters passed 2017 ISO Standard (A_{RMS} $\leq$ 4%), 22/34 passed 2013 FDA guidance (A_{RMS} $\leq$ 3%), 21/34 oximeters passed both

A_{RMS} and differential bias criteria for anticipated ISO standards, and 1/34 passed anticipated FDA criteria. More devices passed with cohorts > 24. Eleven oximeters had more positive bias in participants with dark vs. light (dorsal finger) pigmentation across 70–100% SaO_2. Eighteen devices could pass or fail depending on cohorts selected for analysis.

Conclusions

Pulse oximeters show variable performance across manufacturers and models. Notably, only some devices show more positive bias in people with darker skin. Anticipated updates to ISO and FDA frameworks yield strikingly different assessments and require refinement of cohort sizes and differential bias criteria. Whether new guidelines will translate into improved real-world performance or reduced health disparities is yet to be determined.

Introduction

Pulse oximeters are essential clinical tools that noninvasively estimate the percent of hemoglobin bound with oxygen (SpO_2). However, long-standing concerns over variable performance across manufacturers and worse performance in people with darker skin have only recently received significant attention. Reports of oximeter inaccuracies in people with darker skin pigmentation began in the 1980s, and the COVID-19 pandemic brought this issue to the forefront as numerous studies reported not only positive bias (i.e. SpO_2 overestimating true arterial blood functional oxygen saturation, SaO_2) but also healthcare disparities (i.e. underrecognition or undertreatment of hypoxemia) in patients who self-identified as Black, Asian, Hispanic, or Native American.[1–4] Laboratory studies in healthy participants have shown similar findings for some oximeters, especially at lower SaO_2.[5–8]

In response to these concerns, the International Organization for Standardization (ISO) and US Food and Drug Administration (FDA) began updating pulse oximeter regulatory frameworks to improve performance and reduce potential disparities related to skin pigment. The FDA 510(k) premarket notification guidance for pulse oximeters and ISO 80601-2-61 standards were last updated in 2013 and 2017, respectively, and recommend manufacturers verify pulse oximeter performance using “10 or more healthy participants that vary in age and gender.”[9,10] The 2013 FDA guidance also recommended at least 15% of the participants (i.e., the proportion of the US population in 2013 who identified as African American) should have “darkly pigmented” skin. As previously reported, these recommendations are likely underpowered to ensure equitable device performance.[11] In 2024, the FDA released a discussion paper

proposing larger verification study cohorts (24 participants) with more diversity. In 2025, FDA released a new draft guidance recommending even larger cohorts (150 participants), with more diversity of skin pigment, and additional statistical analyses.[12] The ISO has shared some details of anticipated changes via public forums.[13] Both agencies are expected to imminently release finalized versions of their new regulatory frameworks for pulse oximeters.[13–16]

Considering these upcoming changes and ongoing uncertainty about the extent to which oximeters on the market have bias related to skin color, we sought to independently test the performance of 34 commonly used pulse oximeters with both current and anticipated regulatory frameworks. We hypothesized that relatively few devices would pass anticipated frameworks, and that most but not all pulse oximeters would show more positive bias in people with dark vs. light pigmentation.

Methods

This study was conducted at the University of California San Francisco (UCSF) Hypoxia Laboratory from 2022 to 2024 with UCSF IRB approval (#21-35637, ClinicalTrials.gov ID: NCT06142019). Written informed consent was obtained from all participants.

Oximeter selection

We tested 34 pulse oximeters. Devices were chosen based on popularity in online marketplaces, clinician input from diverse settings, and global health donor procurement lists to capture commonly used devices across varied clinical settings and price points. Purchase prices varied from $10 to $5,999. A list of tested devices, including prices, form factors, wavelengths, and regulatory data is available in Table S4. For three devices (the Shenzhen PC-60NW, Masimo MightySat, and Acare AH-M1/MX), we purchased two of the same model and reported results by year of device purchase. For two additional devices (Masimo Rad97 and Masimo Rad G), we tested each with two different probes and reported each device-probe combination separately. All participants were monitored throughout the protocol with the lab's 'clinical monitor' oximeter (Nellcor PM1000N, Medtronic, USA).

Study demographics and skin pigment assessment

Participants were healthy adults 18–50 years old who were non-smoking, with no history of lung, cardiovascular, kidney, or liver disease and without hemoglobinopathy,

anemia, clotting disorders, or Raynaud's disease. Our enrollment targeted diversity of both skin pigment and sex assigned at birth.

Participants' age, sex, and US National Institutes of Health (NIH) race were self-reported. Participants' height, weight, and finger diameter were measured. Percent modulation of infrared light (an indicator of participant perfusion) was recorded from the clinical monitor and divided by 10 to approximate comparability with Masimo Perfusion Index.[8]

Our approach to skin pigment assessment is provided in the Supplemental Digital Content eMethods. Briefly, two research coordinators assigned subjective skin pigmentation data for each participant using the Monk Skin Tone (MST) scale.[17,18] Coordinators also used the Konica Minolta CM 700-d (KM) spectrophotometer to derive individual typology angle (ITA), a frequently used surrogate for melanin[19,20] at the study participant's dorsal distal phalanx (DP) (the site for study oximeters) and forehead (the recommended site for using MST).[19]

We categorized participants into light, medium or dark bins based on previously proposed cutoffs: 'light' (MST 1–4 and ITA of > 30°), 'medium' (MST 5–7 and ITA between 30° and -30°), 'dark' (MST 8–10 and ITA of < -30° with ≥ 50% of these participants having ITA < -50°).[12,21,22] We refer to participants with both MST 8–10 and ITA of < -50° as 'very dark.' Unless otherwise noted, we used the 'dark' definition above (not 'very dark') for analyses. In cases of discordance between ITA and MST, the participant was binned by ITA.

Controlled desaturation protocol

Each participant underwent a controlled desaturation protocol as previously described and available online and in the Supplemental Digital Content eMethods. [23–25] Briefly, study investigators controlled partial pressures of inspired oxygen, carbon dioxide, and nitrogen to achieve six stable "plateaus" of targeted arterial functional oxygen saturation (SaO_2) between SaO2 ~70% and 100% (i.e. 2 plateaus in each decile 70-80%, 80-90%, and 90-100%). At each plateau, multiple arterial blood samples were collected $\geq$ 20 seconds apart from a radial artery catheter. At the time of each blood sample, pulse oximeter SpO_2's were recorded, and blood samples were immediately analyzed for SaO_2 using Radiometer ABL90 Flex Plus (ABL) (Radiometer, Copenhagen, Denmark) blood gas analyzers. Of note, participants' hands were not warmed by default, and most probes were randomly placed on all five finger digits (see the Supplemental Digital Content eMethods).

Sample size and cohort composition

The cohort size and composition for the primary analysis were based on what had been shared publicly by the FDA and ISO at the time of study design.[13–16,26] These criteria were: 1. $\geq$ 24 unique participants; 2. Each participant must contribute 16 to 30 data points; 3. Pooled SaO_2 data must span at least 73% to 97% SaO_2; 4. $\geq$ 90% of participants must provide $\geq$ 1 data point < 85%; 5. $\geq$ 69% of participants must provide $\geq$ 1 data point in the 70–80% decile; 6. The cohort must include $\geq$ 33% of each sex; 7. Using forehead skin data, $\geq$ 25% of participants must fall into each color bin (i.e. light, medium, or dark). Given uncertainties in anticipated recommendations for cohort size,

we tested all devices in at least 24 participants and continued testing with as many participants as possible based on available resources.

Statistical and sensitivity analyses

The primary analysis used 24 participant cohorts and skin pigment data from the forehead and dorsal distal phalanx (DP). We used two metrics to define a 'passing device' based on what was shared publicly by FDA and ISO at the time of study design (Table 2): 1. accuracy root mean square error (A_{RMS}); and 2. differential bias. A_{RMS} is a performance threshold that measures the square root of the mean of the squared differences in SpO_2 minus SaO_2. Differential bias is defined in two ways: 1. For ITA, as the difference in SpO_2 bias between two theoretical participants with an ITA difference of 100° (i.e., one participant with dark ITA -50° and another with light ITA 50°); and 2. For MST, as the difference between MST bins. We estimated the differential bias and 95% confidence interval (CI) using a linear mixed effects (LME) model. The anticipated ISO standard recommended differential bias thresholds (point estimate $\leq$ 4% for 70–85% SaO_2 and $\leq$ 2% for 85–100% SaO_2) for only ITA, while the anticipated FDA guidance recommended thresholds (95% CI < 3.5% for 70–85% SaO_2 and < 1.5% for 85–100% SaO_2) for both ITA and pairwise comparisons of MST bins 1–4, 5–7, and 8–10.[12,15,16]

We assessed for bias related to pigment by calculating A_{RMS} and error (SpO_2 minus SaO_2) for each pigment bin (forehead and DP).[12] Error and A_{RMS} were summarized as means with 95% CIs. The 95% CIs were determined using bootstrapping (random resampling with replacement) with 1,000 repetitions. Error

across skin color bins was compared using an LME model, which included a random intercept to account for repeated measures within participants. A_{RMS} was computed for each participant and compared between groups of participants based on pigment binning using a Welch's two-sample t-test. A two-sided p-value < 0.05 was considered significant. P-values are reported unadjusted for multiple comparisons in accordance with FDA guidance for independent pairwise analyses, which does not mandate multiplicity adjustments.

We conducted a secondary analysis for each device using the largest cohorts possible that still met cohort diversity criteria. Additional secondary analyses included the impact on device 'passing' of different anatomical pigment assessment sites, different definitions of pigment binning, and different statistical methods for estimating bias, including use of a linear regression (LR) model for differential bias instead of the LME model. The LR model regressed the participant-level bias on the ITA values and estimated the difference in predicted bias between ITA values of -50° and 50° based on the fitted model.

A Monte Carlo simulation explored if participant selection bias could influence 'passing' status. For each device, Monte Carlo simulation created 100 cohorts of 24 participants, with at least six darkly pigmented participants. If a device always passed or always failed in all simulated cohorts, it was determined to be less susceptible to selection bias (i.e. 'cherrypicking' participants).

The SpO_2 from most study oximeters was recorded manually into a Research Electronic Data Capture (REDCap) Database at the time of each arterial blood sample, except for the Nellcor PM1000N and Masimo Rad97 which streamed data directly into

Labview (National Instruments, Austin, TX). Most oximeters used a self-contained interface that displayed SpO_2 data on the oximeter itself and could not directly stream data from the oximeter. All other physiologic data were collected into Labview, and all Labview data were merged into the REDCap Database. All analyses were performed with R 4.3.2 (R Core Team, 2023) and Python v3.9.6 (Python Software Foundation, 2024). De-identified data for this study are openly accessible through the Open Oximetry Data Repository via PhysioNet and accessible via the OpenOximetry.org website.[27]

Results

We completed 348 desaturation studies in 155 participants (30,174 paired SpO_2:SaO_2 samples). Each device was tested in $\geq$ 24 unique participants (median 38, range 24–120) with > 540 SpO_2:SaO_2 pairs approximately equally distributed across SaO_2 deciles, and with > 25% of data points in each decile coming from participants in each skin color bin (Table S3). Demographics are reported in Table 1.

Overall oximeter performance

For cohorts of 24 participants, device A_{RMS} ranged from 1.69 to 8.04 (Figure 1). Of the 12 devices with A_{RMS} > 3% (i.e. 2013 FDA Guidance threshold), eight had 510(k) premarket clearances. All six devices with an A_{RMS} > 4% (i.e. 2017 ISO Standard threshold) reported a European Conformity (CE) marking (Figure 1 and Table S4).

Device SpO_2 bias ranged from -5.71 to 3.48 with a median of 0.97 (IQR: -0.04 – 1.67) (Table S2). We found 24 devices with overall positive bias, and 10 devices with overall negative bias.

A_{RMS} by skin color bin

In cohorts of 24, we analyzed A_{RMS} by DP pigmentation across 70–100% SaO_2 and found four devices with significantly different A_{RMS} between light and dark skin color bins (Table S1 and Figure 2). We found significantly higher A_{RMS} in the dark (DP) skin color bin for two devices at 70–80% SaO_2, four devices at 80–90% SaO_2, and two devices at 90–100% SaO_2. One device showed significantly higher A_{RMS} in the light skin color bin (Table S1).

When analyzing A_{RMS} by forehead pigmentation in cohorts of 24, two devices showed significantly higher A_{RMS} in the dark vs. light color bin at 70–100% SaO_2 (Table S6), with similar but not exact results to the DP site when broken down by SaO_2 deciles.

With cohorts larger than 24, more oximeters demonstrated statistically significant A_{RMS} differences between light and dark bins (Table S1, Table S6, Table S8, Table S9). Two devices demonstrated statistically significant differences in A_{RMS} only when comparing light and 'very dark' color bins (Table S9).

Bias by skin color bin

We found five devices with more positive bias in dark vs. light DP skin color bins across SaO_2 70–100% (median difference in bias (dark minus light) 1.56, IQR: 1.30 – 2.70), and two devices with more positive bias in the light bin (Table S2 and Figure 2).

When analyzing by SaO_2 deciles, we found significantly more positive bias in dark vs light skin color bins for four devices at 70–80% SaO_2 (median difference 4.22, IQR: 3.40 – 4.74), four devices at 80–90% SaO_2 (median difference 2.11 IQR: 1.60 – 2.85) and two devices at 90–100% SaO_2 (median difference 1.36, range: 1.20 – 1.51). Bias by forehead pigmentation also varied (Table S7).

Analysis of larger cohort sizes revealed more devices (11/34) with significant differences in bias when comparing light vs. dark bins at the DP across SaO_2 70–100%. Four devices demonstrated significantly more positive bias when comparing 'very dark' vs light color bins, but not dark vs light (Table S10 and Figure 3).

Figures 4-S13 contain the distribution of A_{RMS} and bias for all devices across all saturation ranges and both anatomical sites.

Differential bias

In cohorts of 24 using DP ITA, differential bias ranged from -1.39 to 9.17 at 70–85% SaO_2 (median 1.50, IQR: 0.00 – 2.53) and -1.12 to 2.18 at 85–100% SaO_2 (median 0.42, IQR: -0.37 – 0.94) (Table S12).

With the larger cohort sizes, DP ITA differential bias ranged from -1.34 to 4.83 at 70–85% SaO_2 (median 1.36, IQR: 0.01 – 2.33) and from -0.92 to 1.72 at 85–100% SaO_2 (median of 0.38, IQR: -0.10 – 0.96) (Table S13). Forehead ITA differential bias was similar. The LME and LR models produced similar results.

Old vs. anticipated regulatory frameworks

With cohorts of 24 participants, 28/34 devices passed 2017 ISO A_{RMS} criteria ($\leq$4%), 22/34 passed 2013 FDA A_{RMS} ($\leq$ 3%) criteria (which is also the anticipated ISO criteria), and 12/34 passed anticipated FDA A_{RMS} criteria (95% UCI < 3%). For differential bias, 1/34 oximeters passed the anticipated FDA criteria, and 32/34 passed the anticipated ISO criteria (Table S12 and Figure 1). Overall, 21/34 oximeters passed both the A_{RMS} and differential bias criteria for anticipated ISO standard, and 1/34 oximeters passed both criteria for anticipated FDA guidance.

With the larger cohort sizes, the number of devices passing both the anticipated A_{RMS} and differential bias criteria increased from 1/34 to 7/34 for FDA, and from 21/34 to 25/34 for ISO (Table S3 and Table S13).

Impact of pigmentation definitions and measurement sites on bias

The use of race, MST or ITA with varied binning thresholds and different anatomical sites yielded different numbers of devices with significant differences in bias (Table S14). Analysis of light vs dark DP ITA in maximum cohort sizes identified the most devices with significant differences in bias (11), while analysis by race (white vs. Black/African American) identified the fewest.

Monte Carlo Simulation

We found numerous devices could ‘pass’ or ‘fail’ A_{RMS} criteria depending on which participants were selected for cohort analysis (Figure S3).

Discussion

We found that many pulse oximeters fail to meet current or anticipated regulatory performance recommendations, and some but not all oximeters had more positive bias in dark vs light participants. Anticipated ISO and FDA regulatory frameworks produced markedly different 'passing' rates for the same devices. Most devices passed the anticipated ISO standard, yet only one device passed the anticipated FDA guidance (in cohorts of 24 healthy adults).

Our findings corroborate prior studies demonstrating variable performance and positive bias in some oximeters on the market.[6,7,23,24,28,29] In several instances, our findings differ from manufacturers' reports (Figure 1 and Table S4), which may be partly explained by methodological differences. For example, we did not warm most participants' hands and used all five digits. This was done in an effort to better reflect clinical reality and contrasts with most prior oximeter regulatory verification studies, which typically only used digits 2-4 and actively warmed participants' hands to increase peripheral perfusion (a factor known to improve device performance and permitted by FDA and ISO).[8] Additionally, we analyzed consecutively enrolled participants. Our Monte Carlo analysis suggests that had we 'cherry-picked' participants for analysis, a practice not explicitly prohibited for regulatory submissions, we could make poorly performing oximeters look better.

Consistent with our hypothesis, we found most (18/34) but not all oximeters exhibited more positive bias (i.e. SpO_2 overestimating SaO_2) in participants with dark vs. light skin (Table S10 and Table S11). Unexpectedly, several devices demonstrated more positive bias in participants with light vs. dark skin.

We confirmed our hypothesis that fewer devices pass anticipated regulatory frameworks, but the contrast in differential bias criteria between FDA and ISO is particularly noteworthy. Nearly all devices we tested passed the anticipated differential bias criteria for ISO, but only one passed anticipated FDA criteria (Figure 1). The anticipated ISO standard performed similarly to the 2013 FDA guidance. This is concerning given that devices cleared by 2013 FDA guidance were used in studies demonstrating oximeter performance and health disparity concerns[4,30] ISO should tighten its recommendations and ensure differential bias thresholds reflect magnitudes of bias (e.g. <2%) seen in clinically relevant SpO2 ranges (e.g. 85-95%).[31–37]The FDA differential bias thresholds necessitate larger study cohorts for some devices to pass. This has generated concern that if studies become too costly or time-consuming, there could be negative global market implications (e.g. decreased device access and increased costs).[13,38]

Increasing the size of verification cohorts (from 10 to 24) and utilizing MST and ITA are good steps toward improving diversity. Our findings suggest that most devices will require cohorts > 24 to pass anticipated FDA criteria, but a cohort size of 150, as proposed in the January 2025 draft FDA guidance, is likely unnecessary. Additionally, we found the anatomical site for skin pigment assessment and definition of “dark” skin pigment also impacted bias, A_{RMS}, and regulatory ‘passing’ rates (Table S1 and Table S15).

Without a better understanding of how performance criteria in the lab correlate with performance (or health disparities) in the clinical setting, the optimal regulatory performance thresholds will remain uncertain.

There was no device characteristic (e.g. cost, number of wavelengths, form factor etc.) that was clearly linked to performance, though the study was not designed for this purpose.

This study had several limitations. First, we generally tested only one oximeter per model. For devices with detachable probes, we replaced probes after approximately 24 participants. Thus, we cannot assess variability related to manufacturing differences or wear and tear. For the three oximeter models, we found different performance across manufactured years despite no obvious differences in hardware or software (Table S16, Figure 1, and Table S5). Our clinical monitor (PM1000N) had reasonably consistent performance in these same participant cohorts, suggesting this was not a methodological problem. Further work is needed to analyze variability across different probes of the same model, a potentially important source of error[13] Second, devices were tested in comparable but not identical cohorts. For example, some cohorts (including those for two devices with the best A_{RMS}) had higher participant perfusion (Table S17). Third, several aspects of our desaturation protocol could have influenced findings. For example, not all oximeters are designed for use on the 1st and 5th digits. Plateau stability was partially defined by stability of the clinical monitor, which may not always align with stability of tested oximeters. To mitigate this, we also used stability of calculated saturation (ScO_2)[23] based on exhaled CO_2 and O_2 partial pressures, to define plateau stability. Another protocol-related limitation is that SaO_2 can vary by 1-3% across co-oximeter (i.e. blood gas analyzer) brands.[39] We could not account for this factor because manufacturers generally do not disclose which co-oximeter was used in testing. Finally, there are also limitations in our skin pigment assessment protocol as

previously reported.[40] Although individuals with ITA < -50° represent a critical group for ensuring devices work equitably across all skin tones, we were unable to enroll a sufficient number of participants in this category to meaningfully assess its impact on bias analyses. Generalizability of our findings to the clinical setting, pediatric patients or devices we did not test should be done with caution.

Conclusion

Clinicians should be aware that many oximeters do not meet current or anticipated regulatory performance recommendations, and some, but not all, devices show more positive bias in people with dark vs light skin pigment. While it may be uncertain to regulators how much differential bias is reasonable to allow, it is clear that some oximeters appear to have minimal or no significant bias related to pigment, and all manufacturers should aspire to this performance goal regardless of the thresholds regulators ultimately propose. Future studies are needed to optimize regulatory frameworks, though many improvements can be made immediately using available data. Despite performance limitations of regulatory-compliant oximeters, clinicians should continue to utilize these essential tools but use great caution when using absolute SpO_2 cutoffs for providing or withholding treatments.

References

1. Ruppel H, Makeneni S, Faerber JA, et al. Evaluating the Accuracy of Pulse Oximetry in Children According to Race. *JAMA Pediatr*. 2023;177(5):540-543. doi:10.1001/jamapediatrics.2023.0071

2. Wong AKI, 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 Netw Open*. 2021;4(11):e2131674. doi:10.1001/jamanetworkopen.2021.31674

3. 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. *Crit Care Med*. 2022;50(2):204-211. doi:10.1097/CCM.0000000000005394

4. Sjoding MW, Dickson RP, Iwashyna TJ, Gay SE, Valley TS. Racial Bias in Pulse Oximetry Measurement. *N Engl J Med*. 2020;383(25):2477-2478. doi:10.1056/NEJMc2029240

5. Zeballos RJ, Weisman IM. Reliability of Noninvasive Oximetry in Black Subjects during Exercise and Hypoxia. *Am Rev Respir Dis*. 1991;144(6):1240-1244. doi:10.1164/ajrccm/144.6.1240

6. Feiner JR, Severinghaus JW, Bickler PE. Dark Skin Decreases the Accuracy of Pulse Oximeters at Low Oxygen Saturation: The Effects of Oximeter Probe Type and Gender. *Anesth Analg*. 2007;105(6):S18. doi:10.1213/01.ane.0000285988.35174.d9

7. Bickler PE, Feiner JR, Severinghaus JW. Effects of Skin Pigmentation on Pulse Oximeter Accuracy at Low Saturation. *Anesthesiology*. 2005;102(4):715-719. doi:10.1097/00000542-200504000-00004

8. 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(3):552. doi:10.1213/ANE.0000000000006755

9. ISO 80601-2-61:2017. ISO. Accessed June 4, 2024. https://www.iso.org/standard/67963.html

10. Pulse Oximeters - Premarket Notification Submissions [510(k)s]: Guidance for Industry and Food and Drug Administration Staff. April 11, 2024. Accessed June 4, 2024. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/pulse-oximeters-premarket-notification-submissions-510ks-guidance-industry-and-food-and-drug

11. Okunlola OE, Lipnick MS, Batchelder PB, Bernstein M, Feiner JR, Bickler PE. Pulse Oximeter Performance, Racial Inequity, and the Work Ahead. *Respir Care*. 2022;67(2):252-257. doi:10.4187/respcare.09795

12. Health C for D and R. Pulse Oximeters for Medical Purposes - Non-Clinical and Clinical Performance Testing, Labeling, and Premarket Submission Recommendations. January 10, 2025. Accessed January 12, 2025. https://www.fda.gov/regulatory-information/search-fda-

guidance-documents/pulse-oximeters-medical-purposes-non-clinical-and-clinical-performance-testing-labeling-and

13. openoximetry.org. Collaborative Community. OpenOximetry. Accessed July 12, 2024. https://openoximetry.org/community/

14. FDA Executive Summary: Performance Evaluation of Pulse Oximeters Taking into Consideration Skin Pigmentation, Race and Ethnicity. Published online February 2, 2024:49. https://www.fda.gov/media/175828/download

15. *Stakeholder Consultation on Pulse Oximeters: Next Steps and Closing Remarks*.; 2024. Accessed June 26, 2024. https://www.youtube.com/watch?v=2MqyMdrSZfs

16. *Stakeholder Consultation on Pulse Oximeters: Updating Global Requirements for Pulse Oximeters*.; 2024. Accessed June 26, 2024. https://www.youtube.com/watch?v=zSl5OWkOyck

17. Hypoxia Lab Skin Color Assessment Protocol. Google Docs. Accessed February 10, 2025. https://docs.google.com/document/d/17SGVjf1phxyhBB3TvetyuOFN_UC5z2humiEHI4dcEro/edit?usp=sharing&usp=embed_facebook

18. Skin Tone Research @ Google AI | undefined. Accessed June 10, 2024. https://www.skintone.google/mste-dataset

19. Vasudevan S, Vogt WC, Weininger S, Pfefer TJ. Melanometry for objective evaluation of skin pigmentation in pulse oximetry studies. *Commun Med*. 2024;4(1):1-19. doi:10.1038/s43856-024-00550-7

20. Stamatas GN, Kollias N. Blood stasis contributions to the perception of skin pigmentation. *J Biomed Opt*. 2004;9(2):315-322. doi:10.1117/1.1647545

21. *Skin Color Subgroup Meeting 3- March 13, 2024*.; 2024. Accessed June 9, 2025. https://www.youtube.com/watch?v=6ELI__1XcbI

22. Del Bino S, Bernerd F. Variations in skin colour and the biological consequences of ultraviolet radiation exposure. *Br J Dermatol*. 2013;169(s3):33-40. doi:10.1111/bjd.12529

23. Leeb G, Auchus I, Law T, et al. The performance of 11 fingertip pulse oximeters during hypoxemia in healthy human participants with varied, quantified skin pigment. *eBioMedicine*. 2024;102. doi:10.1016/j.ebiom.2024.105051

24. Lipnick MS, Feiner JR, Au P, Bernstein M, Bickler PE. The Accuracy of 6 Inexpensive Pulse Oximeters Not Cleared by the Food and Drug Administration: The Possible Global Public Health Implications. *Anesth Analg*. 2016;123(2):338-345. doi:10.1213/ANE.0000000000001300

25. Hypoxia lab study protocol basic pulse ox hypoxemia performance study (Public comment). Google Docs. Accessed July 1, 2024. https://docs.google.com/document/d/1bpVzZg9M0StLoyB_iF07V8at0h8lQhDyYnNCoSMVZQU/edit?usp=sharing&usp=embed_facebook

26. Approach for Improving the Performance Evaluation of Pulse Oximeter Devices Taking into Consideration Skin Pigmentation, Race and Ethnicity: Discussion Paper and Request for Feedback.

27. openoximetry.org. Data Repository. OpenOximetry. Accessed March 21, 2025. https://openoximetry.org/data-repository/

28. Blanchet MA, Mercier G, Delobel A, et al. Accuracy of Multiple Pulse Oximeters in Stable Critically Ill Patients. *Respir Care*. 2023;68(5):565-574. doi:10.4187/respcare.10582

29. 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 Respir Res*. 2021;8(1):e000939. doi:10.1136/bmjresp-2021-000939

30. 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 Intern Med*. 2022;182(7):730. doi:10.1001/jamainternmed.2022.1906

31. Martin D, Johns C, Sorrell L, et al. Effect of skin tone on the accuracy of the estimation of arterial oxygen saturation by pulse oximetry: a systematic review. *Br J Anaesth*. 2024;132(5):945-956. doi:10.1016/j.bja.2024.01.023

32. Graham HR, King C, Rahman AE, et al. Reducing global inequities in medical oxygen access: the Lancet Global Health Commission on medical oxygen security. *Lancet Glob Health*. 2025;13(3):e528-e584. doi:10.1016/S2214-109X(24)00496-0

33. The BOOST II United Kingdom, Australia, and New Zealand Collaborative Groups. Oxygen Saturation and Outcomes in Preterm Infants. *N Engl J Med*. 2013;368(22):2094-2104. doi:10.1056/NEJMoa1302298

34. Oster ME, Pinto NM, Pramanik AK, et al. Newborn Screening for Critical Congenital Heart Disease: A New Algorithm and Other Updated Recommendations: Clinical Report. *Pediatrics*. 2024;155(1):e2024069667. doi:10.1542/peds.2024-069667

35. Siemieniuk RAC, Chu DK, Kim LHY, et al. Oxygen therapy for acutely ill medical patients: a clinical practice guideline. Published online October 24, 2018. doi:10.1136/bmj.k4169

36. World Health Organization. Clinical management of severe acute respiratory infection (SARI) when COVID-19 disease is suspected. Interim guidance. Published online March 13, 2020. Accessed July 21, 2025. https://www.who.int/docs/default-source/coronaviruse/clinical-management-of-novel-cov.pdf

37. Moore KL, Gudelunas K, Lipnick MS, Bickler PE, Hendrickson CM. Pulse Oximeter Bias and Inequities in Retrospective Studies—Now What? *Respir Care*. 2022;67(12):1633-1636. doi:10.4187/respcare.10654

38. Lipnick MS, Ehie O, Igaga EN, Bicker P. Pulse Oximetry and Skin Pigmentation—New Guidance From the FDA. *JAMA*. 2025;333(16):1393-1395. doi:10.1001/jama.2025.1959

39. Gehring H, Duembgen L, Peterlein M, Hagelberg S, Dibbelt L. Hemoximetry as the “Gold Standard”? Error Assessment Based on Differences Among Identical Blood Gas Analyzer Devices of Five Manufacturers. *Anesth Analg*. 2007;105(6):S24. doi:10.1213/01.ane.0000268713.58174.cc

40. Lipnick MS, Chen D, Law T, et al. Comparison of methods for characterizing skin pigment diversity in research cohorts. *Br J Dermatol*. Published online October 10, 2025:ljaf397. doi:10.1093/bjd/ljaf397

Figures

Figure 1

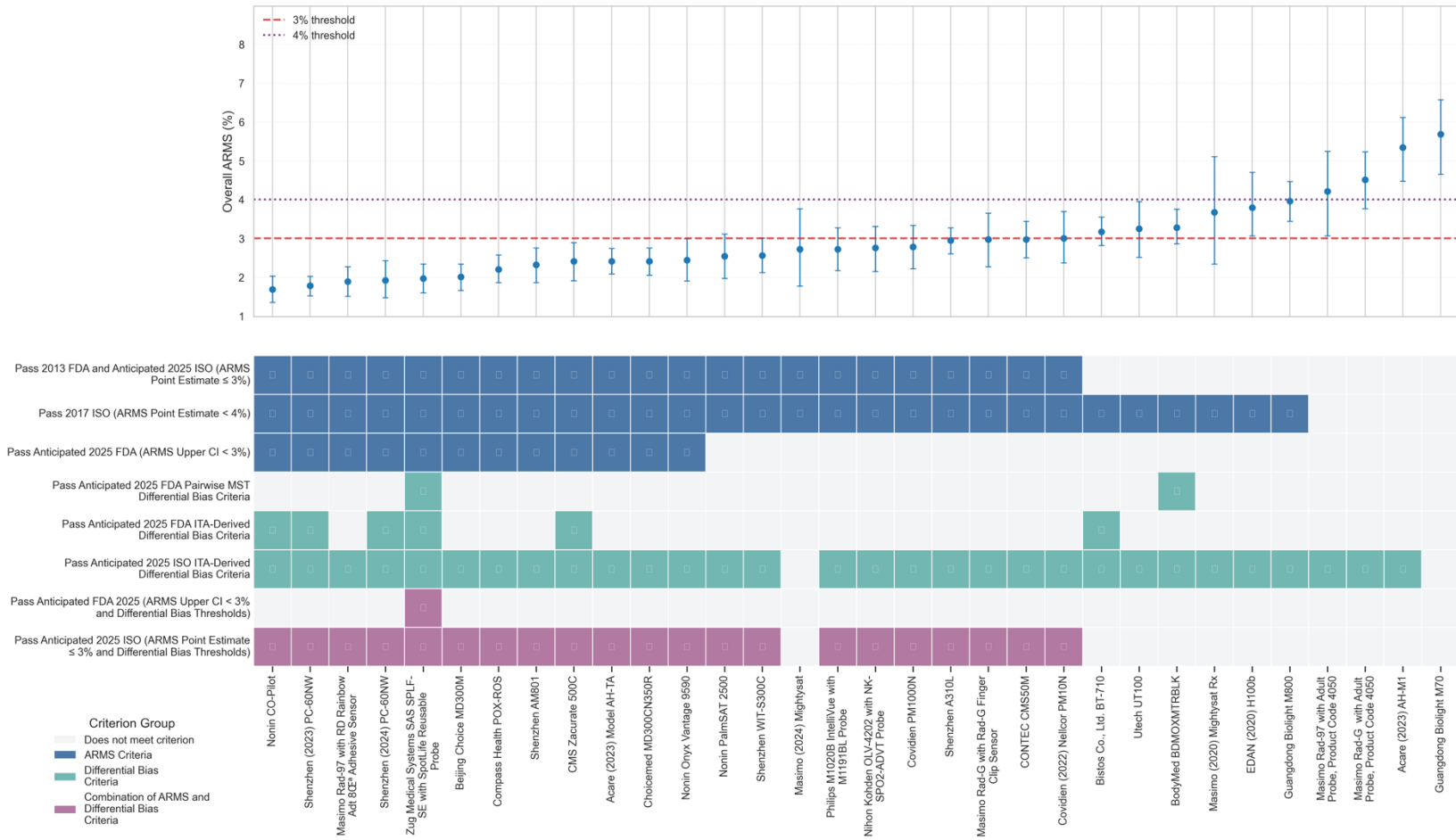

Figure 1. Device conformity with current and anticipated FDA and ISO performance thresholds for 24 participant cohorts and DP ITA

This figure displays devices from lowest to highest accuracy root mean square error (A_{RMS}) with 95% confidence intervals (CI) for each device tested in a cohort of 24 participants. The boxes marked with colors indicate that the device meets the A_{RMS} or differential bias thresholds (differential bias and 95% CI < 3.5% for 70–85% SaO_2 and < 1.5% for 85–100% SaO_2 for 2025 FDA, and differential bias point estimate $\leq$ 4% for 70–85% SaO_2 and $\leq$ 2% for 85–100% SaO_2 for 2025 ISO). Differential bias was calculated with the Monk Skin Tone (MST) Scale and individual typology angle (ITA) as indicated. The 95% CI were determined using bootstrapping (random resampling with replacement) with 1,000 repetitions.

Figure 2

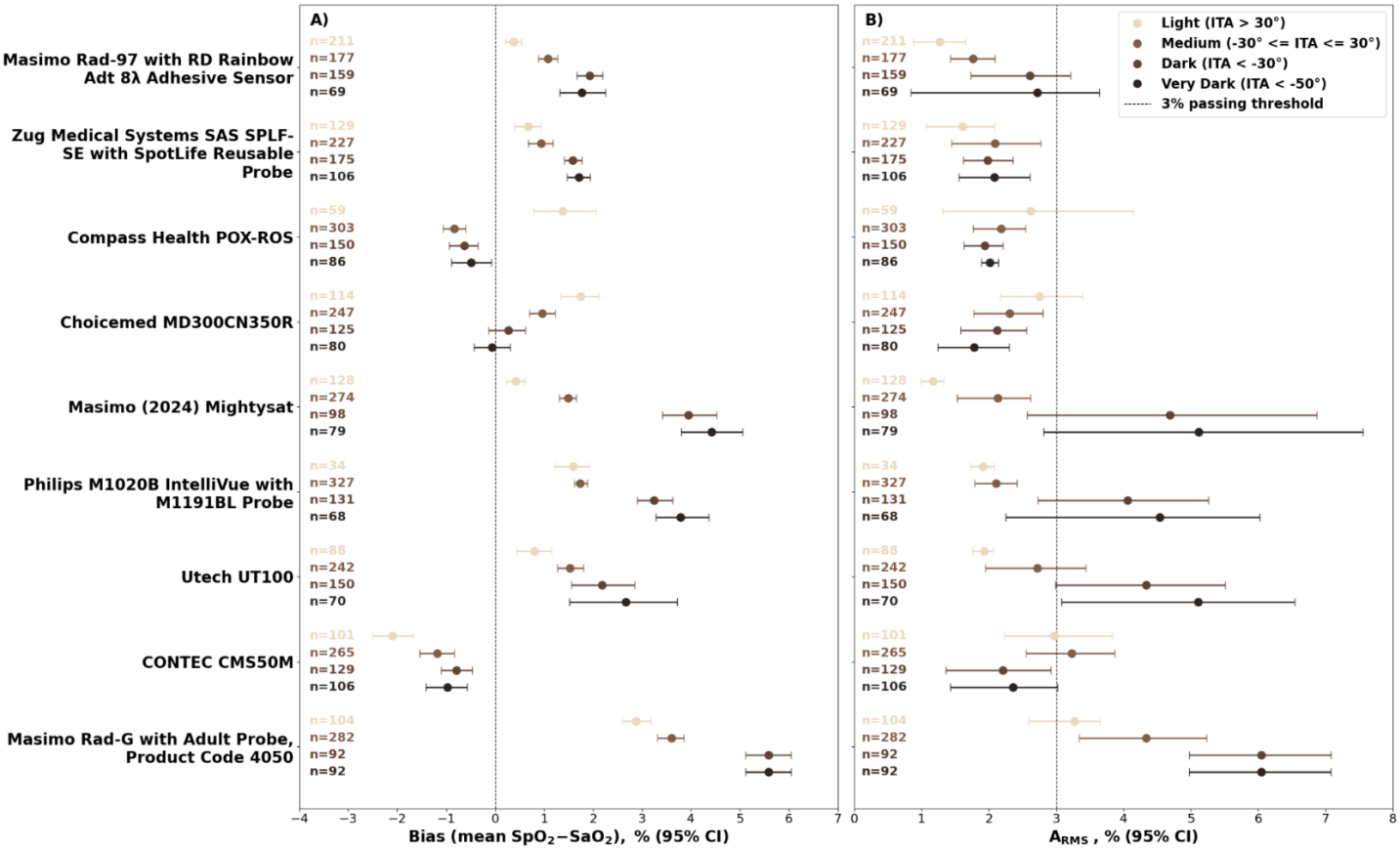

Figure 2. Distribution of bias and A_{RMS} across ITA bins at the DP site over 70-100% saturation range in 24 participant cohorts

This figure shows the distribution of pulse oximeter measured oxygen saturation (SpO_2) bias (mean of SpO_2 minus SaO_2 error, where SaO_2 is functional arterial oxygen saturation) and root mean square error (A_{RMS}) in each skin color bin defined by Individual Typology Angle (ITA) measured at the dorsal distal phalanx (DP) site across the saturation range of 70-100%. Devices shown are selected with statistically significantly differences in either bias or A_{RMS} between light (ITA >30°) and dark (ITA <-30°) bins. Panel A displays the distribution of bias in closed circles with 95% confidence interval (CI) whiskers. Panel B shows the distribution of A_{RMS} in closed circles with 95% CI whiskers. Data are stratified into four ITA bins (top to bottom): ITA > 30°, -30° ≤ ITA ≤ 30°, ITA < -30°, and ITA < -50°; note that ITA < -50° is a subset of ITA < -30°. Sample size (n) for each bin is indicated at left in the corresponding color. The vertical dashed line in Panel A represents bias of 0 and in Panel B represents the A_{RMS} 3% threshold.

Figure 3

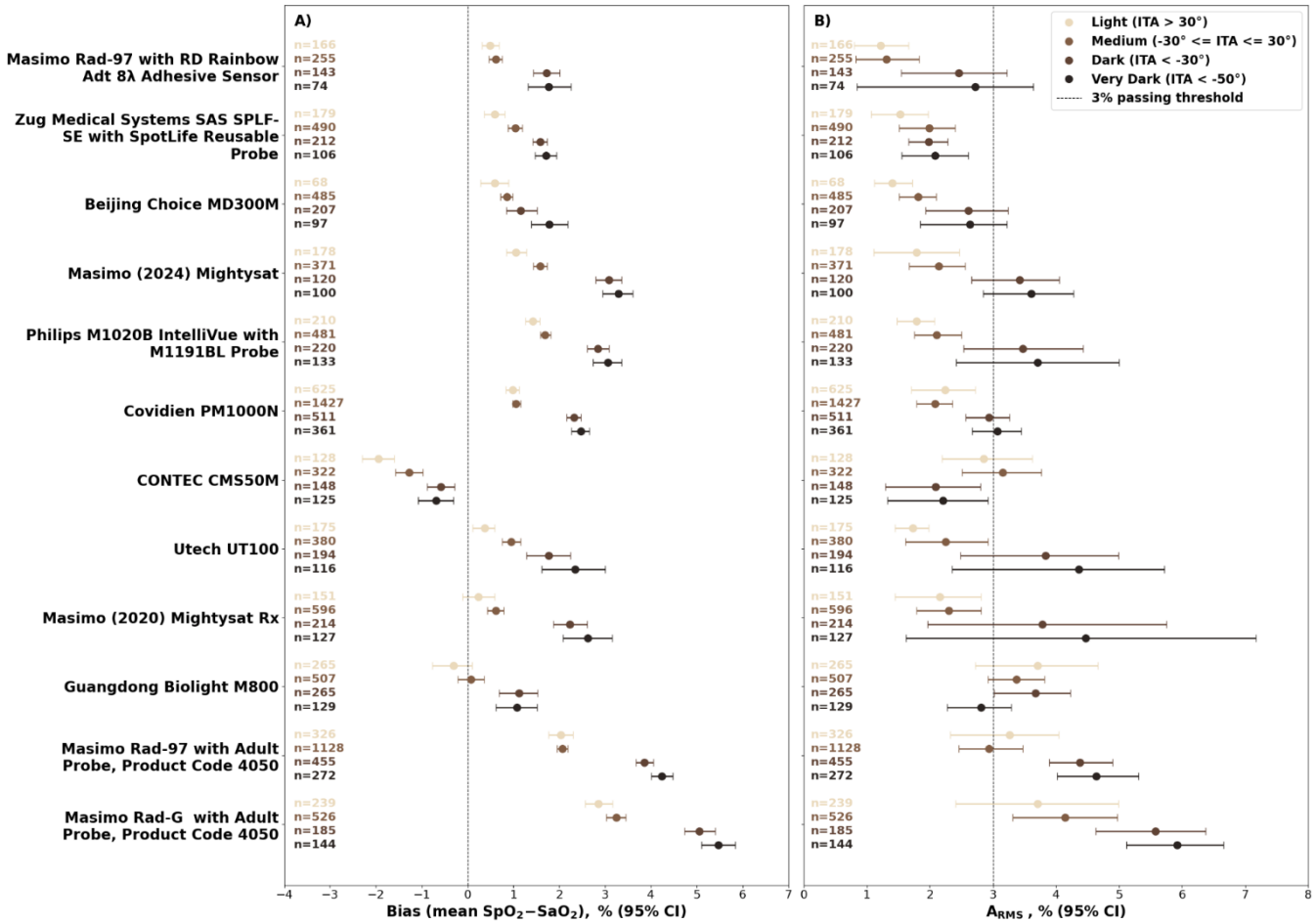

Figure 3. Distribution of bias and A_{RMS} across ITA bins at the DP site over 70-100% saturation range in maximum participant cohorts

This figure shows the distribution of pulse oximeter measured oxygen saturation (SpO_2) bias (mean of SpO_2 minus SaO_2 error, where SaO_2 is functional arterial oxygen saturation) and root mean square error (A_{RMS}) in each skin color bin defined by Individual Typology Angle (ITA) measured at the dorsal distal phalanx (DP) site across the saturation range of 70-100%. Devices shown are selected with statistically significantly differences in either bias or A_{RMS} between light (ITA >30°) and dark (ITA <-30°) bins. Panel A displays the distribution of bias in closed circles with 95% confidence interval (CI) whiskers. Panel B shows the distribution of A_{RMS} in closed circles with 95% CI whiskers. Data are stratified into four ITA bins (top to bottom): ITA > 30°, -30° ≤ ITA ≤ 30°, ITA < -30°, and ITA < -50°; note that ITA < -50° is a subset of ITA < -30°. Sample size (n) for each bin is indicated at left in the corresponding color. The vertical dashed line in Panel A represents bias of 0 and in Panel B represents the A_{RMS} 3% threshold.

Figure 4

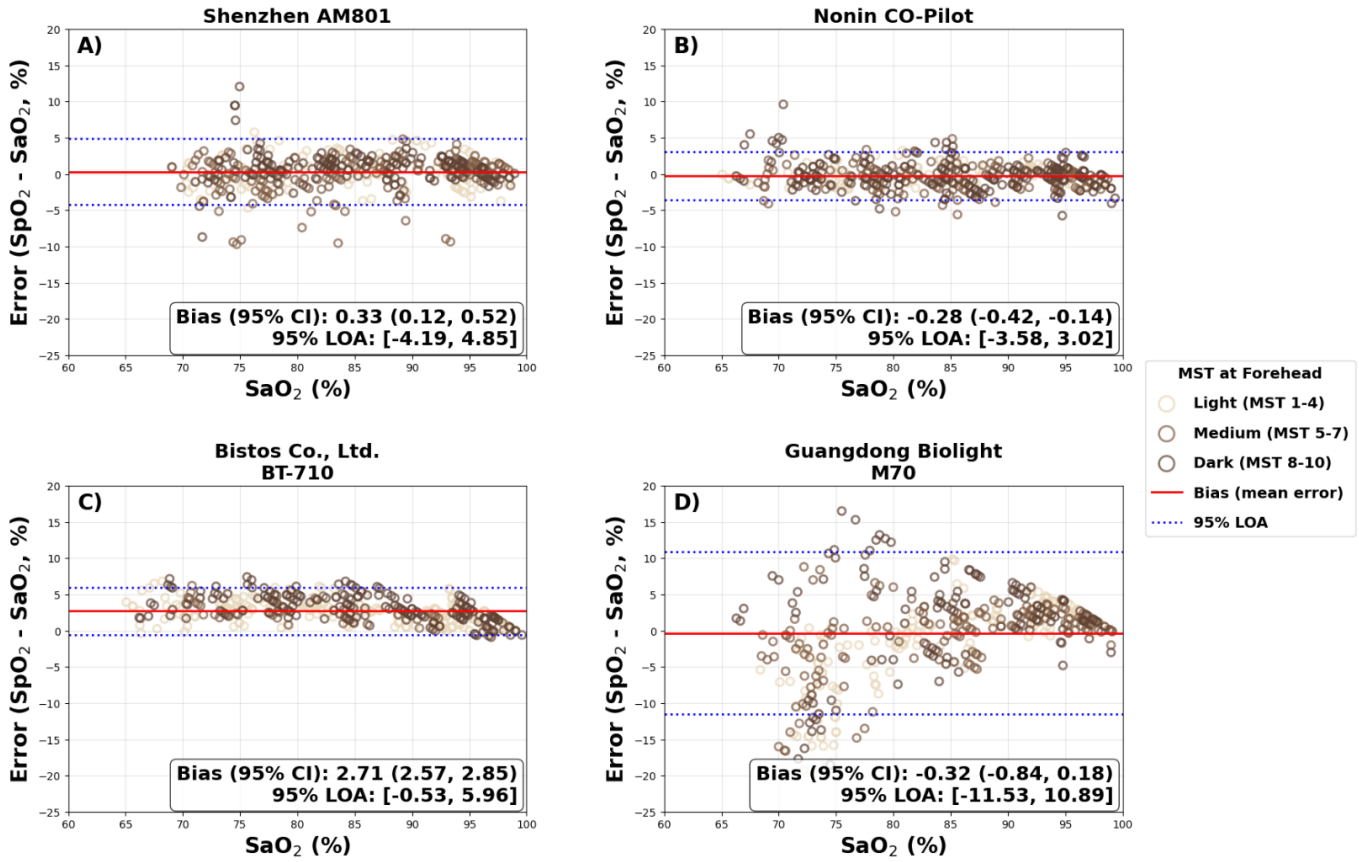

Figure 4. Modified bland altman plots for four selected devices in cohort of 24 participants

This figure shows the modified bland altman (BA) plots for four devices, showing pulse oximeter measured oxygen saturation (SpO_2) bias (mean of SpO_2 minus SaO_2 error, where SaO_2 is functional arterial oxygen saturation) and 95% limits of agreement (LOA). Data points are colored by Monk Skin Tone Scale (MST) at the forehead, as recommended by the U.S. Food and Drug Administration (FDA) draft 510(k) guidance. Panel A: Shenzhen Med-Link Electronics Tech Co., Ltd. AM801 (passing A_{RMS}/failing Individual Typology Angle (ITA) differential bias); Panel B: Nonin Medical, Inc. Co-Pilot H500 (passing A_{RMS}/passing ITA differential bias); Panel C: Bistos Co., Ltd. BT-710 (failing A_{RMS}/passing ITA differential bias); Panel D: Guangdong Biolight Meditech Co., Ltd. M70 (failing A_{RMS}/failing ITA differential bias). In each panel, the red solid horizontal line represents the SpO_2 bias; dashed lines are the 95% LOA, both have accounted for repeated measures.

Figure S1

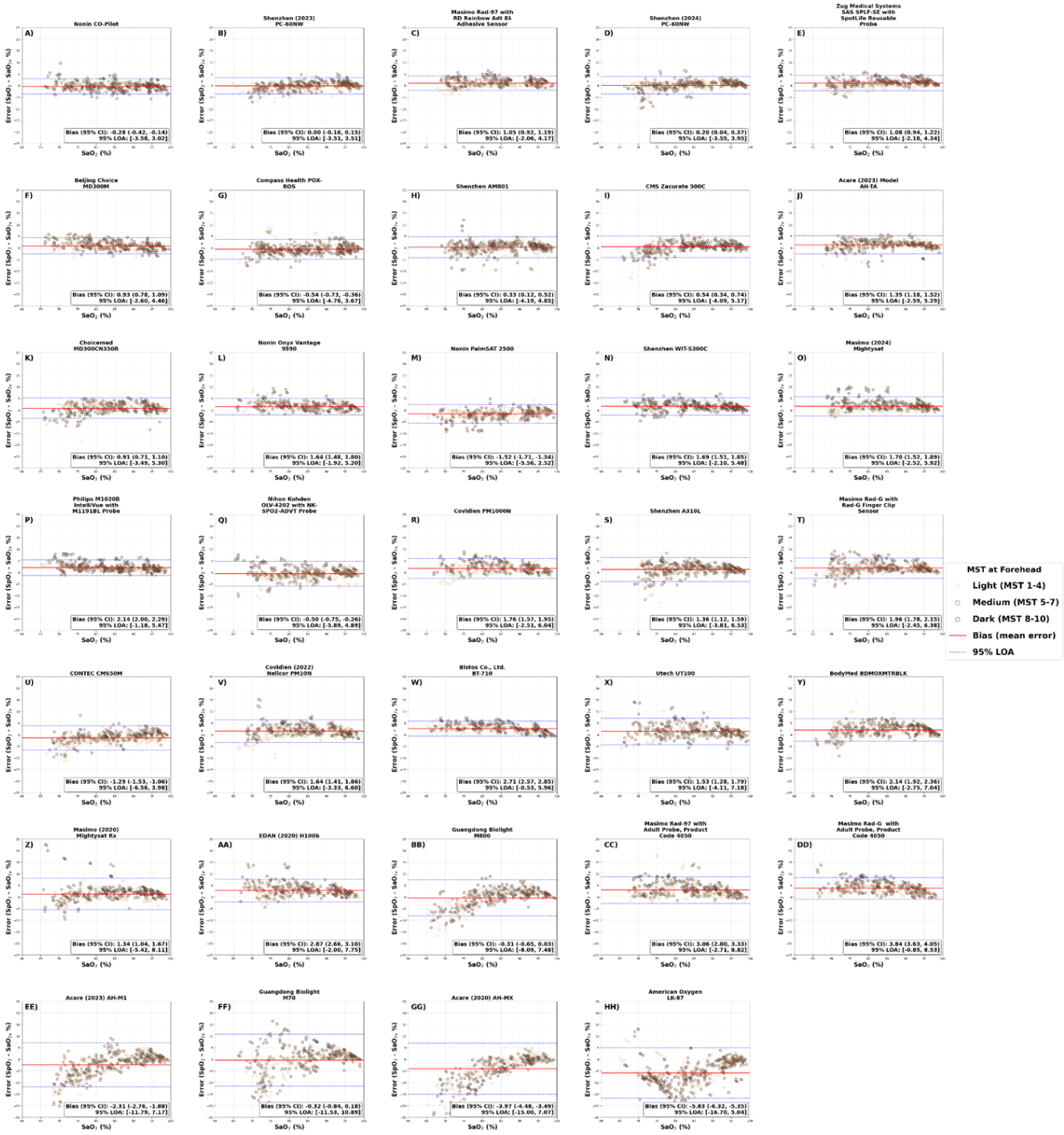

Figure S1. Modified bland altman plots for all 34 devices in cohorts of 24 participants

This figure shows the modified bland altman (BA) plots for all 34 devices, showing pulse oximeter measured oxygen saturation (SpO_2) bias (mean of SpO_2 minus SaO_2 error, where SaO_2 is functional arterial oxygen saturation) and 95% limits of agreement (LOA). Data points are colored by Monk Skin Tone Scale (MST) at the forehead, as recommended by the U.S. Food and Drug Administration (FDA)

draft 510(k) guidance. In each panel, the red solid horizontal line represents the SpO_2 bias; dashed lines are the 95% LOA, both have accounted for repeated measures.

Figure S2

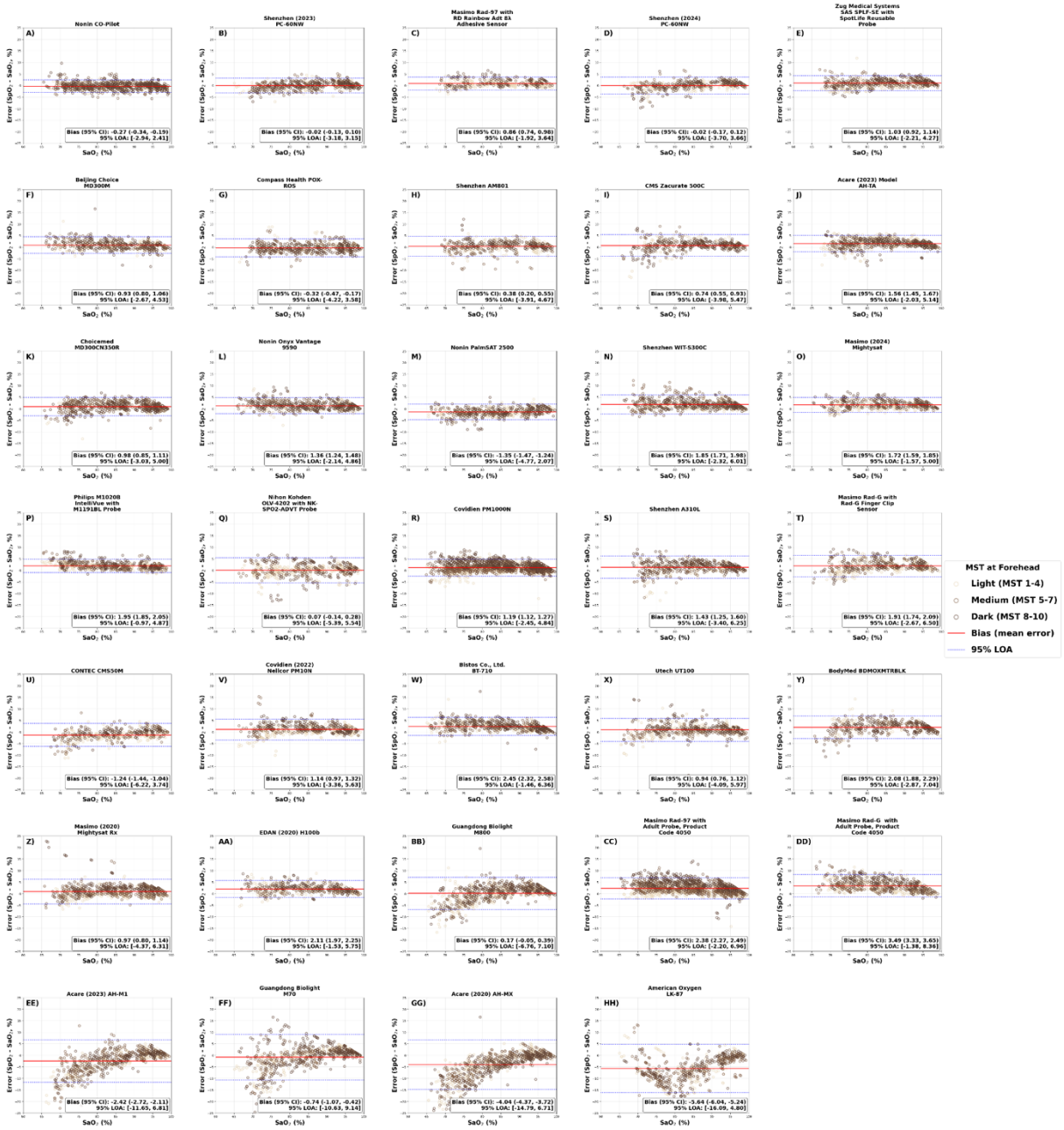

Figure S2. Modified bland altman plots for all 34 devices in maximum participants cohorts

This figure shows the modified bland altman (BA) plots for all 34 devices, showing pulse oximeter measured oxygen saturation (SpO_2) bias (mean of SpO_2 minus SaO_2 error,

where SaO_2 is functional arterial oxygen saturation) and 95% limits of agreement (LOA). Data points are colored by Monk Skin Tone Scale (MST) at the forehead, as recommended by the U.S. Food and Drug Administration (FDA) draft 510(k) guidance. In each panel, the red solid horizontal line represents the SpO_2 bias; dashed lines are the 95% LOA, both have accounted for repeated measures.

Figure S3

Figure S3. Monte Carlo simulation of A_{RMS} *performance in cohorts of 24 participants*

For each of 100 simulations, a cohort of 24 unique participants was randomly sampled without replacement and accuracy root mean square error (A_{RMS}) was calculated; bootstrapping (random resampling with replacement) with 1,000 repetitions was then used to estimate the upper bound of the 95% confidence interval (CI) of A_{RMS} in that simulated cohort. The histogram displays, for each device, the number of simulated cohorts whose upper 95% CI for A_{RMS} fell below the 3% cutoff (red dashed line). Nine devices achieved an upper 95% CI for A_{RMS} below 3% in all 100 simulations, whereas eighteen devices met this criterion in some simulations but not others. Jaccard diversity scores across devices—quantifying the overlap of cohorts across the 100 simulations—had mean values ranging from 0.03 to 0.67 (median 0.35, interquartile range 0.31–0.45), indicating adequate cohort diversity across simulations. All devices had at least thirty unique sessions used in the simulations (median 47.5 sessions).

Figure S4

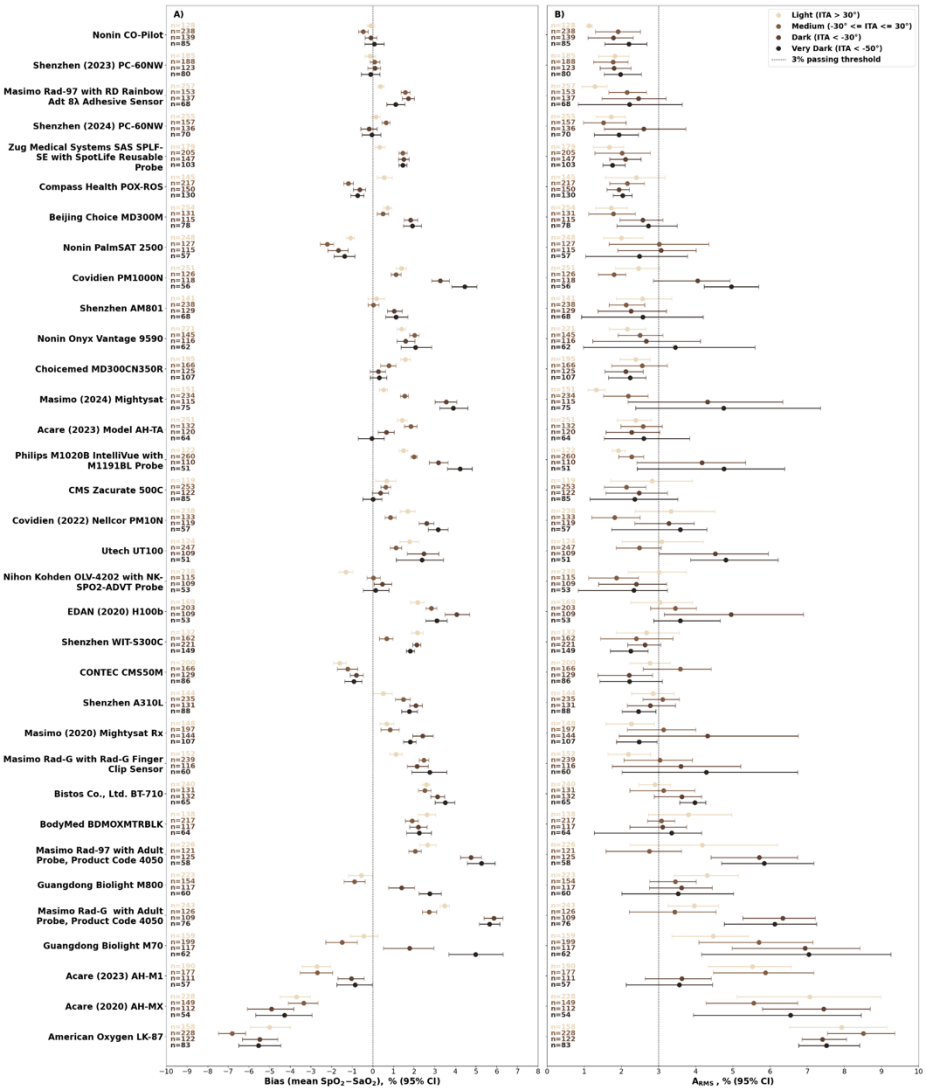

Figure S4. Distribution of bias and A_{RMS} across ITA bins at the forehead over 70–100% saturation range in 24 participants cohorts

This figure shows the distribution of pulse oximeter measured oxygen saturation (SpO_2) bias (mean of SpO_2 minus SaO_2 error, where SaO_2 is functional arterial oxygen saturation) and accuracy root mean square error (A_{RMS}) in each skin color bin defined by Individual Typology Angle (ITA) measured at the forehead across the saturation range of 70–100%. Panel A displays the distribution of bias in closed circles with 95% confidence interval (CI) whiskers. Panel B shows the distribution of A_{RMS} in closed circles with 95% CI whiskers. Data are stratified into four ITA bins (top to bottom): ITA > 30°, -30° ≤ ITA ≤ 30°, ITA < -30°, and ITA < -50°; note that ITA < -50° is a subset of ITA < -30°. Sample size (n) for each bin is indicated at left in the corresponding color. The vertical dashed line in Panel A represents bias of 0 and in Panel B represents the A_{RMS} 3% threshold.

Figure S5

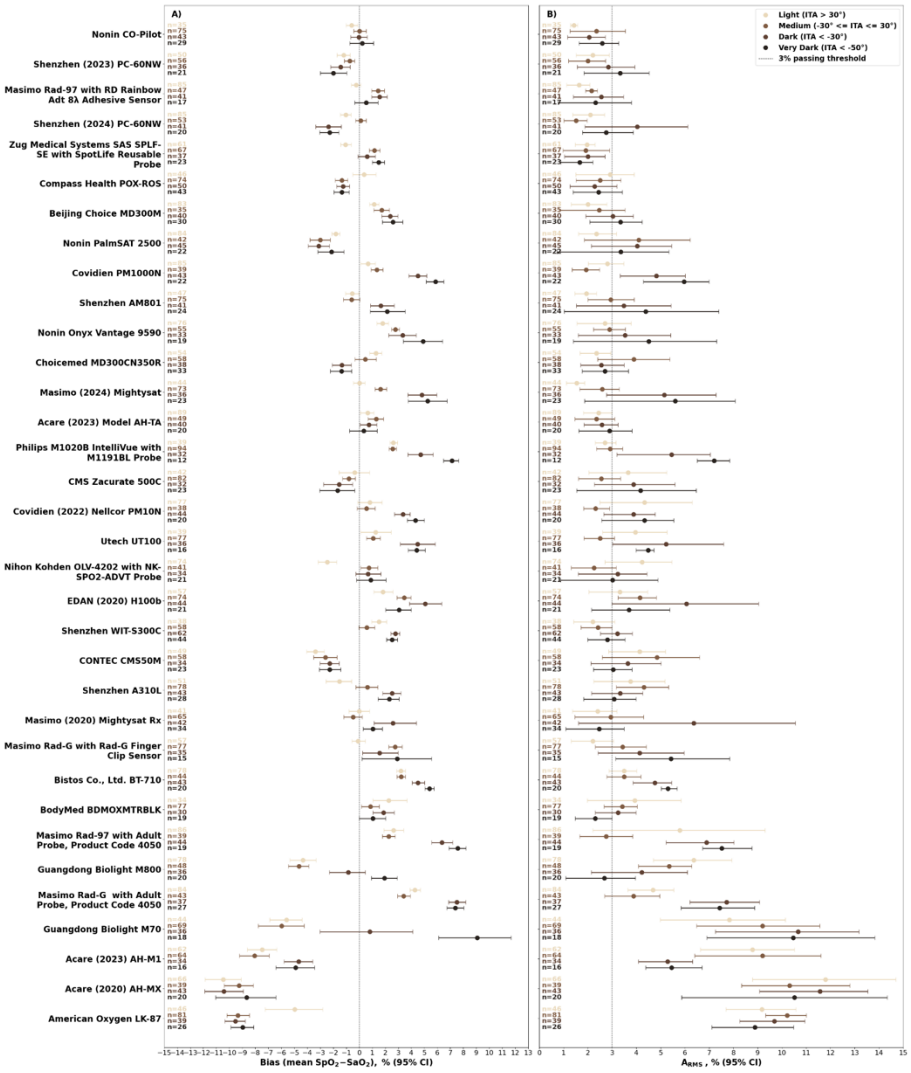

Figure S5. Distribution of bias and A_{RMS} across ITA bins at the forehead over 70–80% saturation range in 24 participants cohorts

This figure shows the distribution of pulse oximeter measured oxygen saturation (SpO_2) bias (mean of SpO_2 minus SaO_2 error, where SaO_2 is functional arterial oxygen saturation) and accuracy root mean square error (A_{RMS}) in each skin color group defined by Individual Typology Angle (ITA) measured at the forehead across the saturation range of 70–80%. Panel A displays the distribution of bias in closed circles with 95% confidence interval (CI) whiskers. Panel B shows the distribution of A_{RMS} in closed circles with 95% CI whiskers. Data are stratified into four ITA bins (top to bottom): ITA > 30°, -30° ≤ ITA ≤ 30°, ITA < -30°, and ITA < -50°; note that ITA < -50° is a subset of ITA < -30°. Sample size (n) for each bin is indicated at left in the corresponding color. The vertical dashed line in Panel A represents bias of 0 and in Panel B represents the A_{RMS} 3% threshold.

Figure S6

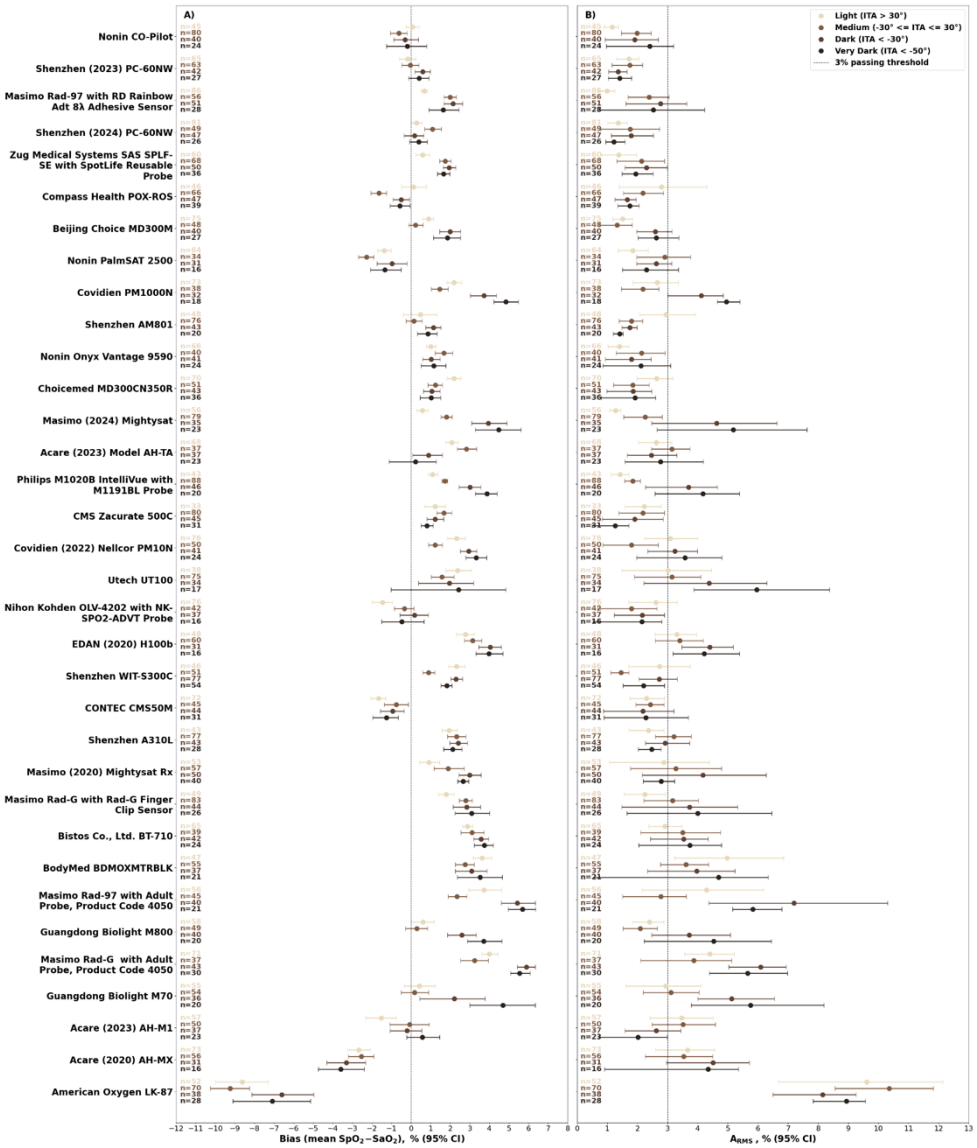

Figure S6. Distribution of bias and A_{RMS} across ITA bins at the forehead over 80–90% saturation range in 24 participants cohorts

This figure shows the distribution of pulse oximeter measured oxygen saturation (SpO_2) bias (mean of SpO_2 minus SaO_2 error, where SaO_2 is functional arterial oxygen saturation) and accuracy root mean square error (A_{RMS}) in each skin color bin defined by Individual Typology Angle (ITA) measured at the forehead across the saturation range of 80–90%. Panel A displays the distribution of bias in closed circles with 95% confidence interval (CI) whiskers. Panel B shows the distribution of A_{RMS} in closed circles with 95% CI whiskers. Data are stratified into four ITA bins (top to bottom): ITA > 30°, -30° ≤ ITA ≤ 30°, ITA < -30°, and ITA < -50°; note that ITA < -50° is a subset of ITA < -30°. Sample size (n) for each bin is indicated at left in the corresponding color. The vertical dashed line in Panel A represents bias of 0 and in Panel B represents the A_{RMS} 3% threshold.

Figure S7

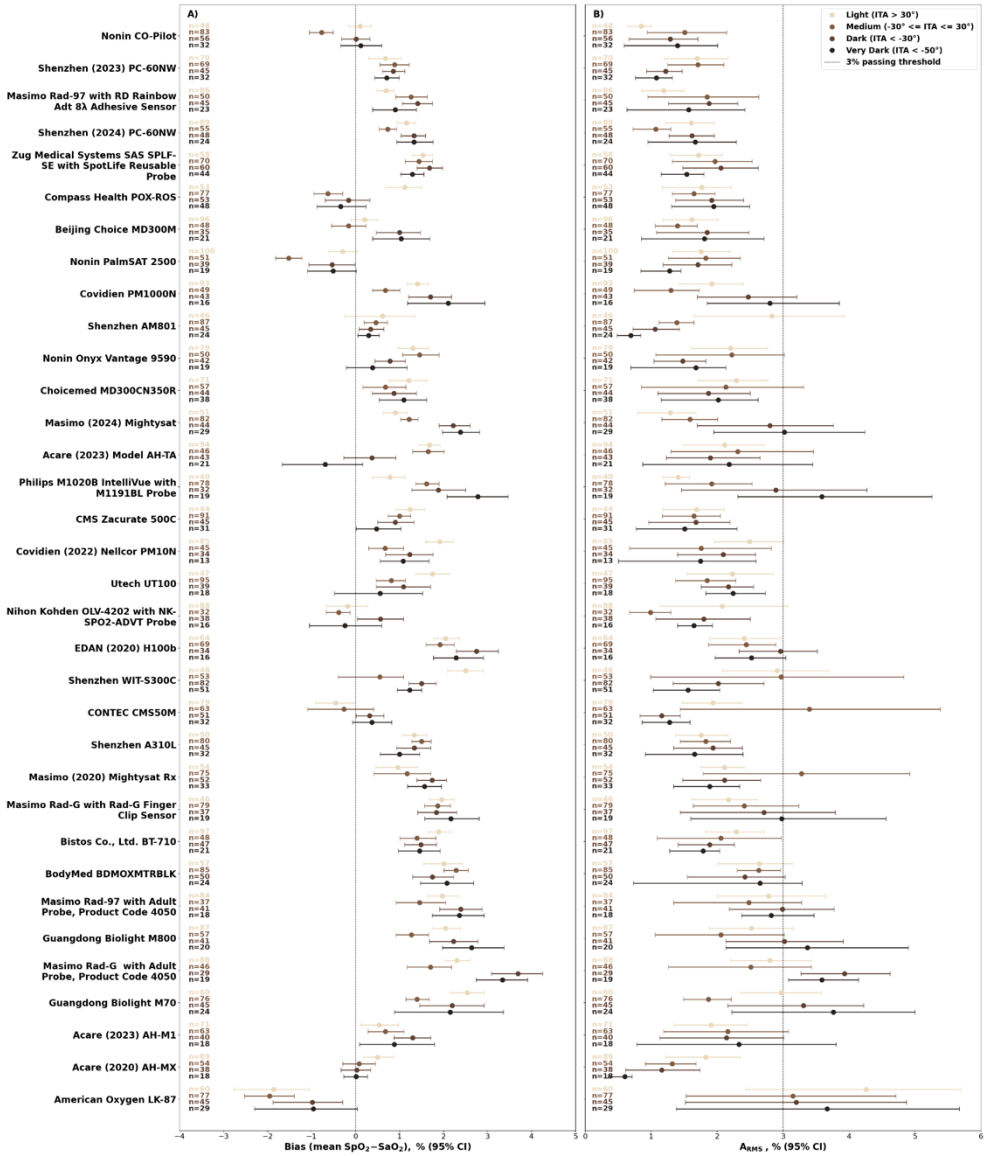

Figure S7. Distribution of bias and A_{RMS} across ITA bins at the forehead over 90–100% saturation range in 24 participants cohorts

This figure shows the distribution of pulse oximeter measured oxygen saturation (SpO_2) bias (mean of SpO_2 minus SaO_2 error, where SaO_2 is functional arterial oxygen saturation) and accuracy root mean square error (A_{RMS}) in each skin color bin defined by Individual Typology Angle (ITA) measured at the forehead across the saturation range of 90–100%. Panel A displays the distribution of bias in closed circles with 95% confidence interval (CI) whiskers. Panel B shows the distribution of A_{RMS} in closed circles with 95% CI whiskers. Data are stratified into four ITA bins (top to bottom): ITA > 30°, -30° ≤ ITA ≤ 30°, ITA < -30°, and ITA < -50°; note that ITA < -50° is a subset of ITA < -30°. Sample size (n) for each bin is indicated at left in the corresponding color. The vertical dashed line in Panel A represents bias of 0 and in Panel B represents the A_{RMS} 3% threshold.

Figure S8

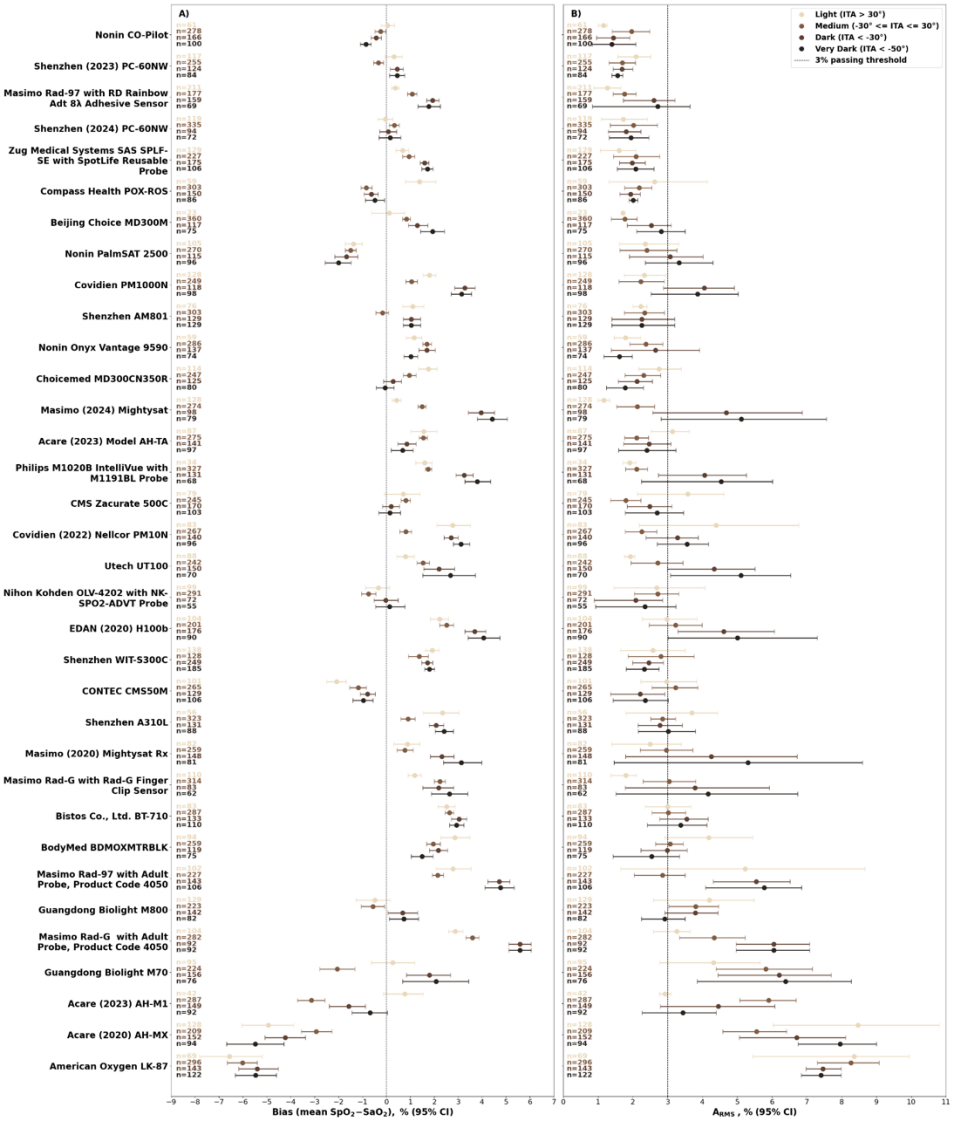

Figure S8. Distribution of bias and A_{RMS} across ITA bins at the DP site over 70–100% saturation range in 24 participants cohorts

This figure shows the distribution of pulse oximeter measured oxygen saturation (SpO_2) bias (mean of SpO_2 minus SaO_2 error, where SaO_2 is functional arterial oxygen saturation) and accuracy root mean square error (A_{RMS}) in each skin color bin defined by Individual Typology Angle (ITA) measured at the dorsal distal phalanx (DP) site across the saturation range of 70–100%. Panel A displays the distribution of bias in closed circles with 95% confidence interval (CI) whiskers. Panel B shows the distribution of A_{RMS} in closed circles with 95% CI whiskers. Data are stratified into four ITA bins (top to bottom): ITA > 30°, -30° ≤ ITA ≤ 30°, ITA < -30°, and ITA < -50°; note that ITA < -50° is a subset of ITA < -30°. Sample size (n) for each bin is indicated at left in the corresponding color. The vertical dashed line in Panel A represents bias of 0 and in Panel B represents the A_{RMS} 3% threshold.

Figure S9

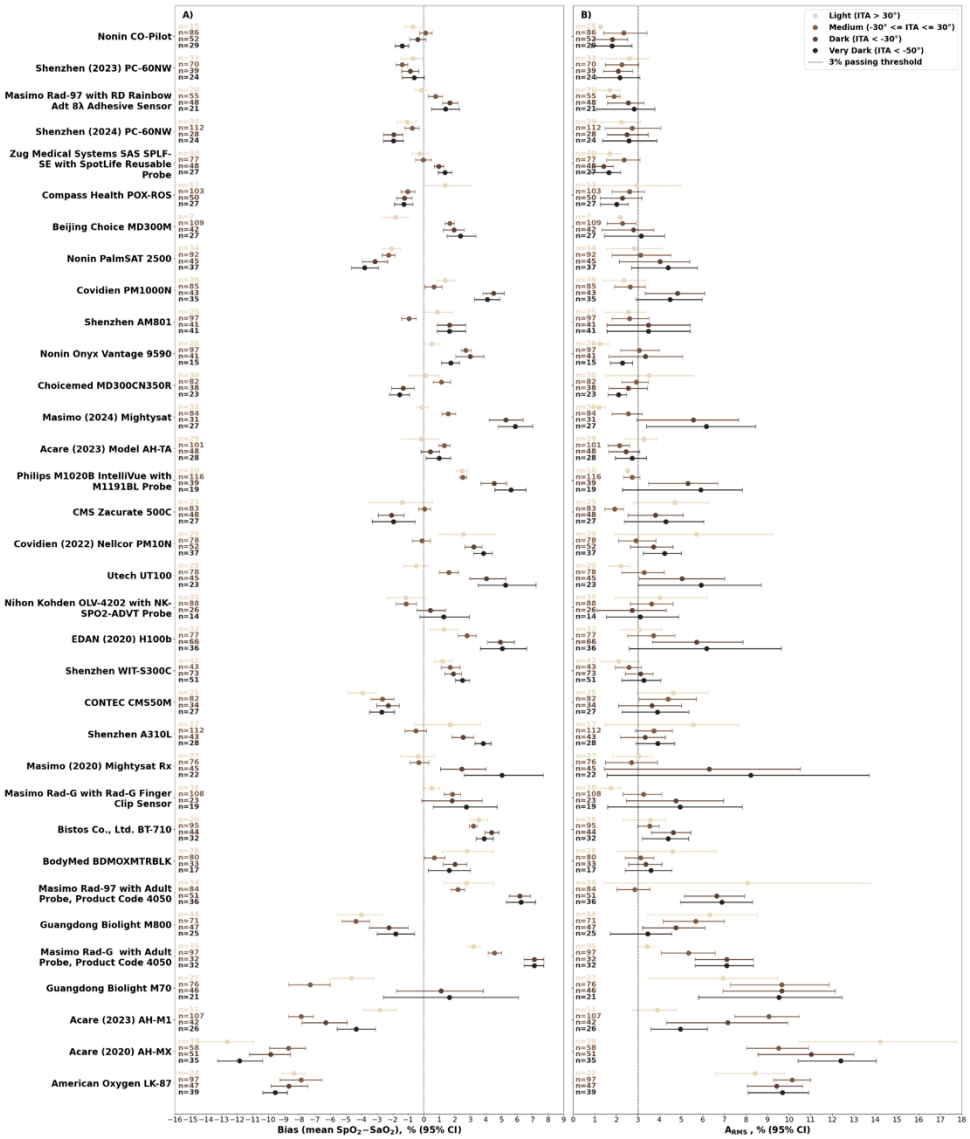

Figure S9. Distribution of bias and A_{RMS} across ITA bins at the DP site over 70–80% saturation range in 24 participants cohorts

This figure shows the distribution of pulse oximeter measured oxygen saturation (SpO_2) bias (mean of SpO_2 minus SaO_2 error, where SaO_2 is functional arterial oxygen saturation) and accuracy root mean square error (A_{RMS}) in each skin color bin defined by Individual Typology Angle (ITA) measured at the dorsal distal phalanx (DP) site across the saturation range of 70–80%. Panel A displays the distribution of bias in closed circles with 95% confidence interval (CI) whiskers. Panel B shows the distribution of A_{RMS} in closed circles with 95% CI whiskers. Data are stratified into four ITA bins (top to bottom): ITA > 30°, -30° ≤ ITA ≤ 30°, ITA < -30°, and ITA < -50°; note that ITA < -50° is a subset of ITA < -30°. Sample size (n) for each bin is indicated at left in the corresponding color. The vertical dashed line in Panel A represents bias of 0 and in Panel B represents the A_{RMS} 3% threshold.

Figure S10

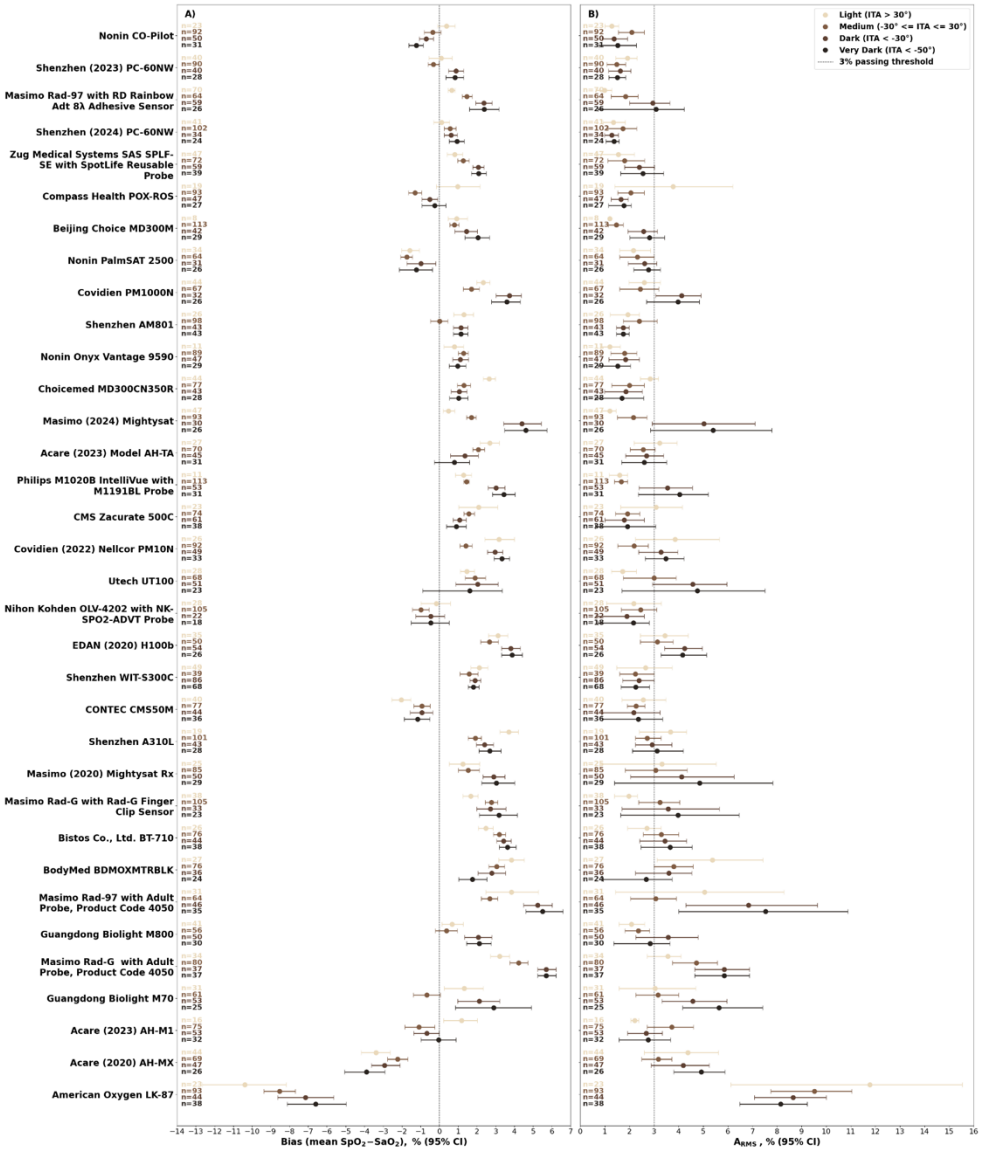

Figure S10. Distribution of bias and A_{RMS} across ITA bins at the DP site over 80–90% saturation range in 24 participants cohorts

This figure shows the distribution of pulse oximeter measured oxygen saturation (SpO_2) bias (mean of SpO_2 minus SaO_2 error, where SaO_2 is functional arterial oxygen saturation) and accuracy root mean square error (A_{RMS}) in each skin color bin defined by Individual Typology Angle (ITA) measured at the dorsal distal phalanx (DP) site across the saturation range of 80–90%. Panel A displays the distribution of bias in closed circles with 95% confidence interval (CI) whiskers. Panel B shows the distribution of A_{RMS} in closed circles with 95% CI whiskers. Data are stratified into four ITA bins (top to bottom): ITA > 30°, -30° ≤ ITA ≤ 30°, ITA < -30°, and ITA < -50°; note that ITA < -50° is a subset of ITA < -30°. Sample size (n) for each bin is indicated at left in the corresponding color. The vertical dashed line in Panel A represents bias of 0 and in Panel B represents the A_{RMS} 3% threshold.

Figure S11

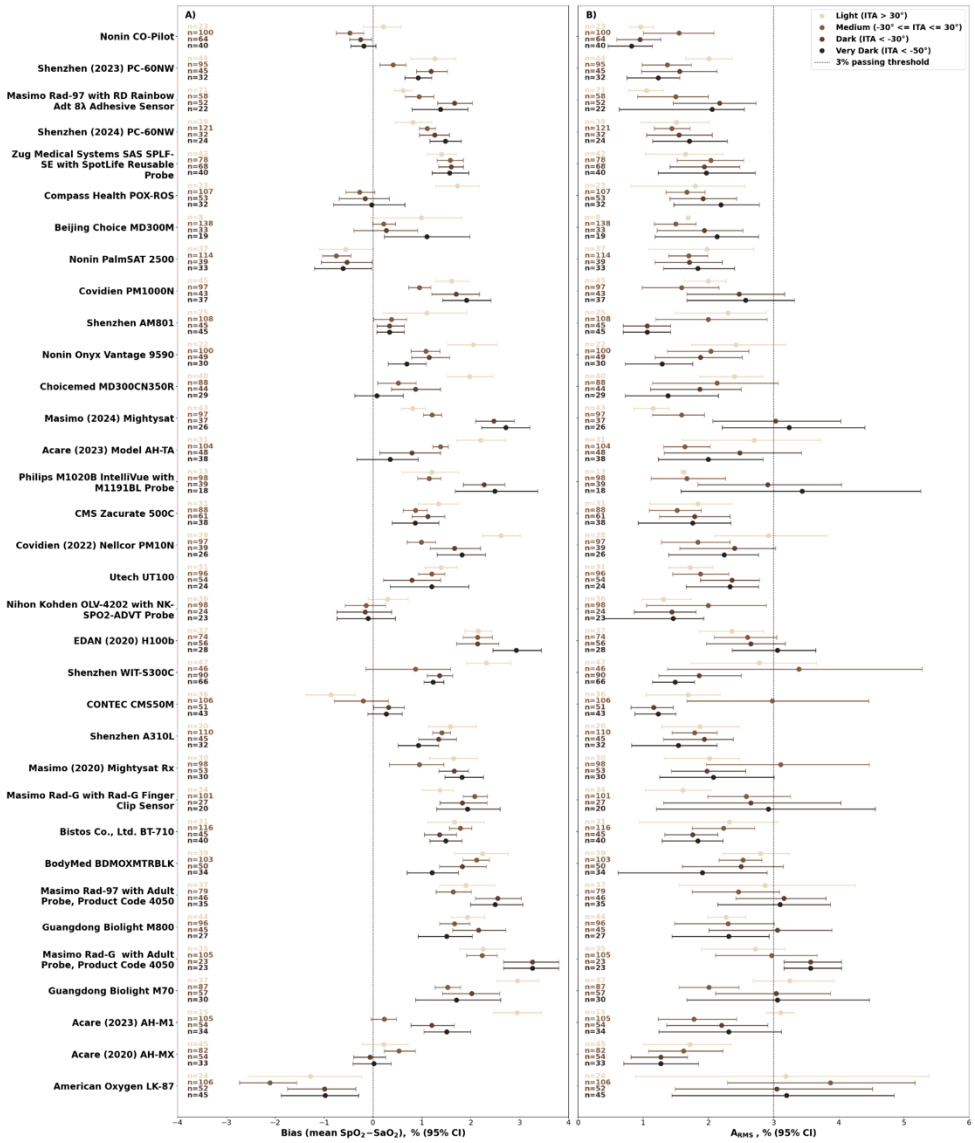

Figure S11. Distribution of bias and A_{RMS} across ITA bins at the DP site over 90–100% saturation range in 24 participants cohorts

This figure shows the distribution of pulse oximeter measured oxygen saturation (SpO_2) bias (mean of SpO_2 minus SaO_2 error, where SaO_2 is functional arterial oxygen saturation) and accuracy root mean square error (A_{RMS}) in each skin color bin defined by Individual Typology Angle (ITA) measured at the dorsal distal phalanx (DP) site across the saturation range of 90–100%. Panel A displays the distribution of bias in closed circles with 95% confidence interval (CI) whiskers. Panel B shows the distribution of A_{RMS} in closed circles with 95% CI whiskers. Data are stratified into four ITA bins (top to bottom): ITA > 30°, -30° ≤ ITA ≤ 30°, ITA < -30°, and ITA < -50°; note that ITA < -50° is a subset of ITA < -30°. Sample size (n) for each bin is indicated at left in the corresponding color. The vertical dashed line in Panel A represents bias of 0 and in Panel B represents the A_{RMS} 3% threshold.

Tables

Table 1. Baseline characteristics of participants

Covariates	Overall
Demographics	
Age (years)	25 [24,29]
BMI (kg/m²)	23 [21,26]
Female	80 (52%)
Male	75 (48%)
Finger Diameter (mm)	10 [9,12]
Percent Modulation	2 [1,5]
NIH Race	
Asian	40 (26%)
Black/African American	36 (23%)
Hispanic/Latino	16 (10%)
White	46 (30%)
Other/Multiple	17 (11%)
MST (Forehead)	
1	5 (1%)
2	41 (12%)
3	17 (5%)
4	57 (17%)
5	74 (22%)
6	37 (11%)
7	20 (6%)
8	65 (19%)
9	21 (6%)
10	4 (1%)

ITA	
ITA (Forehead)	18 [-19,33]
ITA (DP)	11 [-27,27]
ITA Bin (Forehead)	
ITA >30°	111 (32%)
-30 <= ITA <= 30°	158 (45%)
ITA <-30°	79 (23%)
ITA <-50°	42 (12%)

Categorical variables are presented as N (%), while continuous variables are summarized as median [Q1, Q3].

BMI = Body Mass Index; MST = Monk Skin Tone Scale; ITA = Individual Typology Angle; NIH = National Institutes of Health; DP = Dorsal Distal Phalanx

This table presents the descriptive statistics for demographics and skin color data. Data are presented as median [Q1, Q3] for continuous variables and count (%) for categorical variables. Race was classified according to National Institutes of Health (NIH) categories. Skin tone was assessed using the Monk Skin Tone Scale (MST) and the Individual Typology Angle (ITA) at two anatomical sites: forehead and dorsal distal phalanx (DP), with repeated measures taken in participants who completed more than one study visit. Note that participants with ITA <-50° are a subset of participants with ITA <-30°.

Table 2 - Summary of current and anticipated regulatory frameworks for pulse oximetry

	ISO (2017)	Anticipated ISO	FDA (2013)	Anticipated FDA
# participants for verification study cohorts	≥10	≥ 24	≥10	150
Diversity of biological sex		- ≥33% of each biological sex		In each MST group, at least 40% of participants are male, and at least 40% of participants are female
Diversity of skin pigment	Skin color should be described*	**MST at forehead:** ≥25% MST 1-3 (light), 4-7 (medium), 8-10 (dark) **ITA at forehead:** ≥ 25% ITA > 30° (light), 30° ≥ ITA ≥ -30° (medium), < -30° (dark) AND at least 50% of the participants in dark category have an ITA ≤ -50° (If a subject has ITA and MST that do not place this subject in the same light, medium, dark category, use the ITA categorization)	2 of 10, or 15% of total should have dark skin*	**MST at forehead:** ≥25% MST 1-4, 5-7, 8-10 **ITA at the forehead:** at least 50% of the participants in MST group 8-10 have an ITA ≤ -50° at the forehead
A_{RMS} criteria between 70-100% SaO_2	A_{RMS} <4% "upper 95 % and lower 95 % limits of agreement shall at a minimum be provided"	Arms ≤ 3% with 95% CI disclosed	Arms ≤ 3% transmittance Arms ≤ 3.5% reflectance	Arms < 3% including 95% CI for both transmittance and reflectance devices
Statistical assessment of bias due to skin color	None	For SaO2 ranges 70-85% and 85-100%: - Using linear regressions, estimate differential bias between values at an ITA of -50° and +50° at the sensor emitter site - Absolute differential bias ≤ 2% for 85-100% SaO_2, and ≤ 4% for 70-85% SaO_2	None	For SaO2 ranges 70-85% and 85-100%: - Absolute difference in SpO2 bias across ITA at the sensor emitter site AND forehead MST levels < 1.5% for SaO2 > 85% and < 3.5% for SaO2 70-85%

***No specific recommendations provided on how skin color should be measured or described**

Table S1. A_{RMS} at the DP for 24 participant cohorts

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark
Nonin CO-Pilot	70	1.69 (1.35, 2.03)	1.16 (1.00, 1.26)	1.96 (1.41, 2.49)	1.44 (0.96, 1.91)	1.40 (0.83, 2.08)	0.53	0.86	1.26 (1.22, 1.33)	2.34 (1.41, 3.43)	1.82 (0.97, 2.52)	1.80 (0.84, 2.72)	0.48	0.64	1.29 (1.00, 1.57)	2.10 (1.55, 2.60)	1.38 (0.89, 1.93)	1.52 (0.77, 2.28)	0.87	0.94	0.96 (0.79, 1.16)	1.55 (1.00, 2.09)	0.95 (0.59, 1.27)	0.82 (0.46, 1.14)	0.62	0.36
Shenzhen (2023) PC-60NW	21	1.78 (1.52, 2.02)	2.09 (1.57, 2.51)	1.70 (1.31, 2.07)	1.69 (1.43, 2.00)	1.56 (1.40, 1.71)	0.36	0.21	2.61 (1.28, 3.52)	2.26 (1.49, 3.03)	2.09 (1.40, 2.77)	2.17 (1.09, 3.09)	0.72	0.76	1.93 (1.44, 2.31)	1.49 (1.08, 1.84)	1.63 (1.15, 2.06)	1.51 (1.16, 1.84)	0.49	0.37	2.01 (1.65, 2.36)	1.37 (0.98, 1.74)	1.56 (0.97, 2.13)	1.23 (0.75, 1.56)	0.17	**0.04**
Masimo Rad-97 with RD Rainbow Adt 8λ Adult Adhesive Sensor	73.09	1.89 (1.51, 2.27)	1.27 (0.87, 1.65)	1.76 (1.43, 2.09)	2.61 (1.73, 3.21)	2.72 (0.84, 3.64)	**0.046**	0.25	1.70 (1.09, 2.19)	1.89 (1.54, 2.16)	2.56 (1.59, 3.27)	2.82 (1.07, 3.79)	0.24	0.3	0.99 (0.63, 1.28)	1.84 (1.26, 2.35)	2.95 (2.00, 3.65)	3.09 (0.73, 4.23)	**0.02**	0.21	1.05 (0.77, 1.31)	1.50 (0.91, 2.00)	2.17 (1.46, 2.73)	2.06 (0.63, 2.55)	**0.04**	0.28
Shenzhen (2024) PC-60NW	81	1.92 (1.47, 2.42)	1.73 (1.09, 2.42)	2.02 (1.35, 2.71)	1.81 (1.30, 2.23)	1.95 (1.32, 2.46)	0.72	0.53	2.24 (1.02, 3.11)	2.74 (1.47, 4.06)	2.49 (1.58, 3.47)	2.58 (1.35, 3.87)	0.65	0.69	1.35 (0.93, 1.83)	1.74 (1.08, 2.30)	1.29 (0.99, 1.55)	1.38 (1.04, 1.58)	0.99	0.75	1.51 (0.96, 2.01)	1.44 (1.17, 1.72)	1.55 (1.05, 2.06)	1.71 (1.14, 2.29)	0.85	0.6
Zug Medical Systems SAS SPLF-	75	1.97 (1.60, 2.34)	1.61 (1.07, 2.08)	2.09 (1.44, 2.77)	1.98 (1.62, 2.36)	2.08 (1.55, 2.61)	0.24	0.22	1.70 (0.93, 2.22)	2.36 (1.55, 3.10)	1.41 (0.88, 1.87)	1.65 (0.79, 2.20)	0.55	0.91	1.55 (0.66, 2.21)	1.81 (1.12, 2.62)	2.40 (1.80, 3.02)	2.55 (1.65, 3.39)	0.08	0.1	1.65 (1.03, 2.22)	2.04 (1.52, 2.54)	1.94 (1.41, 2.48)	1.97 (1.23, 2.72)	0.45	0.56

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark
SE with SpotLife Reusable Probe																										
Beijing Choice MD300M	40	2.01 (1.66, 2.34)	1.72 (1.72, 1.72)	1.77 (1.38, 2.12)	2.53 (1.84, 3.10)	2.82 (2.11, 3.50)	NA	NA	2.18 (2.18, 2.18)	2.28 (1.56, 2.90)	2.79 (1.32, 3.74)	3.16 (1.44, 4.24)	NA	NA	1.20 (1.20, 1.20)	1.47 (1.14, 1.75)	2.58 (1.94, 3.14)	2.82 (2.02, 3.43)	NA	NA	1.69 (1.69, 1.69)	1.50 (1.17, 1.81)	1.94 (1.21, 2.53)	2.13 (1.18, 2.77)	NA	NA
Compass Health POX-ROS	12	2.20 (1.86, 2.57)	2.62 (1.31, 4.14)	2.18 (1.76, 2.54)	1.94 (1.63, 2.21)	2.01 (1.89, 2.14)	0.73	0.81	2.95 (0.68, 4.99)	2.62 (1.81, 3.31)	2.28 (1.26, 3.18)	2.01 (1.26, 2.55)	0.88	0.85	3.78 (1.40, 6.21)	2.06 (1.52, 2.61)	1.66 (1.26, 1.95)	1.78 (1.15, 2.07)	0.45	0.48	1.80 (0.81, 2.56)	1.67 (1.35, 1.95)	1.92 (1.41, 2.43)	2.19 (1.47, 2.78)	0.84	0.55
Shenzhen AM801	10	2.32 (1.86, 2.75)	2.23 (2.00, 2.40)	2.34 (1.75, 2.91)	2.26 (1.40, 3.20)	2.26 (1.40, 3.20)	0.67	0.67	2.55 (1.45, 3.35)	2.62 (1.79, 3.51)	3.48 (1.57, 5.42)	3.48 (1.57, 5.42)	0.73	0.73	1.94 (1.22, 2.41)	2.41 (1.75, 3.13)	1.75 (1.47, 1.99)	1.75 (1.47, 1.99)	0.87	0.87	2.30 (1.49, 2.89)	2.00 (1.19, 2.90)	1.06 (0.69, 1.42)	1.06 (0.69, 1.42)	0.09	0.09
CMS Zacurate 500C	15	2.41 (1.91, 2.89)	3.58 (2.14, 4.62)	1.80 (1.36, 2.23)	2.49 (1.84, 3.12)	2.70 (1.78, 3.46)	0.35	0.45	4.71 (2.79, 6.27)	1.92 (1.46, 2.33)	3.82 (2.54, 5.09)	4.30 (2.37, 6.05)	0.51	0.73	3.08 (1.64, 4.15)	1.93 (1.43, 2.42)	1.79 (1.01, 2.61)	1.92 (0.62, 3.07)	0.26	0.29	1.84 (1.10, 2.35)	1.52 (1.09, 1.89)	1.79 (1.25, 2.33)	1.76 (0.92, 2.34)	0.81	0.77
Acare (2023) Model AH-TA	56	2.41 (2.08, 2.74)	3.14 (2.52, 3.62)	2.11 (1.76, 2.45)	2.47 (1.74, 3.09)	2.40 (1.59, 3.24)	0.17	0.18	3.29 (2.40, 3.91)	2.15 (1.61, 2.61)	2.45 (1.66, 3.08)	2.73 (1.95, 3.40)	0.22	0.44	3.23 (2.19, 3.94)	2.56 (2.03, 3.05)	2.70 (1.84, 3.39)	2.60 (1.68, 3.51)	0.49	0.45	2.70 (1.60, 3.72)	1.64 (1.31, 2.03)	2.48 (1.32, 3.43)	2.00 (1.23, 2.84)	0.68	0.41
Choicemed MD300CN 350R	14	2.41 (2.05, 2.75)	2.75 (2.17, 3.39)	2.31 (1.77, 2.80)	2.12 (1.58, 2.56)	1.78 (1.24, 2.30)	0.18	0.07	3.52 (1.50, 5.59)	2.93 (2.26, 3.49)	2.55 (1.66, 3.44)	2.10 (1.61, 2.48)	0.76	0.57	2.85 (2.44, 3.18)	2.01 (1.29, 2.60)	1.86 (0.99, 2.52)	1.70 (0.60, 2.58)	**0.04**	0.08	2.40 (1.87, 2.84)	2.13 (1.14, 3.07)	1.87 (1.11, 2.50)	1.38 (0.72, 2.15)	0.23	0.08

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark
Nonin Onyx Vantage 9590	35	2.44 (1.90, 2.99)	1.79 (1.46, 2.22)	2.38 (1.91, 2.86)	2.65 (1.38, 3.92)	1.62 (1.17, 1.98)	0.52	0.59	1.24 (0.94, 1.64)	3.07 (2.20, 3.99)	3.35 (1.65, 5.07)	2.29 (1.71, 2.75)	0.12	**0.045**	1.21 (0.82, 1.63)	1.80 (1.24, 2.30)	1.84 (1.16, 2.40)	1.53 (0.72, 2.05)	0.33	0.67	2.42 (1.74, 3.19)	2.04 (1.37, 2.62)	1.88 (1.18, 2.52)	1.29 (0.72, 1.76)	0.29	0.1
Nonin PalmSAT 2500	36	2.54 (1.97, 3.11)	2.36 (1.61, 3.32)	2.40 (1.63, 3.27)	3.07 (1.90, 4.02)	3.33 (2.36, 4.30)	0.44	0.24	2.83 (1.53, 4.12)	3.13 (1.81, 4.54)	4.03 (2.14, 5.41)	4.40 (2.70, 5.75)	0.45	0.25	2.17 (1.61, 2.86)	2.32 (1.61, 3.00)	2.62 (1.95, 3.11)	2.78 (2.18, 3.26)	0.39	0.23	1.98 (1.08, 2.69)	1.70 (1.39, 1.99)	1.71 (1.18, 2.21)	1.84 (1.31, 2.40)	0.8	1
Shenzhen WIT-S300C	62	2.56 (2.12, 3.01)	2.58 (1.64, 3.50)	2.81 (1.86, 3.76)	2.46 (1.99, 2.88)	2.33 (1.81, 2.75)	0.96	0.8	2.11 (1.25, 3.08)	2.58 (1.96, 3.15)	3.12 (2.42, 3.71)	3.28 (2.26, 4.05)	0.12	0.14	2.66 (1.49, 3.74)	2.25 (1.60, 3.01)	2.39 (1.72, 3.00)	2.26 (1.66, 2.82)	0.73	0.69	2.78 (1.73, 3.66)	3.39 (1.37, 5.28)	1.86 (1.24, 2.50)	1.49 (1.14, 1.79)	0.18	0.1
Masimo (2024) Mightysat	78	2.72 (1.77, 3.76)	1.17 (0.99, 1.33)	2.13 (1.53, 2.62)	4.69 (2.57, 6.87)	5.12 (2.81, 7.56)	0.06	0.08	1.20 (0.82, 1.49)	2.56 (1.81, 3.20)	5.57 (2.95, 7.66)	6.17 (3.40, 8.45)	0.06	0.053	1.20 (0.88, 1.46)	2.16 (1.51, 2.71)	5.03 (2.92, 7.11)	5.41 (2.86, 7.79)	**0.04**	0.07	1.15 (0.85, 1.40)	1.59 (1.14, 1.94)	3.03 (2.07, 4.03)	3.24 (2.21, 4.40)	**0.03**	**0.046**
Philips M1020B IntelliVue with M1191BL Probe	71	2.72 (2.17, 3.27)	1.91 (1.72, 2.08)	2.11 (1.79, 2.42)	4.06 (2.73, 5.26)	4.54 (2.25, 6.02)	**0.04**	0.14	2.53 (2.51, 2.54)	2.74 (2.33, 3.10)	5.31 (3.48, 6.69)	5.91 (2.29, 7.84)	0.09	0.22	1.60 (1.18, 1.93)	1.67 (1.39, 1.92)	3.55 (2.39, 4.56)	4.05 (2.37, 5.21)	0.052	0.07	1.62 (1.57, 1.66)	1.67 (1.12, 2.26)	2.91 (1.84, 4.04)	3.44 (1.58, 5.26)	0.14	0.33
Nihon Kohden OLV-4202 with NK-SPO2-	64.03	2.76 (2.15, 3.30)]	2.69 (1.45, 4.07)	2.72 (2.05, 3.32)	2.08 (0.89, 2.85)	2.35 (0.93, 3.24)	0.62	0.89	4.01 (1.93, 6.22)	3.64 (2.65, 4.63)	2.73 (1.11, 4.31)	3.11 (1.54, 4.89)	0.49	0.72	2.18 (1.07, 3.30)	2.46 (1.67, 3.11)	1.90 (0.64, 2.61)	2.17 (0.73, 2.81)	0.8	0.87	1.31 (0.98, 1.74)	2.00 (1.05, 2.89)	1.44 (0.86, 1.81)	1.46 (0.40, 1.93)	0.84	0.92

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark
ADVT Probe																										
Covidien PM1000N	59	2.78 (2.22, 3.33)	2.33 (1.75, 2.95)	2.23 (1.60, 2.90)	4.05 (2.88, 4.92)	3.86 (2.52, 5.03)	0.08	0.17	2.34 (1.36, 3.37)	2.64 (1.91, 3.35)	4.83 (3.33, 6.10)	4.49 (2.92, 5.97)	**0.03**	0.09	2.60 (1.98, 3.26)	2.45 (1.61, 3.19)	4.12 (3.07, 4.91)	3.98 (2.70, 4.85)	0.11	0.23	2.00 (1.63, 2.27)	1.59 (0.98, 2.16)	2.47 (1.67, 3.17)	2.57 (1.67, 3.32)	0.52	0.49
Shenzhen A310L	11	2.95 (2.60, 3.27)	3.70 (1.80, 4.44)	2.85 (2.51, 3.23)	2.78 (2.15, 3.42)	3.02 (2.16, 3.80)	0.44	0.58	5.57 (1.49, 7.65)	3.74 (2.88, 4.59)	3.33 (2.19, 4.26)	3.92 (2.92, 4.70)	0.41	0.6	3.67 (2.41, 4.32)	2.73 (2.24, 3.28)	2.92 (2.24, 3.72)	3.13 (2.13, 4.19)	0.33	0.49	1.87 (1.28, 2.47)	1.79 (1.44, 2.13)	1.94 (1.31, 2.38)	1.54 (0.82, 2.13)	0.99	0.47
Masimo Rad-G with Rad-G Finger Clip Sensor	55.03	2.97 (2.27, 3.65)	1.80 (1.37, 2.11)	3.05 (2.29, 3.81)	3.79 (1.78, 5.92)	4.16 (1.52, 6.74)	0.32	0.42	1.75 (1.10, 2.22)	3.26 (2.31, 4.10)	4.76 (2.47, 6.97)	4.96 (1.60, 7.84)	0.15	0.3	1.98 (1.39, 2.32)	3.25 (2.38, 4.05)	3.58 (1.70, 5.66)	3.98 (1.65, 6.46)	0.43	0.47	1.61 (1.03, 2.04)	2.58 (1.99, 3.26)	2.65 (1.31, 4.03)	2.92 (1.20, 4.56)	0.38	0.42
CONTEC CMS50M	17	2.97 (2.50, 3.44)	2.97 (2.23, 3.83)	3.23 (2.55, 3.87)	2.21 (1.36, 2.92)	2.36 (1.43, 3.02)	0.21	0.33	4.64 (2.98, 6.26)	4.40 (3.05, 5.70)	3.65 (2.11, 5.02)	3.90 (2.27, 5.36)	0.38	0.52	2.57 (1.70, 3.49)	2.27 (1.91, 2.63)	2.18 (0.86, 3.24)	2.36 (0.89, 3.35)	0.42	0.61	1.69 (1.04, 2.18)	2.98 (1.67, 4.46)	1.16 (0.81, 1.46)	1.23 (0.87, 1.50)	0.26	0.34
Covidien (2022) Nellcor PM10N	41	3.00 (2.37, 3.69)	4.39 (2.18, 6.77)	2.26 (1.78, 2.69)	3.28 (2.38, 3.88)	3.56 (2.70, 4.17)	0.65	0.82	5.70 (1.93, 9.20)	2.92 (2.11, 3.83)	3.73 (2.64, 4.62)	4.23 (3.24, 5.02)	0.65	0.85	3.86 (2.25, 5.66)	2.19 (1.53, 2.76)	3.28 (2.38, 3.97)	3.49 (2.64, 4.22)	0.74	0.92	2.92 (2.09, 3.82)	1.84 (1.28, 2.33)	2.40 (1.56, 3.03)	2.24 (1.39, 2.77)	0.35	0.31
Bistos Co., Ltd. BT-710	61	3.17 (2.82, 3.55)	3.00 (2.35, 3.67)	3.02 (2.55, 3.51)	3.55 (2.77, 4.16)	3.38 (2.41, 4.13)	0.45	0.65	3.57 (2.31, 4.29)	3.54 (2.99, 3.98)	4.64 (3.62, 5.45)	4.40 (3.20, 5.34)	0.27	0.42	2.71 (1.92, 3.29)	3.30 (2.57, 4.01)	3.44 (2.42, 4.33)	3.66 (2.47, 4.55)	0.44	0.33	2.32 (0.94, 3.07)	2.23 (1.75, 2.71)	1.76 (1.33, 2.14)	1.84 (1.29, 2.22)	0.58	0.67

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark
Utech UT100	39	3.25 (2.51, 3.95)	1.93 (1.75, 2.06)	2.72 (1.95, 3.44)	4.34 (2.99, 5.51)	5.11 (3.08, 6.54)	**0.02**	0.08	2.21 (1.61, 2.63)	3.29 (2.25, 4.22)	5.04 (3.06, 7.03)	5.93 (3.02, 8.70)	0.06	0.14	1.73 (1.27, 2.29)	3.01 (1.75, 3.90)	4.58 (2.94, 5.96)	4.77 (1.70, 7.51)	**0.04**	0.25	1.72 (1.39, 2.07)	1.88 (1.45, 2.31)	2.36 (1.88, 2.78)	2.33 (1.66, 2.77)	0.13	0.16
BodyMed BDMOXMTRBLK	2	3.28 (2.86, 3.75)	4.18 (2.92, 5.45)	3.07 (2.65, 3.45)	2.99 (2.23, 3.56)	2.54 (1.43, 3.34)	0.21	0.11	4.59 (2.04, 6.61)	3.12 (2.40, 3.74)	3.37 (2.57, 4.12)	3.61 (2.41, 4.57)	0.67	0.79	5.38 (3.12, 7.43)	3.81 (3.01, 4.59)	3.61 (2.25, 4.54)	2.68 (0.94, 3.74)	0.28	0.12	2.80 (2.22, 3.24)	2.53 (2.16, 2.82)	2.50 (1.60, 3.15)	1.91 (0.61, 2.90)	0.39	0.16
Masimo (2020) Mightysat Rx	7	3.67 (2.34, 5.10)	2.50 (1.40, 3.39)	2.96 (2.21, 3.72)	4.26 (1.79, 6.72)	5.31 (1.46, 8.60)	0.48	0.45	3.04 (1.82, 3.63)	2.71 (1.49, 3.90)	6.30 (1.44, 10.52)	8.23 (1.56, 13.71)	0.62	0.46	3.33 (0.81, 5.53)	3.07 (1.83, 4.35)	4.13 (2.04, 6.26)	4.86 (1.39, 7.83)	0.58	0.57	2.02 (1.32, 2.48)	3.11 (1.97, 4.47)	1.98 (1.43, 2.57)	2.08 (1.25, 3.01)	0.94	0.99
EDAN (2020) H100b	44	3.79 (3.06, 4.70)	2.98 (2.28, 3.84)	3.23 (2.47, 3.99)	4.61 (3.29, 6.07)	5.01 (3.01, 7.30)	0.1	0.25	3.05 (2.16, 4.10)	3.73 (2.51, 4.72)	5.72 (3.66, 7.87)	6.18 (2.60, 9.63)	0.07	0.31	3.44 (2.45, 4.39)	3.14 (2.45, 3.78)	4.24 (3.43, 4.96)	4.17 (3.28, 5.14)	0.28	0.32	2.36 (1.86, 2.85)	2.60 (2.09, 3.05)	2.65 (1.97, 3.18)	3.06 (2.36, 3.65)	0.78	0.2
Guangdong Biolight M800	43	3.96 (3.44, 4.46)	4.20 (2.59, 5.48)	3.81 (3.04, 4.47)	3.80 (2.92, 4.45)	2.92 (2.25, 3.50)	0.81	0.25	6.32 (3.45, 8.53)	5.67 (4.16, 6.99)	4.76 (3.21, 6.11)	3.45 (1.72, 4.57)	0.47	0.18	2.09 (1.58, 2.62)	2.36 (1.82, 2.82)	3.58 (2.26, 4.79)	2.84 (1.37, 3.64)	0.14	0.44	2.27 (1.99, 2.57)	2.30 (1.48, 3.01)	3.06 (2.01, 3.89)	2.31 (1.44, 2.93)	0.36	0.92
Masimo Rad-97 with Adult Probe, Product Code 4050	60	4.21 (3.06, 5.24)	5.23 (1.65, 8.67)	2.85 (2.04, 3.50)	5.55 (4.31, 6.52)	5.78 (4.09, 6.85)	0.46	0.44	8.07 (1.42, 13.75)	2.86 (2.03, 3.55)	6.65 (5.15, 7.94)	6.89 (4.97, 8.32)	0.69	0.66	5.06 (1.43, 8.28)	3.07 (2.04, 3.91)	6.85 (4.30, 9.64)	7.54 (4.00, 10.88)	0.31	0.28	2.87 (1.55, 4.25)	2.46 (1.75, 3.09)	3.16 (2.42, 3.80)	3.10 (2.14, 3.87)	0.56	0.65
Masimo Rad-G	55	4.51 (3.76, 5.23)	3.27 (2.59,	4.34 (3.34,	6.05 (4.97,	6.05 (4.97,	**0.007**	**0.007**	3.43 (3.01,	5.34 (4.09,	7.12 (5.64,	7.12 (5.64,	**0.008**	**0.008**	3.56 (2.73,	4.73 (3.75,	5.86 (4.66,	5.86 (4.66,	**0.02**	**0.02**	2.72 (1.89,	2.97 (2.11,	3.57 (3.16,	3.57 (3.16,	0.12	0.12

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark
with Adult Probe, Product Code 4050			3.65)	5.23)	7.08)	7.08)			3.77)	6.58)	8.35)	8.35)			4.10)	5.58)	6.89)	6.89)			3.17)	3.67)	4.04)	4.04)		
Acare (2023) AH-M1	57	5.34 (4.47, 6.11)	2.93 (2.76, 3.10)	5.91 (5.07, 6.69)	4.46 (2.79, 6.08)	3.44 (2.27, 4.39)	0.28	0.72	3.91 (2.76, 4.79)	9.06 (7.47, 10.48)	7.16 (4.33, 9.93)	4.97 (3.61, 6.21)	0.2	0.52	2.22 (2.06, 2.38)	3.73 (2.72, 4.60)	2.69 (1.92, 3.34)	2.77 (1.58, 3.67)	0.58	0.73	3.11 (2.89, 3.32)	1.78 (1.23, 2.43)	2.20 (1.36, 2.91)	2.31 (1.24, 3.12)	**0.04**	0.13
Guangdong Biolight M70	37	5.68 (4.65, 6.57)	4.32 (2.77, 5.66)	5.82 (4.39, 7.16)	6.21 (4.45, 7.70)	6.39 (3.85, 8.28)	0.23	0.26	6.94 (3.48, 9.45)	9.67 (7.28, 11.85)	9.67 (6.94, 12.14)	9.53 (5.81, 12.46)	0.31	0.32	3.05 (1.58, 4.70)	3.16 (2.26, 4.01)	4.58 (3.33, 5.97)	5.64 (4.16, 7.41)	0.16	0.06	3.25 (2.69, 3.93)	2.01 (1.55, 2.47)	3.04 (2.11, 3.87)	3.06 (1.67, 4.47)	0.51	0.63
Acare (2020) AH-MX	42	6.87 (5.88, 7.90)	8.47 (6.03, 10.80)	5.56 (4.59, 6.42)	6.71 (5.06, 8.11)	7.96 (6.75, 9.00)	0.35	0.97	14.22 (10.45, 17.73)	9.52 (8.03, 10.90)	11.03 (8.55, 13.00)	12.39 (10.41, 14.03)	0.3	0.65	4.38 (2.61, 5.62)	3.18 (2.50, 3.72)	4.19 (2.89, 5.24)	4.93 (3.81, 5.89)	0.97	0.45	1.72 (1.00, 2.35)	1.62 (1.08, 2.22)	1.27 (0.81, 1.68)	1.27 (0.70, 1.85)	0.37	0.4
American Oxygen LK-87	20	8.04 (7.48, 8.62)	8.36 (5.45, 9.94)	8.27 (7.30, 9.08)	7.46 (6.98, 7.99)	7.41 (6.84, 7.99)	0.66	0.64	8.42 (6.61, 9.77)	10.15 (9.29, 11.00)	9.43 (8.06, 10.62)	9.69 (8.10, 10.91)	0.46	0.38	11.77 (6.13, 15.54)	9.52 (7.75, 11.04)	8.66 (7.08, 10.00)	8.15 (6.49, 9.23)	0.43	0.36	3.19 (0.88, 5.38)	3.87 (2.29, 5.17)	3.05 (1.49, 4.52)	3.20 (1.44, 4.85)	0.96	0.93

Accuracy root mean square error (A_{RMS}) is reported for each device in a cohort of 24 participants as the overall point estimate with 95 % confidence interval (CI). Columns include the mean A_{RMS} and 95 % CI within each functional arterial oxygen saturation (SaO_2) decile and within four skin color bins. Skin color bins are defined by Individual Typology Angle (ITA) measured at the dorsal distal phalanx (DP): ITA > 30° (light), -30° ≤ ITA ≤ 30° (medium), ITA < -30° (dark), and ITA <-50° (very dark); note that “very dark” is a subset of the dark bin. P-values test differences in A_{RMS} between light versus dark and light versus very dark bins using a Welch’s two-sample t-test. No correction for multiple pairwise comparisons (e.g., Bonferroni) was applied, as the 2025 draft U.S. Food and Drug Administration (FDA) guidelines do not require adjustment. P-values near 0.05 are reported to three decimal places, and those meeting $p < 0.05$ are bolded. “NA” indicates that the A_{RMS} comparison between the respective skin color bins was not performed for

that device due to insufficient sample size. The 95% CI were determined using bootstrapping (random resampling with replacement) with 1,000 repetitions.

Table S2. Bias (mean of SpO_2 minus SaO_2) at the DP for 24 participant cohorts

Device	Device ID	Overall Bias (95% CI)	70–100% Bias SaO_2 (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI)					
			Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark
Nonin CO-Pilot	70	-0.28 (-0.34, -0.20)	0.05 (-0.23, 0.33)	-0.25 (-0.49, -0.03)	-0.43 (-0.65, -0.22)	-0.86 (-1.10, -0.64)	0.47	0.15	-0.70 (-1.21, -0.11)	0.11 (-0.28, 0.54)	-0.39 (-0.91, 0.13)	-1.38 (-1.84, -0.97)	0.99	0.35	0.38 (-0.10, 0.84)	-0.36 (-0.82, 0.08)	-0.70 (-1.08, -0.33)	-1.24 (-1.63, -0.87)	0.15	**0.02**	0.21 (-0.20, 0.57)	-0.47 (-0.75, -0.19)	-0.25 (-0.48, -0.03)	-0.19 (-0.45, 0.06)	0.3	0.37
Shenzhen (2023) PC-60NW	21	-0.04 (-0.15, 0.08)	0.31 (-0.06, 0.66)	-0.34 (-0.54, -0.13)	0.45 (0.15, 0.70)	0.45 (0.13, 0.75)	0.9	0.93	-0.69 (-1.46, -0.04)	-1.39 (-1.80, -1.02)	-0.86 (-1.42, -0.32)	-0.62 (-1.38, 0.04)	0.77	0.86	0.09 (-0.56, 0.67)	-0.32 (-0.61, -0.03)	0.89 (0.49, 1.28)	0.84 (0.35, 1.28)	0.24	0.41	1.26 (0.77, 1.69)	0.41 (0.14, 0.68)	1.19 (0.89, 1.52)	0.92 (0.65, 1.20)	0.92	0.63
Masimo Rad-97 with RD Rainbow Adt 8λ Adult Adhesive Sensor	73.09	0.86 (0.74, 0.97)	0.37 (0.21, 0.53)	1.08 (0.88, 1.27)	1.93 (1.67, 2.20)	1.77 (1.32, 2.25)	**0.005**	**0.04**	-0.16 (-0.54, 0.24)	0.77 (0.29, 1.19)	1.67 (1.20, 2.19)	1.40 (0.49, 2.30)	**0.01**	0.07	0.65 (0.47, 0.83)	1.47 (1.22, 1.74)	2.37 (1.94, 2.81)	2.39 (1.61, 3.17)	**0.002**	**0.01**	0.62 (0.43, 0.79)	0.94 (0.66, 1.24)	1.67 (1.32, 2.03)	1.38 (0.80, 1.94)	**0.03**	0.19
Shenzhen (2024) PC-60NW	81	-0.02 (-0.17, 0.12)	-0.04 (-0.36, 0.26)	0.33 (0.11, 0.53)	0.08 (-0.29, 0.42)	0.15 (-0.31, 0.59)	0.78	0.68	-1.06 (-1.73, -0.48)	-0.74 (-1.22, -0.29)	-1.93 (-2.59, -1.35)	-1.95 (-2.60, -1.30)	0.25	0.29	0.11 (-0.30, 0.53)	0.57 (0.26, 0.89)	0.62 (0.25, 0.96)	0.93 (0.52, 1.33)	0.32	0.18	0.82 (0.45, 1.20)	1.11 (0.95, 1.28)	1.26 (0.95, 1.56)	1.48 (1.16, 1.81)	0.58	0.43
Zug Medical Systems SAS SPLF-SE with SpotLife Reusable Probe	75	1.08 (0.97, 1.19)	0.67 (0.40, 0.93)	0.94 (0.67, 1.18)	1.59 (1.41, 1.77)	1.71 (1.47, 1.94)	**0.03**	0.053	-0.26 (-0.76, 0.30)	-0.02 (-0.54, 0.49)	0.98 (0.68, 1.31)	1.37 (0.93, 1.81)	0.1	0.14	0.82 (0.42, 1.23)	1.27 (0.97, 1.58)	2.07 (1.77, 2.38)	2.10 (1.71, 2.50)	**0.02**	**0.049**	1.40 (1.11, 1.70)	1.58 (1.31, 1.85)	1.60 (1.34, 1.85)	1.57 (1.21, 1.96)	0.55	0.68

Device	Device ID	Overall Bias (95% CI)	70–100% Bias SaO_2 (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI)					
			Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark
Beijing Choice MD300M	40	0.92 (0.79, 1.05)	0.11 (-0.62, 0.75)	0.84 (0.68, 1.00)	1.29 (0.92, 1.73)	1.93 (1.42, 2.42)	0.57	0.43	-1.83 (-2.71, -1.00)	1.67 (1.37, 1.98)	1.95 (1.26, 2.59)	2.36 (1.50, 3.34)	0.14	0.18	0.92 (0.47, 1.49)	0.79 (0.55, 1.04)	1.45 (0.81, 2.03)	2.06 (1.36, 2.67)	0.87	0.76	0.99 (-0.04, 1.82)	0.22 (-0.01, 0.46)	0.27 (-0.39, 0.91)	1.10 (0.23, 1.98)	0.73	0.98
Compass Health POX-ROS	12	-0.30 (-0.44, -0.17)	1.38 (0.78, 2.06)	-0.84 (-1.07, -0.61)	-0.63 (-0.94, -0.35)	-0.49 (-0.89, -0.08)	**0.03**	0.12	1.37 (-0.01, 3.03)	-1.02 (-1.47, -0.55)	-1.24 (-1.75, -0.76)	-1.28 (-1.88, -0.69)	**0.04**	0.052	0.97 (-0.16, 2.17)	-1.30 (-1.65, -0.95)	-0.51 (-0.93, -0.09)	-0.25 (-0.94, 0.35)	**0.046**	0.19	1.73 (1.28, 2.17)	-0.27 (-0.56, 0.04)	-0.16 (-0.69, 0.33)	-0.03 (-0.81, 0.65)	0.16	0.26
Shenzhen AM801	10	0.38 (0.21, 0.55)	1.10 (0.66, 1.56)	-0.16 (-0.44, 0.09)	1.03 (0.70, 1.42)	1.03 (0.70, 1.42)	0.88	0.88	0.89 (-0.00, 1.86)	-0.94 (-1.42, -0.48)	1.66 (0.85, 2.69)	1.66 (0.85, 2.69)	0.51	0.51	1.30 (0.77, 1.81)	0.01 (-0.48, 0.44)	1.14 (0.74, 1.52)	1.14 (0.74, 1.52)	0.74	0.74	1.10 (0.21, 1.91)	0.38 (0.01, 0.69)	0.34 (0.08, 0.64)	0.34 (0.08, 0.64)	0.32	0.32
CMS Zacurate 500C	15	0.78 (0.58, 0.96)	0.70 (-0.09, 1.40)	0.81 (0.61, 1.00)	0.20 (-0.16, 0.54)	0.14 (-0.33, 0.58)	0.58	0.68	-1.39 (-3.57, 0.58)	0.06 (-0.32, 0.45)	-2.08 (-2.93, -1.27)	-1.94 (-3.33, -0.54)	0.47	0.65	2.10 (1.03, 3.10)	1.57 (1.30, 1.86)	1.07 (0.73, 1.43)	0.90 (0.38, 1.43)	0.73	0.73	1.34 (0.92, 1.74)	0.87 (0.62, 1.11)	1.12 (0.80, 1.47)	0.86 (0.39, 1.35)	0.76	0.6
Acare (2023) Model AH-TA	56	1.54 (1.42, 1.64)	1.56 (1.02, 2.10)	1.54 (1.37, 1.70)	0.85 (0.48, 1.24)	0.68 (0.20, 1.12)	0.46	0.45	-0.18 (-1.40, 0.99)	1.33 (0.95, 1.71)	0.44 (-0.17, 1.01)	0.99 (0.13, 1.74)	0.64	0.48	2.69 (2.17, 3.19)	2.07 (1.77, 2.40)	1.35 (0.59, 2.07)	0.80 (-0.27, 1.61)	0.18	0.1	2.20 (1.71, 2.71)	1.38 (1.22, 1.54)	0.80 (0.14, 1.38)	0.35 (-0.33, 0.92)	0.28	0.08
Choicemed MD300CN350R	14	1.02 (0.89, 1.15)	1.75 (1.34, 2.11)	0.96 (0.70, 1.23)	0.27 (-0.13, 0.62)	-0.06 (-0.43, 0.31)	**0.01**	**<0.001**	0.13 (-0.93, 1.02)	1.13 (0.59, 1.72)	-1.33 (-2.08, -0.61)	-1.55 (-2.19, -0.91)	0.22	0.11	2.65 (2.34, 2.98)	1.30 (0.96, 1.66)	1.06 (0.62, 1.47)	1.03 (0.54, 1.51)	**0.009**	**0.01**	1.98 (1.52, 2.45)	0.52 (0.09, 0.88)	0.87 (0.38, 1.38)	0.08 (-0.37, 0.62)	0.18	**0.02**
Nonin Onyx Vantage 9590	35	1.36 (1.24, 1.47)	1.15 (0.82, 1.48)	1.69 (1.51, 1.88)	1.69 (1.36, 2.04)	1.02 (0.73, 1.30)	0.63	0.9	0.54 (0.06, 0.99)	2.71 (2.41, 3.05)	2.99 (2.05, 3.86)	1.75 (1.13, 2.31)	0.21	**0.046**	0.80 (0.23, 1.29)	1.28 (1.00, 1.54)	1.12 (0.71, 1.55)	0.98 (0.52, 1.41)	0.67	0.87	2.05 (1.52, 2.54)	1.08 (0.78, 1.37)	1.15 (0.79, 1.57)	0.69 (0.31, 1.09)	0.27	0.09
Nonin PalmSAT 2500	36	-1.35 (-1.47, -1.24)	-1.38 (-1.73, -1.02)	-1.50 (-1.72, -1.27)	-1.67 (-2.17, -1.20)	-2.00 (-2.57, -1.48)	0.85	0.63	-2.08 (-2.72, -1.46)	-2.25 (-2.68, -1.85)	-3.13 (-3.94, -2.31)	-3.79 (-4.66, -2.90)	0.55	0.28	-1.59 (-2.01, -1.07)	-1.74 (-2.07, -1.45)	-0.98 (-1.74, -0.20)	-1.23 (-2.14, -0.38)	0.9	0.87	-0.56 (-1.11, 0.02)	-0.75 (-1.03, -0.45)	-0.53 (-1.06, -0.02)	-0.61 (-1.20, -0.03)	0.91	0.99

Device	Device ID	Overall Bias (95% CI)	70–100% Bias SaO_2 (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI)					
			Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark
Shenzhen WIT-S300C	62	1.86 (1.73, 2.01)	1.91 (1.62, 2.20)	1.37 (0.91, 1.74)	1.71 (1.48, 1.93)	1.80 (1.60, 2.01)	0.78	0.89	1.20 (0.66, 1.78)	1.70 (1.11, 2.33)	1.91 (1.37, 2.43)	2.50 (2.03, 2.96)	0.3	0.08	2.12 (1.68, 2.59)	1.59 (1.09, 2.05)	1.90 (1.62, 2.20)	1.81 (1.53, 2.11)	0.78	0.73	2.32 (1.91, 2.82)	0.87 (-0.15, 1.59)	1.36 (1.11, 1.63)	1.23 (1.04, 1.45)	**0.049**	**0.02**
Masimo (2024) Mightysat	78	1.71 (1.58, 1.84)	0.42 (0.23, 0.61)	1.49 (1.31, 1.65)	3.95 (3.42, 4.52)	4.42 (3.80, 5.05)	**0.002**	**0.001**	-0.12 (-0.50, 0.28)	1.58 (1.15, 2.05)	5.29 (4.19, 6.38)	5.88 (4.80, 7.01)	**0.002**	**<0.001**	0.49 (0.20, 0.81)	1.71 (1.45, 1.96)	4.40 (3.42, 5.43)	4.62 (3.48, 5.74)	**0.001**	**0.001**	0.81 (0.58, 1.07)	1.21 (1.03, 1.40)	2.47 (2.10, 2.89)	2.72 (2.22, 3.21)	**0.001**	**<0.001**
Philips M1020B IntelliVue with M1191BL Probe	71	1.90 (1.81, 1.99)	1.60 (1.22, 1.92)	1.74 (1.62, 1.89)	3.25 (2.90, 3.62)	3.79 (3.28, 4.36)	0.21	0.18	2.45 (2.07, 2.83)	2.52 (2.30, 2.77)	4.53 (3.68, 5.32)	5.61 (4.58, 6.57)	0.3	0.18	1.29 (0.85, 1.70)	1.44 (1.28, 1.59)	3.02 (2.60, 3.49)	3.44 (2.83, 4.03)	0.15	0.1	1.20 (0.59, 1.76)	1.15 (0.91, 1.39)	2.27 (1.85, 2.70)	2.49 (1.68, 3.37)	0.26	0.33
Nihon Kohden OLV-4202 with NK-SPO2-ADVT Probe	64.03	0.36 (0.18, 0.53)	-0.34 (-0.88, 0.13)	-0.75 (-1.05, -0.45)	-0.03 (-0.52, 0.49)	0.13 (-0.46, 0.77)	0.8	0.75	-1.15 (-2.32, 0.11)	-1.14 (-1.80, -0.46)	0.44 (-0.45, 1.39)	1.28 (-0.25, 2.93)	0.44	0.44	-0.16 (-0.98, 0.58)	-0.99 (-1.44, -0.55)	-0.46 (-1.26, 0.28)	-0.46 (-1.51, 0.52)	0.93	0.91	0.30 (-0.10, 0.72)	-0.14 (-0.57, 0.26)	-0.16 (-0.74, 0.38)	-0.10 (-0.74, 0.46)	0.45	0.69
Covidien PM1000N	59	1.29 (1.22, 1.37)	1.79 (1.53, 2.06)	1.05 (0.81, 1.29)	3.27 (2.85, 3.70)	3.14 (2.70, 3.55)	0.06	0.11	1.39 (0.82, 2.02)	0.65 (0.08, 1.18)	4.50 (3.81, 5.18)	4.09 (3.26, 4.89)	0.052	0.12	2.33 (1.98, 2.68)	1.70 (1.27, 2.13)	3.73 (3.02, 4.36)	3.59 (2.77, 4.31)	0.08	0.14	1.61 (1.28, 1.95)	0.95 (0.73, 1.18)	1.70 (1.20, 2.18)	1.91 (1.42, 2.41)	0.74	0.41
Shenzhen A310L	11	1.43 (1.25, 1.61)	2.34 (1.53, 3.02)	0.90 (0.60, 1.19)	2.08 (1.78, 2.40)	2.41 (2.03, 2.79)	0.74	0.96	1.71 (-0.61, 3.63)	-0.51 (-1.22, 0.18)	2.53 (1.80, 3.19)	3.83 (3.29, 4.32)	0.44	0.22	3.71 (3.23, 4.21)	1.91 (1.54, 2.24)	2.41 (1.97, 2.88)	2.69 (2.11, 3.28)	0.16	0.38	1.58 (1.12, 2.12)	1.41 (1.22, 1.59)	1.34 (0.93, 1.71)	0.93 (0.51, 1.34)	0.81	0.34

Device	Device ID	Overall Bias (95% CI)	70–100% Bias SaO_2 (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI)					
			Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark
Masimo Rad-G with Radg-G Finger Clip Sensor	55.03	1.94 (1.77, 2.10)	1.18 (0.90, 1.45)	2.23 (2.00, 2.46)	2.18 (1.51, 2.81)	2.64 (1.87, 3.39)	0.39	0.29	0.51 (-0.04, 1.03)	1.85 (1.32, 2.36)	1.83 (-0.14, 3.77)	2.74 (0.60, 4.72)	0.94	0.62	1.67 (1.26, 2.06)	2.77 (2.44, 3.11)	2.72 (1.98, 3.54)	3.17 (2.13, 4.15)	0.31	0.25	1.37 (1.01, 1.65)	2.08 (1.84, 2.34)	1.83 (1.37, 2.33)	1.93 (1.30, 2.60)	0.31	0.27
CONTEC CMS50M	17	-1.24 (-1.44, -1.04)	-2.09 (-2.51, -1.68)	-1.18 (-1.54, -0.84)	-0.79 (-1.10, -0.47)	-0.97 (-1.41, -0.57)	**0.02**	0.053	-3.95 (-4.88, -3.04)	-2.66 (-3.40, -1.91)	-2.27 (-3.03, -1.56)	-2.69 (-3.48, -1.88)	0.39	0.61	-2.04 (-2.54, -1.53)	-0.94 (-1.37, -0.50)	-0.94 (-1.57, -0.36)	-1.17 (-1.87, -0.51)	0.14	0.26	-0.86 (-1.38, -0.36)	-0.20 (-0.79, 0.31)	0.32 (0.01, 0.64)	0.27 (-0.11, 0.60)	0.07	0.12
Covidien (2022) Nellcor PM10N	41	1.17 (0.99, 1.34)	2.77 (2.10, 3.51)	0.81 (0.56, 1.05)	2.70 (2.41, 3.00)	3.12 (2.79, 3.48)	0.93	0.81	2.55 (0.90, 4.56)	-0.12 (-0.74, 0.43)	3.22 (2.64, 3.74)	3.85 (3.20, 4.40)	0.71	0.49	3.18 (2.42, 4.00)	1.41 (1.09, 1.74)	2.96 (2.56, 3.37)	3.33 (2.92, 3.74)	0.75	0.99	2.62 (2.23, 3.01)	0.99 (0.70, 1.28)	1.67 (1.17, 2.20)	1.82 (1.31, 2.30)	0.3	0.24
Bistos Co., Ltd. BT-710	61	2.46 (2.33, 2.58)	2.51 (2.16, 2.86)	2.63 (2.45, 2.81)	3.04 (2.71, 3.35)	2.93 (2.63, 3.23)	0.44	0.58	3.54 (2.94, 4.12)	3.19 (2.93, 3.46)	4.36 (3.92, 4.83)	3.89 (3.37, 4.48)	0.21	0.36	2.48 (2.07, 2.87)	3.20 (2.88, 3.53)	3.42 (3.05, 3.83)	3.64 (3.19, 4.08)	0.44	0.31	1.67 (1.12, 2.27)	1.79 (1.56, 2.02)	1.36 (1.05, 1.71)	1.49 (1.16, 1.82)	0.56	0.77
Utech UT100	39	1.02 (0.85, 1.19)	0.80 (0.44, 1.15)	1.53 (1.27, 1.80)	2.19 (1.56, 2.85)	2.67 (1.51, 3.72)	0.19	0.1	-0.48 (-1.29, 0.27)	1.61 (1.02, 2.23)	4.03 (2.94, 5.28)	5.26 (3.50, 7.23)	**0.01**	**0.009**	1.47 (1.12, 1.86)	1.90 (1.38, 2.46)	2.04 (0.86, 3.12)	1.62 (-0.90, 3.36)	0.59	0.88	1.39 (1.06, 1.72)	1.20 (0.93, 1.47)	0.80 (0.21, 1.38)	1.20 (0.35, 1.96)	0.38	0.69
BodyMed BDMOXMTR BLK	2	2.12 (1.92, 2.32)	2.86 (2.27, 3.49)	1.95 (1.68, 2.25)	2.17 (1.79, 2.55)	1.49 (1.03, 1.94)	0.4	0.19	2.79 (1.18, 4.49)	0.69 (0.03, 1.36)	2.01 (1.26, 2.77)	1.64 (0.29, 2.99)	0.96	0.99	3.85 (3.16, 4.52)	3.06 (2.64, 3.48)	2.79 (2.06, 3.52)	1.76 (1.03, 2.53)	0.23	0.07	2.24 (1.67, 2.77)	2.12 (1.84, 2.38)	1.83 (1.36, 2.32)	1.21 (0.70, 1.75)	0.58	0.1
Masimo (2020) Mightysat Rx	7	0.92 (0.77, 1.08)	0.87 (0.30, 1.38)	0.77 (0.43, 1.12)	2.32 (1.83, 2.82)	3.13 (2.38, 3.98)	0.24	0.16	-0.36 (-1.49, 0.68)	-0.31 (-0.89, 0.32)	2.46 (1.07, 4.02)	5.04 (2.60, 7.70)	0.3	0.19	1.26 (0.51, 2.15)	1.53 (1.01, 2.13)	2.89 (2.32, 3.50)	3.03 (2.25, 4.02)	0.46	0.47	1.65 (1.15, 2.13)	0.95 (0.33, 1.45)	1.66 (1.35, 1.95)	1.82 (1.47, 2.26)	0.85	0.72

Device	Device ID	Overall Bias (95% CI)	70–100% Bias SaO_2 (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI)					
			Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark
EDAN (2020) H100b	44	2.34 (2.19, 2.51)	2.22 (1.83, 2.59)	2.52 (2.22, 2.81)	3.69 (3.28, 4.16)	4.06 (3.41, 4.74)	0.12	0.14	1.31 (0.40, 2.23)	2.79 (2.21, 3.36)	4.92 (4.10, 5.83)	5.06 (3.63, 6.61)	0.06	0.19	3.12 (2.61, 3.65)	2.67 (2.20, 3.15)	3.80 (3.32, 4.32)	3.88 (3.31, 4.42)	0.27	0.26	2.15 (1.88, 2.43)	2.14 (1.84, 2.44)	2.14 (1.71, 2.57)	2.93 (2.45, 3.44)	0.97	0.16
Guangdong Biolight M800	43	0.24 (0.04, 0.45)	-0.49 (-1.27, 0.15)	-0.57 (-1.06, -0.09)	0.68 (0.06, 1.30)	0.73 (0.12, 1.34)	0.25	0.25	-4.01 (-5.54, -2.65)	-4.36 (-5.25, -3.49)	-2.23 (-3.50, -1.01)	-1.79 (-3.01, -0.59)	0.41	0.31	0.68 (0.14, 1.27)	0.38 (-0.21, 0.95)	2.08 (1.34, 2.79)	2.12 (1.45, 2.76)	0.1	0.09	1.93 (1.60, 2.28)	1.67 (1.36, 1.98)	2.16 (1.63, 2.72)	1.51 (0.92, 2.03)	0.62	0.23
Masimo Rad-97 with Adult Probe, Product Code 4050	60	2.49 (2.39, 2.59)	2.78 (2.08, 3.53)	2.14 (1.91, 2.38)	4.71 (4.24, 5.16)	4.77 (4.12, 5.34)	0.3	0.36	2.77 (1.27, 4.48)	2.20 (1.74, 2.64)	6.17 (5.50, 6.87)	6.27 (5.30, 7.18)	0.59	0.63	3.84 (2.47, 5.28)	2.68 (2.24, 3.11)	5.25 (4.49, 6.01)	5.51 (4.61, 6.59)	0.2	0.19	1.90 (1.36, 2.49)	1.64 (1.29, 2.01)	2.55 (2.09, 3.03)	2.50 (1.99, 3.06)	0.5	0.58
Masimo Rad-G with Adult Probe, Product Code 4050	55	3.48 (3.31, 3.63)	2.88 (2.60, 3.19)	3.60 (3.31, 3.86)	5.58 (5.11, 6.05)	5.58 (5.11, 6.05)	**<0.001**	**<0.001**	3.20 (2.81, 3.62)	4.56 (4.13, 5.01)	7.12 (6.44, 7.73)	7.12 (6.44, 7.73)	**<0.001**	**<0.001**	3.21 (2.72, 3.73)	4.22 (3.76, 4.71)	5.70 (5.24, 6.23)	5.70 (5.24, 6.23)	**0.002**	**0.002**	2.25 (1.78, 2.70)	2.23 (1.91, 2.54)	3.26 (2.67, 3.80)	3.26 (2.67, 3.80)	0.2	0.2
Acare (2023) AH-M1	57	-2.36 (-2.66, -2.09)	0.77 (-0.14, 1.53)	-3.14 (-3.72, -2.58)	-1.58 (-2.39, -0.89)	-0.69 (-1.44, 0.03)	0.06	0.1	-2.81 (-3.94, -1.74)	-7.88 (-8.69, -7.10)	-6.30 (-7.83, -4.93)	-4.34 (-5.57, -3.11)	0.41	0.45	1.18 (0.22, 2.03)	-1.10 (-1.84, -0.25)	-0.67 (-1.37, -0.02)	-0.05 (-0.98, 0.89)	0.1	0.27	2.95 (2.46, 3.44)	0.23 (-0.04, 0.48)	1.20 (0.78, 1.66)	1.51 (1.04, 2.00)	0.18	0.28
Guangdong Biolight M70	37	-0.48 (-0.81, -0.14)	0.26 (-0.63, 1.18)	-2.06 (-2.79, -1.32)	1.79 (0.84, 2.67)	2.08 (0.68, 3.43)	0.51	0.55	-4.65 (-6.04, -3.22)	-7.30 (-8.69, -6.02)	1.12 (-1.76, 3.82)	1.65 (-2.60, 6.10)	0.15	0.13	1.33 (0.25, 2.34)	-0.67 (-1.39, 0.04)	2.12 (0.97, 3.21)	2.89 (0.85, 4.91)	0.73	0.68	2.96 (2.54, 3.38)	1.53 (1.27, 1.79)	2.02 (1.42, 2.59)	1.71 (0.87, 2.61)	0.26	0.28
Acare (2020) AH-MX	42	-3.98 (-4.31, -3.65)	-4.94 (-6.04, -	-2.94 (-3.57, -	-4.23 (-5.09, -	-5.48 (-6.69, -	0.48	0.78	-12.65 (-14.48, -	-8.70 (-9.92, -	-9.84 (-11.22, -	-11.84 (-13.25, -	0.21	0.61	-3.39 (-4.17, -	-2.22 (-2.77, -	-2.92 (-3.63, -	-3.89 (-5.06, -	0.78	0.65	0.22 (-0.22,	0.53 (0.23,	-0.06 (-0.40,	0.02 (-0.41,	0.55	0.64

Device	Device ID	Overall Bias (95% CI)	70–100% Bias SaO_2 (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI)					
			Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark	Light	Medium	Dark	Very Dark	P-value: Light vs Dark	P-value: Light vs Very Dark
			3.89)	2.29)	3.39)	4.28)			10.94)	7.58)	8.56)	10.37)			2.63)	1.68)	2.10)	2.91)			0.72)	0.86)	0.26)	0.37)		
American Oxygen LK-87	20	-5.71 (-6.11, -5.31)	-6.57 (-7.82, -5.19)	-6.02 (-6.66, -5.41)	-5.41 (-6.18, -4.52)	-5.47 (-6.32, -4.61)	0.29	0.37	-8.34 (-9.10, -7.65)	-7.89 (-9.23, -6.55)	-8.67 (-9.82, -7.47)	-9.54 (-10.36, -8.77)	0.82	0.36	-10.39 (-12.65, -8.18)	-8.53 (-9.37, -7.68)	-7.14 (-8.63, -5.64)	-6.61 (-8.13, -4.97)	0.13	0.07	-1.28 (-2.55, -0.23)	-2.11 (-2.73, -1.56)	-0.99 (-1.75, -0.35)	-0.98 (-1.88, -0.29)	0.83	0.85

This table reports the overall bias (mean of pulse oximeter measured oxygen saturation (SpO_2) - functional arterial oxygen saturation(SaO_2)) for each device in a cohort of 24 participants. The bias is presented with 95% confidence interval (CI) overall, and separately for each SaO_2 decile and for each skin color bin. Skin color bins were defined using the Individual Typology Angle (ITA) measured at the dorsal distal phalanx (DP): ITA > 30° for light, -30° ≤ ITA ≤ 30° for medium, ITA < -30° for dark, and ITA < -50° for dark; note that “very dark” is a subset of the dark bin. P-values indicate whether the SpO_2 bias differs significantly between light versus dark bins and between light versus very dark bins using a linear mixed-effects model with SpO_2- SaO_2 error as the outcome, skin color bin as a fixed-effect predictor, and participant ID as a random-effect grouping factor. No correction for multiple pairwise comparisons (e.g., Bonferroni) was applied, as the 2025 draft U.S. Food and Drug Administration (FDA) guidelines do not require adjustment. P-values near the 0.05 threshold are reported to three decimal places, and those meeting $p < 0.05$ are bolded. The 95% CI were determined using bootstrapping (random resampling with replacement) with 1,000 repetitions.

Table S3. Number of SaO_2 measurements (24 participant cohorts, max participant cohorts) stratified by skin color measured at the forehead

			70–100% SaO_2				70–80% SaO_2				80–90% SaO_2				90–100% SaO_2			
Device	**Device ID**	**Maximum Participant Cohort**	**Light**	**Medium**	**Dark**	**Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**
Nonin CO-Pilot	70	64	128, 401	238, 592	139, 389	85, 273	35, 119	75, 196	43, 106	29, 76	45, 138	80, 186	40, 126	24, 88	48, 144	83, 210	56, 157	32, 109
Shenzhen (2023) PC-60NW	21	35	185, 239	188, 299	123, 221	80, 134	50, 70	56, 86	36, 67	21, 40	65, 82	63, 105	42, 78	27, 45	70, 87	69, 108	45, 76	32, 49
Masimo Rad-97 with RD Rainbow Adt 8λ Adult Adhesive Sensor	73.09	24	257, 214	153, 207	137, 143	68, 72	85, 71	47, 68	41, 42	17, 18	86, 71	56, 68	51, 53	28, 30	86, 72	50, 71	45, 48	23, 24
Shenzhen (2024) PC-60NW	81	28	255, 236	157, 248	136, 157	70, 91	85, 84	53, 81	41, 48	20, 27	81, 73	49, 82	47, 55	26, 34	89, 79	55, 85	48, 54	24, 30
Zug Medical Systems SAS SPLF-SE with SpotLife Reusable Probe	75	40	179, 323	205, 357	147, 215	103, 131	61, 102	67, 105	37, 61	23, 35	60, 108	68, 127	50, 76	36, 50	58, 113	70, 125	60, 78	44, 46
Beijing Choice MD300MMd300m/Md300k2	40	36	254, 269	131, 309	115, 182	78, 121	83, 82	35, 94	40, 59	30, 41	75, 85	48, 106	40, 62	27, 41	96, 102	48, 109	35, 61	21, 39
Compass Health POX-ROS	12	32	145, 171	217, 305	150, 215	130, 174	46, 54	74, 101	50, 69	43, 57	46, 56	66, 98	47, 72	39, 54	53, 61	77, 106	53, 74	48, 63
Shenzhen AM801	10	28	141, 191	238, 255	129, 157	68, 94	47, 64	75, 74	41, 53	24, 35	48, 61	76, 88	43, 52	20, 28	46, 66	87, 93	45, 52	24, 31
CMS Zacurate 500C	15	28	119, 144	253, 220	122, 235	85, 150	42, 47	82, 70	32, 67	23, 43	33, 44	80, 70	45, 84	31, 52	44, 53	91, 80	45, 84	31, 55
Acare (2023) Model AH-TA	56	48	251, 383	132, 388	120, 244	64, 145	89, 124	49, 126	40, 74	20, 40	68, 113	37, 123	37, 79	23, 49	94, 146	46, 139	43, 91	21, 56

			70–100% SaO_2				70–80% SaO_2				80–90% SaO_2				90–100% SaO_2			
Device	**Device ID**	**Maximum Participant Cohort**	**Light**	**Medium**	**Dark**	**Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**
Choicemed MD300CN350R	14	45	195, 244	166, 423	125, 273	107, 214	54, 62	58, 135	38, 81	33, 65	70, 91	51, 132	43, 87	36, 66	71, 91	57, 156	44, 105	38, 83
Nonin Onyx Vantage 9590	35	40	221, 251	145, 357	116, 227	62, 153	76, 84	55, 126	33, 73	19, 52	66, 72	40, 101	41, 74	24, 51	79, 95	50, 130	42, 80	19, 50
Nonin PalmSAT 2500	36	40	248, 296	127, 328	115, 202	57, 123	84, 89	42, 113	45, 70	22, 42	64, 93	34, 100	31, 59	16, 38	100, 114	51, 115	39, 73	19, 43
Shenzhen WIT-S300C	62	43	132, 260	162, 426	221, 241	149, 149	38, 75	58, 137	62, 67	44, 44	46, 88	51, 135	77, 84	54, 54	48, 97	53, 154	82, 90	51, 51
Masimo (2024) Mightysat	78	31	151, 250	234, 239	115, 180	75, 97	44, 82	73, 74	36, 51	23, 24	56, 85	79, 79	35, 57	23, 30	51, 83	82, 86	44, 72	29, 43
Philips M1020B IntelliVue with M1191BL Probe	71	40	122, 334	260, 346	110, 221	51, 113	39, 113	94, 117	32, 69	12, 37	43, 105	88, 116	46, 75	20, 37	40, 116	78, 113	32, 77	19, 39
Nihon Kohden OLV-4202 with NK-SPO2-ADVT Probe	64.03	32	238, 277	115, 261	109, 173	53, 86	74, 96	41, 85	34, 51	21, 27	76, 87	42, 88	37, 59	16, 28	88, 94	32, 88	38, 63	16, 31
Covidien PM1000N	59	120	251, 983	126, 969	118, 633	56, 360	85, 308	39, 319	43, 180	22, 110	73, 315	38, 312	32, 215	18, 115	93, 360	49, 338	43, 238	16, 135
Shenzhen A310L	11	36	144, 230	235, 320	131, 220	88, 132	51, 73	78, 96	43, 70	28, 43	43, 67	77, 112	43, 76	28, 42	50, 90	80, 112	45, 74	32, 47
Masimo Rad-G with Rad-G Finger Clip Sensor	55.03	32	152, 275	239, 271	116, 172	60, 89	57, 96	77, 87	35, 49	15, 23	49, 83	83, 93	44, 61	26, 34	46, 96	79, 91	37, 62	19, 32
CONTEC CMS50M	17	28	200, 212	166, 238	129, 148	86, 86	49, 55	58, 80	34, 39	23, 23	72, 72	45, 69	44, 51	31, 31	79, 85	63, 89	51, 58	32, 32
Covidien (2022) Nellcor PM10N	41	32	238, 252	133, 260	119, 178	57, 93	77, 79	38, 77	44, 59	20, 29	76, 82	50, 88	41, 60	24, 35	85, 91	45, 95	34, 59	13, 29
Bistos Co., Ltd. BT-710	61	44	240, 324	131, 349	132, 262	65, 150	78, 101	44, 118	43, 73	20, 38	65, 95	39, 103	42, 97	24, 59	97, 128	48, 128	47, 92	21, 53
Utech UT100	39	36	124, 225	247, 348	109, 176	51, 117	39, 63	77, 106	36, 56	16, 36	38, 78	75, 116	34, 56	17, 39	47, 84	95, 126	39, 64	18, 42

			70–100% SaO_2				70–80% SaO_2				80–90% SaO_2				90–100% SaO_2			
Device	**Device ID**	**Maximum Participant Cohort**	**Light**	**Medium**	**Dark**	**Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**
BodyMed BDMOXMTRBLK	2	30	138, 162	217, 244	117, 201	64, 129	34, 42	77, 87	30, 54	19, 38	47, 55	55, 64	37, 66	21, 43	57, 65	85, 93	50, 81	24, 48
Masimo (2020) Mightysat Rx	7	44	148, 266	197, 462	144, 233	107, 155	41, 80	65, 142	42, 70	34, 50	53, 89	57, 151	50, 77	40, 54	54, 97	75, 169	52, 86	33, 51
EDAN (2020) H100b	44	32	169, 261	203, 259	109, 159	53, 78	57, 79	74, 91	44, 48	21, 25	48, 84	60, 72	31, 51	16, 25	64, 98	69, 96	34, 60	16, 28
Guangdong Biolight M800	43	48	223, 370	154, 405	117, 262	60, 140	78, 119	48, 131	36, 77	20, 41	58, 113	49, 136	40, 91	20, 51	87, 138	57, 138	41, 94	20, 48
Masimo Rad-97 with Adult Probe, Product Code 4050	60	88	226, 720	121, 737	125, 474	58, 248	86, 236	39, 242	44, 141	19, 70	56, 217	45, 234	40, 154	21, 86	84, 267	37, 261	41, 179	18, 92
Masimo Rad-G with Adult Probe, Product Code 4050	55	44	243, 342	126, 379	109, 229	76, 146	84, 104	43, 134	37, 70	27, 45	71, 106	37, 115	43, 76	30, 48	88, 132	46, 130	29, 83	19, 53
Acare (2023) AH-M1	57	44	190, 325	177, 373	111, 233	57, 133	62, 96	64, 125	34, 66	16, 33	57, 105	50, 114	37, 81	23, 52	71, 124	63, 134	40, 86	18, 48
Guangdong Biolight M70	37	42	159, 278	199, 337	117, 278	62, 202	44, 81	69, 118	36, 80	18, 59	55, 93	54, 95	36, 91	20, 69	60, 104	76, 124	45, 107	24, 74
Acare (2020) AH-MX	42	51	228, 291	149, 511	112, 291	54, 141	66, 84	39, 170	43, 88	20, 36	73, 92	56, 165	31, 95	16, 53	89, 115	54, 176	38, 108	18, 52
American Oxygen LK-87	20	32	158, 184	228, 345	122, 174	83, 112	46, 53	81, 115	39, 55	26, 36	52, 63	70, 111	38, 56	28, 37	60, 68	77, 119	45, 63	29, 39

Number of SpO_2:SaO_2 data pairs—where SpO_2 is pulse oximeter measured oxygen saturation and SaO_2 is functional arterial oxygen saturation—reported across four skin color bins defined by Individual Typology Angle (ITA) (ITA > 30° for light, -30° ≤ ITA ≤ 30° for medium, ITA < -30° for dark, and ITA < -50° for dark; note that “very dark” is a subset of the dark bin) and stratified by SaO_2 deciles. Each table cell shows the data pair count for the 24 participant cohort and, after the comma, the count for the maximum cohort size

possible that still meets anticipated 2025 U.S. Food and Drug Administration (FDA) and International Organization for Standardization (ISO) enrollment criteria.

Table S4. Complete list of devices, prices, specifications, and regulatory markings

Device	Device ID	Oximeter Type	Probe Name (if applicable)	Cost	510k	CE Marking	ISO 80601-2-61
Nonin CO-Pilot	70	Handheld	8330AA Adult Reusable Sensor *	$5,999	✓	✓	✓
Shenzhen (2023) PC-60NW	21	Fingertip		$40	✓	✓	✓
Masimo Rad-97	73.09	Tabletop	RD Rainbow Adt 8λ Adult Adhesive Sensor *	$1,730	✓	✓	✓
Shenzhen (2024) PC-60NW	81	Fingertip		$40	✓		✓
Zug Medical Systems SAS SPLF-SE	75	Tabletop	SpotLife Reusable Adult SpO2 Probe	$1,125			
Beijing Choice MD300M	40	Handheld	M-50E Probe	$259	✓	✓	
Compass Health POX-ROS	12	Fingertip		$21			
Shenzhen AM801	10	Fingertip		$30	✓		✓
CMS Zacurate 500C	15	Fingertip		$23			
Acare (2023) Model AH-TA	56	Handheld	ASANR-D3 Probe	$250	✓	✓	✓
Choicemed MD300CN350R	14	Fingertip		$30		✓	
Nonin Onyx Vantage 9590	35	Fingertip		$200	✓	✓	✓

Nonin PalmSAT 2500	36	Handheld	8000AA-1M Probe	$500	✓	✓	✓
Shenzhen WIT-S300C	62	Handheld	WTL403-01 Probe	$250	✓	✓	
Masimo (2024) Mightysat	78	Fingertip		$299	✓	✓	✓
Philips M1020B IntelliVue	71	Tabletop	M1191BL Probe	$4,470	✓	✓	
Nihon Kohden OLV-4202	64.03	Tabletop	NK-SPO2-ADVT Probe			✓	
Covidien PM1000N	59	Tabletop	DS100A probe	$2,000	✓		
Shenzhen A310L	11	Fingertip		$10	✓		
Masimo Rad-G Pulse Oximeter	55.03	Handheld	Masimo Rad-G Reusable Finger Clip Sensor	$485	✓	✓	✓
CONTEC CMS50M	17	Fingertip		$20		✓	
Covidien (2022) Nellcor PM10N	41	Handheld	DS100A Probe	$500	✓	✓	✓
Bistos Co., Ltd. BT-710	61	Handheld	U403-01 Probe	$350		✓	✓
Utech UT100	39	Handheld	CSS029B-1 Probe	$190			
BodyMed BDMOXMTRBLK	2	Fingertip		$25			

Masimo (2020) Mightysat Rx	7	Fingertip		$299	✓	✓	✓
EDAN (2020) H100b	44	Handheld	Default DB9 Adult Reusable Finger Sensor	$268	✓	✓	✓
Guangdong Biolight M800	43	Handheld	Default Probe	$329	✓	✓	
Masimo Rad-97	60	Tabletop	Reusable Adult Probe, Product Code 4050	$1,730	✓	✓	✓
Masimo Rad-G	55	Handheld	Reusable Adult Probe, Product Code 4050	$485	✓	✓	
Acare (2023) AH-M1	57	Handheld	ASANR-D3 Probe	$250	✓	✓	
Guangdong Biolight M70	37	Fingertip		$60	✓	✓	
Acare (2020) AH-MX	42	Handheld	ASANR-D3 Probe	$250	✓	✓	
American Oxygen LK-87	20	Fingertip		$25		✓	

This table provides details on each of the 34 tested devices, including device type, probe name, cost (in U.S. dollars), and applicable regulatory markings or reported regulatory compliance. Regulatory compliance and markings were found by searching manufacturer websites, product literature, the FDA 510(k) Premarket Notification Database, and by looking at the labels on the devices themselves. Device probes with > 2 wavelengths are denoted by an asterisk (*).

Table S5. Device conformity with current and anticipated FDA and ISO performance thresholds for maximum cohort sizes and DP ITA

		A_{RMS} Criteria				Differential Bias Criteria			Combination of A_{RMS} and Differential Bias Criteria		Significant Bias Difference (Dark - Light) at DP in at Least One Decile	Significant A_{RMS} Difference (Dark - Light) at DP in at Least One Decile	Significant Bias Difference (Very Dark - Light) at DP in at Least One Decile	Significant A_{RMS} Difference (Very Dark - Light) at DP in at Least One Decile
Device	**Device ID**	**Overall A_{RMS} (95% CI)**	**Pass 2013 FDA and Anticipated 2025 ISO (A_{RMS} Point Estimate ≤ 3%)**	**Pass 2017 ISO (A_{RMS} Point Estimate ≤ 4%)**	**Pass Anticipated 2025 FDA (A_{RMS} Upper CI < 3%)**	**Pass Anticipated 2025 FDA Pairwise MST Differential Bias Criteria**	**Pass Anticipated 2025 FDA ITA-Derived Differential Bias Criteria**	**Pass Anticipated 2025 ISO ITA-Derived Differential Bias Criteria**	**Pass Anticipated FDA 2025 (A_{RMS} Upper CI <3% and Differential Bias Thresholds)**	**Pass Anticipated 2025 ISO (A_{RMS} Point Estimate ≤3% and Differential Bias Thresholds)**				
Nonin CO-Pilot	70	1.39 (1.19, 1.60)	✓	✓	✓	✓	✓	✓	✓	✓				
Shenzhen (2023) PC-60NW	21	1.61 (1.44, 1.80)	✓	✓	✓	✓	✓	✓	✓	✓				✓ -0.69
Masimo Rad-97 with RD Rainbow Adt 8λ Adult Adhesive Sensor	73.09	1.64 (1.03, 1.91)	✓	✓	✓			✓		✓	✓ 2.09		✓ 2.87	
Shenzhen (2024) PC-60NW	81	1.87 (1.45, 2.28)	✓	✓	✓	✓		✓		✓	✓ -1.23			
Zug Medical Systems SAS SPLF-SE with SpotLife Reusable Probe	75	1.94 (1.66, 2.22)	✓	✓	✓	✓	✓	✓	✓	✓	✓ 1.62	✓ 0.80	✓ 1.48	
Compass Health POX-ROS	12	2.01 (1.73, 2.28)	✓	✓	✓			✓		✓				

		A_{RMS} Criteria				Differential Bias Criteria			Combination of A_{RMS} and Differential Bias Criteria		Significant Bias Difference (Dark - Light) at DP in at Least One Decile	Significant A_{RMS} Difference (Dark - Light) at DP in at Least One Decile	Significant Bias Difference (Very Dark - Light) at DP in at Least One Decile	Significant A_{RMS} Difference (Very Dark - Light) at DP in at Least One Decile
Device	**Device ID**	**Overall A_{RMS} (95% CI)**	**Pass 2013 FDA and Anticipated 2025 ISO (A_{RMS} Point Estimate ≤ 3%)**	**Pass 2017 ISO (A_{RMS} Point Estimate ≤ 4%)**	**Pass Anticipated 2025 FDA (A_{RMS} Upper CI < 3%)**	**Pass Anticipated 2025 FDA Pairwise MST Differential Bias Criteria**	**Pass Anticipated 2025 FDA ITA-Derived Differential Bias Criteria**	**Pass Anticipated 2025 ISO ITA-Derived Differential Bias Criteria**	**Pass Anticipated FDA 2025 (A_{RMS} Upper CI <3% and Differential Bias Thresholds)**	**Pass Anticipated 2025 ISO (A_{RMS} Point Estimate ≤3% and Differential Bias Thresholds)**				
Beijing Choice MD300M	40	2.05 (1.75, 2.37)	✓	✓	✓		✓	✓		✓	✓ 2.26	✓ 1.05		✓ 1.37
Nonin PalmSAT 2500	36	2.20 (1.87, 2.53)	✓	✓	✓		✓	✓		✓				
Covidien PM1000N	59	2.21 (2.06, 2.59)	✓	✓	✓	✓	✓	✓	✓	✓	✓ 3.01	✓ 1.23	✓ 3.23	✓ 1.53
Shenzhen AM801	10	2.21 (1.79, 2.61)	✓	✓	✓	✓	✓	✓	✓	✓		✓ -0.98		✓ -0.98
Nonin Onyx Vantage 9590	35	2.24 (1.84, 2.68)	✓	✓	✓	✓	✓	✓	✓	✓		✓ 1.80	✓ -1.08	✓ 0.85
Choicemed MD300CN350R	14	2.27 (2.03, 2.49)	✓	✓	✓	✓		✓		✓			✓ -1.69	
Masimo (2024) Mightysat	78	2.39 (1.98, 2.77)	✓	✓	✓			✓		✓	✓ 3.97	✓ 2.34	✓ 4.35	✓ 2.69
Acare (2023) Model AH-TA	56	2.40 (2.15, 2.65)	✓	✓	✓	✓	✓	✓	✓	✓				

		A_{RMS} Criteria				Differential Bias Criteria			Combination of A_{RMS} and Differential Bias Criteria		Significant Bias Difference (Dark - Light) at DP in at Least One Decile	Significant A_{RMS} Difference (Dark - Light) at DP in at Least One Decile	Significant Bias Difference (Very Dark - Light) at DP in at Least One Decile	Significant A_{RMS} Difference (Very Dark - Light) at DP in at Least One Decile
Device	**Device ID**	**Overall A_{RMS} (95% CI)**	**Pass 2013 FDA and Anticipated 2025 ISO (A_{RMS} Point Estimate ≤ 3%)**	**Pass 2017 ISO (A_{RMS} Point Estimate ≤ 4%)**	**Pass Anticipated 2025 FDA (A_{RMS} Upper CI < 3%)**	**Pass Anticipated 2025 FDA Pairwise MST Differential Bias Criteria**	**Pass Anticipated 2025 FDA ITA-Derived Differential Bias Criteria**	**Pass Anticipated 2025 ISO ITA-Derived Differential Bias Criteria**	**Pass Anticipated FDA 2025 (A_{RMS} Upper CI <3% and Differential Bias Thresholds)**	**Pass Anticipated 2025 ISO (A_{RMS} Point Estimate ≤3% and Differential Bias Thresholds)**				
Philips M1020B IntelliVue with M1191BL Probe	71	2.45 **(1.93, 3.15)**	✓	✓				✓		✓	✓ 1.82	✓ 2.03	✓ 1.45	✓ 1.97
CMS Zacurate 500C	15	2.51 (2.03, 2.96)	✓	✓	✓		✓	✓		✓				
Covidien (2022) Nellcor PM10N	41	2.55 **(1.92, 3.25)**	✓	✓				✓		✓				
Utech UT100	39	2.72 **(2.14, 3.33)**	✓	✓		✓		✓		✓	✓ 2.99	✓ 2.56	✓ 4.07	
Nihon Kohden OLV-4202 with NK-SPO2-ADVT Probe	64.03	2.76 **(2.14, 4.21)**	✓	✓				✓		✓		✓ 2.14		✓ 2.25
EDAN (2020) H100b	44	2.80 **(2.45, 3.18)**	✓	✓		✓		✓		✓				
Shenzhen WIT-S300C	62	2.81 **(2.20, 3.38)**	✓	✓		✓	✓	✓		✓		✓ 1.28	✓ 2.07	
CONTEC CMS50M	17	2.82 **(2.36, 3.23)**	✓	✓				✓		✓	✓			

		A_{RMS} Criteria				Differential Bias Criteria			Combination of A_{RMS} and Differential Bias Criteria		Significant Bias Difference (Dark - Light) at DP in at Least One Decile	Significant A_{RMS} Difference (Dark - Light) at DP in at Least One Decile	Significant Bias Difference (Very Dark - Light) at DP in at Least One Decile	Significant A_{RMS} Difference (Very Dark - Light) at DP in at Least One Decile
Device	**Device ID**	**Overall A_{RMS} (95% CI)**	**Pass 2013 FDA and Anticipated 2025 ISO (A_{RMS} Point Estimate ≤ 3%)**	**Pass 2017 ISO (A_{RMS} Point Estimate ≤ 4%)**	**Pass Anticipated 2025 FDA (A_{RMS} Upper CI < 3%)**	**Pass Anticipated 2025 FDA Pairwise MST Differential Bias Criteria**	**Pass Anticipated 2025 FDA ITA-Derived Differential Bias Criteria**	**Pass Anticipated 2025 ISO ITA-Derived Differential Bias Criteria**	**Pass Anticipated FDA 2025 (A_{RMS} Upper CI <3% and Differential Bias Thresholds)**	**Pass Anticipated 2025 ISO (A_{RMS} Point Estimate ≤3% and Differential Bias Thresholds)**				
											1.72			
Shenzhen A310L	11	2.84 **(2.47, 3.18)**	✓	✓		✓		✓		✓			✓ 2.89	
Masimo (2020) Mightysat Rx	7	2.88 **(2.00, 4.11)**	✓	✓				✓		✓			✓ 3.28	
Masimo Rad-G with Rad-G Finger Clip Sensor	55.03	**3.00 (2.43, 3.54)**	✓	✓				✓		✓				
Bistos Co., Ltd. BT-710	61	**3.15 (2.86, 3.42)**		✓		✓	✓	✓					✓ 1.54	
BodyMed BDMOXMTRBLK	2	**3.26 (2.92, 3.64)**		✓		✓		✓					✓ -1.17	✓ -0.84
Masimo Rad-97 with Adult Probe, Product Code 4050	60	**3.33 (2.90, 3.77)**		✓				✓			✓ 3.01	✓ 2.32	✓ 3.77	✓ 2.70

		A_{RMS} Criteria				Differential Bias Criteria			Combination of A_{RMS} and Differential Bias Criteria		Significant Bias Difference (Dark - Light) at DP in at Least One Decile	Significant A_{RMS} Difference (Dark - Light) at DP in at Least One Decile	Significant Bias Difference (Very Dark - Light) at DP in at Least One Decile	Significant A_{RMS} Difference (Very Dark - Light) at DP in at Least One Decile
Device	**Device ID**	**Overall A_{RMS} (95% CI)**	**Pass 2013 FDA and Anticipated 2025 ISO (A_{RMS} Point Estimate ≤ 3%)**	**Pass 2017 ISO (A_{RMS} Point Estimate ≤ 4%)**	**Pass Anticipated 2025 FDA (A_{RMS} Upper CI < 3%)**	**Pass Anticipated 2025 FDA Pairwise MST Differential Bias Criteria**	**Pass Anticipated 2025 FDA ITA-Derived Differential Bias Criteria**	**Pass Anticipated 2025 ISO ITA-Derived Differential Bias Criteria**	**Pass Anticipated FDA 2025 (A_{RMS} Upper CI <3% and Differential Bias Thresholds)**	**Pass Anticipated 2025 ISO (A_{RMS} Point Estimate ≤3% and Differential Bias Thresholds)**				
Guangdong Biolight M800	43	**3.53 (3.17, 3.88)**		✓		✓		✓			✓ 1.09	✓ 1.30	✓ 1.09	✓ -2.67
Masimo Rad-G with Adult Probe, Product Code 4050	55	**4.27 (3.70, 4.89)**						✓			✓ 3.35	✓ 2.51	✓ 3.35	✓ 2.84
Guangdong Biolight M70	37	**5.08 (4.30, 5.91)**				✓						✓ 2.29		✓ 2.52
Acare (2023) AH-M1	57	**5.28 (4.60, 5.96)**						✓						
Acare (2020) AH-MX	42	**6.80 (6.05, 7.48)**						✓						
American Oxygen LK-87	20	**7.75 (7.23, 8.32)**						✓						

This table lists devices from lowest to highest accuracy root mean square error (A_{RMS}) with 95% confidence intervals (CI) for each device tested in a cohort that is as large as possible while satisfying enrollment criteria in the anticipated 2025 U.S. Food and Drug Administration (FDA) and 2025 International Organization for Standardization (ISO) guidelines. The boxes with check marks (✓) indicate that the device meets the A_{RMS} or differential bias thresholds (differential bias and 95% CI < 3.5% for 70–85% SaO_2 and <

1.5% for 85–100% SaO_2 for 2025 FDA, and differential bias point estimate $\leq$ 4% for 70–85% SaO_2 and $\leq$ 2% for 85–100% SaO_2 for 2025 ISO). Differential bias was calculated with the Monk Skin Tone (MST) Scale and individual typology angle (ITA) as indicated. Finally, the table indicates whether each device has demonstrated statistically significant difference in bias or A_{RMS} between light versus dark and light versus very dark skin color bins in at least one of the functional arterial oxygen saturation (SaO_2) deciles. Skin color bins were defined using the Individual Typology Angle (ITA) measured at the dorsal distal phalanx (DP): ITA > 30° for light, -30° ≤ ITA ≤ 30° for medium, ITA < -30° for dark, and ITA <-50° for very dark; note that “very dark” is a subset of the dark bin. Devices that have FDA 510(k) (2013 FDA) are highlighted in grey. The A_{RMS} values and their 95% CIs are shown in bold if they fail to meet any of the A_{RMS} criteria. The 95% CI were determined using bootstrapping (random resampling with replacement) with 1,000 repetitions.

Table S6. A_{RMS} at the forehead for 24 participant cohorts

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Nonin CO-Pilot	70	1.69 (1.35, 2.03)	1.14 (1.06, 1.22)	1.91 (1.32, 2.50)	1.79 (1.09, 2.31)	2.20 (1.55, 2.69)	0.19	0.1	1.43 (1.28, 1.57)	2.35 (1.23, 3.58)	2.05 (1.08, 2.69)	2.59 (1.64, 3.17)	0.48	0.16	1.16 (0.91, 1.36)	1.99 (1.46, 2.46)	1.91 (0.92, 2.68)	2.41 (0.97, 3.20)	0.33	0.21	0.85 (0.66, 1.01)	1.51 (0.93, 2.13)	1.29 (0.71, 1.73)	1.40 (0.59, 2.01)	0.32	0.38
Shenzhen (2023) PC-60NW	21	1.78 (1.52, 2.02)	1.83 (1.45, 2.18)	1.78 (1.31, 2.19)	1.80 (1.45, 2.27)	1.98 (1.54, 2.54)	1	0.64	2.20 (1.51, 2.87)	2.01 (1.22, 2.73)	2.84 (1.58, 3.99)	3.33 (1.83, 4.51)	0.56	0.28	1.71 (1.33, 2.02)	1.75 (1.21, 2.18)	1.36 (1.02, 1.67)	1.41 (1.06, 1.80)	0.29	0.45	1.70 (1.21, 2.17)	1.71 (1.29, 2.08)	1.22 (0.93, 1.46)	1.08 (0.75, 1.32)	0.29	0.16
Masimo Rad-97 with RD Rainbow Adt 8λ Adult Adhesive Sensor	73.09	1.89 (1.51, 2.27)	1.29 (0.95, 1.60)	2.15 (1.64, 2.64)	2.46 (1.54, 3.22)	2.22 (0.84, 3.64)	0.12	0.56	1.64 (1.13, 2.06)	2.15 (1.88, 2.38)	2.56 (1.44, 3.47)	2.32 (0.78, 3.79)	0.31	0.71	0.99 (0.72, 1.24)	2.39 (1.66, 3.04)	2.77 (1.69, 3.66)	2.53 (0.73, 4.23)	0.06	0.46	1.19 (0.84, 1.49)	1.85 (0.92, 2.55)	1.88 (1.31, 2.33)	1.57 (0.63, 2.42)	0.12	0.65
Shenzhen (2024) PC-60NW	81	1.92 (1.47, 2.42)	1.73 (1.34, 2.10)	1.52 (0.98, 2.14)	2.60 (1.51, 3.73)	1.94 (1.27, 2.46)	0.3	0.56	2.11 (1.44, 2.74)	1.51 (1.03, 1.97)	4.03 (1.85, 6.11)	2.76 (1.78, 3.87)	0.24	0.38	1.36 (1.01, 1.67)	1.76 (0.79, 2.75)	1.79 (1.12, 2.54)	1.22 (0.95, 1.58)	0.41	0.83	1.61 (1.25, 1.95)	1.07 (0.78, 1.32)	1.62 (1.24, 1.98)	1.67 (0.95, 2.29)	0.87	0.89
Zug Medical Systems SAS SPLF-SE with SpotLife Reusable Probe	75	1.97 (1.6, 2.34)	1.68 (1.23, 2.06)	2.02 (1.29, 2.79)	2.11 (1.68, 2.49)	1.76 (1.50, 2.10)	0.17	0.55	1.97 (1.52, 2.29)	1.91 (1.00, 2.88)	2.00 (1.15, 2.71)	1.67 (0.74, 2.20)	0.8	0.48	1.38 (0.81, 1.97)	2.14 (1.29, 2.91)	2.30 (1.58, 3.00)	1.94 (1.48, 2.52)	0.06	0.12	1.72 (1.26, 2.08)	1.97 (1.26, 2.55)	2.06 (1.45, 2.58)	1.54 (1.24, 1.80)	0.43	0.79

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Beijing Choice MD300M	40	2.01 (1.66, 2.34)	1.73 (1.30, 2.14)	1.79 (1.15, 2.38)	2.58 (1.97, 3.13)	2.73 (1.76, 3.38)	**0.04**	0.1	2.01 (1.31, 2.72)	2.47 (0.82, 3.50)	3.03 (1.96, 3.86)	3.34 (1.85, 4.24)	0.12	0.12	1.51 (1.16, 1.83)	1.32 (0.73, 1.82)	2.59 (1.98, 3.14)	2.63 (2.02, 3.37)	**0.01**	0.052	1.62 (1.18, 2.01)	1.40 (1.06, 1.68)	1.85 (1.08, 2.49)	1.81 (0.85, 2.71)	0.6	0.83
Compass Health POX-ROS	12	2.2 (1.86, 2.57)	2.40 (1.55, 3.19)	2.16 (1.68, 2.57)	1.94 (1.63, 2.21)	2.04 (1.78, 2.28)	0.54	0.71	2.92 (1.42, 3.97)	2.51 (1.39, 3.33)	2.28 (1.26, 3.18)	2.44 (1.35, 3.33)	0.56	0.71	2.80 (1.39, 4.28)	2.18 (1.51, 2.85)	1.66 (1.26, 1.95)	1.75 (1.35, 2.05)	0.38	0.46	1.77 (1.17, 2.25)	1.65 (1.29, 1.94)	1.92 (1.41, 2.43)	1.95 (1.33, 2.51)	0.68	0.69
Shenzhen AM801	10	2.32 (1.86, 2.75)	2.57 (1.88, 3.36)	2.13 (1.66, 2.65)	2.26 (1.40, 3.20)	2.58 (0.93, 4.20)	0.51	0.81	1.94 (1.48, 2.35)	2.95 (1.98, 3.92)	3.48 (1.57, 5.42)	4.38 (1.03, 7.39)	0.4	0.56	2.95 (2.10, 3.90)	1.80 (1.37, 2.17)	1.75 (1.47, 1.99)	1.41 (1.20, 1.52)	0.07	**0.03**	2.83 (1.65, 3.96)	1.39 (1.09, 1.62)	1.06 (0.69, 1.42)	0.69 (0.48, 0.84)	0.06	**0.03**
CMS Zacurate 500C	15	2.41 (1.91, 2.89)	2.83 (1.74, 3.91)	2.14 (1.57, 2.70)	2.48 (1.65, 3.29)	2.36 (1.15, 3.52)	0.73	0.62	3.66 (2.11, 5.29)	2.56 (1.66, 3.37)	3.88 (2.22, 5.60)	4.17 (1.76, 6.53)	0.88	0.88	2.22 (1.61, 2.81)	2.18 (1.39, 2.91)	1.91 (0.85, 2.86)	1.26 (0.51, 1.71)	0.43	0.11	1.69 (1.20, 2.12)	1.65 (1.15, 2.08)	1.68 (0.96, 2.21)	1.51 (0.78, 2.30)	0.81	0.59
Acare (2023) Model AH-TA	56	2.41 (2.08, 2.74)	2.39 (1.92, 2.81)	2.59 (1.99, 3.07)	2.28 (1.57, 3.04)	2.60 (1.54, 3.84)	0.77	0.88	2.44 (1.84, 2.98)	2.36 (1.48, 3.15)	2.58 (1.82, 3.23)	2.89 (1.63, 3.82)	0.66	0.49	2.62 (2.03, 3.15)	3.14 (2.44, 3.69)	2.46 (1.63, 3.31)	2.77 (1.59, 4.18)	0.79	0.93	2.11 (1.53, 2.67)	2.31 (1.30, 3.44)	1.90 (1.20, 2.66)	2.18 (0.87, 3.45)	0.7	0.97
Choicemed MD300CN350R	14	2.41 (2.05, 2.75)	2.39 (1.95, 2.78)	2.56 (1.74, 3.23)	2.12 (1.58, 2.56)	2.24 (1.73, 2.66)	0.51	0.77	2.36 (1.75, 2.95)	3.90 (2.33, 5.47)	2.55 (1.66, 3.44)	2.71 (1.77, 3.67)	0.77	0.6	2.64 (1.99, 3.19)	1.84 (1.25, 2.40)	1.86 (0.99, 2.52)	1.92 (0.84, 2.61)	0.16	0.24	2.29 (1.72, 2.83)	2.13 (0.87, 3.31)	1.87 (1.11, 2.50)	2.02 (1.16, 2.67)	0.42	0.64
Nonin Onyx Vantage 9590	35	2.44 (1.90, 2.99)	2.16 (1.69, 2.70)	2.50 (1.92, 3.14)	2.67 (1.22, 4.10)	3.45 (0.98, 5.59)	0.86	0.63	2.70 (1.49, 3.81)	2.90 (2.23, 3.50)	3.53 (1.64, 5.39)	4.51 (1.40, 7.30)	0.57	0.5	1.40 (1.01, 1.70)	2.14 (1.20, 2.89)	1.80 (0.90, 2.45)	2.12 (0.85, 3.10)	0.56	0.42	2.20 (1.63, 2.75)	2.22 (1.11, 3.00)	1.48 (1.02, 1.82)	1.68 (0.69, 2.13)	0.14	0.47

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Nonin PalmSAT 2500	36	2.54 (1.97, 3.11)	2.00 (1.54, 2.58)	3.02 (1.66, 4.35)	3.07 (1.90, 4.02)	2.49 (1.04, 3.78)	0.2	0.72	2.36 (1.64, 3.18)	4.10 (1.83, 6.20)	4.03 (2.14, 5.41)	3.36 (0.73, 5.34)	0.24	0.68	1.85 (1.37, 2.30)	2.91 (1.98, 3.76)	2.62 (1.95, 3.11)	2.30 (1.50, 3.36)	0.08	0.55	1.76 (1.36, 2.18)	1.83 (1.25, 2.30)	1.71 (1.18, 2.21)	1.28 (0.84, 1.45)	0.94	0.24
Shenzhen WIT-S300C	62	2.56 (2.12, 3.01)	2.68 (1.82, 3.56)	2.40 (1.44, 3.41)	2.64 (2.20, 3.05)	2.25 (1.71, 2.71)	0.96	0.52	2.20 (1.40, 3.13)	2.42 (1.71, 3.00)	3.22 (2.53, 3.83)	2.81 (1.95, 3.54)	0.11	0.38	2.73 (1.66, 3.77)	1.46 (1.14, 1.71)	2.72 (2.03, 3.30)	2.20 (1.49, 2.89)	0.96	0.51	2.91 (2.02, 3.69)	2.97 (0.96, 4.82)	2.02 (1.35, 2.69)	1.56 (1.06, 2.00)	0.11	**0.04**
Masimo (2024) Mightysat	78	2.72 (1.77, 3.76)	1.33 (1.11, 1.54)	2.19 (1.51, 2.73)	4.32 (2.31, 6.31)	4.75 (2.39, 7.36)	0.07	0.15	1.54 (1.09, 1.88)	2.59 (1.72, 3.33)	5.15 (2.77, 7.28)	5.60 (1.87, 8.07)	0.06	0.17	1.27 (1.06, 1.44)	2.26 (1.56, 2.82)	4.63 (2.57, 6.62)	5.18 (2.40, 7.63)	0.046	0.1	1.29 (0.82, 1.68)	1.59 (1.17, 1.97)	2.80 (1.81, 3.75)	3.02 (1.92, 4.24)	0.046	0.09
Philips M1020B IntelliVue with M1191BL Probe	71	2.72 (2.17, 3.27)	1.92 (1.74, 2.10)	2.28 (1.93, 2.59)	4.17 (2.24, 5.44)	4.76 (2.43, 6.39)	0.11	0.15	2.71 (2.28, 3.13)	2.92 (2.25, 3.41)	5.45 (3.22, 7.01)	7.20 (6.50, 7.84)	0.17	0.07	1.42 (1.11, 1.70)	1.84 (1.57, 2.07)	3.70 (2.17, 4.73)	4.17 (2.58, 5.39)	0.06	0.08	1.41 (1.19, 1.59)	1.92 (1.22, 2.55)	2.89 (1.52, 4.26)	3.59 (2.31, 5.26)	0.23	0.18
Nihon Kohden OLV-4202 with NK-SPO2-ADVT Probe	64.03	2.76 (2.15, 3.30)]	3.01 (2.17, 3.74)	1.87 (1.12, 2.48)	2.40 (1.31, 3.20)	2.34 (0.84, 3.24)	0.41	0.54	4.24 (2.81, 5.57)	2.25 (1.34, 3.19)	3.23 (1.52, 4.35)	3.02 (0.87, 4.89)	0.41	0.47	2.61 (1.74, 3.34)	1.80 (0.70, 2.65)	2.16 (1.23, 2.89)	2.15 (0.52, 2.81)	0.58	0.72	2.08 (1.14, 3.09)	0.99 (0.67, 1.30)	1.80 (1.01, 2.48)	1.65 (1.40, 1.93)	0.87	0.98
Covidien PM1000N	59	2.78 (2.22, 3.33)	2.46 (1.83, 3.05)	1.80 (1.35, 2.12)	4.05 (2.88, 4.92)	4.96 (4.23, 5.69)	0.08	**0.005**	2.80 (1.98, 3.50)	1.93 (1.39, 2.48)	4.83 (3.33, 6.10)	5.97 (4.28, 6.99)	0.06	**0.03**	2.66 (1.76, 3.42)	2.18 (1.43, 2.72)	4.12 (3.07, 4.91)	4.95 (4.65, 5.40)	0.09	**<0.001**	1.92 (1.44, 2.40)	1.30 (0.70, 1.70)	2.47 (1.67, 3.17)	2.80 (1.85, 3.85)	0.3	0.25

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Shenzhen A310L	11	2.95 (2.60, 3.27)	2.85 (2.28, 3.47)	3.11 (2.55, 3.59)	2.78 (2.15, 3.42)	2.46 (2.02, 2.93)	0.88	0.44	3.76 (2.28, 5.23)	4.31 (3.06, 5.39)	3.33 (2.19, 4.26)	3.08 (1.83, 3.98)	0.8	0.67	2.36 (1.65, 2.91)	3.21 (2.65, 3.74)	2.92 (2.24, 3.72)	2.47 (2.03, 2.78)	0.28	0.57	1.76 (1.35, 2.16)	1.83 (1.43, 2.19)	1.94 (1.31, 2.38)	1.66 (0.91, 2.39)	0.77	0.72
Masimo Rad-G with Rad-G Finger Clip Sensor	55.03	2.97 (2.27, 3.65)	2.19 (1.63, 2.78)	3.04 (2.11, 3.88)	3.60 (1.71, 5.19)	4.29 (2.02, 6.74)	0.33	0.39	2.21 (1.31, 3.03)	3.43 (2.32, 4.35)	4.13 (2.23, 5.96)	5.42 (3.14, 7.84)	0.17	0.15	2.25 (1.48, 2.92)	3.17 (2.23, 4.01)	3.73 (1.46, 5.32)	4.00 (1.65, 6.46)	0.4	0.49	2.17 (1.62, 2.63)	2.41 (1.62, 3.22)	2.71 (1.41, 3.69)	2.98 (1.60, 4.56)	0.61	0.58
CONTEC CMS50M	17	2.97 (2.50, 3.44)	2.77 (2.26, 3.32)	3.59 (2.59, 4.37)	2.21 (1.36, 2.92)	2.22 (1.43, 3.11)	0.22	0.33	4.14 (2.95, 5.22)	4.85 (2.67, 6.61)	3.65 (2.11, 5.02)	3.04 (2.23, 3.82)	0.61	0.33	2.30 (1.76, 2.90)	2.43 (1.96, 2.87)	2.18 (0.86, 3.24)	2.28 (0.89, 3.69)	0.55	0.69	1.94 (1.49, 2.33)	3.40 (1.46, 5.39)	1.16 (0.81, 1.46)	1.28 (0.86, 1.64)	**0.03**	0.11
Covidien (2022) Nellcor PM10N	41	3.00 (2.37, 3.69)	3.34 (2.29, 4.48)	1.82 (1.18, 2.52)	3.28 (2.36, 3.99)	3.59 (1.74, 4.30)	0.9	0.68	4.34 (2.40, 6.10)	2.32 (1.76, 2.92)	3.88 (2.74, 4.77)	4.33 (2.57, 5.54)	0.97	0.62	3.09 (2.15, 3.95)	1.80 (0.84, 2.70)	3.24 (2.36, 4.02)	3.58 (1.97, 4.80)	0.64	0.52	2.49 (1.88, 3.01)	1.76 (0.65, 2.83)	2.09 (1.40, 2.58)	1.75 (0.50, 2.59)	0.46	0.35
Bistos Co., Ltd. BT-710	61	3.17 (2.82, 3.55)	2.90 (2.44, 3.32)	3.14 (2.22, 3.97)	3.63 (2.88, 4.19)	3.98 (3.57, 4.27)	0.16	**0.006**	3.49 (2.88, 4.00)	3.50 (2.79, 4.19)	4.76 (3.86, 5.48)	5.30 (5.03, 5.68)	0.07	**<0.001**	2.91 (2.32, 3.47)	3.50 (2.06, 4.69)	3.54 (2.51, 4.33)	3.74 (2.04, 4.79)	0.41	0.43	2.29 (1.83, 2.71)	2.06 (0.96, 2.96)	1.89 (1.38, 2.23)	1.79 (1.28, 2.04)	0.33	0.27
Utech UT100	39	3.25 (2.51, 3.95)	3.09 (2.09, 4.21)	2.49 (1.80, 3.03)	4.53 (2.79, 5.85)	4.81 (3.86, 6.22)	0.26	0.12	3.96 (2.71, 5.25)	2.51 (1.82, 3.12)	5.23 (3.04, 7.50)	4.48 (3.99, 4.72)	0.56	0.31	3.01 (1.58, 4.42)	3.15 (1.93, 4.06)	4.38 (2.20, 6.46)	5.96 (3.87, 8.38)	0.41	0.14	2.23 (1.57, 2.86)	1.85 (1.34, 2.28)	2.17 (1.77, 2.54)	2.24 (1.83, 2.73)	0.96	0.81

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
BodyMed BDMOXMTRBLK	2	3.28 (2.86, 3.75)	3.80 (2.71, 4.96)	3.08 (2.65, 3.41)	3.11 (2.20, 3.74)	3.35 (1.27, 4.16)	0.4	0.69	3.93 (1.99, 5.84)	3.42 (2.71, 4.01)	3.25 (2.25, 4.07)	2.31 (1.48, 2.99)	0.9	0.38	4.98 (3.15, 6.85)	3.61 (2.78, 4.38)	3.97 (2.21, 5.22)	4.69 (0.57, 6.34)	0.42	0.82	2.64 (1.98, 3.12)	2.63 (2.29, 2.94)	2.42 (1.39, 3.06)	2.65 (0.73, 3.29)	0.63	0.94
Masimo (2020) Mightysat Rx	7	3.67 (2.34, 5.10)	2.27 (1.57, 2.91)	3.14 (2.17, 3.96)	4.32 (1.93, 6.73)	2.48 (1.87, 2.96)	0.33	0.56	2.40 (1.41, 3.21)	2.94 (1.51, 4.24)	6.36 (1.60, 10.49)	2.47 (1.09, 3.49)	0.36	0.87	2.87 (1.07, 4.39)	3.27 (1.74, 4.71)	4.17 (2.17, 6.24)	2.79 (2.19, 3.23)	0.34	0.54	2.11 (1.76, 2.42)	3.28 (1.80, 4.91)	2.11 (1.42, 2.70)	1.89 (1.34, 2.41)	0.79	0.47
EDAN (2020) H100b	44	3.79 (3.06, 4.7)	3.04 (2.25, 3.94)	3.45 (2.80, 4.02)	4.95 (3.24, 6.88)	3.59 (2.87, 4.66)	0.15	0.4	3.32 (2.09, 4.53)	4.15 (3.25, 4.84)	6.06 (3.15, 9.01)	3.69 (2.15, 5.38)	0.22	0.7	3.31 (2.59, 3.96)	3.40 (2.50, 4.20)	4.40 (3.50, 5.16)	4.22 (3.18, 5.39)	0.09	0.27	2.41 (1.90, 2.95)	2.44 (1.84, 2.91)	2.96 (2.37, 3.48)	2.52 (1.97, 3.04)	0.21	0.68
Guangdong Biolight M800	43	3.96 (3.44, 4.46)	4.31 (3.39, 5.10)	3.45 (2.76, 4.01)	3.62 (2.72, 4.41)	3.53 (2.01, 5.02)	0.42	0.52	6.36 (4.72, 7.85)	5.35 (4.06, 6.33)	4.22 (2.19, 6.13)	2.68 (1.09, 3.96)	0.14	**0.03**	2.40 (1.88, 2.91)	2.09 (1.48, 2.66)	3.72 (2.42, 4.99)	4.53 (2.22, 6.45)	0.16	0.25	2.52 (1.91, 3.15)	2.06 (1.08, 3.02)	3.02 (2.15, 3.89)	3.37 (2.13, 4.90)	0.39	0.46
Masimo Rad-97 with Adult Probe, Product Code 4050	60	4.21 (3.06, 5.24)	4.18 (2.19, 6.15)	2.75 (1.66, 3.62)	5.71 (4.33, 6.77)	5.84 (4.69, 7.18)	0.07	0.052	5.79 (2.24, 9.28)	2.76 (1.66, 3.85)	6.89 (5.29, 8.12)	7.52 (6.73, 8.76)	0.11	**0.03**	4.29 (2.32, 6.29)	2.78 (1.54, 3.66)	7.20 (4.33, 10.30)	5.83 (5.16, 6.80)	0.12	**0.03**	2.78 (1.93, 3.54)	2.48 (1.37, 3.28)	2.99 (2.20, 3.76)	2.82 (2.37, 3.47)	0.55	0.55
Masimo Rad-G with Adult Probe, Product Code 4050	55	4.51 (3.76, 5.23)	3.96 (3.30, 4.60)	3.44 (2.24, 4.55)	6.34 (5.27, 7.22)	6.13 (4.77, 7.25)	**0.006**	0.06	4.69 (3.72, 5.62)	3.89 (2.72, 4.98)	7.72 (6.20, 9.07)	7.43 (5.84, 8.87)	**0.01**	**0.04**	4.40 (3.59, 5.29)	3.87 (2.21, 5.15)	6.09 (5.03, 6.94)	5.66 (4.39, 6.98)	**0.03**	0.19	2.80 (2.15, 3.37)	2.51 (1.32, 3.43)	3.93 (3.27, 4.62)	3.59 (3.08, 4.14)	**0.02**	0.06

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Acare (2023) AH-M1	57	5.34 (4.47, 6.11)	5.53 (4.26, 6.63)	5.88 (4.47, 7.27)	3.63 (2.63, 4.45)	3.56 (2.13, 4.45)	0.07	0.12	8.78 (6.60, 10.58)	9.20 (6.20, 11.73)	5.29 (3.93, 6.30)	5.45 (4.39, 6.70)	0.06	0.09	3.47 (2.43, 4.55)	3.51 (2.42, 4.57)	2.62 (1.59, 3.53)	2.02 (0.72, 2.98)	0.28	0.16	1.91 (1.32, 2.47)	2.16 (1.12, 3.04)	2.14 (1.10, 2.93)	2.33 (0.78, 3.80)	0.88	0.9
Guangdong Biolight M70	37	5.68 (4.65, 6.57)	4.47 (3.43, 5.42)	5.70 (3.86, 7.14)	6.94 (5.14, 8.46)	7.04 (4.16, 9.25)	0.06	0.22	7.83 (5.28, 10.07)	9.20 (6.25, 11.47)	10.66 (7.71, 13.24)	10.46 (6.91, 13.83)	0.17	0.26	2.94 (1.65, 4.26)	3.11 (2.18, 4.06)	5.13 (4.00, 6.58)	5.76 (3.78, 8.19)	**0.02**	0.16	2.97 (2.29, 3.59)	1.87 (1.48, 2.17)	3.31 (2.32, 4.25)	3.76 (2.22, 5.00)	0.68	0.44
Acare (2020) AH-MX	42	6.87 (5.88, 7.9)	7.07 (5.34, 8.82)	5.56 (4.37, 6.74)	7.44 (5.79, 8.81)	6.55 (3.94, 8.45)	0.55	0.94	11.79 (9.04, 14.49)	10.32 (8.36, 12.81)	11.57 (9.15, 13.54)	10.51 (5.86, 14.34)	0.86	0.77	3.66 (2.73, 4.53)	3.53 (2.33, 4.48)	4.51 (2.96, 5.60)	4.35 (0.90, 5.35)	0.44	0.76	1.83 (1.26, 2.33)	1.32 (0.93, 1.67)	1.16 (0.61, 1.73)	0.60 (0.35, 0.71)	0.13	**0.004**
American Oxygen LK-87	20	8.04 (7.48, 8.62)	7.93 (6.63, 9.13)	8.51 (7.56, 9.42)	7.41 (6.84, 7.99)	7.52 (6.77, 8.41)	0.71	0.82	9.17 (7.60, 10.53)	10.21 (9.32, 11.02)	9.69 (8.10, 10.91)	8.88 (7.10, 10.62)	0.62	0.87	9.61 (6.94, 12.01)	10.36 (8.53, 11.84)	8.15 (6.49, 9.23)	8.94 (7.75, 9.57)	0.59	0.97	4.26 (2.60, 5.85)	3.15 (1.56, 4.71)	3.20 (1.44, 4.85)	3.67 (1.39, 5.67)	0.41	0.7

Accuracy root mean square error (A_{RMS}) is reported for each device in a cohort of 24 participants as the overall point estimate with 95 % confidence interval (CI). Columns include the mean A_{RMS} and 95 % CI within each functional arterial oxygen saturation (SaO_2) decile and within four skin color bins. Skin color bins are defined by Individual Typology Angle (ITA) measured at the forehead: ITA > 30° (light), -30° ≤ ITA ≤ 30° (medium), ITA < -30° (dark), and ITA <-50° (very dark); note that “very dark” is a subset of the dark bin. P-values test differences in A_{RMS} between light versus dark and light versus very dark bins using a Welch’s two-sample t-test. No correction for multiple pairwise comparisons (e.g., Bonferroni) was applied, as the 2025 draft U.S. Food and Drug Administration (FDA) guidelines do not require adjustment. P-values near 0.05 are reported to three decimal places, and those meeting $p < 0.05$ are bolded. The 95% CI were determined using bootstrapping (random resampling with replacement) with 1,000 repetitions.

Table S7. Bias (mean of SpO_2 minus SaO_2) at the forehead for 24 participant cohorts

Device	Device ID	Overall Bias (95% CI)	70–100% SaO_2 Bias (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Nonin CO-Pilot	70	-0.28 (-0.34, -0.2)	-0.10 (-0.29, 0.09)	-0.47 (-0.69, -0.22)	-0.09 (-0.38, 0.19)	0.07 (-0.39, 0.55)	0.96	0.79	-0.59 (-1.02, -0.13)	0.02 (-0.43, 0.52)	-0.04 (-0.67, 0.60)	0.23 (-0.73, 1.09)	0.71	0.62	0.07 (-0.25, 0.40)	-0.63 (-1.06, -0.21)	-0.30 (-0.89, 0.37)	-0.19 (-1.25, 0.79)	0.76	1	0.10 (-0.16, 0.35)	-0.77 (-1.05, -0.51)	0.01 (-0.32, 0.33)	0.12 (-0.34, 0.60)	0.86	0.95
Shenzhen (2023) PC-60NW	21	-0.04 (-0.15, 0.08)	-0.14 (-0.38, 0.13)	0.09 (-0.14, 0.34)	0.10 (-0.24, 0.38)	-0.10 (-0.56, 0.34)	0.81	0.91	-1.21 (-1.72, -0.72)	-0.74 (-1.16, -0.36)	-1.43 (-2.21, -0.71)	-2.00 (-3.00, -0.97)	0.44	0.22	-0.18 (-0.61, 0.22)	-0.04 (-0.47, 0.39)	0.59 (0.20, 0.97)	0.41 (-0.13, 0.91)	0.3	0.56	0.68 (0.30, 1.04)	0.89 (0.55, 1.22)	0.86 (0.61, 1.12)	0.71 (0.43, 0.99)	0.71	0.97
Masimo Rad-97 with RD Rainbow Adt 8λ Adult Adhesive Sensor	73.09	0.86 (0.74, 0.97)	0.38 (0.23, 0.54)	1.58 (1.36, 1.80)	1.72 (1.43, 2.01)	1.11 (0.68, 1.56)	**0.009**	0.2	-0.24 (-0.61, 0.11)	1.45 (0.96, 1.93)	1.55 (0.96, 2.15)	0.52 (-0.35, 1.45)	**0.01**	0.17	0.67 (0.51, 0.85)	1.99 (1.67, 2.31)	2.15 (1.67, 2.62)	1.65 (0.91, 2.44)	**0.004**	0.15	0.69 (0.49, 0.87)	1.26 (0.91, 1.63)	1.41 (1.06, 1.75)	0.90 (0.39, 1.38)	0.11	0.7
Shenzhen (2024) PC-60NW	81	-0.02 (-0.17, 0.12)	0.15 (-0.08, 0.36)	0.64 (0.44, 0.84)	-0.18 (-0.59, 0.21)	-0.05 (-0.53, 0.39)	0.44	0.64	-1.03 (-1.46, -0.64)	0.11 (-0.28, 0.53)	-2.36 (-3.35, -1.38)	-2.28 (-3.05, -1.57)	0.14	0.12	0.27 (-0.01, 0.56)	1.09 (0.70, 1.54)	0.18 (-0.36, 0.64)	0.39 (-0.08, 0.82)	0.78	0.86	1.16 (0.94, 1.37)	0.73 (0.54, 0.93)	1.33 (1.03, 1.59)	1.33 (0.93, 1.76)	0.72	0.8
Zug Medical Systems SAS SPLF-SE with SpotLife Reusable Probe	75	1.08 (0.97, 1.19)	0.33 (0.09, 0.58)	1.45 (1.26, 1.66)	1.50 (1.25, 1.75)	1.46 (1.26, 1.65)	**0.003**	**0.005**	-1.07 (-1.44, -0.61)	1.17 (0.76, 1.58)	0.59 (-0.11, 1.23)	1.48 (1.01, 1.94)	0.13	**0.01**	0.59 (0.25, 0.96)	1.75 (1.46, 2.05)	1.95 (1.64, 2.27)	1.66 (1.35, 1.97)	**0.01**	**0.046**	1.53 (1.30, 1.75)	1.44 (1.13, 1.75)	1.68 (1.40, 1.98)	1.29 (1.02, 1.55)	0.6	0.61

Device	Device ID	Overall Bias (95% CI)	70–100% SaO_2 Bias (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Beijing Choice MD300M	40	0.92 (0.79, 1.05)	0.72 (0.53, 0.92)	0.49 (0.21, 0.76)	1.83 (1.51, 2.17)	1.92 (1.48, 2.36)	0.052	0.1	1.15 (0.80, 1.52)	1.73 (1.12, 2.32)	2.37 (1.72, 2.97)	2.59 (1.77, 3.35)	0.22	0.22	0.89 (0.60, 1.15)	0.23 (-0.13, 0.61)	2.00 (1.46, 2.52)	1.86 (1.14, 2.53)	0.13	0.34	0.21 (-0.11, 0.51)	-0.16 (-0.55, 0.24)	1.00 (0.47, 1.47)	1.04 (0.39, 1.69)	0.25	0.3
Compass Health POX-ROS	12	-0.30 (-0.44, -0.17)	0.56 (0.21, 0.94)	-1.18 (-1.41, -0.95)	-0.63 (-0.94, -0.35)	-0.74 (-1.06, -0.43)	0.1	0.09	0.37 (-0.50, 1.28)	-1.34 (-1.85, -0.90)	-1.24 (-1.75, -0.76)	-1.35 (-1.95, -0.79)	0.07	0.07	0.12 (-0.49, 0.78)	-1.64 (-2.06, -1.26)	-0.51 (-0.93, -0.09)	-0.57 (-1.07, -0.05)	0.29	0.3	1.12 (0.68, 1.50)	-0.63 (-0.95, -0.29)	-0.16 (-0.69, 0.33)	-0.34 (-0.88, 0.24)	0.15	0.08
Shenzhen AM801	10	0.38 (0.21, 0.55)	0.18 (-0.23, 0.59)	0.03 (-0.22, 0.30)	1.03 (0.70, 1.42)	1.12 (0.62, 1.69)	0.33	0.45	-0.54 (-1.05, -0.02)	-0.59 (-1.22, 0.04)	1.66 (0.85, 2.69)	2.15 (0.84, 3.52)	**0.03**	**0.02**	0.48 (-0.40, 1.31)	0.15 (-0.26, 0.56)	1.14 (0.74, 1.52)	0.86 (0.33, 1.33)	0.63	0.89	0.61 (-0.24, 1.35)	0.46 (0.19, 0.72)	0.34 (0.08, 0.64)	0.30 (0.05, 0.54)	0.73	0.79
CMS Zacurate 500C	15	0.78 (0.58, 0.96)	0.68 (0.14, 1.13)	0.63 (0.39, 0.87)	0.38 (-0.03, 0.77)	0.02 (-0.48, 0.44)	0.73	0.43	-0.35 (-1.55, 0.80)	-0.80 (-1.30, -0.32)	-1.56 (-2.74, -0.52)	-1.66 (-3.04, -0.36)	0.39	0.31	1.23 (0.68, 1.75)	1.67 (1.31, 2.07)	1.23 (0.81, 1.66)	0.81 (0.51, 1.12)	0.87	0.75	1.24 (0.92, 1.57)	1.00 (0.74, 1.25)	0.90 (0.50, 1.33)	0.48 (0.01, 1.03)	0.65	0.28
Acare (2023) Model AH-TA	56	1.54 (1.42, 1.64)	1.42 (1.18, 1.67)	1.85 (1.53, 2.15)	0.66 (0.25, 1.04)	-0.04 (-0.71, 0.56)	0.23	**0.049**	0.64 (0.09, 1.12)	1.30 (0.67, 1.87)	0.75 (0.06, 1.35)	0.34 (-0.75, 1.37)	0.69	0.88	2.08 (1.76, 2.39)	2.83 (2.36, 3.34)	0.89 (0.07, 1.60)	0.22 (-1.13, 1.28)	0.1	**0.03**	1.68 (1.44, 1.92)	1.65 (1.30, 2.01)	0.37 (-0.27, 0.91)	-0.69 (-1.67, 0.16)	**0.03**	**0.003**
Choicemed MD300CN350R	14	1.02 (0.89, 1.15)	1.58 (1.33, 1.83)	0.78 (0.38, 1.13)	0.27 (-0.13, 0.62)	0.31 (-0.12, 0.68)	**0.01**	**0.03**	1.29 (0.80, 1.73)	0.46 (-0.33, 1.31)	-1.33 (-2.08, -0.61)	-1.36 (-2.24, -0.59)	**0.001**	**0.003**	2.19 (1.83, 2.54)	1.24 (0.86, 1.60)	1.06 (0.62, 1.47)	1.02 (0.46, 1.51)	0.11	0.13	1.21 (0.77, 1.62)	0.68 (0.16, 1.14)	0.87 (0.38, 1.38)	1.10 (0.54, 1.61)	0.55	0.8
Nonin Onyx Vantage 9590	35	1.36 (1.24, 1.47)	1.39 (1.17, 1.59)	2.02 (1.78, 2.24)	1.59 (1.17, 2.04)	2.07 (1.37, 2.85)	0.82	0.43	1.80 (1.35, 2.25)	2.77 (2.48, 3.08)	3.34 (2.26, 4.36)	4.90 (3.39, 6.40)	0.57	0.25	1.01 (0.77, 1.26)	1.68 (1.23, 2.13)	1.02 (0.59, 1.48)	1.16 (0.51, 1.78)	0.97	0.72	1.31 (0.96, 1.66)	1.46 (1.07, 1.90)	0.78 (0.44, 1.13)	0.39 (-0.21, 1.17)	0.32	0.3

Device	Device ID	Overall Bias (95% CI)	70–100% SaO_2 Bias (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Nonin PalmSAT 2500	36	-1.35 (-1.47, -1.24)	-1.08 (-1.29, -0.88)	-2.21 (-2.54, -1.90)	-1.67 (-2.17, -1.20)	-1.37 (-1.87, -0.87)	0.45	0.63	-1.80 (-2.12, -1.50)	-3.00 (-3.80, -2.24)	-3.13 (-3.94, -2.31)	-2.14 (-3.19, -1.18)	0.2	0.59	-1.37 (-1.71, -1.02)	-2.27 (-2.69, -1.91)	-0.98 (-1.74, -0.20)	-1.34 (-2.07, -0.51)	0.95	0.82	-0.29 (-0.61, 0.03)	-1.52 (-1.82, -1.22)	-0.53 (-1.06, -0.02)	-0.51 (-1.09, 0.02)	0.73	0.8
Shenzhen WIT-S300C	62	1.86 (1.73, 2.01)	2.16 (1.87, 2.44)	0.67 (0.31, 0.98)	2.13 (1.94, 2.33)	1.82 (1.62, 2.03)	0.95	0.59	1.51 (0.95, 2.10)	0.58 (-0.02, 1.19)	2.77 (2.43, 3.11)	2.52 (2.09, 2.93)	**0.04**	0.19	2.33 (1.93, 2.75)	0.90 (0.59, 1.21)	2.30 (2.02, 2.62)	1.82 (1.52, 2.10)	0.94	0.51	2.50 (2.09, 2.91)	0.55 (-0.39, 1.09)	1.50 (1.21, 1.83)	1.23 (0.95, 1.50)	**0.04**	**0.01**
Masimo (2024) Mightysat	78	1.71 (1.58, 1.84)	0.52 (0.31, 0.71)	1.55 (1.34, 1.73)	3.56 (3.01, 4.07)	3.90 (3.24, 4.61)	**0.003**	**0.004**	0.01 (-0.44, 0.44)	1.63 (1.20, 2.11)	4.81 (3.73, 5.95)	5.25 (3.75, 6.75)	**0.003**	**0.006**	0.58 (0.28, 0.87)	1.81 (1.53, 2.10)	3.95 (3.10, 4.89)	4.47 (3.28, 5.61)	**0.002**	**0.001**	0.90 (0.63, 1.17)	1.22 (1.02, 1.42)	2.22 (1.90, 2.60)	2.38 (1.97, 2.82)	**0.007**	**0.004**
Philips M1020B IntelliVue with M1191BL Probe	71	1.90 (1.81, 1.99)	1.48 (1.27, 1.69)	1.99 (1.85, 2.15)	3.18 (2.73, 3.65)	4.23 (3.62, 4.82)	**0.04**	**<0.001**	2.62 (2.36, 2.93)	2.56 (2.28, 2.84)	4.73 (3.74, 5.68)	7.12 (6.47, 7.65)	0.09	**<0.001**	1.10 (0.87, 1.34)	1.72 (1.57, 1.86)	3.01 (2.44, 3.56)	3.88 (3.27, 4.40)	**0.01**	**<0.001**	0.78 (0.39, 1.13)	1.61 (1.37, 1.90)	1.88 (1.28, 2.50)	2.78 (2.08, 3.47)	0.14	**0.01**
Nihon Kohden OLV-4202 with NK-SPO2-ADVT Probe	64.03	0.36 (0.18, 0.53)	-1.30 (-1.65, -0.96)	0.03 (-0.28, 0.36)	0.46 (0.06, 0.91)	0.13 (-0.47, 0.78)	0.08	0.3	-2.46 (-3.16, -1.73)	0.74 (0.11, 1.41)	0.67 (-0.27, 1.65)	0.88 (-0.22, 2.05)	**0.03**	0.06	-1.46 (-1.99, -0.96)	-0.34 (-0.85, 0.15)	0.17 (-0.57, 0.88)	-0.47 (-1.50, 0.66)	0.11	0.4	-0.18 (-0.66, 0.27)	-0.38 (-0.67, -0.13)	0.57 (0.04, 1.09)	-0.24 (-1.05, 0.59)	0.59	0.78
Covidien PM1000N	59	1.29 (1.22, 1.37)	1.39 (1.12, 1.64)	1.13 (0.89, 1.37)	3.27 (2.85, 3.70)	4.46 (3.82, 5.04)	**0.02**	**0.001**	0.66 (0.12, 1.24)	1.36 (0.91, 1.81)	4.50 (3.81, 5.18)	5.86 (5.15, 6.49)	**0.02**	**0.001**	2.20 (1.82, 2.60)	1.46 (1.03, 1.89)	3.73 (3.02, 4.36)	4.85 (4.22, 5.47)	**0.04**	**0.002**	1.41 (1.17, 1.66)	0.68 (0.38, 1.01)	1.70 (1.20, 2.18)	2.11 (1.18, 2.94)	0.49	0.36

Device	Device ID	Overall Bias (95% CI)	70–100% SaO_2 Bias (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Shenzhen A310L	11	1.43 (1.25, 1.61)	0.51 (0.04, 0.95)	1.48 (1.09, 1.81)	2.08 (1.78, 2.40)	1.77 (1.40, 2.16)	0.06	0.17	-1.52 (-2.56, -0.60)	0.63 (-0.26, 1.43)	2.53 (1.80, 3.19)	2.30 (1.45, 3.06)	**0.01**	**0.04**	1.95 (1.58, 2.35)	2.33 (1.86, 2.79)	2.41 (1.97, 2.88)	2.12 (1.66, 2.59)	0.52	0.89	1.34 (1.06, 1.62)	1.50 (1.28, 1.72)	1.34 (0.93, 1.71)	1.00 (0.56, 1.46)	0.98	0.52
Masimo Rad-G with Rad-G Finger Clip Sensor	55.03	1.94 (1.77, 2.1)	1.13 (0.83, 1.43)	2.47 (2.25, 2.71)	2.14 (1.66, 2.69)	2.76 (1.89, 3.59)	0.31	0.22	-0.10 (-0.57, 0.47)	2.75 (2.26, 3.30)	1.57 (0.24, 2.98)	2.92 (0.20, 5.56)	0.64	0.67	1.79 (1.39, 2.19)	2.79 (2.46, 3.13)	2.85 (2.14, 3.54)	3.10 (2.24, 4.01)	0.31	0.23	1.96 (1.68, 2.24)	1.87 (1.57, 2.16)	1.84 (1.41, 2.30)	2.17 (1.57, 2.81)	0.65	0.45
CONTEC CMS50M	17	-1.24 (-1.44, -1.04)	-1.61 (-1.91, -1.29)	-1.22 (-1.72, -0.73)	-0.79 (-1.10, -0.47)	-0.92 (-1.36, -0.52)	0.25	0.44	-3.37 (-4.06, -2.70)	-2.61 (-3.51, -1.71)	-2.27 (-3.03, -1.56)	-2.27 (-3.08, -1.43)	0.59	0.43	-1.66 (-2.05, -1.29)	-0.76 (-1.37, -0.14)	-0.94 (-1.57, -0.36)	-1.26 (-1.95, -0.66)	0.46	0.84	-0.45 (-0.91, -0.03)	-0.26 (-1.09, 0.41)	0.32 (0.01, 0.64)	0.37 (-0.07, 0.82)	0.21	0.28
Covidien (2022) Nellcor PM10N	41	1.17 (0.99, 1.34)	1.69 (1.32, 2.06)	0.85 (0.59, 1.13)	2.61 (2.24, 2.96)	3.17 (2.68, 3.65)	0.33	0.28	0.82 (-0.12, 1.75)	0.56 (-0.17, 1.22)	3.37 (2.71, 3.91)	4.33 (3.68, 5.01)	0.09	0.1	2.32 (1.85, 2.77)	1.23 (0.90, 1.60)	2.94 (2.51, 3.36)	3.33 (2.80, 3.86)	0.43	0.38	1.91 (1.59, 2.23)	0.67 (0.30, 1.09)	1.23 (0.68, 1.76)	1.08 (0.56, 1.67)	0.39	0.46
Bistos Co., Ltd. BT-710	61	2.46 (2.33, 2.58)	2.59 (2.41, 2.77)	2.52 (2.21, 2.82)	3.14 (2.81, 3.48)	3.51 (3.02, 3.98)	0.18	**0.046**	3.21 (2.88, 3.55)	3.22 (2.90, 3.56)	4.52 (4.07, 5.02)	5.39 (5.05, 5.76)	**0.03**	**0.003**	2.88 (2.62, 3.15)	3.11 (2.54, 3.72)	3.58 (3.19, 3.96)	3.74 (3.23, 4.20)	0.31	0.23	1.89 (1.65, 2.15)	1.40 (1.01, 1.82)	1.49 (1.11, 1.84)	1.46 (0.97, 1.92)	0.23	0.28
Utech UT100	39	1.02 (0.85, 1.19)	1.78 (1.31, 2.24)	1.12 (0.84, 1.39)	2.48 (1.66, 3.19)	2.39 (1.14, 3.42)	0.55	0.64	1.25 (0.12, 2.45)	1.08 (0.56, 1.60)	4.48 (3.15, 5.85)	4.42 (3.76, 5.06)	0.18	0.15	2.38 (1.77, 3.07)	1.57 (1.03, 2.19)	1.96 (0.38, 3.20)	2.42 (-1.02, 4.84)	0.72	0.86	1.75 (1.37, 2.14)	0.81 (0.47, 1.13)	1.09 (0.47, 1.70)	0.56 (-0.48, 1.52)	0.22	0.14
BodyMed BDMOXMTR BLK	2	2.12 (1.92, 2.32)	2.62 (2.21, 3.04)	1.91 (1.58, 2.21)	2.21 (1.79, 2.63)	2.25 (1.64, 2.83)	0.56	0.72	2.26 (1.04, 3.66)	0.86 (0.15, 1.53)	1.87 (1.04, 2.68)	1.05 (0.00, 2.04)	0.85	0.66	3.63 (3.17, 4.13)	2.76 (2.26, 3.22)	3.09 (2.26, 3.87)	3.53 (2.36, 4.66)	0.39	0.8	2.01 (1.55, 2.43)	2.29 (2.00, 2.56)	1.75 (1.30, 2.23)	2.08 (1.47, 2.68)	0.7	0.91

Device	Device ID	Overall Bias (95% CI)	70–100% SaO_2 Bias (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Masimo (2020) Mightysat Rx	7	0.92 (0.77, 1.08)	0.67 (0.34, 1.02)	0.84 (0.39, 1.27)	2.42 (1.93, 2.93)	1.81 (1.49, 2.10)	0.09	0.13	-0.01 (-0.80, 0.79)	-0.47 (-1.20, 0.22)	2.59 (1.15, 4.40)	1.04 (0.29, 1.78)	0.27	0.44	0.91 (0.42, 1.46)	1.90 (1.18, 2.71)	2.99 (2.45, 3.57)	2.66 (2.38, 2.94)	0.17	0.29	0.96 (0.46, 1.42)	1.17 (0.42, 1.71)	1.74 (1.39, 2.07)	1.57 (1.18, 1.95)	0.32	0.5
EDAN (2020) H100b	44	2.34 (2.19, 2.51)	2.18 (1.83, 2.49)	2.84 (2.57, 3.11)	4.06 (3.50, 4.69)	3.10 (2.56, 3.60)	0.054	0.31	1.81 (1.12, 2.59)	3.45 (2.89, 3.97)	5.07 (3.85, 6.33)	3.05 (2.04, 4.00)	0.1	0.5	2.77 (2.31, 3.24)	3.15 (2.73, 3.61)	4.05 (3.46, 4.61)	3.97 (3.31, 4.69)	0.06	0.18	2.05 (1.78, 2.36)	1.92 (1.60, 2.24)	2.75 (2.29, 3.24)	2.29 (1.77, 2.91)	0.21	0.78
Guangdong Biolight M800	43	0.24 (0.04, 0.45)	-0.56 (-1.17, -0.02)	-0.88 (-1.40, -0.37)	1.40 (0.76, 2.02)	2.76 (2.24, 3.32)	0.051	**0.005**	-4.33 (-5.34, -3.33)	-4.64 (-5.45, -3.90)	-0.85 (-2.29, 0.46)	1.93 (0.93, 2.92)	0.09	**0.004**	0.61 (0.02, 1.18)	0.29 (-0.29, 0.84)	2.59 (1.86, 3.33)	3.71 (2.87, 4.65)	**0.03**	**0.007**	2.05 (1.74, 2.38)	1.27 (0.92, 1.66)	2.23 (1.68, 2.78)	2.64 (1.97, 3.38)	0.74	0.28
Masimo Rad-97 with Adult Probe, Product Code 4050	60	2.49 (2.39, 2.59)	2.66 (2.27, 3.07)	2.05 (1.74, 2.36)	4.76 (4.25, 5.26)	5.27 (4.57, 5.92)	0.08	0.13	2.63 (1.90, 3.42)	2.26 (1.77, 2.75)	6.36 (5.56, 7.17)	7.57 (6.90, 8.20)	0.15	0.16	3.73 (2.95, 4.62)	2.35 (1.89, 2.84)	5.43 (4.61, 6.35)	5.69 (4.96, 6.36)	0.06	0.15	1.97 (1.64, 2.34)	1.46 (0.92, 2.05)	2.39 (1.91, 2.88)	2.36 (1.74, 2.92)	0.66	0.77
Masimo Rad-G with Adult Probe, Product Code 4050	55	3.48 (3.31, 3.63)	3.48 (3.25, 3.71)	2.73 (2.40, 3.09)	5.86 (5.39, 6.30)	5.65 (5.16, 6.15)	**<0.001**	**0.002**	4.27 (3.89, 4.69)	3.40 (2.94, 3.93)	7.51 (6.88, 8.19)	7.39 (6.73, 8.04)	**<0.001**	**0.003**	4.00 (3.61, 4.43)	3.24 (2.52, 3.95)	5.89 (5.43, 6.35)	5.54 (5.07, 6.07)	**0.008**	0.08	2.30 (2.03, 2.59)	1.70 (1.17, 2.18)	3.69 (3.09, 4.25)	3.34 (2.74, 3.91)	**0.048**	0.2
Acare (2023) AH-M1	57	-2.36 (-2.66, -2.09)	-2.70 (-3.42, -2.06)	-2.69 (-3.52, -1.95)	-1.03 (-1.70, -0.42)	-0.86 (-1.74, -0.03)	0.12	0.23	-7.48 (-8.61, -6.35)	-8.05 (-9.23, -6.94)	-4.67 (-5.79, -3.58)	-4.90 (-6.42, -3.44)	0.06	0.22	-1.53 (-2.33, -0.78)	-0.08 (-1.09, 0.92)	-0.21 (-1.08, 0.54)	0.58 (-0.22, 1.46)	0.39	0.26	0.54 (0.11, 0.98)	0.68 (0.28, 1.10)	1.30 (0.87, 1.71)	0.88 (0.09, 1.79)	0.31	0.58

Device	Device ID	Overall Bias (95% CI)	70–100% SaO_2 Bias (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Guangdong Biolight M70	37	-0.48 (-0.81, -0.14)	-0.44 (-1.09, 0.24)	-1.48 (-2.28, -0.75)	1.78 (0.51, 2.96)	4.98 (3.67, 6.31)	0.31	**0.002**	-5.59 (-6.88, -4.41)	-5.97 (-7.78, -4.25)	0.81 (-3.02, 4.12)	9.06 (6.11, 11.66)	0.11	**<0.001**	0.43 (-0.36, 1.23)	0.18 (-0.51, 0.90)	2.21 (0.45, 3.78)	4.69 (3.00, 6.35)	0.43	**0.03**	2.54 (2.16, 2.92)	1.40 (1.14, 1.67)	2.20 (1.46, 2.92)	2.15 (0.88, 3.36)	0.65	0.72
Acare (2020) AH-MX	42	-3.98 (-4.31, -3.65)	-3.69 (-4.48, -3.02)	-3.35 (-4.10, -2.66)	-4.91 (-6.07, -3.83)	-4.27 (-5.67, -2.94)	0.35	0.75	-10.46 (-11.88, -9.08)	-9.25 (-10.39, -8.15)	-10.43 (-11.90, -8.91)	-8.67 (-11.04, -6.41)	0.91	0.76	-2.67 (-3.24, -2.09)	-2.54 (-3.21, -1.91)	-3.30 (-4.31, -2.33)	-3.59 (-4.74, -2.39)	0.46	0.54	0.50 (0.18, 0.86)	0.08 (-0.29, 0.45)	0.03 (-0.33, 0.34)	0.01 (-0.27, 0.27)	0.29	0.47
American Oxygen LK-87	20	-5.71 (-6.11, -5.31)	-4.99 (-5.92, -4.01)	-6.81 (-7.48, -6.16)	-5.47 (-6.32, -4.61)	-5.54 (-6.50, -4.46)	0.68	0.72	-4.96 (-7.24, -2.82)	-9.34 (-10.18, -8.46)	-9.54 (-10.36, -8.77)	-8.98 (-9.87, -8.13)	0.07	0.22	-8.62 (-10.00, -7.29)	-9.24 (-10.25, -8.26)	-6.61 (-8.13, -4.97)	-7.09 (-9.11, -5.14)	0.34	0.53	-1.86 (-2.77, -1.04)	-1.95 (-2.52, -1.39)	-0.98 (-1.88, -0.29)	-0.96 (-2.30, 0.04)	0.25	0.32

This table reports the overall bias (mean of pulse oximeter measured oxygen saturation (SpO_2) - functional arterial oxygen saturation(SaO_2)) for each device in a cohort of 24 participants. The bias is presented with 95% confidence interval (CI) overall, and separately for each SaO_2 decile and for each skin color bin. Skin color bins were defined using the Individual Typology Angle (ITA) measured at the forehead: ITA > 30° for light, -30° ≤ ITA ≤ 30° for medium, ITA < -30° for dark, and ITA < -50° for dark; note that "very dark" is a subset of the dark bin. P-values indicate whether the SpO_2 bias differs significantly between light versus dark bins and between light versus very dark bins using a linear mixed-effects model with SpO_2-SaO_2 error as the outcome, skin color bin as a fixed-effect predictor, and participant ID as a random-effect grouping factor. No correction for multiple pairwise comparisons (e.g., Bonferroni) was applied, as the 2025 draft U.S. Food and Drug Administration (FDA) guidelines do not require adjustment. P-values near the 0.05 threshold are reported to three decimal places, and those meeting $p < 0.05$ are bolded. The 95% CI were determined using bootstrapping (random resampling with replacement) with 1,000 repetitions.

Table S8. A_{RMS} at the forehead for maximum participant cohorts

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Nonin CO-Pilot	70	1.39 (1.19, 1.60)	1.04 (0.94, 1.14)	1.58 (1.21, 1.96)	1.43 (1.09, 1.76)	1.53 (1.11, 1.91)	0.09	0.09	1.24 (1.05, 1.45)	1.82 (1.20, 2.54)	1.74 (1.25, 2.17)	1.95 (1.29, 2.50)	0.24	0.15	1.09 (0.94, 1.22)	1.68 (1.34, 2.00)	1.51 (1.06, 1.94)	1.63 (1.06, 2.21)	0.2	0.18	0.90 (0.74, 1.05)	1.31 (0.92, 1.70)	1.17 (0.90, 1.41)	1.17 (0.88, 1.46)	0.17	0.19
Shenzhen (2023) PC-60NW	21	1.61 (1.44, 1.80)	1.64 (1.36, 1.89)	1.66 (1.33, 1.96)	1.57 (1.26, 1.96)	1.75 (1.29, 2.26)	0.72	0.81	1.83 (1.45, 2.16)	1.90 (1.29, 2.52)	2.39 (1.48, 3.29)	2.83 (1.55, 3.97)	0.53	0.28	1.49 (1.23, 1.74)	1.68 (1.26, 2.01)	1.25 (1.02, 1.48)	1.32 (1.03, 1.62)	0.27	0.54	1.63 (1.22, 2.00)	1.53 (1.21, 1.84)	1.08 (0.86, 1.28)	0.97 (0.73, 1.19)	0.09	0.051
Masimo Rad-97 with RD Rainbow Adt 8λ Adult Adhesive Sensor	73.09	1.64 (1.03, 1.91)	1.25 (0.88, 1.62)	1.30 (0.73, 1.90)	2.46 (1.48, 3.20)	2.22 (0.84, 3.64)	0.11	0.55	1.57 (1.07, 2.02)	1.58 (0.94, 2.20)	2.56 (1.40, 3.47)	2.32 (0.78, 3.79)	0.28	0.68	1.01 (0.66, 1.29)	1.27 (0.68, 1.89)	2.77 (1.62, 3.63)	2.53 (0.73, 4.23)	0.06	0.46	1.20 (0.78, 1.52)	0.99 (0.47, 1.53)	1.88 (1.26, 2.31)	1.57 (0.63, 2.42)	0.12	0.64
Shenzhen (2024) PC-60NW	81	1.87 (1.45, 2.28)	1.69 (1.21, 2.13)	1.51 (1.01, 1.97)	2.59 (1.66, 3.59)	2.11 (1.62, 2.51)	0.17	0.22	2.14 (1.33, 2.80)	1.61 (1.03, 2.18)	4.03 (2.12, 5.86)	3.12 (1.89, 3.94)	0.14	0.15	1.23 (0.92, 1.55)	1.65 (0.94, 2.38)	1.71 (1.14, 2.40)	1.21 (0.97, 1.50)	0.27	0.79	1.54 (1.08, 1.97)	1.15 (0.89, 1.40)	1.60 (1.28, 1.91)	1.63 (1.11, 2.12)	0.59	0.66
Zug Medical Systems SAS SPLF-SE with SpotLife Reusable Probe	75	1.94 (1.66, 2.22)	1.66 (1.37, 1.94)	2.00 (1.39, 2.57)	2.13 (1.84, 2.42)	1.89 (1.71, 2.10)	**0.03**	0.12	1.93 (1.68, 2.13)	2.37 (1.09, 3.20)	1.95 (1.38, 2.48)	1.88 (1.35, 2.28)	0.78	0.79	1.54 (1.05, 1.98)	1.89 (1.42, 2.44)	2.26 (1.77, 2.80)	2.08 (1.77, 2.50)	**0.02**	**0.03**	1.56 (1.21, 1.86)	1.70 (1.26, 2.15)	2.16 (1.72, 2.60)	1.69 (1.39, 1.99)	0.053	0.42

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Compass Health POX-ROS	12	2.01 (1.73, 2.28)	2.27 (1.42, 3.01)	1.89 (1.50, 2.29)	1.84 (1.58, 2.07)	1.89 (1.61, 2.14)	0.59	0.68	2.77 (1.43, 3.77)	1.61 (1.35, 1.84)	2.04 (1.27, 2.79)	2.23 (1.34, 3.05)	0.43	0.61	2.63 (1.26, 4.06)	2.10 (1.56, 2.65)	1.71 (1.46, 1.92)	1.75 (1.49, 1.96)	0.56	0.61	1.66 (1.08, 2.12)	1.72 (1.40, 2.01)	1.82 (1.33, 2.26)	1.74 (1.14, 2.29)	0.63	0.83
Beijing Choice MD300M	40	2.05 (1.75, 2.37)	1.82 (1.47, 2.19)	1.71 (1.36, 2.07)	2.70 (2.02, 3.40)	2.97 (2.04, 3.79)	0.054	0.08	2.03 (1.36, 2.79)	2.48 (1.73, 3.14)	3.83 (2.26, 5.46)	4.43 (2.43, 6.38)	0.09	0.1	1.77 (1.41, 2.10)	1.22 (0.95, 1.51)	2.46 (2.05, 2.93)	2.56 (2.09, 3.09)	**0.03**	**0.03**	1.69 (1.32, 2.04)	1.11 (0.86, 1.36)	1.93 (1.22, 2.51)	2.00 (0.90, 2.72)	0.68	0.78
Nonin PalmSAT 2500	36	2.20 (1.87, 2.53)	2.24 (1.72, 2.75)	1.87 (1.55, 2.21)	2.57 (1.72, 3.36)	2.14 (1.25, 2.96)	0.74	0.78	2.87 (2.08, 3.60)	2.11 (1.59, 2.68)	3.39 (2.08, 4.55)	2.89 (1.41, 4.12)	0.77	0.87	2.16 (1.63, 2.62)	2.05 (1.73, 2.34)	2.20 (1.67, 2.72)	1.95 (1.28, 2.67)	0.9	0.73	1.61 (1.24, 1.99)	1.50 (1.20, 1.75)	1.47 (1.02, 1.88)	1.19 (0.89, 1.45)	0.6	0.2
Covidien PM1000N	59	2.21 (2.06, 2.59)	1.92 (1.61, 2.20)	1.92 (1.57, 2.26)	2.94 (2.54, 3.31)	3.12 (2.53, 3.71)	**<0.001**	**0.002**	2.17 (1.78, 2.49)	2.19 (1.57, 2.84)	3.85 (3.31, 4.35)	3.98 (3.25, 4.69)	**<0.001**	**<0.001**	1.89 (1.55, 2.26)	2.14 (1.72, 2.51)	3.02 (2.54, 3.43)	3.26 (2.61, 3.82)	**<0.001**	**0.002**	1.77 (1.42, 2.10)	1.54 (1.30, 1.77)	1.83 (1.46, 2.15)	1.80 (1.32, 2.26)	0.65	0.84
Shenzhen AM801	10	2.21 (1.79, 2.61)	2.42 (1.84, 3.09)	1.99 (1.57, 2.52)	2.23 (1.48, 3.07)	2.45 (1.10, 3.69)	0.6	0.83	1.89 (1.43, 2.25)	2.88 (1.87, 3.79)	3.39 (1.74, 5.15)	4.05 (1.22, 6.44)	0.28	0.41	2.77 (2.07, 3.60)	1.68 (1.34, 1.98)	1.71 (1.45, 1.94)	1.41 (1.26, 1.51)	**0.04**	**0.01**	2.61 (1.56, 3.58)	1.31 (1.04, 1.54)	1.01 (0.67, 1.33)	0.66 (0.50, 0.78)	**0.03**	**0.01**
Nonin Onyx Vantage 9590	35	2.24 (1.84, 2.68)	2.13 (1.64, 2.62)	2.07 (1.68, 2.47)	2.54 (1.41, 3.51)	2.96 (1.33, 4.20)	0.83	0.57	2.76 (1.68, 3.77)	2.52 (2.05, 2.96)	3.20 (1.85, 4.46)	3.68 (1.70, 5.25)	0.66	0.49	1.37 (0.98, 1.66)	1.74 (1.19, 2.22)	1.95 (1.16, 2.65)	2.21 (1.15, 3.04)	0.28	0.21	2.04 (1.45, 2.58)	1.72 (1.14, 2.23)	1.81 (1.08, 2.60)	2.05 (0.96, 3.07)	0.51	0.8
Choicemed MD300CN350R	14	2.27 (2.03, 2.49)	2.22 (1.74,	2.39 (1.99,	2.05 (1.70,	2.04 (1.72,	0.7	0.75	2.28 (1.65,	3.26 (2.49,	2.50 (1.91,	2.56 (1.91,	0.56	0.51	2.47 (1.82,	2.25 (1.81,	2.14 (1.53,	1.98 (1.35,	0.43	0.31	2.12 (1.54,	1.83 (1.26,	1.57 (1.09,	1.62 (0.96,	0.18	0.25

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
			2.63)	2.75)	2.37)	2.31)			2.85)	4.06)	3.09)	3.32)			3.00)	2.64)	2.70)	2.41)			2.60)	2.47)	1.98)	2.14)		
Masimo (2024) Mightysat	78	2.39 (1.98, 2.77)	2.05 (1.44, 2.62)	2.15 (1.52, 2.64)	2.92 (2.09, 3.62)	3.18 (2.52, 3.99)	0.1	**0.04**	2.27 (1.48, 2.91)	2.47 (1.73, 3.05)	3.56 (2.33, 4.48)	3.88 (2.35, 5.03)	0.12	0.12	2.13 (1.47, 2.71)	2.31 (1.58, 2.90)	3.33 (2.28, 4.20)	3.85 (2.81, 4.81)	0.08	**0.04**	1.75 (1.27, 2.19)	1.57 (1.14, 1.96)	2.27 (1.70, 2.76)	2.50 (2.04, 3.08)	0.14	0.05
Acare (2023) Model AH-TA	56	2.40 (2.15, 2.65)	2.47 (2.05, 2.86)	2.30 (1.99, 2.60)	2.38 (1.96, 2.84)	2.27 (1.72, 2.95)	0.91	0.68	2.66 (1.99, 3.32)	2.43 (1.83, 3.01)	2.67 (2.06, 3.24)	2.39 (1.50, 3.01)	0.79	0.75	2.79 (2.23, 3.24)	2.87 (2.43, 3.33)	2.55 (2.04, 3.05)	2.49 (1.74, 3.29)	0.65	0.57	2.01 (1.57, 2.45)	1.51 (1.27, 1.75)	1.92 (1.49, 2.35)	1.96 (1.18, 2.65)	0.86	0.81
Philips M1020B IntelliVue with M1191BL Probe	71	2.45 (1.93, 3.15)	1.89 (1.65, 2.12)	2.21 (1.78, 2.65)	3.45 (2.43, 4.46)	4.40 (3.41, 5.53)	**0.03**	**0.01**	2.23 (1.92, 2.54)	2.92 (2.22, 3.58)	4.40 (3.08, 5.65)	5.58 (4.08, 6.96)	**0.02**	**0.01**	1.55 (1.27, 1.82)	1.76 (1.46, 2.07)	3.02 (2.17, 3.83)	3.82 (2.96, 4.73)	**0.02**	**0.009**	1.90 (1.39, 2.34)	1.68 (1.19, 2.24)	2.62 (1.71, 3.48)	3.38 (2.16, 4.52)	0.21	0.08
CMS Zacurate 500C	15	2.51 (2.03, 2.96)	2.65 (1.59, 3.66)	2.36 (1.45, 3.14)	2.46 (1.94, 2.97)	2.52 (1.62, 3.22)	0.97	0.96	3.41 (1.81, 4.91)	3.15 (1.90, 4.10)	4.03 (2.70, 5.05)	4.45 (2.59, 5.83)	0.52	0.42	2.11 (1.50, 2.65)	2.52 (1.34, 3.54)	1.86 (1.44, 2.22)	1.55 (1.14, 1.83)	0.57	0.25	1.62 (1.15, 2.01)	1.25 (0.92, 1.57)	1.40 (0.97, 1.78)	1.37 (0.79, 1.92)	0.38	0.37
Covidien (2022) Nellcor PM10N	41	2.55 (1.92, 3.25)	3.07 (1.89, 4.33)	1.47 (1.17, 1.77)	3.10 (2.38, 3.71)	3.21 (1.67, 4.26)	0.64	0.68	4.14 (2.03, 6.16)	1.71 (1.44, 1.99)	3.61 (2.69, 4.44)	3.96 (2.57, 4.98)	0.87	0.62	2.60 (1.54, 3.65)	1.54 (1.03, 1.99)	3.07 (2.31, 3.75)	3.14 (1.55, 4.49)	0.26	0.49	2.23 (1.57, 2.83)	1.13 (0.75, 1.52)	2.11 (1.40, 2.60)	1.56 (0.62, 2.37)	0.9	0.33
Utech UT100	39	2.72 (2.14, 3.33)	2.45 (1.60, 3.29)	1.96 (1.47, 2.40)	3.87 (2.35, 5.17)	3.67 (2.25, 4.95)	0.14	0.21	3.37 (2.24, 4.47)	2.36 (1.54, 3.18)	4.38 (2.40, 6.42)	3.38 (2.20, 4.35)	0.58	0.9	2.32 (1.18, 3.39)	2.33 (1.25, 3.23)	3.82 (2.08, 5.56)	4.52 (2.36, 6.44)	0.19	0.14	1.88 (1.39, 2.31)	1.52 (1.12, 1.90)	2.07 (1.65, 2.46)	2.06 (1.52, 2.52)	0.46	0.56

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Nihon Kohden OLV-4202 with NK-SPO2-ADVT Probe	64.03	2.76 (2.14, 4.21)	2.40 (1.59, 3.12)	1.74 (1.25, 2.17)	4.34 (2.76, 6.05)	5.08 (2.55, 7.69)	0.07	0.2	3.28 (1.57, 4.85)	2.06 (1.36, 2.69)	5.72 (3.15, 8.29)	6.89 (2.83, 10.73)	0.11	0.27	1.71 (0.98, 2.53)	1.86 (1.15, 2.42)	3.91 (2.73, 5.11)	4.33 (2.61, 6.27)	**0.007**	0.08	2.39 (1.47, 3.33)	1.17 (0.93, 1.40)	2.78 (1.79, 3.56)	2.96 (1.76, 4.11)	0.46	0.37
EDAN (2020) H100b	44	2.80 (2.45, 3.18)	3.02 (2.40, 3.69)	2.44 (1.93, 2.92)	3.08 (2.57, 3.65)	3.29 (2.37, 4.30)	0.7	0.62	3.42 (2.53, 4.26)	2.99 (2.14, 3.71)	3.31 (2.65, 4.07)	3.51 (2.36, 4.85)	0.81	0.78	3.45 (2.66, 4.17)	2.42 (1.91, 2.86)	3.60 (2.84, 4.33)	3.85 (2.80, 4.93)	0.66	0.55	2.30 (1.72, 2.83)	1.77 (1.33, 2.22)	2.25 (1.86, 2.59)	2.26 (1.61, 2.81)	0.79	0.89
Shenzhen WIT-S300C	62	2.81 (2.20, 3.38)	2.16 (1.56, 2.79)	3.18 (2.17, 4.05)	2.74 (2.29, 3.15)	2.25 (1.70, 2.72)	0.09	0.61	1.97 (1.43, 2.52)	3.68 (2.10, 5.06)	3.37 (2.67, 3.95)	2.81 (1.94, 3.56)	**0.01**	0.19	2.21 (1.53, 2.93)	3.21 (2.12, 4.16)	2.88 (2.20, 3.48)	2.20 (1.47, 2.91)	0.15	0.86	2.21 (1.44, 2.89)	2.58 (1.71, 3.50)	2.07 (1.48, 2.70)	1.56 (1.05, 2.04)	0.95	0.34
CONTEC CMS50M	17	2.82 (2.36, 3.23)	2.75 (2.22, 3.32)	3.33 (2.55, 4.03)	2.09 (1.27, 2.71)	2.22 (1.43, 3.10)	0.13	0.35	4.07 (2.87, 5.17)	4.54 (2.84, 5.97)	3.42 (1.91, 4.71)	3.04 (2.23, 3.82)	0.45	0.36	2.29 (1.76, 2.91)	2.38 (1.96, 2.73)	2.06 (0.91, 3.00)	2.28 (0.89, 3.69)	0.41	0.7	1.87 (1.38, 2.31)	3.01 (1.48, 4.70)	1.14 (0.86, 1.38)	1.28 (0.86, 1.64)	**0.049**	0.18
Shenzhen A310L	11	2.84 (2.47, 3.18)	2.52 (2.03, 3.03)	2.93 (2.48, 3.32)	2.94 (2.22, 3.71)	2.99 (1.85, 4.09)	0.45	0.64	3.69 (2.53, 4.73)	3.76 (2.81, 4.80)	3.65 (2.43, 4.87)	3.94 (2.14, 5.57)	0.93	0.91	2.12 (1.64, 2.57)	3.19 (2.71, 3.66)	2.86 (2.35, 3.34)	2.69 (2.00, 3.38)	**0.045**	0.21	1.51 (1.15, 1.87)	1.97 (1.67, 2.27)	1.87 (1.46, 2.22)	1.65 (1.00, 2.25)	0.27	0.86
Masimo (2020) Mightysat Rx	7	2.88 (2.00, 4.11)	2.11 (1.67, 2.56)	2.35 (1.69, 2.94)	3.67 (1.94, 5.57)	2.32 (1.84, 2.78)	0.28	0.45	2.56 (1.93, 3.11)	1.83 (1.39, 2.29)	5.25 (1.81, 8.54)	2.36 (1.43, 3.15)	0.45	0.77	2.27 (0.99, 3.46)	2.60 (1.67, 3.63)	3.52 (2.03, 5.20)	2.53 (1.97, 3.00)	0.15	0.17	1.83 (1.44, 2.13)	2.59 (1.62, 3.67)	2.13 (1.69, 2.56)	1.87 (1.50, 2.27)	0.29	0.67

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Masimo Rad-G with Rad-G Finger Clip Sensor	55.03	3.00 (2.43, 3.54)	2.82 (2.02, 3.48)	2.74 (1.74, 3.71)	3.78 (2.52, 4.94)	3.92 (2.15, 5.96)	0.24	0.51	3.06 (2.14, 3.84)	2.87 (1.75, 3.92)	4.10 (2.62, 5.47)	4.79 (2.62, 7.09)	0.31	0.31	2.63 (1.82, 3.33)	2.93 (1.89, 3.88)	3.91 (2.46, 5.09)	3.75 (1.79, 5.67)	0.16	0.48	2.83 (1.98, 3.52)	2.34 (1.47, 3.17)	3.27 (2.45, 3.92)	2.90 (1.70, 4.05)	0.32	0.84
Bistos Co., Ltd. BT-710	61	3.15 (2.86, 3.42)	2.84 (2.47, 3.18)	3.23 (2.66, 3.74)	3.39 (2.84, 3.84)	3.60 (2.88, 4.13)	0.17	0.15	3.36 (2.80, 3.86)	3.69 (3.04, 4.37)	4.95 (3.88, 6.17)	5.69 (3.86, 7.34)	0.08	0.12	2.83 (2.42, 3.24)	3.17 (2.56, 3.79)	3.17 (2.61, 3.72)	3.18 (2.46, 3.89)	0.39	0.48	2.34 (1.90, 2.70)	2.80 (1.68, 3.88)	1.97 (1.64, 2.23)	1.91 (1.51, 2.30)	0.27	0.27
BodyMed BDMOXMTRBLK	2	3.26 (2.92, 3.64)	3.79 (2.82, 4.86)	3.14 (2.77, 3.47)	3.00 (2.41, 3.43)	3.08 (2.20, 3.77)	0.23	0.34	4.00 (2.32, 5.69)	3.63 (2.85, 4.26)	3.50 (2.60, 4.43)	3.33 (2.07, 4.71)	0.9	0.72	4.88 (3.24, 6.64)	3.57 (2.82, 4.27)	3.55 (2.43, 4.53)	3.81 (2.16, 5.14)	0.2	0.35	2.60 (2.02, 3.07)	2.57 (2.22, 2.90)	2.13 (1.44, 2.67)	2.03 (0.96, 2.75)	0.26	0.26
Masimo Rad-97 with Adult Probe, Product Code 4050	60	3.33 (2.90, 3.77)	2.55 (2.15, 2.95)	2.95 (2.32, 3.59)	4.80 (4.20, 5.37)	5.15 (4.46, 5.80)	**<0.001**	**<0.001**	2.63 (2.14, 3.04)	3.56 (2.92, 4.16)	5.56 (4.87, 6.35)	5.82 (4.85, 6.78)	**<0.001**	**<0.001**	2.77 (2.24, 3.25)	2.95 (2.27, 3.63)	5.40 (4.21, 6.80)	5.14 (4.11, 5.90)	**<0.001**	**0.002**	2.34 (1.82, 2.85)	2.23 (1.51, 2.88)	3.16 (2.59, 3.73)	3.61 (2.78, 4.49)	**0.009**	**0.009**
Guangdong Biolight M800	43	3.53 (3.17, 3.88)	3.69 (2.89, 4.41)	3.28 (2.85, 3.69)	3.66 (2.97, 4.24)	3.53 (2.21, 4.50)	0.84	0.85	5.67 (4.21, 6.90)	4.49 (3.63, 5.27)	5.22 (3.32, 6.76)	4.70 (1.89, 7.36)	0.57	0.4	2.15 (1.79, 2.47)	2.75 (1.83, 3.56)	3.24 (2.38, 4.12)	3.82 (2.58, 5.12)	**0.04**	0.07	2.02 (1.66, 2.41)	2.40 (1.86, 2.88)	2.73 (2.05, 3.31)	2.78 (1.77, 3.82)	0.11	0.29
Masimo Rad-G with Adult Probe, Product Code 4050	55	4.27 (3.70, 4.89)	3.53 (2.85, 4.21)	4.22 (3.13, 5.25)	5.58 (4.63, 6.42)	5.62 (4.53, 6.61)	**0.008**	**0.02**	4.31 (3.37, 5.20)	5.02 (3.83, 6.19)	6.85 (5.64, 7.91)	6.77 (5.60, 7.90)	**0.006**	0.01	3.88 (2.99, 4.70)	4.06 (2.98, 5.08)	5.42 (4.50, 6.25)	5.43 (4.34, 6.48)	**0.03**	**0.048**	2.49 (1.92, 3.05)	3.18 (2.05, 4.27)	3.44 (2.77, 3.99)	3.41 (2.73, 3.98)	**0.046**	**0.048**

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Guangdong Biolight M70	37	5.08 (4.30, 5.91)	3.77 (2.90, 4.53)	5.71 (4.10, 7.25)	5.55 (4.20, 6.69)	5.21 (3.53, 6.74)	0.06	0.21	6.36 (4.23, 8.27)	9.20 (6.59, 11.45)	8.78 (6.40, 10.92)	8.38 (6.12, 10.72)	0.1	0.14	2.07 (1.59, 2.60)	2.63 (1.86, 3.38)	4.19 (3.14, 5.14)	3.87 (2.40, 5.40)	**0.005**	0.08	2.35 (1.95, 2.71)	1.87 (1.58, 2.11)	2.56 (1.85, 3.23)	2.47 (1.38, 3.42)	0.92	0.85
Acare (2023) AH-M1	57	5.28 (4.60, 5.96)	4.81 (3.99, 5.69)	5.53 (4.61, 6.43)	4.92 (3.23, 6.57)	4.11 (2.26, 5.82)	0.78	0.39	8.12 (6.66, 9.24)	8.88 (7.12, 10.60)	8.44 (5.01, 11.29)	7.17 (3.92, 10.34)	0.81	0.53	3.20 (2.41, 3.99)	2.96 (2.22, 3.77)	3.21 (2.20, 4.14)	2.49 (1.32, 3.46)	0.84	0.28	1.70 (1.27, 2.12)	1.87 (1.30, 2.44)	1.83 (1.10, 2.45)	2.05 (0.91, 2.97)	0.97	0.71
Acare (2020) AH-MX	42	6.80 (6.05, 7.48)	6.59 (5.04, 8.29)	6.74 (5.77, 7.79)	6.80 (5.39, 8.08)	5.79 (4.43, 7.03)	0.76	0.68	10.98 (8.56, 13.57)	10.95 (9.62, 12.38)	11.34 (8.83, 13.33)	10.10 (7.71, 12.33)	0.75	0.84	3.55 (2.77, 4.35)	3.59 (3.07, 4.08)	4.48 (3.57, 5.24)	4.51 (3.14, 5.45)	0.18	0.32	1.80 (1.32, 2.25)	1.56 (1.23, 1.90)	1.53 (0.90, 2.14)	1.95 (0.94, 2.80)	0.28	0.96
American Oxygen LK-87	20	7.75 (7.23, 8.32)	7.82 (6.60, 8.84)	8.00 (7.21, 8.78)	7.32 (6.65, 7.92)	7.16 (6.16, 8.08)	0.65	0.55	9.37 (8.11, 10.59)	10.02 (9.26, 10.75)	9.18 (7.95, 10.40)	8.70 (7.27, 10.14)	0.89	0.59	9.32 (6.59, 11.63)	9.43 (8.06, 10.80)	8.38 (6.32, 9.97)	8.16 (6.05, 9.50)	0.69	0.68	4.03 (2.25, 5.41)	2.73 (1.58, 4.03)	2.90 (1.50, 4.29)	3.32 (1.20, 5.18)	0.36	0.61

Accuracy root mean square error (A_{RMS}) is reported for each device in a cohort that is as large as possible while satisfying the enrollment criteria in the anticipated 2025 U.S. Food and Drug Administration (FDA) and 2025 International Organization for Standardization (ISO) guidelines as the overall point estimate with 95 % confidence interval (CI). Columns include the mean A_{RMS} and 95 % CI within each functional arterial oxygen saturation (SaO_2) decile and within four skin color bins. Skin color bins are defined by Individual Typology Angle (ITA) measured at the forehead: ITA > 30° (light), -30° ≤ ITA ≤ 30° (medium), ITA < -30° (dark), and ITA <-50° (very dark); note that “very dark” is a subset of the dark bin. P-values test differences in A_{RMS} between light versus dark and light versus very dark bins using a Welch’s two-sample t-test. No correction for multiple pairwise comparisons (e.g., Bonferroni) was applied, as the 2025 draft FDA guidelines do not require adjustment. P-values near 0.05 are reported to three decimal places, and

those meeting $p < 0.05$ are bolded. The 95% CI were determined using bootstrapping (random resampling with replacement) with 1,000 repetitions.

Table S9. A_{RMS} at the DP for maximum participant cohorts

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Nonin CO-Pilot	70	1.39 (1.19, 1.60)	1.07 (0.95, 1.18)	1.61 (1.29, 1.94)	1.32 (1.06, 1.57)	1.30 (1.01, 1.65)	0.16	0.29	1.16 (0.93, 1.40)	1.83 (1.30, 2.45)	1.73 (1.19, 2.16)	1.81 (1.25, 2.37)	0.11	0.09	1.18 (1.01, 1.34)	1.72 (1.37, 2.04)	1.31 (1.04, 1.59)	1.36 (1.04, 1.76)	0.63	0.52	0.96 (0.79, 1.11)	1.34 (1.00, 1.69)	1.03 (0.80, 1.22)	0.97 (0.75, 1.17)	0.8	0.92
Shenzhen (2023) PC-60NW	21	1.61 (1.44, 1.80)	1.89 (1.45, 2.27)	1.61 (1.30, 1.89)	1.51 (1.23, 1.79)	1.43 (1.22, 1.60)	0.22	0.15	2.24 (1.19, 3.14)	2.16 (1.54, 2.76)	1.85 (1.28, 2.40)	1.94 (1.13, 2.67)	0.74	0.81	1.85 (1.41, 2.19)	1.41 (1.10, 1.69)	1.46 (1.08, 1.82)	1.40 (1.08, 1.70)	0.23	0.2	1.79 (1.42, 2.13)	1.30 (1.01, 1.59)	1.36 (0.91, 1.81)	1.10 (0.74, 1.42)	0.1	**0.02**
Masimo Rad-97 with RD Rainbow Adt 8λ Adult Adhesive Sensor	73.09	1.64 (1.03, 1.91)	1.22 (0.80, 1.66)	1.31 (0.82, 1.83)	2.46 (1.54, 3.22)	2.72 (0.84, 3.64)	0.1	0.24	1.63 (1.03, 2.17)	1.54 (0.98, 2.06)	2.56 (1.44, 3.47)	2.82 (1.07, 3.79)	0.32	0.29	1.00 (0.54, 1.35)	1.23 (0.74, 1.77)	2.77 (1.69, 3.66)	3.09 (0.73, 4.23)	0.053	0.21	1.02 (0.64, 1.35)	1.15 (0.59, 1.59)	1.88 (1.31, 2.33)	2.06 (0.63, 2.55)	0.07	0.26
Shenzhen (2024) PC-60NW	81	1.87 (1.45, 2.28)	1.59 (1.11, 2.12)	1.98 (1.21, 2.80)	2.05 (1.59, 2.37)	1.95 (1.32, 2.46)	0.14	0.34	1.96 (1.08, 2.73)	2.74 (1.32, 4.22)	2.94 (2.08, 3.59)	2.58 (1.35, 3.87)	0.1	0.5	1.34 (0.93, 1.72)	1.69 (0.96, 2.39)	1.37 (1.12, 1.59)	1.38 (1.04, 1.58)	0.72	0.67	1.37 (0.99, 1.79)	1.34 (1.02, 1.69)	1.62 (1.22, 2.00)	1.71 (1.14, 2.29)	0.33	0.41
Zug Medical Systems SAS SPLF-SE with SpotLife Reusable Probe	75	1.94 (1.66, 2.22)	1.53 (1.07, 1.97)	1.99 (1.51, 2.40)	1.98 (1.66, 2.28)	2.08 (1.55, 2.61)	0.09	0.14	1.73 (1.30, 2.14)	2.43 (1.75, 3.11)	1.34 (0.91, 1.76)	1.65 (0.79, 2.20)	0.24	0.79	1.48 (0.82, 2.07)	1.77 (1.32, 2.23)	2.28 (1.81, 2.78)	2.55 (1.65, 3.39)	**0.03**	0.07	1.47 (0.88, 2.01)	1.75 (1.39, 2.06)	2.06 (1.59, 2.49)	1.97 (1.18, 2.72)	0.09	0.32

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Compass Health POX-ROS	12	2.01 (1.73, 2.28)	2.29 (1.17, 3.31)	1.94 (1.54, 2.29)	1.86 (1.62, 2.07)	1.87 (1.64, 2.06)	0.8	0.84	2.59 (0.95, 3.94)	1.93 (1.38, 2.49)	2.15 (1.40, 2.82)	1.95 (1.40, 2.43)	0.83	0.77	3.15 (1.24, 4.90)	1.99 (1.50, 2.51)	1.74 (1.51, 1.94)	1.83 (1.56, 2.00)	0.47	0.53	1.55 (0.84, 2.11)	1.69 (1.35, 1.99)	1.75 (1.25, 2.19)	1.84 (1.21, 2.36)	0.65	0.59
Beijing Choice MD300M	40	2.05 (1.75, 2.37)	1.40 (1.12, 1.72)	1.81 (1.51, 2.10)	2.61 (1.93, 3.24)	2.63 (1.85, 3.22)	**0.01**	**0.04**	1.51 (0.67, 2.18)	2.38 (1.78, 2.94)	3.67 (2.11, 5.20)	2.99 (1.60, 3.91)	0.08	0.15	1.32 (1.20, 1.54)	1.53 (1.26, 1.79)	2.37 (1.91, 2.81)	2.69 (2.04, 3.31)	**0.004**	**0.02**	1.33 (0.79, 1.69)	1.43 (1.13, 1.71)	1.89 (1.22, 2.45)	1.92 (0.91, 2.64)	0.43	0.5
Nonin PalmSAT 2500	36	2.20 (1.87, 2.53)	2.44 (1.59, 3.12)	1.75 (1.49, 2.03)	2.65 (1.96, 3.37)	2.95 (1.97, 3.90)	0.78	0.53	3.14 (1.69, 4.15)	2.05 (1.72, 2.41)	3.41 (2.31, 4.44)	3.93 (2.42, 5.24)	0.89	0.53	2.30 (1.74, 2.76)	1.92 (1.54, 2.28)	2.32 (1.89, 2.70)	2.44 (1.76, 3.03)	0.95	0.94	1.61 (0.96, 2.25)	1.41 (1.18, 1.62)	1.67 (1.32, 2.02)	1.60 (1.03, 2.09)	0.65	0.86
Covidien PM1000N	59	2.21 (2.06, 2.59)	2.24 (1.70, 2.72)	2.08 (1.79, 2.36)	2.94 (2.57, 3.26)	3.07 (2.65, 3.43)	**0.008**	**0.005**	2.56 (1.67, 3.37)	2.47 (2.02, 2.95)	3.79 (3.21, 4.41)	4.09 (3.44, 4.74)	**0.001**	**<0.001**	2.26 (1.73, 2.73)	2.21 (1.89, 2.51)	3.17 (2.66, 3.61)	3.27 (2.71, 3.77)	**0.01**	**0.01**	1.99 (1.51, 2.45)	1.63 (1.41, 1.84)	1.87 (1.51, 2.21)	1.87 (1.43, 2.30)	0.93	0.93
Shenzhen AM801	10	2.21 (1.79, 2.61)	2.12 (1.91, 2.31)	2.22 (1.58, 2.78)	2.23 (1.47, 3.06)	2.23 (1.45, 3.07)	0.82	0.82	2.57 (1.67, 3.28)	2.43 (1.62, 3.32)	3.39 (1.69, 5.12)	3.39 (1.68, 5.13)	0.64	0.64	2.00 (1.58, 2.29)	2.31 (1.63, 3.05)	1.71 (1.44, 1.94)	1.71 (1.44, 1.95)	0.41	0.41	1.99 (1.40, 2.49)	1.96 (1.09, 2.87)	1.01 (0.64, 1.34)	1.01 (0.66, 1.34)	**0.02**	**0.02**
Nonin Onyx Vantage 9590	35	2.24 (1.84, 2.68)	1.79 (1.46, 2.22)	2.15 (1.80, 2.54)	2.48 (1.54, 3.37)	1.55 (1.31, 1.77)	0.49	0.4	1.24 (0.94, 1.64)	2.81 (2.26, 3.42)	3.04 (1.86, 4.25)	2.09 (1.67, 2.44)	**0.03**	**0.045**	1.21 (0.82, 1.63)	1.64 (1.22, 2.06)	1.91 (1.25, 2.50)	1.35 (0.99, 1.66)	0.25	0.77	2.42 (1.74, 3.19)	1.76 (1.29, 2.21)	1.96 (1.25, 2.61)	1.21 (0.93, 1.46)	0.27	0.1
Choicemed MD300CN350R	14	2.27 (2.03, 2.49)	2.45 (1.76, 3.08)	2.24 (1.86, 2.56)	2.14 (1.81, 2.45)	1.97 (1.61, 2.31)	0.64	0.41	3.29 (1.42, 5.21)	2.85 (2.41, 3.24)	2.49 (1.89, 3.12)	2.71 (1.97, 3.51)	0.86	0.95	2.55 (2.00, 2.97)	2.17 (1.67, 2.61)	2.31 (1.75, 2.83)	1.82 (1.18, 2.37)	0.41	0.09	2.12 (1.54, 2.57)	1.82 (1.16, 2.44)	1.70 (1.18, 2.11)	1.25 (0.65, 1.73)	0.28	0.06

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Masimo (2024) Mightysat	78	2.39 (1.98, 2.77)	1.79 (1.11, 2.47)	2.14 (1.67, 2.56)	3.42 (2.66, 4.05)	3.61 (2.84, 4.28)	**0.007**	**0.005**	1.84 (1.03, 2.67)	2.47 (1.89, 3.00)	4.18 (2.90, 5.16)	4.53 (3.40, 5.41)	**0.02**	**0.009**	1.92 (1.09, 2.63)	2.24 (1.67, 2.71)	3.94 (3.05, 4.75)	4.09 (3.04, 4.93)	**0.006**	**0.01**	1.68 (1.10, 2.26)	1.60 (1.29, 1.91)	2.61 (2.12, 3.05)	2.72 (2.15, 3.23)	**0.02**	**0.02**
Acare (2023) Model AH-TA	56	2.40 (2.15, 2.65)	2.85 (2.23, 3.43)	2.22 (1.97, 2.47)	2.35 (1.87, 2.84)	2.48 (1.83, 2.98)	0.32	0.51	3.36 (2.33, 4.30)	2.22 (1.80, 2.67)	2.71 (2.01, 3.34)	2.69 (1.77, 3.47)	0.33	0.34	3.04 (2.31, 3.67)	2.79 (2.39, 3.20)	2.53 (2.01, 3.03)	2.72 (2.03, 3.27)	0.38	0.66	2.13 (1.36, 2.96)	1.68 (1.43, 1.91)	1.82 (1.43, 2.24)	1.98 (1.46, 2.48)	0.69	0.98
Philips M1020B IntelliVue with M1191BL Probe	71	2.45 (1.93, 3.15)	1.79 (1.48, 2.07)	2.11 (1.75, 2.50)	3.47 (2.53, 4.43)	3.71 (2.38, 5.05)	**0.01**	0.06	2.47 (1.89, 3.04)	2.71 (2.20, 3.25)	4.50 (3.33, 5.67)	4.50 (2.82, 6.18)	**0.02**	0.11	1.28 (1.03, 1.51)	1.73 (1.42, 2.02)	2.99 (2.15, 3.80)	3.25 (2.19, 4.30)	**0.004**	**0.02**	1.24 (0.78, 1.62)	1.67 (1.24, 2.13)	2.63 (1.73, 3.47)	3.06 (1.74, 4.18)	**0.01**	**0.04**
CMS Zacurate 500C	15	2.51 (2.03, 2.96)	3.77 (2.59, 4.59)	1.55 (1.29, 1.82)	2.48 (1.95, 2.95)	2.57 (1.97, 3.08)	0.17	0.19	4.87 (3.36, 6.11)	1.96 (1.34, 2.53)	4.07 (2.85, 5.05)	4.25 (2.80, 5.33)	0.41	0.49	3.54 (2.10, 4.55)	1.68 (1.37, 1.97)	1.80 (1.33, 2.16)	1.87 (1.38, 2.30)	0.13	0.15	1.77 (1.22, 2.20)	1.25 (0.98, 1.56)	1.37 (1.00, 1.73)	1.41 (0.96, 1.80)	0.26	0.3
Covidien (2022) Nellcor PM10N	41	2.55 (1.92, 3.25)	3.84 (1.59, 6.05)	1.95 (1.55, 2.31)	2.96 (2.04, 3.65)	3.56 (2.66, 4.17)	0.83	0.75	4.89 (1.36, 8.14)	2.44 (1.80, 3.24)	3.76 (2.72, 4.59)	4.23 (3.24, 5.02)	0.93	0.71	3.32 (1.62, 5.03)	1.81 (1.25, 2.24)	2.91 (1.98, 3.62)	3.49 (2.67, 4.28)	0.95	0.56	2.79 (1.77, 3.59)	1.54 (1.10, 1.89)	1.83 (1.08, 2.38)	2.24 (1.48, 2.77)	0.21	0.54
Utech UT100	39	2.72 (2.14, 3.33)	1.73 (1.44, 1.98)	2.25 (1.61, 2.92)	3.83 (2.48, 4.99)	4.36 (2.32, 5.79)	**0.03**	0.06	2.28 (1.64, 2.84)	2.95 (1.95, 3.84)	4.39 (2.40, 6.29)	4.95 (2.40, 7.40)	0.15	0.18	1.39 (1.01, 1.81)	2.42 (1.44, 3.30)	3.95 (2.30, 5.41)	4.20 (1.99, 6.31)	**0.03**	0.11	1.48 (1.23, 1.74)	1.67 (1.28, 2.07)	2.16 (1.69, 2.59)	2.24 (1.64, 2.66)	0.07	0.1
Nihon Kohden OLV-4202 with	64.03	2.76 (2.14, 4.21)	2.08 (1.24, 2.99)	2.01 (1.45, 2.52)	3.27 (2.46, 3.96)	3.33 (2.49, 4.10)	0.06	0.06	3.73 (1.37, 6.00)	2.25 (1.60, 2.85)	3.93 (2.59, 4.99)	4.24 (2.82, 5.48)	0.51	0.4	1.10 (0.66, 1.46)	2.09 (1.39, 2.63)	3.24 (2.54, 3.85)	3.35 (2.61, 4.10)	**0.001**	**0.002**	1.55 (1.19, 1.89)	1.83 (1.11, 2.59)	2.51 (1.69, 3.29)	2.11 (1.44, 2.65)	0.12	0.27

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
NK-SPO2-ADVT Probe																										
EDAN (2020) H100b	44	2.80 (2.45, 3.18)	2.76 (2.17, 3.24)	2.70 (2.13, 3.34)	4.31 (2.73, 6.06)	4.41 (2.50, 6.65)	0.21	0.36	3.17 (2.45, 3.79)	3.09 (2.23, 3.96)	5.45 (3.02, 7.93)	5.60 (2.70, 8.77)	0.21	0.38	3.31 (2.30, 4.38)	2.76 (2.26, 3.28)	3.87 (2.96, 4.73)	3.70 (2.78, 4.77)	0.35	0.49	2.20 (1.50, 2.79)	2.06 (1.62, 2.52)	2.36 (1.69, 2.99)	2.52 (1.73, 3.27)	0.71	0.52
Shenzhen WIT-S300C	62	2.81 (2.20, 3.38)	2.16 (1.48, 2.87)	3.35 (2.37, 4.27)	2.56 (2.05, 3.04)	2.48 (1.89, 2.95)	0.28	0.4	1.96 (1.36, 2.60)	3.73 (2.17, 5.19)	3.24 (2.51, 3.88)	3.42 (2.49, 4.17)	**0.03**	0.06	2.25 (1.48, 3.05)	3.47 (2.43, 4.42)	2.58 (1.89, 3.24)	2.53 (1.79, 3.16)	0.51	0.49	2.21 (1.38, 2.92)	2.76 (1.91, 3.65)	1.91 (1.31, 2.54)	1.61 (1.18, 1.98)	0.69	0.41
CONTEC CMS50M	17	2.82 (2.36, 3.23)	2.85 (2.19, 3.62)	3.15 (2.51, 3.77)	2.09 (1.29, 2.80)	2.21 (1.34, 2.88)	0.15	0.24	4.49 (3.18, 5.91)	4.26 (2.93, 5.49)	3.42 (1.94, 4.79)	3.60 (1.96, 5.04)	0.24	0.34	2.48 (1.76, 3.29)	2.28 (1.95, 2.57)	2.06 (0.88, 3.04)	2.20 (0.94, 3.24)	0.32	0.47	1.60 (1.01, 2.10)	2.82 (1.60, 4.20)	1.14 (0.86, 1.38)	1.20 (0.91, 1.44)	0.27	0.36
Shenzhen A310L	11	2.84 (2.47, 3.18)	2.79 (1.67, 3.58)	2.76 (2.47, 3.05)	2.93 (2.19, 3.70)	3.12 (2.07, 4.07)	0.73	0.64	4.54 (2.61, 5.94)	3.41 (2.66, 4.21)	3.58 (2.26, 4.81)	4.14 (2.62, 5.49)	0.45	0.84	2.70 (1.67, 3.45)	2.81 (2.38, 3.22)	2.89 (2.38, 3.36)	2.89 (2.16, 3.54)	0.55	0.62	1.42 (0.93, 1.85)	1.90 (1.61, 2.17)	1.90 (1.45, 2.27)	1.50 (1.04, 1.84)	0.19	0.76
Masimo (2020) Mightysat Rx	7	2.88 (2.00, 4.11)	2.16 (1.44, 2.81)	2.30 (1.79, 2.81)	3.78 (1.96, 5.75)	4.47 (1.61, 7.20)	0.3	0.38	2.71 (1.92, 3.28)	1.97 (1.49, 2.39)	5.44 (1.82, 8.88)	6.82 (1.64, 11.37)	0.54	0.44	2.59 (0.75, 4.26)	2.44 (1.62, 3.30)	3.64 (2.08, 5.36)	4.09 (1.52, 6.56)	0.27	0.39	1.76 (1.20, 2.21)	2.48 (1.72, 3.38)	2.08 (1.59, 2.49)	2.00 (1.43, 2.66)	0.34	0.54
Masimo Rad-G with Rad-G Finger Clip Sensor	55.03	3.00 (2.43, 3.54)	2.70 (1.73, 3.69)	2.89 (2.22, 3.48)	4.07 (2.31, 5.55)	4.31 (2.02, 6.74)	0.26	0.47	3.01 (1.80, 4.08)	2.91 (2.18, 3.64)	4.75 (2.97, 6.47)	5.01 (1.96, 7.84)	0.19	0.45	2.58 (1.44, 3.70)	3.01 (2.27, 3.67)	4.01 (2.10, 5.51)	4.19 (1.65, 6.46)	0.23	0.41	2.55 (1.26, 3.51)	2.74 (2.13, 3.27)	3.22 (1.93, 4.13)	3.22 (1.80, 4.56)	0.33	0.42

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Bistos Co., Ltd. BT-710	61	3.15 (2.86, 3.42)	3.01 (2.58, 3.49)	3.06 (2.62, 3.46)	3.38 (2.84, 3.85)	3.51 (2.79, 4.04)	0.46	0.34	3.39 (2.37, 4.13)	3.66 (3.15, 4.15)	4.84 (3.56, 6.08)	4.92 (3.93, 5.50)	0.23	0.07	2.79 (2.31, 3.15)	3.07 (2.59, 3.56)	3.15 (2.61, 3.67)	3.74 (2.97, 4.38)	0.46	0.09	2.62 (1.78, 3.24)	2.51 (1.82, 3.28)	2.16 (1.77, 2.56)	1.93 (1.60, 2.19)	0.52	0.34
BodyMed BDMOXMTRBLK	2	3.26 (2.92, 3.64)	4.11 (3.08, 5.22)	3.12 (2.69, 3.47)	2.93 (2.41, 3.32)	2.54 (1.96, 3.03)	0.12	**0.049**	4.57 (2.55, 6.43)	3.35 (2.59, 3.97)	3.57 (2.73, 4.43)	3.22 (2.45, 3.94)	0.61	0.46	5.19 (3.34, 7.19)	3.76 (2.87, 4.54)	3.31 (2.45, 4.06)	2.87 (2.08, 3.58)	0.15	0.09	2.72 (2.23, 3.14)	2.49 (2.11, 2.79)	2.18 (1.43, 2.72)	1.88 (1.15, 2.51)	0.1	**0.04**
Masimo Rad-97 with Adult Probe, Product Code 4050	60	3.33 (2.9, 3.77)	3.26 (2.32, 4.04)	2.94 (2.46, 3.47)	4.38 (3.89, 4.90)	4.64 (3.95, 5.30)	**0.01**	**0.005**	3.02 (1.87, 4.03)	3.42 (2.88, 3.92)	5.34 (4.77, 5.95)	5.72 (4.88, 6.54)	**<0.001**	**<0.001**	3.40 (2.38, 4.28)	3.02 (2.49, 3.55)	4.74 (4.18, 5.29)	5.00 (4.36, 5.55)	**0.009**	**0.002**	3.07 (2.17, 3.85)	2.31 (1.80, 2.82)	2.77 (2.36, 3.16)	2.88 (2.37, 3.36)	0.99	0.75
Guangdong Biolight M800	43	3.53 (3.17, 3.88)	3.71 (2.72, 4.66)	3.37 (2.92, 3.82)	3.67 (3.02, 4.23)	2.81 (2.27, 3.29)	0.85	0.26	5.72 (4.03, 7.29)	4.84 (3.83, 5.76)	5.04 (3.30, 6.49)	3.05 (1.72, 3.97)	0.59	**0.048**	1.98 (1.53, 2.41)	2.67 (1.96, 3.35)	3.28 (2.41, 4.11)	2.80 (1.95, 3.46)	**0.03**	0.16	2.03 (1.77, 2.30)	2.21 (1.71, 2.65)	2.87 (2.15, 3.48)	2.52 (1.64, 3.21)	0.11	0.5
Masimo Rad-G with Adult Probe, Product Code 4050	55	4.27 (3.70, 4.89)	3.71 (2.41, 4.99)	4.14 (3.31, 4.97)	5.58 (4.63, 6.37)	5.92 (5.12, 6.63)	**0.01**	**0.003**	4.21 (2.89, 5.45)	5.16 (4.09, 6.26)	6.72 (5.61, 7.56)	7.05 (6.11, 7.89)	**0.004**	**0.002**	3.84 (2.59, 5.08)	4.16 (3.31, 4.91)	5.46 (4.47, 6.30)	5.77 (4.76, 6.63)	**0.03**	**0.01**	3.15 (1.74, 4.57)	2.82 (2.15, 3.45)	3.33 (2.73, 3.81)	3.52 (3.03, 3.97)	0.29	0.15
Guangdong Biolight M70	37	5.08 (4.30, 5.91)	3.70 (2.57, 4.70)	5.10 (3.95, 6.23)	5.71 (4.22, 7.05)	5.03 (3.41, 6.55)	0.1	0.27	5.64 (3.31, 7.51)	8.51 (6.56, 10.31)	8.80 (6.26, 11.01)	7.75 (5.07, 10.16)	0.1	0.22	1.90 (1.59, 2.21)	2.69 (1.93, 3.43)	4.19 (2.96, 5.32)	4.42 (2.80, 5.69)	**0.007**	**0.03**	2.50 (2.11, 2.83)	1.90 (1.54, 2.27)	2.64 (1.88, 3.41)	2.45 (1.60, 3.30)	0.73	0.56

Device	Device ID	Overall Device A_{RMS} (95% Confidence Interval)	70–100% SaO_2 A_{RMS} (Mean, 95% CI)						70–80% SaO_2 A_{RMS} (Mean, 95% CI)						80–90% SaO_2 A_{RMS} (Mean, 95% CI)						90–100% SaO_2 A_{RMS} (Mean, 95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Acare (2023) AH-M1	57	5.28 (4.60, 5.96)	4.26 (3.13, 5.33)	5.63 (4.79, 6.54)	4.63 (3.39, 5.79)	4.40 (3.03, 5.51)	0.93	0.96	7.49 (5.13, 9.26)	9.28 (7.64, 10.95)	7.62 (5.49, 9.47)	7.45 (5.03, 9.55)	0.92	0.95	2.99 (2.08, 3.88)	3.35 (2.59, 4.07)	2.77 (2.17, 3.30)	2.93 (2.13, 3.54)	0.7	0.94	1.87 (0.93, 2.61)	1.68 (1.25, 2.19)	1.94 (1.40, 2.53)	2.00 (1.22, 2.71)	0.82	0.78
Acare (2020) AH-MX	42	6.80 (6.05, 7.48)	7.86 (5.62, 9.75)	6.08 (5.36, 6.72)	6.90 (5.52, 8.14)	6.57 (5.08, 7.99)	0.55	0.47	12.54 (9.50, 15.41)	10.11 (8.83, 11.25)	11.71 (9.56, 13.68)	10.81 (8.87, 12.70)	0.76	0.53	3.96 (2.84, 4.92)	3.44 (3.01, 3.83)	4.33 (3.46, 5.08)	4.43 (3.28, 5.45)	0.57	0.47	1.72 (1.31, 2.10)	1.42 (1.09, 1.76)	1.86 (1.21, 2.45)	1.43 (0.94, 1.84)	0.95	0.36
American Oxygen LK-87	20	7.75 (7.23, 8.32)	8.02 (6.23, 9.48)	7.94 (7.21, 8.61)	7.24 (6.62, 7.85)	7.16 (6.43, 7.86)	0.58	0.54	9.08 (7.53, 10.32)	9.88 (9.12, 10.63)	9.26 (8.02, 10.37)	9.46 (8.14, 10.75)	0.9	0.78	10.70 (6.33, 13.85)	9.17 (7.76, 10.43)	8.21 (6.46, 9.66)	7.67 (5.93, 8.98)	0.42	0.33	2.81 (0.88, 4.68)	3.37 (2.08, 4.37)	2.88 (1.45, 4.27)	2.99 (1.36, 4.52)	0.86	0.83

Accuracy root mean square error (A_{RMS}) is reported for each device in a cohort that is as large as possible while satisfying the enrollment criteria in the anticipated 2025 U.S. Food and Drug Administration (FDA) and 2025 International Organization for Standardization (ISO) guidelines as the overall point estimate with 95 % confidence interval (CI). Columns include the mean A_{RMS} and 95 % CI within each functional arterial oxygen saturation (SaO_2) decile and within four skin color bins. Skin color bins are defined by Individual Typology Angle (ITA) measured at the dorsal distal phalanx (DP): ITA > 30° (light), -30° ≤ ITA ≤ 30° (medium), ITA < -30° (dark), and ITA <-50° (very dark); note that “very dark” is a subset of the dark bin. P-values test differences in A_{RMS} between light versus dark and light versus very dark bins using a Welch’s two-sample t-test. No correction for multiple pairwise comparisons (e.g., Bonferroni) was applied, as the 2025 draft FDA guidelines do not require adjustment. P-values near 0.05 are reported to three decimal places, and those meeting $p < 0.05$ are bolded. The 95% CI were determined using bootstrapping (random resampling with replacement) with 1,000 repetitions.

Table S10. Bias (mean of SpO_2 minus SaO_2) at the DP for maximum participant cohorts

Device	Device ID	Overall Bias (95% CI)	70–100% SaO_2 Bias (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Nonin CO-Pilot	70	-0.28 (-0.34, -0.20)	-0.14 (-0.24, -0.02)	-0.32 (-0.44, -0.20)	-0.32 (-0.46, -0.20)	-0.42 (-0.57, -0.27)	0.53	0.35	-0.15 (-0.37, 0.05)	-0.13 (-0.34, 0.10)	-0.40 (-0.69, -0.10)	-0.86 (-1.18, -0.53)	0.66	0.24	-0.11 (-0.33, 0.11)	-0.39 (-0.63, -0.16)	-0.27 (-0.52, -0.05)	-0.35 (-0.64, -0.08)	0.65	0.56	0.03 (-0.32, 0.20)	-0.21 (-0.84, 0.14)	-0.19 (-0.65, 0.13)	-0.26 (-0.48, 0.11)	0.54	0.81
Shenzhen (2023) PC-60NW	21	-0.04 (-0.15, 0.08)	0.19 (-0.05, 0.42)	-0.27 (-0.43, -0.11)	0.16 (-0.04, 0.37)	0.10 (-0.17, 0.33)	0.88	0.77	-0.65 (-1.13, -0.24)	-1.18 (-1.54, -0.84)	-0.72 (-1.12, -0.33)	-0.78 (-1.28, -0.32)	0.86	0.79	0.15 (-0.30, 0.59)	-0.26 (-0.51, -0.01)	0.43 (0.12, 0.73)	0.34 (-0.06, 0.69)	0.51	0.77	-0.15 (-0.31, 0.02)	-0.43 (-0.60, -0.28)	-0.31 (-0.47, -0.15)	-0.20 (-0.38, -0.01)	0.6	0.45
Masimo Rad-97 with RD Rainbow Adt 8λ Adult Adhesive Sensor	73.09	0.86 (0.74, 0.97)	0.49 (0.32, 0.68)	0.62 (0.47, 0.76)	1.72 (1.43, 2.01)	1.77 (1.32, 2.25)	**0.049**	0.09	0.22 (-0.22, 0.66)	0.48 (0.16, 0.82)	1.55 (0.96, 2.15)	1.40 (0.49, 2.30)	0.08	0.13	0.67 (0.46, 0.89)	0.82 (0.63, 1.04)	2.15 (1.67, 2.62)	2.39 (1.61, 3.17)	**0.02**	**0.03**	0.96 (0.63, 1.29)	0.38 (0.16, 0.58)	0.71 (0.45, 0.97)	0.62 (0.36, 0.89)	0.09	0.21
Shenzhen (2024) PC-60NW	81	-0.02 (-0.17, 0.12)	-0.23 (-0.45, 0.01)	0.13 (-0.09, 0.33)	-0.09 (-0.43, 0.23)	0.15 (-0.31, 0.59)	0.7	0.43	-1.02 (-1.43, -0.62)	-0.93 (-1.43, -0.48)	-2.35 (-2.93, -1.85)	-1.95 (-2.60, -1.30)	**0.01**	0.17	-0.23 (-0.55, 0.12)	0.32 (-0.00, 0.66)	0.51 (0.16, 0.85)	0.93 (0.52, 1.33)	0.15	0.06	0.56 (0.34, 0.79)	0.56 (0.36, 0.77)	1.41 (1.06, 1.75)	1.38 (0.80, 1.94)	0.12	0.21
Zug Medical Systems SAS SPLF-SE with SpotLife Reusable Probe	75	1.08 (0.97, 1.19)	0.59 (0.36, 0.81)	1.04 (0.88, 1.19)	1.58 (1.42, 1.73)	1.71 (1.47, 1.94)	**0.006**	**0.03**	-0.25 (-0.72, 0.23)	0.53 (0.15, 0.94)	0.86 (0.55, 1.16)	1.37 (0.93, 1.81)	0.15	0.15	0.82 (0.52, 1.15)	1.18 (0.98, 1.39)	1.97 (1.71, 2.20)	2.10 (1.71, 2.50)	**0.008**	**0.03**	0.60 (0.30, 0.89)	1.00 (0.83, 1.18)	1.35 (1.12, 1.58)	1.48 (1.16, 1.81)	0.1	0.36

Device	Device ID	Overall Bias (95% CI)	70–100% SaO_2 Bias (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Compass Health POX-ROS	12	-0.30 (-0.44, -0.17)	0.38 (-0.04, 0.83)	-0.23 (-0.44, -0.02)	-0.73 (-0.94, -0.52)	-0.69 (-0.96, -0.44)	0.07	0.12	0.43 (-0.36, 1.38)	-0.30 (-0.64, 0.10)	-1.14 (-1.58, -0.75)	-1.12 (-1.54, -0.71)	0.06	0.07	-0.09 (-0.85, 0.70)	-0.46 (-0.84, -0.04)	-0.68 (-1.07, -0.34)	-0.62 (-1.05, -0.16)	0.14	0.25	1.10 (0.87, 1.39)	1.33 (1.17, 1.49)	1.72 (1.51, 1.97)	1.57 (1.21, 1.96)	0.21	0.21
Beijing Choice MD300M	40	0.92 (0.79, 1.05)	0.59 (0.28, 0.89)	0.86 (0.72, 0.99)	1.16 (0.85, 1.51)	1.78 (1.39, 2.19)	0.54	0.29	-0.19 (-0.92, 0.49)	1.43 (1.11, 1.75)	2.17 (1.59, 2.80)	2.24 (1.51, 2.99)	**0.04**	0.13	1.00 (0.66, 1.34)	0.86 (0.66, 1.05)	1.13 (0.67, 1.58)	2.05 (1.39, 2.62)	0.9	0.51	0.76 (0.26, 1.24)	0.04 (-0.23, 0.35)	-0.38 (-0.74, 0.01)	-0.39 (-0.83, 0.05)	0.47	0.96
Nonin PalmSAT 2500	36	-1.35 (-1.47, -1.24)	-1.58 (-1.83, -1.30)	-1.07 (-1.20, -0.92)	-1.67 (-1.93, -1.41)	-1.66 (-2.04, -1.29)	0.99	0.99	-2.24 (-2.80, -1.71)	-1.56 (-1.77, -1.32)	-2.61 (-3.11, -2.11)	-3.11 (-3.83, -2.41)	0.89	0.59	-1.89 (-2.21, -1.54)	-1.55 (-1.77, -1.31)	-1.42 (-1.77, -1.02)	-1.15 (-1.69, -0.59)	0.65	0.61	0.83 (0.41, 1.24)	0.40 (0.20, 0.58)	0.13 (-0.34, 0.58)	0.87 (0.27, 1.47)	0.58	0.85
Covidien PM1000N	59	1.29 (1.22, 1.37)	0.98 (0.83, 1.12)	1.06 (0.98, 1.16)	2.32 (2.16, 2.47)	2.47 (2.27, 2.66)	**<0.001**	**0.001**	0.52 (0.22, 0.82)	0.97 (0.78, 1.16)	3.31 (3.04, 3.58)	3.68 (3.37, 3.97)	**<0.001**	**<0.001**	1.56 (1.33, 1.80)	1.55 (1.40, 1.69)	2.67 (2.44, 2.89)	2.76 (2.46, 3.03)	**0.003**	**0.003**	-0.60 (-1.01, -0.21)	-0.31 (-0.53, -0.08)	-0.92 (-1.22, -0.64)	-0.40 (-0.88, 0.05)	0.34	0.39
Shenzhen AM801	10	0.38 (0.21, 0.55)	0.63 (0.31, 0.95)	-0.09 (-0.35, 0.15)	1.07 (0.77, 1.40)	1.07 (0.77, 1.40)	0.31	0.31	0.11 (-0.62, 0.82)	-0.91 (-1.41, -0.45)	1.78 (1.09, 2.56)	1.78 (1.09, 2.56)	0.07	0.07	0.85 (0.29, 1.33)	0.29 (-0.17, 0.71)	1.13 (0.76, 1.48)	1.13 (0.76, 1.48)	0.6	0.6	0.82 (0.58, 1.04)	0.71 (0.57, 0.81)	1.18 (1.00, 1.37)	1.24 (1.02, 1.49)	0.27	0.27
Nonin Onyx Vantage 9590	35	1.36 (1.24, 1.47)	1.15 (0.82, 1.48)	1.25 (1.09, 1.40)	1.60 (1.38, 1.84)	1.05 (0.89, 1.20)	0.59	0.84	0.54 (0.06, 0.99)	1.96 (1.64, 2.28)	2.47 (2.03, 3.00)	1.67 (1.35, 1.94)	0.14	**0.02**	0.80 (0.23, 1.29)	0.84 (0.62, 1.07)	1.16 (0.84, 1.48)	0.79 (0.52, 1.08)	0.61	0.97	0.90 (0.38, 1.41)	0.25 (-0.11, 0.59)	0.30 (0.07, 0.58)	0.30 (0.07, 0.58)	0.31	**0.02**
Choicemed MD300CN350R	14	1.02 (0.89, 1.15)	1.39 (1.07, 1.69)	1.09 (0.92, 1.26)	0.66 (0.42, 0.89)	0.14 (-0.15, 0.45)	0.11	**0.01**	0.09 (-0.77, 0.86)	1.33 (0.93, 1.69)	-0.33 (-0.82, 0.16)	-1.00 (-1.72, -0.28)	0.72	0.28	2.10 (1.79, 2.44)	1.46 (1.22, 1.73)	1.33 (0.94, 1.74)	0.95 (0.57, 1.35)	0.1	**0.04**	2.05 (1.52, 2.54)	0.92 (0.72, 1.12)	1.17 (0.89, 1.45)	0.77 (0.55, 1.00)	0.27	0.06

Device	Device ID	Overall Bias (95% CI)	70–100% SaO_2 Bias (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Masimo (2024) Mightysat	78	1.71 (1.58, 1.84)	1.06 (0.85, 1.29)	1.59 (1.44, 1.73)	3.08 (2.79, 3.36)	3.29 (2.95, 3.60)	**0.002**	**0.001**	0.72 (0.23, 1.20)	1.69 (1.34, 2.04)	3.69 (2.99, 4.36)	4.03 (3.36, 4.70)	**0.002**	**<0.001**	1.18 (0.76, 1.55)	1.79 (1.57, 2.02)	3.50 (3.00, 4.01)	3.56 (2.98, 4.18)	**0.001**	**0.002**	1.50 (1.08, 1.90)	0.56 (0.32, 0.80)	0.89 (0.62, 1.18)	0.37 (0.08, 0.68)	**0.02**	**0.02**
Acare (2023) Model AH-TA	56	1.54 (1.42, 1.64)	1.77 (1.42, 2.09)	1.61 (1.48, 1.73)	1.28 (1.06, 1.50)	1.30 (0.96, 1.59)	0.42	0.54	1.13 (0.18, 2.06)	1.26 (1.01, 1.51)	1.00 (0.46, 1.44)	1.28 (0.63, 1.83)	0.81	0.93	2.43 (2.04, 2.83)	2.32 (2.10, 2.55)	1.85 (1.44, 2.19)	1.70 (1.07, 2.26)	0.21	0.27	1.27 (0.98, 1.54)	1.29 (1.14, 1.44)	2.34 (2.04, 2.66)	2.48 (2.13, 2.83)	0.22	0.3
Philips M1020B IntelliVue with M1191BL Probe	71	1.90 (1.81, 1.99)	1.42 (1.26, 1.57)	1.69 (1.59, 1.81)	2.84 (2.61, 3.08)	3.06 (2.74, 3.36)	**0.003**	**0.003**	2.27 (2.03, 2.53)	2.38 (2.18, 2.59)	3.92 (3.43, 4.43)	3.96 (3.38, 4.59)	**0.02**	0.053	1.07 (0.87, 1.25)	1.44 (1.30, 1.58)	2.58 (2.28, 2.87)	2.81 (2.42, 3.22)	**0.001**	**<0.001**	1.66 (1.34, 2.01)	1.39 (1.25, 1.52)	1.01 (0.73, 1.27)	0.98 (0.55, 1.37)	**0.005**	**0.003**
CMS Zacurate 500C	15	0.78 (0.58, 0.96)	1.50 (0.82, 2.13)	0.91 (0.74, 1.06)	0.37 (0.07, 0.65)	0.44 (0.09, 0.78)	0.2	0.27	0.51 (-1.43, 2.26)	0.78 (0.44, 1.15)	-0.74 (-1.62, 0.07)	-0.39 (-1.47, 0.65)	0.28	0.38	2.79 (1.87, 3.67)	1.23 (0.98, 1.48)	1.00 (0.68, 1.30)	0.96 (0.59, 1.31)	0.18	0.21	0.79 (0.58, 1.03)	1.25 (1.08, 1.42)	2.02 (1.75, 2.32)	2.33 (1.93, 2.83)	0.14	0.13
Covidien (2022) Nellcor PM10N	41	1.17 (0.99, 1.34)	1.78 (1.20, 2.46)	0.59 (0.40, 0.77)	2.21 (1.91, 2.48)	3.12 (2.79, 3.48)	0.7	0.31	1.18 (-0.22, 2.94)	-0.12 (-0.54, 0.27)	3.33 (2.84, 3.81)	3.85 (3.20, 4.40)	0.17	0.14	2.11 (1.41, 2.93)	1.13 (0.89, 1.38)	2.39 (1.99, 2.79)	3.33 (2.92, 3.74)	0.87	0.41	1.32 (0.97, 1.67)	0.74 (0.52, 0.94)	0.66 (0.41, 0.91)	0.58 (0.31, 0.87)	0.22	0.73
Utech UT100	39	1.02 (0.85, 1.19)	0.37 (0.11, 0.59)	0.95 (0.75, 1.16)	1.77 (1.28, 2.24)	2.35 (1.62, 3.00)	**0.03**	**0.003**	-1.08 (-1.62, -0.55)	0.77 (0.32, 1.28)	3.21 (2.23, 4.21)	3.74 (2.54, 5.04)	**0.001**	**0.001**	0.88 (0.62, 1.16)	1.13 (0.79, 1.52)	1.55 (0.59, 2.42)	1.75 (0.31, 2.77)	0.19	0.11	2.07 (1.50, 2.58)	0.70 (0.49, 0.94)	0.86 (0.46, 1.29)	1.82 (1.31, 2.30)	0.41	0.66
Nihon Kohden OLV-4202 with	64.03	0.36 (0.18, 0.53)	0.58 (0.24, 0.90)	-0.04 (-0.24, 0.15)	1.34 (0.88, 1.82)	1.50 (0.91, 2.06)	0.53	0.51	0.16 (-0.94, 1.12)	0.04 (-0.27, 0.41)	2.27 (1.21, 3.22)	3.11 (1.74, 4.30)	0.23	0.25	0.39 (0.07, 0.68)	-0.19 (-0.55, 0.12)	1.22 (0.23, 2.12)	1.14 (0.03, 2.12)	0.6	0.6	1.15 (0.92, 1.38)	0.94 (0.73, 1.17)	0.80 (0.31, 1.25)	1.47 (0.81, 2.05)	0.92	0.79

Device	Device ID	Overall Bias (95% CI)	70–100% SaO_2 Bias (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
NK-SPO2-ADVT Probe																										
EDAN (2020) H100b	44	2.34 (2.19, 2.51)	2.06 (1.81, 2.33)	2.00 (1.81, 2.20)	3.36 (2.92, 3.83)	3.47 (2.93, 4.10)	0.12	0.14	1.98 (1.29, 2.61)	2.20 (1.81, 2.60)	4.81 (3.75, 5.86)	4.97 (3.71, 6.48)	0.08	0.14	2.40 (2.00, 2.83)	2.38 (2.12, 2.67)	3.34 (2.90, 3.75)	3.30 (2.90, 3.68)	0.17	0.25	1.07 (0.76, 1.40)	0.05 (-0.25, 0.34)	0.73 (0.06, 1.39)	0.66 (0.02, 1.22)	0.8	0.45
Shenzhen WIT-S300C	62	1.86 (1.73, 2.01)	1.52 (1.32, 1.72)	2.11 (1.84, 2.39)	1.81 (1.60, 2.00)	1.92 (1.70, 2.11)	0.41	0.3	1.30 (0.96, 1.67)	2.34 (1.82, 2.88)	2.00 (1.47, 2.53)	2.62 (2.14, 3.11)	0.16	**0.02**	1.68 (1.34, 2.00)	2.55 (2.17, 2.94)	2.05 (1.74, 2.34)	2.00 (1.69, 2.32)	0.44	0.45	1.79 (1.51, 2.09)	1.50 (1.29, 1.73)	1.95 (1.59, 2.31)	2.21 (1.86, 2.62)	0.73	0.6
CONTEC CMS50M	17	-1.24 (-1.44, -1.04)	-1.94 (-2.30, -1.60)	-1.27 (-1.57, -0.98)	-0.58 (-0.88, -0.28)	-0.69 (-1.08, -0.31)	**0.009**	**0.02**	-3.63 (-4.40, -2.79)	-2.61 (-3.20, -2.00)	-1.86 (-2.63, -1.11)	-2.13 (-2.95, -1.30)	0.32	0.46	-2.11 (-2.50, -1.73)	-1.12 (-1.47, -0.74)	-0.70 (-1.23, -0.21)	-0.84 (-1.45, -0.23)	**0.04**	0.08	1.55 (1.27, 1.87)	1.54 (1.13, 1.84)	1.42 (1.18, 1.70)	1.33 (1.12, 1.54)	**0.04**	0.07
Shenzhen A310L	11	1.43 (1.25, 1.61)	1.17 (0.74, 1.57)	1.13 (0.87, 1.35)	2.13 (1.87, 2.41)	2.33 (1.97, 2.66)	0.22	0.22	0.75 (-0.60, 1.84)	-0.28 (-0.87, 0.32)	2.70 (2.02, 3.39)	3.85 (3.11, 4.57)	0.09	**0.04**	1.80 (1.15, 2.45)	2.10 (1.78, 2.37)	2.46 (2.10, 2.82)	2.48 (2.08, 2.88)	0.44	0.52	-0.66 (-1.05, -0.28)	-0.26 (-0.80, 0.14)	0.39 (0.11, 0.68)	0.36 (0.04, 0.66)	0.56	0.59
Masimo (2020) Mightysat Rx	7	0.92 (0.77, 1.08)	0.24 (-0.11, 0.59)	0.62 (0.43, 0.79)	2.23 (1.87, 2.61)	2.62 (2.08, 3.15)	**0.02**	**0.03**	-1.03 (-1.72, -0.34)	0.06 (-0.22, 0.34)	2.35 (1.38, 3.45)	3.67 (2.14, 5.40)	0.06	**0.048**	0.65 (0.17, 1.23)	1.12 (0.85, 1.39)	2.61 (2.21, 3.06)	2.56 (1.92, 3.26)	0.12	0.21	0.98 (0.72, 1.25)	1.51 (1.32, 1.69)	1.26 (0.95, 1.59)	0.80 (0.46, 1.12)	0.13	0.24
Masimo Rad-G with Rad-G Finger Clip Sensor	55.03	1.94 (1.77, 2.10)	1.40 (1.00, 1.75)	1.89 (1.70, 2.10)	2.96 (2.40, 3.49)	3.25 (2.63, 3.96)	0.2	0.23	0.54 (-0.16, 1.35)	1.05 (0.63, 1.51)	2.87 (1.29, 4.37)	3.57 (1.67, 5.40)	0.5	0.44	1.84 (1.38, 2.32)	2.50 (2.24, 2.77)	3.31 (2.64, 3.91)	3.61 (2.77, 4.46)	0.14	0.19	1.09 (0.69, 1.51)	0.60 (0.27, 0.90)	1.78 (1.51, 2.04)	1.77 (1.48, 2.10)	0.33	0.43

Device	Device ID	Overall Bias (95% CI)	70–100% SaO_2 Bias (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Bistos Co., Ltd. BT-710	61	2.46 (2.33, 2.50)]	2.60 (2.33, 2.85)	2.34 (2.19, 2.51)	2.65 (2.40, 2.91)	3.16 (2.89, 3.43)	0.97	0.22	3.21 (2.64, 3.69)	2.97 (2.65, 3.29)	3.73 (3.04, 4.45)	4.69 (4.14, 5.30)	0.46	**0.04**	2.61 (2.31, 2.94)	2.77 (2.55, 3.00)	2.85 (2.56, 3.13)	3.56 (3.22, 3.89)	0.73	0.07	1.89 (1.46, 2.38)	2.18 (1.93, 2.42)	2.61 (2.12, 3.13)	2.50 (1.82, 3.15)	0.3	0.52
BodyMed BDMOXMTRBLK	2	2.12 (1.92, 2.32)	3.00 (2.48, 3.54)	2.07 (1.81, 2.33)	1.68 (1.37, 2.02)	1.78 (1.50, 2.06)	0.13	0.09	3.16 (1.84, 4.47)	1.10 (0.43, 1.78)	1.17 (0.31, 1.94)	1.96 (1.17, 2.69)	0.54	0.86	3.88 (3.37, 4.41)	3.07 (2.67, 3.46)	2.35 (1.80, 2.89)	2.29 (1.82, 2.76)	0.07	**0.03**	2.04 (1.58, 2.46)	1.53 (1.25, 1.75)	1.50 (1.17, 1.80)	1.75 (1.53, 1.97)	0.2	**0.04**
Masimo Rad-97 with Adult Probe, Product Code 4050	60	2.49 (2.39, 2.59)	2.04 (1.77, 2.30)	2.07 (1.95, 2.18)	3.86 (3.67, 4.05)	4.24 (4.01, 4.48)	**0.001**	**0.001**	2.02 (1.64, 2.42)	2.70 (2.49, 2.91)	5.17 (4.89, 5.47)	5.64 (5.23, 6.03)	**<0.001**	**<0.001**	2.73 (2.28, 3.17)	2.35 (2.16, 2.55)	4.40 (4.12, 4.67)	4.92 (4.65, 5.20)	**0.002**	**0.001**	2.24 (1.77, 2.71)	2.10 (1.83, 2.33)	1.49 (1.17, 1.83)	1.28 (0.91, 1.68)	0.18	0.14
Guangdong Biolight M800	43	0.24 (0.04, 0.45)	-0.31 (-0.77, 0.10)	0.07 (-0.21, 0.36)	1.12 (0.69, 1.53)	1.08 (0.62, 1.51)	**0.04**	0.06	-3.55 (-4.48, -2.57)	-2.58 (-3.19, -1.93)	-0.93 (-1.86, 0.08)	-0.65 (-1.70, 0.23)	0.1	0.11	0.72 (0.34, 1.08)	1.04 (0.71, 1.42)	2.09 (1.63, 2.53)	1.99 (1.46, 2.50)	**0.01**	**0.04**	1.30 (0.79, 1.75)	1.21 (1.04, 1.38)	2.29 (2.08, 2.51)	2.51 (2.29, 2.74)	0.3	0.95
Masimo Rad-G with Adult Probe, Product Code 4050	55	3.48 (3.31, 3.63)	2.85 (2.56, 3.16)	3.24 (3.03, 3.45)	5.06 (4.73, 5.40)	5.47 (5.10, 5.84)	**0.008**	**0.002**	4.00 (3.49, 4.57)	4.31 (3.93, 4.68)	6.74 (6.33, 7.17)	7.01 (6.59, 7.47)	**0.001**	**<0.001**	2.80 (2.36, 3.23)	3.49 (3.13, 3.83)	5.36 (4.98, 5.78)	5.60 (5.23, 6.05)	**0.01**	**0.004**	1.71 (1.50, 1.93)	1.55 (1.34, 1.75)	2.05 (1.67, 2.41)	1.74 (1.22, 2.22)	0.31	0.23
Guangdong Biolight M70	37	-0.48 (-0.81, -0.14)	-0.44 (-1.06, 0.14)	-1.59 (-2.12, -1.06)	0.61 (0.05, 1.19)	0.27 (-0.42, 1.00)	0.48	0.64	-4.28 (-5.33, -3.25)	-6.00 (-7.03, -5.00)	-1.41 (-3.00, 0.23)	-2.22 (-4.20, 0.00)	0.26	0.34	0.27 (-0.21, 0.75)	-0.39 (-0.83, 0.05)	1.29 (0.62, 2.01)	1.04 (0.08, 2.22)	0.48	0.69	1.75 (1.39, 2.18)	1.92 (1.66, 2.19)	2.75 (2.33, 3.16)	3.19 (2.71, 3.63)	0.32	0.15
Acare (2023) AH-M1	57	-2.36 (-2.66, -2.09)	-1.81 (-2.43, -	-2.71 (-3.17, -	-2.06 (-2.57, -	-2.03 (-2.71, -	0.95	0.92	-6.28 (-7.40, -	-7.70 (-8.54, -	-6.39 (-7.35, -	-6.88 (-8.14, -	0.8	0.86	-1.14 (-1.83, -	-1.14 (-1.67, -	-1.19 (-1.69, -	-1.32 (-1.93, -	0.8	0.93	2.21 (1.93,	1.42 (1.21,	1.59 (1.20,	1.31 (0.82,	0.89	0.81

Device	Device ID	Overall Bias (95% CI)	70–100% SaO_2 Bias (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI)					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
			1.14)	2.28)	1.60)	1.43)			5.18)	6.96)	5.43)	5.78)			0.50)	0.63)	0.68)	0.64)			2.51)	1.63)	1.94)	1.80)		
Acare (2020) AH-MX	42	-3.98 (-4.31, -3.65)	-4.50 (-5.35, -3.75)	-3.62 (-4.03, -3.24)	-4.22 (-4.87, -3.57)	-4.25 (-5.02, -3.54)	0.67	0.77	-11.27 (-12.66, -9.92)	-9.18 (-9.84, -8.51)	-10.18 (-11.33, -8.98)	-10.00 (-11.14, -8.81)	0.55	0.56	-3.13 (-3.65, -2.58)	-2.58 (-2.91, -2.23)	-3.12 (-3.66, -2.55)	-3.25 (-3.96, -2.64)	0.92	0.66	0.74 (0.28, 1.18)	0.54 (0.33, 0.76)	0.73 (0.42, 1.05)	0.78 (0.39, 1.18)	0.21	0.13
American Oxygen LK-87	20	-5.71 (-6.11, -5.31)	-6.31 (-7.44, -5.25)	-5.78 (-6.30, -5.27)	-5.19 (-5.91, -4.39)	-5.22 (-6.00, -4.44)	0.23	0.29	-8.83 (-9.59, -8.09)	-7.76 (-8.82, -6.68)	-8.50 (-9.48, -7.42)	-9.17 (-9.90, -8.51)	0.8	0.81	-9.37 (-11.19, -7.60)	-8.15 (-8.83, -7.50)	-6.50 (-7.88, -5.08)	-5.97 (-7.43, -4.65)	0.15	0.08	0.37 (0.03, 0.74)	0.20 (0.02, 0.40)	-0.27 (-0.61, 0.05)	-0.31 (-0.68, -0.02)	0.78	0.78

This table reports the overall bias (mean of pulse oximeter measured oxygen saturation (SpO_2) - functional arterial oxygen saturation(SaO_2)) for each device in a cohort that is as large as possible while satisfying the enrollment criteria in the anticipated 2025 U.S. Food and Drug Administration (FDA) and 2025 International Organization for Standardization (ISO) guidelines. The bias is presented with 95% confidence interval (CI) overall, and separately for each SaO_2 decile and for each skin color bin. Skin color bins were defined using the Individual Typology Angle (ITA) measured at the dorsal distal phalanx (DP): ITA > 30° for light, -30° ≤ ITA ≤ 30° for medium, ITA < -30° for dark, and ITA < -50° for dark; note that “very dark” is a subset of the dark bin. P-values indicate whether the SpO_2 bias differs significantly between light versus dark bins and between light versus very dark bins using a linear mixed-effects model with SpO_2-SaO_2 error as the outcome, skin color bin as a fixed-effect predictor, and participant ID as a random-effect grouping factor. No correction for multiple pairwise comparisons (e.g., Bonferroni) was applied, as the 2025 draft FDA guidelines do not require adjustment. P-values near the 0.05 threshold are reported to three decimal places, and those meeting $p < 0.05$ are bolded. The 95% CI were determined using bootstrapping (random resampling with replacement) with 1,000 repetitions.

Table S11. Bias (mean of SpO_2 minus SaO_2) at the forehead for maximum participant cohorts

Device	Device ID	Overall Bias (95% CI)	70–100% SaO_2 Bias (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Nonin CO-Pilot	70	-0.28 (-0.34, -0.20)	-0.23 (-0.33, -0.13)	-0.34 (-0.46, -0.23)	-0.20 (-0.34, -0.07)	-0.11 (-0.27, 0.06)	0.86	0.63	-0.26 (-0.45, -0.05)	-0.11 (-0.32, 0.13)	-0.26 (-0.56, 0.06)	-0.07 (-0.48, 0.34)	0.85	0.45	-0.22 (-0.40, -0.05)	-0.42 (-0.64, -0.14)	-0.16 (-0.43, 0.12)	-0.07 (-0.40, 0.26)	0.79	0.59	-0.20 (-0.34, -0.06)	-0.48 (-0.65, -0.33)	-0.20 (-0.39, -0.03)	-0.19 (-0.40, 0.03)	0.95	0.86
Shenzhen (2023) PC-60NW	21	-0.04 (-0.15, 0.08)	-0.04 (-0.24, 0.17)	0.12 (-0.07, 0.30)	-0.15 (-0.36, 0.04)	-0.24 (-0.50, 0.02)	0.63	0.54	-1.09 (-1.44, -0.75)	-0.57 (-0.93, -0.24)	-1.12 (-1.56, -0.66)	-1.33 (-2.02, -0.69)	0.54	0.32	0.10 (-0.21, 0.42)	0.04 (-0.26, 0.34)	0.15 (-0.13, 0.44)	0.07 (-0.30, 0.51)	0.99	0.83	0.68 (0.35, 0.99)	0.74 (0.49, 0.98)	0.38 (0.16, 0.60)	0.38 (0.12, 0.64)	0.44	0.51
Masimo Rad-97 with RD Rainbow Adt 8λ Adult Adhesive Sensor	73.09	0.86 (0.74, 0.97)	0.47 (0.31, 0.63)	0.67 (0.52, 0.83)	1.72 (1.43, 2.01)	1.11 (0.68, 1.56)	**0.03**	0.29	0.04 (-0.30, 0.42)	0.74 (0.39, 1.06)	1.55 (0.96, 2.15)	0.52 (-0.35, 1.45)	**0.03**	0.26	0.70 (0.52, 0.87)	0.83 (0.60, 1.07)	2.15 (1.67, 2.62)	1.65 (0.91, 2.44)	**0.01**	0.21	0.67 (0.45, 0.89)	0.45 (0.27, 0.65)	1.41 (1.06, 1.75)	0.90 (0.39, 1.38)	0.13	0.7
Shenzhen (2024) PC-60NW	81	-0.02 (-0.17, 0.12)	-0.11 (-0.34, 0.11)	0.25 (0.06, 0.42)	-0.31 (-0.71, 0.07)	-0.30 (-0.76, 0.11)	0.6	0.56	-1.25 (-1.66, -0.89)	-0.42 (-0.74, -0.09)	-2.53 (-3.36, -1.68)	-2.61 (-3.35, -1.96)	0.15	**0.049**	0.04 (-0.25, 0.32)	0.49 (0.16, 0.84)	-0.00 (-0.48, 0.42)	0.05 (-0.35, 0.46)	0.91	0.98	0.98 (0.73, 1.24)	0.67 (0.48, 0.85)	1.34 (1.10, 1.59)	1.36 (1.05, 1.68)	0.41	0.52
Zug Medical Systems SAS SPLF-SE with SpotLife	75	1.08 (0.97, 1.19)	0.54 (0.38, 0.71)	1.33 (1.18, 1.49)	1.26 (1.01, 1.49)	1.16 (0.88, 1.41)	0.06	0.15	-0.41 (-0.76, -0.02)	1.40 (1.02, 1.79)	0.04 (-0.49, 0.55)	0.56 (-0.07, 1.14)	0.81	0.43	0.77 (0.52, 1.01)	1.39 (1.18, 1.62)	1.66 (1.32, 2.00)	1.39 (0.98, 1.76)	**0.049**	0.23	1.19 (1.00, 1.37)	1.22 (1.03, 1.43)	1.82 (1.58, 2.07)	1.37 (1.09, 1.64)	0.07	0.62

Device	Device ID	Overall Bias (95% CI)	70–100% SaO_2 Bias (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Reusable Probe																										
Compass Health POX-ROS	12	-0.30 (-0.44, -0.17)	0.41 (0.08, 0.75)	-0.50 (-0.71, -0.28)	-0.61 (-0.84, -0.39)	-0.71 (-0.97, -0.47)	0.07	0.07	0.14 (-0.60, 0.97)	-0.39 (-0.71, -0.07)	-0.93 (-1.38, -0.54)	-1.09 (-1.65, -0.63)	0.12	0.1	0.05 (-0.48, 0.63)	-0.72 (-1.16, -0.30)	-0.54 (-0.89, -0.15)	-0.72 (-1.14, -0.32)	0.21	0.2	0.98 (0.62, 1.34)	-0.39 (-0.68, -0.09)	-0.37 (-0.77, 0.05)	-0.36 (-0.78, 0.06)	0.05	0.047
Beijing Choice MD300M	40	0.92 (0.79, 1.05)	0.98 (0.78, 1.15)	0.55 (0.38, 0.75)	1.46 (1.14, 1.79)	1.39 (0.86, 1.88)	0.32	0.53	1.02 (0.62, 1.41)	1.41 (0.98, 1.82)	2.34 (1.70, 3.01)	2.59 (1.74, 3.56)	0.07	0.08	1.31 (1.04, 1.57)	0.36 (0.15, 0.58)	1.47 (0.93, 1.98)	1.17 (0.50, 1.84)	0.72	0.9	0.66 (0.34, 0.96)	0.00 (-0.21, 0.21)	0.57 (0.08, 1.01)	0.29 (-0.45, 0.89)	0.89	0.72
Nonin PalmSAT 2500	36	-1.35 (-1.47, -1.24)	-1.37 (-1.57, -1.15)	-1.35 (-1.50, -1.20)	-1.33 (-1.63, -1.05)	-1.26 (-1.55, -0.99)	0.99	1	-2.34 (-2.71, -2.02)	-1.49 (-1.78, -1.21)	-2.50 (-3.11, -1.92)	-1.92 (-2.55, -1.35)	0.95	0.71	-1.78 (-2.07, -1.47)	-1.76 (-1.95, -1.56)	-0.99 (-1.43, -0.52)	-1.27 (-1.67, -0.87)	0.45	0.64	-0.28 (-0.58, -0.01)	-0.85 (-1.10, -0.62)	-0.45 (-0.79, -0.13)	-0.57 (-0.91, -0.24)	0.69	0.65
Covidien PM1000N	59	1.29 (1.22, 1.37)	0.64 (0.52, 0.76)	1.04 (0.94, 1.14)	2.32 (2.18, 2.46)	2.45 (2.26, 2.65)	**<0.001**	**<0.001**	0.04 (-0.19, 0.27)	1.19 (1.02, 1.36)	3.48 (3.20, 3.73)	3.70 (3.36, 4.03)	**<0.001**	**<0.001**	1.17 (1.00, 1.34)	1.43 (1.25, 1.60)	2.56 (2.39, 2.77)	2.70 (2.43, 2.99)	**<0.001**	**<0.001**	0.69 (0.53, 0.86)	0.54 (0.39, 0.68)	1.17 (1.01, 1.34)	1.16 (0.94, 1.36)	0.09	0.14
Shenzhen AM801	10	0.38 (0.21, 0.55)	0.12 (-0.22, 0.44)	0.15 (-0.09, 0.40)	1.07 (0.77, 1.40)	1.16 (0.75, 1.59)	0.17	0.25	-0.63 (-1.07, -0.20)	-0.53 (-1.18, 0.12)	1.78 (1.09, 2.56)	2.19 (1.27, 3.21)	**0.003**	**0.002**	0.42 (-0.34, 1.15)	0.50 (0.15, 0.84)	1.13 (0.76, 1.48)	0.92 (0.48, 1.32)	0.52	0.78	0.60 (-0.05, 1.17)	0.37 (0.11, 0.60)	0.30 (0.07, 0.58)	0.25 (0.03, 0.47)	0.66	0.69
Nonin Onyx Vantage 9590	35	1.36 (1.24, 1.47)	1.20 (0.97, 1.40)	1.31 (1.16, 1.47)	1.57 (1.31, 1.84)	1.78 (1.41, 2.16)	0.45	0.29	1.42 (0.90, 1.94)	1.98 (1.67, 2.26)	2.64 (2.06, 3.26)	3.14 (2.41, 3.84)	0.34	0.23	0.95 (0.71, 1.20)	0.84 (0.54, 1.14)	1.14 (0.77, 1.52)	1.24 (0.76, 1.72)	0.64	0.49	1.19 (0.87, 1.49)	1.03 (0.81, 1.26)	0.93 (0.65, 1.22)	0.83 (0.41, 1.30)	0.64	0.71

Device	Device ID	Overall Bias (95% CI)	70–100% SaO_2 Bias (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Choicemed MD300CN350 R	14	1.02 (0.89, 1.15)	1.38 (1.16, 1.59)	1.14 (0.93, 1.35)	0.51 (0.29, 0.75)	0.53 (0.26, 0.79)	**0.03**	0.06	1.12 (0.70, 1.52)	1.05 (0.58, 1.49)	-0.33 (-0.88, 0.21)	-0.65 (-1.25, -0.01)	**0.03**	**0.008**	1.90 (1.58, 2.21)	1.60 (1.33, 1.87)	1.14 (0.76, 1.50)	1.22 (0.84, 1.58)	0.1	0.19	1.04 (0.66, 1.38)	0.80 (0.57, 1.05)	0.64 (0.39, 0.93)	0.89 (0.60, 1.17)	0.34	0.67
Masimo (2024) Mightysat	78	1.71 (1.58, 1.84)	1.37 (1.15, 1.58)	1.54 (1.36, 1.73)	2.42 (2.16, 2.66)	2.82 (2.52, 3.13)	0.06	**0.03**	1.34 (0.90, 1.74)	1.60 (1.15, 2.05)	2.61 (2.02, 3.24)	3.12 (2.36, 3.89)	0.09	**0.03**	1.38 (1.04, 1.71)	1.85 (1.56, 2.19)	2.74 (2.32, 3.19)	3.34 (2.75, 3.95)	**0.03**	**0.008**	1.38 (1.13, 1.60)	1.20 (1.02, 1.40)	2.00 (1.76, 2.25)	2.25 (1.96, 2.55)	0.11	0.054
Acare (2023) Model AH-TA	56	1.54 (1.42, 1.64)	1.56 (1.37, 1.76)	1.78 (1.64, 1.94)	1.19 (0.91, 1.43)	0.53 (0.17, 0.88)	0.44	**0.047**	0.91 (0.46, 1.38)	1.71 (1.37, 2.02)	0.88 (0.33, 1.46)	-0.32 (-1.04, 0.34)	0.91	0.16	2.23 (1.95, 2.52)	2.47 (2.26, 2.68)	1.71 (1.25, 2.15)	1.19 (0.55, 1.81)	0.25	0.08	1.60 (1.40, 1.79)	1.24 (1.08, 1.39)	0.99 (0.66, 1.30)	0.57 (0.09, 1.04)	0.17	**0.04**
Philips M1020B IntelliVue with M1191BL Probe	71	1.90 (1.81, 1.99)	1.55 (1.44, 1.66)	1.78 (1.67, 1.92)	2.72 (2.47, 2.98)	3.87 (3.55, 4.21)	**0.005**	**<0.001**	2.01 (1.84, 2.18)	2.45 (2.19, 2.71)	3.78 (3.24, 4.25)	5.07 (4.60, 5.59)	**0.003**	**<0.001**	1.36 (1.20, 1.51)	1.52 (1.38, 1.68)	2.50 (2.13, 2.82)	3.54 (3.20, 3.93)	**0.002**	**<0.001**	1.26 (1.02, 1.51)	1.35 (1.17, 1.55)	1.86 (1.52, 2.17)	2.76 (2.31, 3.24)	0.18	**0.01**
CMS Zacurate 500C	15	0.78 (0.58, 0.96)	0.71 (0.26, 1.10)	1.08 (0.80, 1.40)	0.54 (0.23, 0.83)	0.58 (0.19, 0.94)	0.63	0.72	-0.26 (-1.38, 0.78)	0.77 (0.13, 1.40)	-0.32 (-1.22, 0.59)	-0.03 (-1.26, 1.24)	0.43	0.48	1.20 (0.77, 1.65)	1.78 (1.30, 2.26)	1.13 (0.82, 1.46)	1.04 (0.73, 1.34)	0.96	0.92	1.18 (0.90, 1.47)	0.76 (0.54, 0.97)	0.63 (0.36, 0.90)	0.64 (0.33, 0.96)	0.19	0.26
Covidien (2022) Nellcor PM10N	41	1.17 (0.99, 1.34)	0.71 (0.34, 1.06)	0.84 (0.71, 0.98)	2.33 (2.05, 2.63)	2.39 (1.97, 2.83)	0.054	0.14	-0.56 (-1.38, 0.36)	0.79 (0.48, 1.12)	3.17 (2.66, 3.58)	3.78 (3.23, 4.29)	**0.002**	**0.01**	1.40 (0.96, 1.86)	1.14 (0.92, 1.35)	2.58 (2.20, 3.00)	2.60 (2.06, 3.21)	0.12	0.27	1.22 (0.86, 1.59)	0.61 (0.43, 0.80)	1.09 (0.61, 1.55)	0.32 (-0.13, 0.80)	0.95	0.6
Utech UT100	39	1.02	1.00	0.49	2.12	1.90	0.13	0.24	0.28 (-	0.04 (-	3.38	2.75	0.06	0.09	1.21	0.83	1.86	2.02	0.32	0.28	1.35	0.55	1.21	1.05	0.66	0.52

Device	Device ID	Overall Bias (95% CI)	70–100% SaO_2 Bias (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
		(0.85, 1.19)	(0.69, 1.28)	(0.29, 0.68)	(1.61, 2.57)	(1.30, 2.39)			0.45, 1.06)	0.38, 0.44)	(2.48, 4.43)	(2.10, 3.40)			(0.80, 1.66)	(0.47, 1.18)	(0.83, 2.69)	(0.58, 3.12)			(1.06, 1.64)	(0.30, 0.81)	(0.78, 1.66)	(0.49, 1.59)		
Nihon Kohden OLV-4202 with NK-SPO2-ADVT Probe	64.03	0.36 (0.18, 0.53)	-0.27 (-0.58, -0.01)	-0.03 (-0.22, 0.18)	0.83 (0.16, 1.50)	-0.84 (-2.01, 0.26)	0.38	0.73	-1.26 (-1.73, -0.76)	0.08 (-0.36, 0.53)	0.83 (-0.89, 2.37)	-1.34 (-4.61, 1.33)	0.22	0.89	-0.15 (-0.53, 0.19)	-0.23 (-0.61, 0.15)	0.85 (-0.23, 1.84)	-0.90 (-2.57, 0.70)	0.38	0.69	0.63 (0.11, 1.10)	0.08 (-0.16, 0.30)	0.82 (0.13, 1.47)	-0.34 (-1.45, 0.72)	0.9	0.45
EDAN (2020) H100b	44	2.34 (2.19, 2.51)	2.26 (2.01, 2.51)	1.80 (1.62, 1.99)	2.48 (2.21, 2.79)	2.78 (2.41, 3.20)	0.67	0.45	2.40 (1.82, 2.99)	2.06 (1.65, 2.48)	2.88 (2.35, 3.39)	3.03 (2.17, 3.84)	0.5	0.57	2.65 (2.31, 3.01)	2.11 (1.81, 2.41)	2.77 (2.20, 3.35)	3.41 (2.88, 4.01)	0.95	0.35	1.77 (1.51, 2.07)	1.29 (1.08, 1.51)	1.89 (1.58, 2.22)	1.91 (1.47, 2.36)	0.85	0.84
Shenzhen WIT-S300C	62	1.86 (1.73, 2.01)	1.56 (1.36, 1.73)	1.79 (1.54, 2.05)	2.23 (2.04, 2.42)	1.82 (1.62, 2.03)	0.06	0.5	1.48 (1.14, 1.83)	1.89 (1.39, 2.40)	2.85 (2.50, 3.21)	2.52 (2.09, 2.93)	**0.002**	0.08	1.65 (1.34, 1.96)	2.17 (1.82, 2.55)	2.44 (2.14, 2.75)	1.82 (1.52, 2.10)	0.11	0.73	1.53 (1.23, 1.86)	1.38 (1.01, 1.68)	1.59 (1.31, 1.89)	1.23 (0.95, 1.50)	0.96	0.57
CONTEC CMS50M	17	-1.24 (-1.44, -1.04)	-1.55 (-1.85, -1.22)	-1.38 (-1.74, -1.01)	-0.58 (-0.88, -0.28)	-0.92 (-1.36, -0.52)	0.14	0.47	-3.24 (-3.85, -2.58)	-2.60 (-3.28, -1.83)	-1.86 (-2.63, -1.11)	-2.27 (-3.08, -1.43)	0.44	0.55	-1.75 (-2.07, -1.42)	-1.11 (-1.56, -0.63)	-0.70 (-1.23, -0.21)	-1.26 (-1.95, -0.66)	0.24	0.8	-0.29 (-0.70, 0.11)	-0.47 (-1.13, 0.00)	0.39 (0.11, 0.68)	0.37 (-0.07, 0.82)	0.2	0.35
Shenzhen A310L	11	1.43 (1.25, 1.61)	0.46 (0.12, 0.78)	1.62 (1.36, 1.88)	2.15 (1.89, 2.41)	2.11 (1.75, 2.50)	**0.008**	**0.04**	-1.06 (-1.88, -0.28)	0.65 (-0.06, 1.31)	2.82 (2.11, 3.50)	3.11 (2.21, 4.08)	**0.002**	**0.01**	1.36 (0.99, 1.73)	2.47 (2.12, 2.80)	2.42 (2.05, 2.78)	2.34 (1.90, 2.78)	0.08	0.23	1.05 (0.83, 1.26)	1.61 (1.40, 1.82)	1.24 (0.93, 1.55)	1.01 (0.65, 1.38)	0.69	0.86
Masimo (2020) Mightysat Rx	7	0.92 (0.77, 1.08)	0.13 (-0.13, 0.38)	0.76 (0.55, 0.96)	2.13 (1.80, 2.48)	1.77 (1.53, 1.97)	**0.004**	**0.003**	-1.10 (-1.63, -0.55)	0.36 (0.07, 0.64)	2.16 (1.31, 3.16)	1.27 (0.74, 1.82)	**0.03**	**0.02**	0.55 (0.20, 0.89)	1.30 (0.98, 1.67)	2.50 (2.08, 2.93)	2.38 (2.12, 2.63)	**0.03**	0.054	0.79 (0.48, 1.11)	0.61 (0.19, 0.96)	1.75 (1.50, 2.01)	1.59 (1.28, 1.90)	0.07	0.21

Device	Device ID	Overall Bias (95% CI)	70–100% SaO_2 Bias (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Masimo Rad-G with Rad-G Finger Clip Sensor	55.03	1.94 (1.77, 2.10)	1.44 (1.16, 1.73)	2.05 (1.83, 2.28)	2.57 (2.13, 2.99)	2.64 (2.01, 3.25)	0.21	0.33	0.19 (-0.35, 0.74)	1.94 (1.49, 2.42)	1.61 (0.54, 2.77)	2.50 (0.72, 4.26)	0.58	0.63	2.04 (1.69, 2.41)	2.40 (2.07, 2.75)	3.22 (2.76, 3.73)	3.04 (2.36, 3.71)	0.09	0.28	2.29 (1.97, 2.63)	1.79 (1.49, 2.10)	2.72 (2.31, 3.12)	2.24 (1.75, 2.79)	0.34	0.87
Bistos Co., Ltd. BT-710	61	2.46 (2.33, 2.50)]	2.43 (2.26, 2.60)	2.28 (2.03, 2.52)	2.70 (2.44, 2.97)	2.67 (2.30, 3.11)	0.45	0.59	2.83 (2.44, 3.19)	2.86 (2.46, 3.29)	4.14 (3.58, 4.72)	4.56 (3.62, 5.62)	**0.047**	0.05	2.69 (2.45, 2.91)	2.74 (2.45, 3.04)	2.88 (2.57, 3.17)	2.84 (2.42, 3.23)	0.53	0.71	1.93 (1.72, 2.16)	1.38 (0.97, 1.72)	1.36 (1.04, 1.64)	1.12 (0.65, 1.47)	0.06	**0.03**
BodyMed BDMOXMTRBLK	2	2.12 (1.92, 2.32)	2.75 (2.39, 3.13)	2.05 (1.73, 2.33)	1.69 (1.35, 2.04)	1.36 (0.88, 1.84)	0.17	0.14	2.67 (1.56, 3.87)	1.27 (0.58, 1.93)	1.05 (0.15, 1.88)	0.16 (-0.97, 1.15)	0.62	0.27	3.68 (3.25, 4.11)	2.81 (2.36, 3.21)	2.53 (1.95, 3.12)	2.45 (1.60, 3.19)	0.13	0.23	2.03 (1.63, 2.45)	2.25 (1.98, 2.51)	1.44 (1.13, 1.78)	1.33 (0.89, 1.78)	0.26	0.23
Masimo Rad-97 with Adult Probe, Product Code 4050	60	2.49 (2.39, 2.59)	1.74 (1.60, 1.87)	2.06 (1.91, 2.21)	3.87 (3.58, 4.14)	3.94 (3.51, 4.31)	**<0.001**	**<0.001**	1.76 (1.52, 1.98)	3.06 (2.81, 3.31)	5.36 (5.00, 5.74)	5.54 (5.08, 6.00)	**<0.001**	**<0.001**	2.26 (2.00, 2.48)	2.26 (2.03, 2.49)	4.67 (4.32, 5.02)	4.93 (4.53, 5.34)	**<0.001**	**<0.001**	1.26 (1.02, 1.47)	0.94 (0.73, 1.14)	1.83 (1.39, 2.19)	1.37 (0.58, 2.16)	0.11	0.31
Guangdong Biolight M800	43	0.24 (0.04, 0.45)	-0.44 (-0.84, -0.10)	0.25 (-0.07, 0.57)	1.19 (0.77, 1.58)	2.48 (2.10, 2.88)	**0.02**	**<0.001**	-3.57 (-4.32, -2.81)	-2.41 (-3.03, -1.74)	-0.65 (-1.71, 0.48)	2.48 (1.69, 3.53)	0.07	**<0.001**	0.52 (0.15, 0.87)	1.33 (0.94, 1.71)	2.02 (1.55, 2.50)	2.96 (2.46, 3.47)	**0.01**	**<0.001**	1.49 (1.27, 1.69)	1.77 (1.52, 2.00)	1.95 (1.61, 2.28)	1.94 (1.46, 2.42)	0.18	0.2
Masimo Rad-G with Adult Probe, Product Code 4050	55	3.48 (3.31, 3.63)	2.92 (2.73, 3.12)	3.21 (2.96, 3.47)	4.83 (4.49, 5.22)	5.09 (4.71, 5.50)	**0.002**	**0.002**	3.83 (3.45, 4.22)	4.39 (3.96, 4.83)	6.53 (6.03, 7.11)	6.78 (6.26, 7.32)	**0.001**	**0.001**	3.17 (2.82, 3.55)	3.24 (2.88, 3.66)	5.09 (4.66, 5.53)	5.37 (4.94, 5.82)	**0.02**	**0.02**	1.90 (1.66, 2.15)	1.77 (1.45, 2.10)	2.66 (2.18, 3.08)	2.81 (2.38, 3.30)	0.1	0.06

Device	Device ID	Overall Bias (95% CI)	70–100% SaO_2 Bias (95% CI)						70–80% SaO_2 Bias (95% CI)						80–90% SaO_2 Bias (95% CI)						90–100% SaO_2 Bias (95% CI					
			Light	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**	**P-value: Light vs Dark**	**P-value: Light vs Very Dark**
Guangdong Biolight M70	37	-0.48 (-0.81, -0.14)	-0.60 (-1.07, -0.15)	-1.59 (-2.15, -1.04)	0.20 (-0.45, 0.84)	0.68 (-0.03, 1.37)	0.5	0.28	-4.54 (-5.36, -3.74)	-6.04 (-7.28, -4.86)	-2.07 (-3.90, -0.21)	-0.42 (-2.55, 1.67)	0.29	0.09	0.09 (-0.33, 0.49)	-0.00 (-0.43, 0.45)	0.75 (-0.08, 1.59)	1.14 (0.25, 2.05)	0.58	0.36	1.95 (1.71, 2.20)	1.43 (1.24, 1.63)	1.44 (1.05, 1.85)	1.13 (0.63, 1.61)	0.32	0.17
Acare (2023) AH-M1	57	-2.36 (-2.66, -2.09)	-2.47 (-2.92, -2.04)	-2.33 (-2.85, -1.80)	-2.25 (-2.91, -1.67)	-1.72 (-2.42, -1.03)	0.69	0.39	-7.04 (-7.81, -6.28)	-7.37 (-8.24, -6.49)	-6.71 (-8.10, -5.40)	-5.82 (-7.49, -4.09)	0.71	0.41	-1.68 (-2.15, -1.17)	-0.34 (-0.84, 0.22)	-1.62 (-2.20, -1.03)	-0.91 (-1.56, -0.23)	0.83	0.37	0.37 (0.09, 0.67)	0.86 (0.63, 1.10)	0.66 (0.31, 0.98)	0.39 (-0.13, 0.93)	0.53	0.81
Acare (2020) AH-MX	42	-3.98 (-4.31, -3.65)	-3.49 (-4.11, -2.88)	-4.09 (-4.53, -3.64)	-4.25 (-4.90, -3.68)	-3.64 (-4.38, -2.92)	0.4	0.88	-9.70 (-10.90, -8.51)	-9.98 (-10.66, -9.34)	-9.73 (-10.98, -8.42)	-7.70 (-9.60, -5.66)	1	0.44	-2.75 (-3.24, -2.25)	-2.57 (-2.96, -2.21)	-3.52 (-4.06, -2.99)	-3.55 (-4.36, -2.83)	0.21	0.3	0.40 (0.08, 0.71)	0.19 (-0.02, 0.41)	-0.34 (-0.66, -0.08)	-0.81 (-1.38, -0.37)	0.06	**0.03**
American Oxygen LK-87	20	-5.71 (-6.11, -5.31)	-5.05 (-5.93, -4.17)	-6.26 (-6.79, -5.79)	-5.29 (-6.06, -4.53)	-5.18 (-6.10, -4.25)	0.77	0.89	-5.62 (-7.67, -3.42)	-8.84 (-9.64, -8.00)	-8.84 (-9.53, -8.14)	-8.62 (-9.30, -7.92)	0.15	0.29	-8.33 (-9.46, -7.18)	-8.41 (-9.17, -7.72)	-6.54 (-7.99, -5.03)	-6.17 (-7.87, -4.54)	0.32	0.31	-1.73 (-2.50, -0.98)	-1.71 (-2.10, -1.33)	-0.91 (-1.62, -0.33)	-0.73 (-1.74, 0.12)	0.2	0.21

This table reports the overall bias (mean of pulse oximeter measured oxygen saturation (SpO_2) - functional arterial oxygen saturation(SaO_2)) for each device in a cohort that is as large as possible while satisfying the enrollment criteria in the anticipated 2025 U.S. Food and Drug Administration (FDA) and 2025 International Organization for Standardization (ISO) guidelines. The bias is presented with 95% confidence interval (CI) overall, and separately for each SaO_2 decile and for each skin color bin. Skin color bins were defined using the Individual Typology Angle (ITA) measured at the dorsal distal phalanx (DP): ITA > 30° for light, -30° ≤ ITA ≤ 30° for medium, ITA < -30° for dark, and ITA < -50° for dark; note that “very dark” is a subset of the dark bin. P-values indicate whether the SpO_2 bias differs significantly between light versus dark bins and between light versus very dark bins using a linear mixed-effects model with SpO_2-SaO_2 error as the outcome, skin color bin as a fixed-effect predictor, and participant ID as a random-effect grouping factor. No correction for multiple pairwise comparisons (e.g., Bonferroni) was applied, as the 2025 draft FDA

guidelines do not require adjustment. P-values near the 0.05 threshold are reported to three decimal places, and those meeting $p < 0.05$ are bolded. The 95% CI were determined using bootstrapping (random resampling with replacement) with 1,000 repetitions.

Table S12. Differential bias for 24 participant cohorts

Device	Device ID	Pass FDA 2025 Differential Bias Thresholds	Pass Anticipated ISO Differential Bias Thresholds	DP ITA-Derived Differential Bias (70–85% SaO_2)	MST 1–4 vs 5–7 Differential Bias (70–85% SaO_2)	MST 5–7 vs 8–10 Differential Bias (70–85% SaO_2)	MST 1–4 vs 8–10 Differential Bias (70–85% SaO_2)	DP ITA-Derived Differential Bias (85–100% SaO_2)	MST 1–4 vs 5–7 Differential Bias (85–100% SaO_2)	MST 5–7 vs 8–10 Differential Bias (85–100% SaO_2)	MST 1–4 vs 8–10 Differential Bias (85–100% SaO_2)
Nonin CO-Pilot	70		✓	-0.41 (-2.07, 1.26)	-0.39 (-2.20, 1.42)	0.57 (-1.02, 2.16)	0.18 (-1.55, 1.91)	-0.32 (-1.47, 0.83)	-0.70 **(-1.94, 0.53)**	0.28 (-0.81, 1.36)	-0.43 **(-1.61, 0.75)**
Shenzhen (2023) PC-60NW	21		✓	0.19 (-0.97, 1.35)	-0.03 (-1.20, 1.13)	-0.02 (-1.26, 1.22)	-0.06 (-1.27, 1.15)	0.29 (-0.75, 1.33)	-0.18 (-1.19, 0.83)	0.66 **(-0.42, 1.73)**	0.48 **(-0.57, 1.52)**
Masimo Rad-97 with RD Rainbow Adt 8λ Adult Adhesive Sensor	73.09		✓	1.87 (0.50, 3.24)	2.21 (0.93, 3.50)	-0.67 (-2.24, 0.90)	1.54 (0.18, 2.91)	1.13 **(0.18, 2.08)**	1.28 **(0.33, 2.22)**	-0.48 **(-1.64, 0.69)**	0.80 **(-0.21, 1.81)**
Shenzhen (2024) PC-60NW	81		✓	-0.37 (-2.28, 1.53)	1.02 (-0.44, 2.47)	-1.98 **(-3.65, -0.30)**	-0.96 (-2.49, 0.57)	0.06 (-0.88, 1.00)	-0.19 (-0.99, 0.62)	0.18 (-0.75, 1.10)	-0.01 (-0.85, 0.83)
Zug Medical Systems SAS SPLF-SE with SpotLife Reusable Probe	75	✓	✓	1.75 (0.37, 3.13)	1.54 (0.21, 2.87)	-0.24 (-1.69, 1.21)	1.30 (-0.04, 2.64)	0.46 (-0.53, 1.44)	0.50 (-0.47, 1.47)	-0.04 (-1.09, 1.01)	0.46 (-0.50, 1.43)
Beijing Choice MD300M	40		✓	0.95 (-0.60, 2.50)	0.91 (-0.49, 2.32)	0.11 (-1.39, 1.61)	1.02 (-0.45, 2.48)	0.78 **(-0.57, 2.14)**	-0.40 **(-1.66, 0.87)**	0.68 **(-0.67, 2.03)**	0.28 **(-1.03, 1.60)**
Compass Health POX-ROS	12		✓	-0.98 (-2.81, 0.86)	-0.75 (-2.67, 1.17)	0.02 (-1.69, 1.73)	-0.73 (-2.84, 1.38)	-0.43 **(-1.99, 1.13)**	-1.10 **(-2.65, 0.45)**	0.29 **(-1.09, 1.68)**	-0.81 **(-2.52, 0.90)**
Shenzhen AM801	10		✓	2.14 **(0.61, 3.68)**	-0.86 (-2.44, 0.73)	1.81 (0.22, 3.40)	0.95 (-0.78, 2.68)	-0.02 **(-1.54, 1.49)**	-0.19 **(-1.71, 1.33)**	0.39 **(-1.13, 1.90)**	0.20 **(-1.45, 1.85)**
CMS Zacurate 500C	15		✓	-0.86 (-2.97, 1.26)	-0.6 (-2.80, 1.59)	-0.66 (-2.78, 1.45)	-1.27 **(-3.52, 0.99)**	-0.29 (-1.39, 0.82)	-0.66 **(-1.79, 0.48)**	0.26 (-0.84, 1.36)	-0.39 **(-1.57, 0.78)**

Device	Device ID	Pass FDA 2025 Differential Bias Thresholds	Pass Anticipated ISO Differential Bias Thresholds	DP ITA-Derived Differential Bias (70–85% SaO_2)	MST 1–4 vs 5–7 Differential Bias (70–85% SaO_2)	MST 5–7 vs 8–10 Differential Bias (70–85% SaO_2)	MST 1–4 vs 8–10 Differential Bias (70–85% SaO_2)	DP ITA-Derived Differential Bias (85–100% SaO_2)	MST 1–4 vs 5–7 Differential Bias (85–100% SaO_2)	MST 5–7 vs 8–10 Differential Bias (85–100% SaO_2)	MST 1–4 vs 8–10 Differential Bias (85–100% SaO_2)
Acare (2023) Model AH-TA	56		✓	-0.03 (-1.66, 1.59)	0.29 (-1.25, 1.84)	-0.44 (-2.21, 1.33)	-0.15 (-1.85, 1.55)	-0.66 **(-2.06, 0.73)**	-0.09 (-1.33, 1.14)	-1.27 **(-2.68, 0.14)**	-1.37 **(-2.72, -0.02)**
Choicemed MD300CN350R	14		✓	-1.39 (-2.96, 0.17)	-0.56 (-2.04, 0.93)	-1.05 (-2.71, 0.60)	-1.61 (-3.27, 0.06)	-1.12 **(-2.57, 0.33)**	0.28 **(-1.20, 1.75)**	-0.47 **(-2.12, 1.18)**	-0.19 **(-1.84, 1.46)**
Nonin Onyx Vantage 9590	35		✓	0.27 (-1.43, 1.97)	1.07 (-0.53, 2.67)	-0.50 (-2.24, 1.24)	0.57 (-1.02, 2.17)	-0.59 **(-1.88, 0.70)**	0.63 **(-0.63, 1.89)**	-0.49 **(-1.85, 0.88)**	0.14 (-1.12, 1.40)
Nonin PalmSAT 2500	36		✓	-0.97 (-2.89, 0.95)	**-2.12 (-3.96, -0.28)**	1.15 (-0.88, 3.17)	-0.98 (-2.62, 0.67)	0.1 **(-1.36, 1.57)**	-1.82 **(-3.13, -0.52)**	**1.55 (0.11, 2.99)**	-0.28 (-1.45, 0.90)
Shenzhen WIT-S300C	62		✓	0.76 (-0.80, 2.33)	-0.63 (-2.76, 1.51)	1.16 (-0.79, 3.11)	0.53 (-1.07, 2.14)	-0.47 **(-1.53, 0.58)**	**-1.64 (-2.94, -0.33)**	1.01 **(-0.18, 2.20)**	-0.63 **(-1.60, 0.35)**
Masimo (2024) Mightysat	78			**4.63 (2.91, 6.35)**	1.58 (-0.12, 3.28)	2.38 **(0.62, 4.14)**	**3.96 (2.15, 5.77)**	**2.18 (1.04, 3.32)**	0.61 **(-0.52, 1.73)**	1.15 **(-0.01, 2.32)**	1.76 **(0.56, 2.96)**
Philips M1020B IntelliVue with M1191BL Probe	71		✓	1.88 (0.56, 3.21)	0.00 (-1.47, 1.47)	1.33 (-0.14, 2.81)	1.33 (-0.27, 2.92)	1.69 **(0.66, 2.72)**	0.44 **(-0.69, 1.57)**	1.10 **(-0.04, 2.23)**	**1.54 (0.31, 2.76)**
Nihon Kohden OLV-4202 with NK-SPO2-ADVT Probe	64.03		✓	2.28 **(-1.12, 5.67)**	1.22 **(-1.81, 4.24)**	1.52 **(-1.77, 4.82)**	2.74 **(-0.70, 6.19)**	0.75 **(-1.13, 2.63)**	-0.09 **(-1.77, 1.59)**	1.01 **(-0.83, 2.84)**	0.92 **(-1.00, 2.84)**
Covidien PM1000N	59		✓	2.91 **(0.77, 5.04)**	0.18 (-2.36, 2.72)	2.17 **(-0.62, 4.96)**	2.35 **(0.27, 4.44)**	0.9 **(-0.22, 2.03)**	-0.67 **(-1.92, 0.58)**	1.38 **(0.00, 2.75)**	0.70 **(-0.33, 1.73)**
Shenzhen A310L	11		✓	2.00 **(-0.73, 4.74)**	-0.29 (-2.79, 2.21)	2.23 **(-0.39, 4.85)**	**1.95 (-0.93, 4.83)**	-0.63 (-1.41, 0.15)	-0.22 (-0.98, 0.53)	-0.15 (-0.94, 0.65)	-0.37 (-1.24, 0.50)

Device	Device ID	Pass FDA 2025 Differential Bias Thresholds	Pass Anticipated ISO Differential Bias Thresholds	DP ITA-Derived Differential Bias (70–85% SaO_2)	MST 1–4 vs 5–7 Differential Bias (70–85% SaO_2)	MST 5–7 vs 8–10 Differential Bias (70–85% SaO_2)	MST 1–4 vs 8–10 Differential Bias (70–85% SaO_2)	DP ITA-Derived Differential Bias (85–100% SaO_2)	MST 1–4 vs 5–7 Differential Bias (85–100% SaO_2)	MST 5–7 vs 8–10 Differential Bias (85–100% SaO_2)	MST 1–4 vs 8–10 Differential Bias (85–100% SaO_2)
Masimo Rad-G with Rad-G Finger Clip Sensor	55.03		✓	1.24 **(-1.58, 4.06)**	1.59 **(-0.58, 3.76)**	-1.07 **(-3.50, 1.36)**	0.52 (-1.91, 2.95)	0.99 **(-0.65, 2.62)**	0.29 **(-1.04, 1.63)**	0.29 **(-1.21, 1.79)**	0.59 **(-0.91, 2.08)**
CONTEC CMS50M	17		✓	0.60 (-1.38, 2.57)	1.50 (-0.33, 3.33)	-0.61 (-2.73, 1.51)	0.89 (-0.96, 2.74)	0.9 **(-0.5, 2.3)**	0.69 **(-0.68, 2.05)**	0.20 **(-1.38, 1.77)**	0.88 **(-0.49, 2.25)**
Covidien (2022) Nellcor PM10N	41		✓	2.54 **(0.33, 4.75)**	0.96 (-1.39, 3.31)	1.46 **(-1.11, 4.03)**	2.42 **(0.18, 4.66)**	0.28 **(-1.18, 1.73)**	-0.21 **(-1.74, 1.32)**	-0.12 **(-1.80, 1.56)**	-0.33 **(-1.79, 1.14)**
Bistos Co., Ltd. BT-710	61		✓	1.49 (0.31, 2.68)	0.44 (-1.12, 2.00)	0.75 (-0.90, 2.40)	1.19 (0.09, 2.29)	0.38 (-0.71, 1.47)	**-0.45** **(-1.85, 0.95)**	0.77 **(-0.71, 2.25)**	0.32 (-0.66, 1.30)
Utech UT100	39		✓	3.05 **(0.41, 5.69)**	-0.12 (-2.71, 2.47)	1.23 **(-1.39, 3.84)**	1.11 **(-1.51, 3.72)**	-0.39 **(-1.72, 0.94)**	**-0.43** **(-1.57, 0.70)**	-0.44 (-1.59, 0.71)	**-0.88** **(-2.02, 0.27)**
BodyMed BDMOXMTRBLK	2		✓	0.09 (-2.52, 2.70)	-0.74 (-3.02, 1.54)	0.92 (-1.36, 3.21)	0.19 (-2.29, 2.66)	-0.58 **(-1.7, 0.54)**	0.49 (-0.51, 1.49)	-0.22 (-1.22, 0.79)	0.27 (-0.81, 1.36)
Masimo (2020) Mightysat Rx	7		✓	3.48 **(0.03, 6.92)**	-0.41 **(-3.74, 2.92)**	2.43 **(-0.90, 5.76)**	2.03 **(-1.30, 5.36)**	0.95 **(-0.8, 2.7)**	0.63 **(-0.95, 2.21)**	0.68 **(-0.91, 2.26)**	1.31 **(-0.28, 2.89)**
EDAN (2020) H100b	44		✓	2.51 **(0.24, 4.79)**	0.56 (-2.06, 3.18)	0.96 **(-1.71, 3.62)**	1.51 **(-0.69, 3.72)**	0.67 **(-0.33, 1.67)**	0.13 (-0.99, 1.24)	0.21 (-0.93, 1.35)	0.34 (-0.59, 1.27)
Guangdong Biolight M800	43		✓	2.85 **(0.20, 5.50)**	-0.81 **(-3.62, 2.01)**	2.38 **(-0.56, 5.33)**	1.57 **(-1.02, 4.17)**	0.66 **(-0.86, 2.18)**	-0.76 **(-2.29, 0.78)**	1.03 **(-0.57, 2.64)**	0.28 **(-1.13, 1.69)**
Masimo Rad-97 with Adult Probe, Product Code 4050	60		✓	3.43 **(0.28, 6.57)**	-1.68 **(-4.91, 1.55)**	**4.26** **(0.76, 7.76)**	2.58 **(-0.65, 5.81)**	1.42 **(-0.06, 2.89)**	-0.20 **(-1.79, 1.39)**	1.33 **(-0.40, 3.05)**	1.13 **(-0.47, 2.72)**

Device	Device ID	Pass FDA 2025 Differential Bias Thresholds	Pass Anticipated ISO Differential Bias Thresholds	DP ITA-Derived Differential Bias (70–85% SaO_2)	MST 1–4 vs 5–7 Differential Bias (70–85% SaO_2)	MST 5–7 vs 8–10 Differential Bias (70–85% SaO_2)	MST 1–4 vs 8–10 Differential Bias (70–85% SaO_2)	DP ITA-Derived Differential Bias (85–100% SaO_2)	MST 1–4 vs 5–7 Differential Bias (85–100% SaO_2)	MST 5–7 vs 8–10 Differential Bias (85–100% SaO_2)	MST 1–4 vs 8–10 Differential Bias (85–100% SaO_2)
Masimo Rad-G with Adult Probe, Product Code 4050	55		✓	3.81 **(2.08, 5.54)**	-0.66 (-2.56, 1.24)	3.23 **(1.17, 5.29)**	2.57 **(0.67, 4.47)**	1.47 **(0.09, 2.86)**	-0.60 **(-1.90, 0.69)**	**1.91** **(0.49, 3.32)**	1.30 **(-0.00, 2.61)**
Acare (2023) AH-M1	57		✓	1.51 **(-1.76, 4.79)**	0.67 **(-2.38, 3.73)**	1.61 **(-1.47, 4.70)**	2.28 **(-1.06, 5.62)**	1.12 **(-1.36, 3.59)**	0.16 **(-2.29, 2.62)**	0.23 **(-2.24, 2.69)**	0.39 **(-2.28, 3.06)**
Guangdong Biolight M70	37			**9.17 (3.15, 15.19)**	1.60 **(-4.18, 7.38)**	**4.59** **(-1.88, 11.05)**	**6.19 (-0.28, 12.65)**	0.51 **(-1.4, 2.42)**	-0.94 **(-2.54, 0.67)**	1.23 **(-0.57, 3.03)**	0.29 **(-1.51, 2.10)**
Acare (2020) AH-MX	42		✓	-0.27 **(-3.62, 3.09)**	-0.54 **(-4.12, 3.03)**	-1.71 **(-5.50, 2.08)**	-2.25 **(-5.35, 0.85)**	-0.86 **(-2.15, 0.42)**	0.34 **(-1.03, 1.72)**	-1.18 **(-2.65, 0.29)**	-0.84 **(-2.03, 0.35)**
American Oxygen LK-87	20		✓	-0.6 **(-4.13, 2.93)**	-1.38 **(-5.01, 2.25)**	0.71 **(-2.90, 4.32)**	-0.67 **(-4.61, 3.27)**	1.79 **(-0.29, 3.87)**	0.23 **(-1.90, 2.37)**	1.39 **(-0.76, 3.54)**	1.63 **(-0.71, 3.97)**

Point estimates and 95% confidence intervals (CI) for each differential bias criterion are reported per the anticipated 2025 U.S. Food and Drug Administration (FDA) and International Organization for Standardization (ISO) guidelines; a check mark (✓) indicates devices meeting the overall differential bias success criteria in each guideline. Columns display Individual Typology Angle (ITA)-derived differential bias at functional arterial oxygen saturation (SaO_2) ranges 70–85% and 85–100%, as well as MST-derived differential bias at those same ranges. Point estimates (and their CIs) are bolded if they fail to meet any criterion (point estimate and 95% CI < 3.5% for 70–85% SaO_2 and < 1.5% for 85–100% SaO_2 for 2025 FDA, and point estimate $\leq$ 4% for 70–85% SaO_2 and $\leq$ 2% for 85–100% SaO_2 for 2025 ISO). Data reflect devices tested in a cohort size of 24 participants. 95% CIs for differential bias are calculated as described in eMethods, Supplemental Digital Content.

Table S13. Differential bias for maximum participant cohorts

Device	Device ID	Pass FDA 2025 Differential Bias Thresholds	Pass Anticipated ISO Differential Bias Thresholds	DP ITA-Derived Differential Bias (70–85% SaO_2)	MST 1–4 vs 5–7 Differential Bias (70–85% SaO_2)	MST 5–7 vs 8–10 Differential Bias (70–85% SaO_2)	MST 1–4 vs 8–10 Differential Bias (70–85% SaO_2)	DP ITA-Derived Differential Bias (85–100% SaO_2)	MST 1–4 vs 5–7 Differential Bias (85–100% SaO_2)	MST 5–7 vs 8–10 Differential Bias (85–100% SaO_2)	MST 1–4 vs 8–10 Differential Bias (85–100% SaO_2)
Nonin CO-Pilot	70	✓	✓	0.06 (-0.65, 0.77)	-0.05 (-0.73, 0.64)	0.26 (-0.40, 0.92)	0.21 (-0.48, 0.91)	0.00 (-0.60, 0.61)	-0.12 (-0.71, 0.47)	0.08 (-0.49, 0.65)	-0.04 (-0.64, 0.55)
Shenzhen (2023) PC-60NW	21	✓	✓	0.21 (-0.72, 1.14)	0.06 (-0.83, 0.96)	-0.21 (-1.09, 0.66)	-0.15 (-1.09, 0.79)	-0.02 (-0.82, 0.78)	-0.31 (-1.11, 0.49)	0.19 (-0.59, 0.97)	-0.12 (-0.97, 0.73)
Masimo Rad-97 with RD Rainbow Adt 8λ Adult Adhesive Sensor	73.09		✓	1.45 (0.12, 2.78)	0.66 (-0.82, 2.14)	0.59 (-1.20, 2.38)	1.25 (-0.23, 2.73)	1.11 **(0.19, 2.03)**	0.85 **(-0.15, 1.84)**	0.12 (-1.08, 1.33)	0.97 **(-0.02, 1.96)**
Shenzhen (2024) PC-60NW	81		✓	-0.43 (-2.01, 1.15)	1.07 (-0.12, 2.25)	-1.77 (-3.05, -0.48)	-0.70 (-1.93, 0.53)	0.68 **(-0.29, 1.65)**	0.07 (-0.76, 0.90)	0.50 (-0.40, 1.40)	0.57 (-0.29, 1.43)
Zug Medical Systems SAS SPLF-SE with SpotLife Reusable Probe	75	✓	✓	1.31 (0.04, 2.59)	0.85 (-0.26, 1.96)	-0.57 (-1.81, 0.67)	0.28 (-0.87, 1.42)	0.62 (-0.16, 1.40)	-0.16 (-0.84, 0.51)	0.74 (-0.01, 1.48)	0.58 (-0.11, 1.27)
Compass Health POX-ROS	12		✓	-1.12 (-2.42, 0.17)	0.21 (-1.24, 1.67)	-0.68 (-1.88, 0.52)	-0.47 (-2.00, 1.06)	-0.92 **(-2.09, 0.26)**	-0.58 **(-1.93, 0.77)**	-0.49 **(-1.60, 0.62)**	-1.07 **(-2.49, 0.34)**
Beijing Choice MD300M	40		✓	1.08 (-0.13, 2.28)	0.42 (-0.65, 1.48)	0.29 (-0.82, 1.41)	0.71 (-0.39, 1.80)	0.23 (-0.87, 1.32)	-0.88 **(-1.77, 0.02)**	0.53 (-0.41, 1.46)	-0.35 (-1.27, 0.57)
Nonin PalmSAT 2500	36		✓	-0.39 (-1.59, 0.81)	0.00 (-1.21, 1.21)	-0.22 (-1.50, 1.06)	-0.22 (-1.34, 0.90)	-0.12 (-1.03, 0.79)	-0.91 (-1.77, -0.05)	0.63 **(-0.28, 1.54)**	-0.27 (-1.07, 0.52)
Covidien PM1000N	59	✓	✓	2.6 (1.82, 3.38)	0.55 (-0.11, 1.20)	1.93 (1.21, 2.66)	2.48 (1.75, 3.21)	0.66 (-0.01, 1.33)	0.01 (-0.53, 0.54)	0.58 (-0.01, 1.16)	0.58 (-0.01, 1.18)
Shenzhen AM801	10	✓	✓	2.10 (0.99, 3.21)	-0.42 (-1.51, 0.67)	2.17 (0.97, 3.37)	1.75 (0.53, 2.97)	-0.21 (-1.42, 1.00)	-0.12 (-1.33, 1.08)	0.03 (-1.30, 1.37)	-0.09 (-1.45, 1.27)
Nonin Onyx Vantage 9590	35	✓	✓	0.44 (-1.04, 1.92)	0.51 (-0.88, 1.90)	0.02 (-1.29, 1.34)	0.53 (-0.81, 1.87)	-0.19 (-1.21, 0.83)	0.09 (-0.87, 1.05)	0.15 (-0.75, 1.06)	0.25 (-0.68, 1.18)
Choicemed MD300CN350R	14		✓	-1.11 (-2.40, 0.17)	0.06 (-1.18, 1.31)	-0.88 (-2.01, 0.25)	-0.82 (-2.12, 0.49)	-0.82 **(-1.83, 0.20)**	0.24 (-0.76, 1.24)	-0.48 (-1.39, 0.43)	-0.24 (-1.29, 0.81)

Device	Device ID	Pass FDA 2025 Differential Bias Thresholds	Pass Anticipated ISO Differential Bias Thresholds	DP ITA-Derived Differential Bias (70–85% SaO_2)	MST 1–4 vs 5–7 Differential Bias (70–85% SaO_2)	MST 5–7 vs 8–10 Differential Bias (70–85% SaO_2)	MST 1–4 vs 8–10 Differential Bias (70–85% SaO_2)	DP ITA-Derived Differential Bias (85–100% SaO_2)	MST 1–4 vs 5–7 Differential Bias (85–100% SaO_2)	MST 5–7 vs 8–10 Differential Bias (85–100% SaO_2)	MST 1–4 vs 8–10 Differential Bias (85–100% SaO_2)
Masimo (2024) Mightysat	78		✓	2.73 **(1.28, 4.18)**	0.84 (-0.59, 2.27)	0.86 (-0.68, 2.40)	1.71 (0.23, 3.18)	1.20 **(0.29, 2.11)**	0.13 (-0.72, 0.98)	0.66 **(-0.25, 1.57)**	0.79 **(-0.08, 1.67)**
Acare (2023) Model AH-TA	56	✓	✓	-0.03 (-1.25, 1.19)	0.22 (-0.87, 1.32)	-0.18 (-1.35, 0.99)	0.04 (-1.10, 1.19)	-0.55 (-1.28, 0.18)	-0.29 (-0.95, 0.37)	-0.35 (-1.06, 0.35)	-0.64 (-1.33, 0.04)
Philips M1020B IntelliVue with M1191BL Probe	71		✓	1.66 (0.79, 2.54)	0.34 (-0.59, 1.28)	0.96 (-0.12, 2.04)	1.30 (0.36, 2.23)	1.41 **(0.76, 2.06)**	0.14 (-0.63, 0.91)	0.69 **(-0.20, 1.58)**	0.83 **(0.06, 1.60)**
CMS Zacurate 500C	15		✓	-1.34 (-3.40, 0.72)	0.30 (-2.18, 2.78)	-1.36 (-3.41, 0.70)	-1.06 (-3.46, 1.35)	-0.63 (-1.44, 0.18)	-0.51 (-1.49, 0.48)	-0.17 (-0.98, 0.65)	-0.67 **(-1.63, 0.28)**
Covidien (2022) Nellcor PM10N	41		✓	3.22 **(1.53, 4.91)**	1.89 (0.46, 3.31)	1.45 (-0.06, 2.97)	3.34 **(1.95, 4.73)**	0.42 **(-0.92, 1.75)**	0.35 **(-0.91, 1.62)**	0.05 (-1.29, 1.39)	0.40 **(-0.82, 1.63)**
Utech UT100	39		✓	3.16 **(1.30, 5.01)**	0.12 (-1.69, 1.93)	1.97 **(0.04, 3.89)**	2.09 **(0.27, 3.91)**	0.20 (-0.75, 1.15)	-0.57 (-1.42, 0.27)	0.23 (-0.67, 1.13)	-0.34 (-1.20, 0.51)
Nihon Kohden OLV-4202 with NK-SPO2-ADVT Probe	64.03		✓	1.77 **(-0.52, 4.05)**	0.66 **(-2.22, 3.53)**	-0.29 **(-3.61, 3.03)**	0.37 (-2.51, 3.24)	0.95 **(-0.72, 2.62)**	-0.45 **(-2.31, 1.41)**	0.53 **(-1.61, 2.68)**	0.08 **(-1.78, 1.94)**
EDAN (2020) H100b	44		✓	2.06 **(0.47, 3.64)**	-0.46 [-1.69, 0.77]	1.13 (-0.18, 2.44)	0.68 (-0.53, 1.88)	0.47 (-0.35, 1.28)	-0.43 (-1.22, 0.36)	0.26 (-0.58, 1.09)	-0.17 (-0.94, 0.59)
Shenzhen WIT-S300C	62	✓	✓	1.03 (-0.53, 2.59)	0.83 (-0.69, 2.35)	0.21 (-1.31, 1.74)	1.05 (-0.51, 2.60)	0.20 (-0.75, 1.15)	0.15 (-0.77, 1.06)	-0.01 (-0.93, 0.90)	0.13 (-0.80, 1.07)
CONTEC CMS50M	17		✓	0.70 (-1.11, 2.50)	0.83 (-0.86, 2.53)	0.16 (-1.80, 2.11)	0.99 (-0.80, 2.77)	1.11 **(-0.07, 2.3)**	0.46 **(-0.68, 1.61)**	0.58 **(-0.74, 1.90)**	1.05 **(-0.15, 2.24)**
Shenzhen A310L	11		✓	2.25 **(0.37, 4.12)**	0.58 (-1.35, 2.51)	1.41 (-0.48, 3.29)	1.99 **(0.03, 3.95)**	0.02 (-0.72, 0.76)	0.37 (-0.37, 1.10)	-0.01 (-0.73, 0.70)	0.35 (-0.40, 1.10)
Masimo (2020) Mightysat Rx	7		✓	3.46 **(1.52, 5.40)**	0.90 (-0.78, 2.57)	1.73 **(-0.10, 3.57)**	2.63 **(0.73, 4.53)**	1.43 **(0.13, 2.72)**	0.34 (-0.70, 1.37)	1.22 **(0.09, 2.36)**	**1.56 (0.38, 2.74)**

Device	Device ID	Pass FDA 2025 Differential Bias Thresholds	Pass Anticipated ISO Differential Bias Thresholds	DP ITA-Derived Differential Bias (70–85% SaO_2)	MST 1–4 vs 5–7 Differential Bias (70–85% SaO_2)	MST 5–7 vs 8–10 Differential Bias (70–85% SaO_2)	MST 1–4 vs 8–10 Differential Bias (70–85% SaO_2)	DP ITA-Derived Differential Bias (85–100% SaO_2)	MST 1–4 vs 5–7 Differential Bias (85–100% SaO_2)	MST 5–7 vs 8–10 Differential Bias (85–100% SaO_2)	MST 1–4 vs 8–10 Differential Bias (85–100% SaO_2)
Masimo Rad-G with Rad-G Finger Clip Sensor	55.03		✓	1.91 **(-0.47, 4.28)**	1.07 (-0.90, 3.03)	-0.07 (-2.32, 2.18)	1.00 (-1.10, 3.10)	0.96 **(-0.59, 2.52)**	-0.21 (-1.44, 1.02)	1.18 **(-0.23, 2.60)**	0.98 **(-0.34, 2.30)**
Bistos Co., Ltd. BT-710	61	✓	✓	1.40 (0.25, 2.55)	-0.09 (-1.11, 0.94)	1.04 (-0.03, 2.10)	0.95 (-0.09, 1.99)	0.34 (-0.48, 1.16)	-0.14 (-0.88, 0.60)	0.18 (-0.57, 0.94)	0.04 (-0.69, 0.78)
BodyMed BDMOXMTRBLK	2		✓	-0.50 (-2.70, 1.70)	-0.69 (-2.95, 1.57)	0.11 (-1.97, 2.19)	-0.58 (-2.85, 1.69)	-0.82 **(-1.69, 0.05)**	0.38 (-0.54, 1.31)	-0.55 (-1.40, 0.30)	-0.16 (-1.09, 0.76)
Masimo Rad-97 with Adult Probe, Product Code 4050	60		✓	3.22 **(2.28, 4.17)**	1.19 (0.42, 1.97)	2.07 (1.19, 2.95)	3.26 **(2.35, 4.17)**	1.41 **(0.4, 2.43)**	0.26 (-0.60, 1.11)	1.13 **(0.16, 2.10)**	1.39 **(0.38, 2.39)**
Guangdong Biolight M800	43		✓	2.35 **(0.12, 4.58)**	0.29 (-1.74, 2.32)	1.00 (-1.22, 3.22)	1.29 (-0.71, 3.28)	0.69 **(-0.32, 1.7)**	0.37 (-0.53, 1.28)	0.18 (-0.80, 1.17)	0.56 (-0.32, 1.44)
Masimo Rad-G with Adult Probe, Product Code 4050	55		✓	3.30 **(1.76, 4.83)**	0.25 (-1.26, 1.75)	1.94 **(0.23, 3.66)**	2.19 **(0.54, 3.84)**	1.17 **(-0.23, 2.57)**	-0.03 (-1.28, 1.22)	1.13 **(-0.30, 2.55)**	1.10 **(-0.28, 2.47)**
Guangdong Biolight M70	37			**4.83 (0.61, 9.05)**	0.68 **(-3.31, 4.66)**	1.59 **(-2.40, 5.58)**	2.26 **(-1.73, 6.25)**	0.13 (-0.98, 1.24)	-0.49 (-1.48, 0.49)	0.25 (-0.74, 1.23)	-0.25 (-1.23, 0.73)
Acare (2023) AH-M1	57		✓	0.48 (-2.05, 3.01)	0.86 (-1.36, 3.08)	-0.16 (-2.59, 2.28)	0.70 (-1.83, 3.23)	0.74 **(-0.77, 2.26)**	0.53 **(-0.82, 1.88)**	-0.13 **(-1.61, 1.36)**	0.41 **(-1.13, 1.94)**
Acare (2020) AH-MX	42		✓	-0.20 (-2.43, 2.04)	-0.31 (-2.35, 1.73)	-1.17 (-3.32, 0.98)	-1.48 **(-3.64, 0.68)**	-0.77 **(-1.81, 0.27)**	0.05 (-0.89, 0.99)	-1.08 **(-2.07, -0.10)**	-1.03 **(-2.02, -0.05)**
American Oxygen LK-87	20		✓	-0.01 (-2.89, 2.86)	-0.35 (-3.10, 2.40)	0.30 (-2.43, 3.04)	-0.05 (-3.08, 2.98)	1.72 **(0.02, 3.42)**	0.23 **(-1.39, 1.84)**	1.23 **(-0.39, 2.85)**	1.46 **(-0.34, 3.25)**

Point estimates and 95 % confidence intervals (CI) for each differential bias criterion are reported per the anticipated 2025 U.S. Food and Drug Administration (FDA) and International Organization for Standardization (ISO) guidelines; a check mark (✓) indicates devices meeting the overall differential bias success criteria in each guideline. Columns display Individual Typology Angle (ITA)-derived differential bias at functional arterial oxygen saturation (SaO_2) ranges 70–85% and 85–100%, as well as MST-derived

differential bias at those same ranges. Point estimates (and their CIs) are bolded if they fail to meet any criterion (95% CI < 3.5% for 70–85% SaO_2 and < 1.5% for 85–100% SaO_2 for 2025 FDA, and point estimate $\leq$ 4% for 70–85% SaO_2 and $\leq$ 2% for 85–100% SaO_2 for 2025 ISO). Data reflect devices tested in a cohort size that is as large as possible while satisfying enrollment criteria in the anticipated 2025 FDA and 2025 ISO guidelines. 95% CIs for differential bias are calculated as described in eMethods, Supplemental Digital Content.

Table S14. Count of devices with statistically significant differences in bias (mean of SpO_2 minus SaO_2) across different definitions of dark and anatomical sites in maximum participant cohorts

	ITA>30° vs ITA<-30° with half <-50° (Light vs. Dark)		**ITA>30° vs ITA<-50° (Light vs. Very Dark)**		**MST 1–3 vs 8–10**	**MST 1–4 vs 8–10**	**White vs Black/African American**
	Forehead	DP	Forehead	DP	Forehead	Forehead	Overall
SaO_2 70–100%	9	11	9	10	9	8	7
SaO_2 70–80%	12	8	12	10	7	10	7
SaO_2 80–90%	9	9	6	10	9	9	8

SaO_2 90–100%	0	3	5	4	4	2	1

Number of devices showing a statistically significant bias difference between dark and light skin color bins at the forehead and dorsal distal phalanx (DP) sites, stratified by saturation ranges. Bias is defined as the mean SpO_2 - SaO_2 error, where SpO_2 is pulse oximeter measured oxygen saturation and SaO_2 is functional arterial oxygen saturation. Comparisons between light versus dark bins were made using a linear mixed-effects model (SpO_2 - SaO_2 error as the outcome, skin color bin as a fixed effect, participant ID as a random intercept) with $\alpha = 0.05$. Light and dark bins are defined by either Individual Typology Angle (ITA), Monk Skin Tone Scale (MST), or National Institutes of Health (NIH) race. No correction for multiple comparisons (e.g., Bonferroni) was applied, per the 2025 draft U.S. Food and Drug Administration (FDA) guidelines. Results are based on cohorts that are as large as possible while satisfying enrollment criteria in the anticipated 2025 FDA and 2025 ISO (International Organization for Standardization) guidelines.

Table S15. Count of devices with 95% CI of difference in bias (mean of SpO_2 minus SaO_2) exceeding pre-defined thresholds across different definitions of dark and anatomical sites in maximum participant cohorts

	ITA > 30° vs ITA < -30° with half < -50° (Light vs. Dark)		**ITA > 30° vs ITA < -50° (Light vs. Very Dark)**		**MST 1–3 vs 8–10**	**MST 1–4 vs 8–10**	**White vs Black/African American**
	Forehead	DP	Forehead	DP	Forehead	Forehead	Overall
SaO_2 70–85%	14	18	19	19	17	11	11
SaO_2 85–100%	17	22	20	25	18	15	17

Number of devices for which the 95% confidence interval (CI) of the difference in bias (mean SpO_2 - SaO_2, where SpO_2 is pulse oximeter measured oxygen saturation and SaO_2 is functional arterial oxygen saturation) exceeded the pre-defined thresholds of 1.5% for the SaO_2 range 85–100% and 3.5% for the SaO_2 range 70–85%. Light and dark bins are defined by either Individual Typology Angle (ITA), Monk Skin Tone Scale (MST), or National Institutes of Health (NIH) race. The 95% confidence interval is formed by taking the estimated effect from the mixed-effects model (SpO_2 – SaO_2 error as the outcome, skin color bin as a fixed

effect, participant ID as a random intercept) and adding and subtracting the t-critical value (for two-sided 2.5% with the model's residual degrees of freedom) times its standard error. Results are based on cohorts that are as large as possible while satisfying enrollment criteria in the anticipated 2025 FDA and 2025 ISO (International Organization for Standardization) guidelines.

Table S16. Reproducibility of device conformity with with regulatory frameworks

				A_{RMS} Criteria				Differential Bias Criteria			Combination of A_{RMS} and Differential Bias Criteria	
Device	**Device ID**	**Number of Data Pairs**	**Cohort Size**	**Overall A_{RMS} (95% CI)**	**Pass 2013 FDA and Anticipated 2025 ISO (A_{RMS} Point Estimate ≤ 3%)**	**Pass 2017 ISO (A_{RMS} Point Estimate ≤ 4%)**	**Pass Anticipated 2025 FDA (A_{RMS} Upper CI < 3%)**	**Pass Anticipated 2025 FDA Pairwise MST Differential Bias Criteria**	**Pass Anticipated 2025 FDA ITA-Derived Differential Bias Criteria**	**Pass Anticipated 2025 ISO ITA-Derived Differential Bias Criteria**	**Pass Anticipated FDA 2025 (A_{RMS} Upper CI <3% and Differential Bias Thresholds)**	**Pass Anticipated 2025 ISO (A_{RMS} Point Estimate ≤3% and Differential Bias Thresholds)**
Shenzhen PC-60NW (2023)	21	761	35	1.61 (1.44, 1.80)	✓	✓	✓	✓	✓	✓	✓	✓
Shenzhen PC-60NW (2024)	81	638	28	1.87 (1.45, 2.28)	✓	✓	✓	✓		✓		✓
Shenzhen PC-60NW (2024 + 2023)	81+21	1450	64	1.73 (1.52, 1.95)	✓	✓	✓	✓	✓	✓	✓	✓
Masimo Mightysat Rx (2020)	7	665	31	2.39 (1.98, 2.88)	✓	✓				✓		✓
	78	946	44	2.88	✓	✓	✓			✓		✓

Masimo Mightysat (2024)				(2.00, 4.11)								
Masimo Mightysat (2020 + 2024)	7+78	1419	64	2.84 (2.06, 3.70)	✓	✓			✓	✓		✓
AcareAH-MX (2020)	42	919	44	5.28 (4.60, 5.96)						✓		
Acare AH-M1 (2023)	57	1093	51	6.80 (6.05, 7.48)						✓		
Acare AH-M1 (2023 +2024)	42+57	1878	64	6.02 (5.42, 6.63)					✓	✓		

This table reports the accuracy root mean square error (A_{RMS}) with 95% confidence intervals (CI) for three pulse oximeter models tested with two devices each. Results are based on cohorts that are as large as possible while satisfying the enrollment criteria outlined in the anticipated 2025 U.S. Food and Drug Administration (FDA) and 2025 International Organization for Standardization (ISO) guidelines. Each device's performance is shown in a separate row, and an additional row presents results from the aggregated cohort combining data from the same model across different purchased years. The aggregated cohort size is not simply the sum of the two individual device cohort sizes. This is because the combined cohort must still meet all inclusion constraints—including balanced representation across skin tone categories (≥25% in each), sex assigned at birth (≥33% of each). These constraints may not be simultaneously satisfiable at full scale when combining the two datasets, resulting in a smaller feasible cohort. A check mark (✓) indicates devices meeting the overall ARMS and/or differential bias success criteria in each guideline. The 95% CIs for A_{RMS} were calculated using bootstrapping (random resampling with replacement) with 1,000 iterations.

Table S17. ITA and pulsatility amplitude for participant pigment bins in all device cohorts

			24 Participant Cohort Size				Maximum Participant Cohort			
Device	**Device ID**	**Metric**	**Light**	**Medium**	**Dark**	**Very Dark**	**Light**	**Medium**	**Dark**	**Very Dark**
Nonin CO-Pilot	70	ITA (Forehead)	42.32 [39.37, 43.50]	17.97 [-11.97, 20.78]	-53.32 [-66.82, -47.91]	-66.82 [-71.89, -54.63]	40.97 [35.96, 45.29]	18.03 [9.18, 24.39]	-56.14 [-59.99, -48.05]	-58.25 [-66.82, -56.14]
		ITA (DP)	20.94 [19.90, 30.22]	11.33 [-28.79, 19.25]	-57.66 [-66.69, -37.96]	-57.66 [-84.81, -45.82]	33.11 [23.79, 38.64]	10.44 [-10.51, 19.25]	-58.91 [-66.83, -38.43]	-60.16 [-66.83, -45.82]
		MST (Forehead)	3 [1, 3]	5 [5, 8]	9 [8, 10]	9 [8, 10]	4 [1, 5]	5 [2, 8]	8 [8, 10]	8 [8, 10]
		MST (DP)	2 [2, 4]	5 [2, 8]	9 [8, 10]	9 [8, 10]	2 [1, 5]	5 [2, 8]	8 [7, 10]	8 [8, 10]
		%Mod	2.70 [1.15, 4.30]	1.00 [0.70, 4.10]	4.50 [1.40, 6.70]	4.50 [2.10, 7.40]	3.10 [2.00, 4.50]	2.70 [0.90, 5.10]	2.50 [1.20, 5.50]	2.60 [1.20, 5.60]
Shenzhen (2023) PC-60NW	21	ITA (Forehead)	39.37 [31.95, 43.48]	16.59 [3.56, 20.71]	-58.94 [-60.45, -43.66]	-60.45 [-71.89, -58.94]	39.37 [32.79, 43.50]	18.03 [3.56, 25.95]	-51.48 [-60.45, -43.66]	-60.45 [-64.00, -54.63]
		ITA (DP)	26.31 [18.64, 30.22]	8.24 [-1.84, 21.30]	-68.65 [-68.79, -37.96]	-68.65 [-84.81, -60.10]	27.27 [18.64, 33.29]	8.24 [-10.51, 21.82]	-61.80 [-68.65, -37.96]	-61.80 [-68.65, -44.20]
		MST (Forehead)	2 [1, 5]	5 [2, 8]	8 [8, 10]	8 [8, 10]	2 [1, 5]	5 [2, 8]	8 [8, 10]	8 [8, 10]
		MST (DP)	2 [1, 6]	5 [3, 8]	9 [8, 9]	9 [8, 9]	2 [1, 6]	5 [3, 8]	8 [8, 9]	8 [8, 9]
		%Mod	2.50 [1.10, 3.40]	3.90 [0.70, 6.10]	1.60 [0.70, 3.20]	1.60 [0.40, 2.80]	2.50 [1.50, 3.40]	3.10 [0.90, 5.40]	2.50 [1.20, 6.30]	2.50 [1.50, 6.30]
Masimo Rad-97 with RD Rainbow Adt 8λ Adult Adhesive Sensor	73.09	ITA (Forehead)	41.16 [37.97, 44.48]	2.52 [-3.65, 20.77]	-53.55 [-68.03, -46.24]	-68.03 [-68.06, -53.55]	41.96 [39.23, 44.48]	20.77 [16.85, 24.83]	-53.55 [-68.03, -46.24]	-68.03 [-68.06, -53.55]
		ITA (DP)	39.68 [33.11, 44.68]	4.10 [-13.01, 22.28]	-60.53 [-75.10, -36.92]	-60.53 [-82.47, -36.92]	39.68 [34.58, 41.46]	10.64 [-4.70, 24.11]	-60.53 [-75.10, -36.92]	-60.53 [-82.47, -36.92]
		MST (Forehead)	4 [2, 4]	6 [4, 7]	8 [7, 9]	9 [8, 9]	2 [2, 4]	4 [2, 6]	8 [7, 9]	9 [8, 9]

		MST (DP)	2 [1, 4]	7 [2, 7]	8 [7, 10]	10 [7, 10]	2 [1, 4]	4 [2, 7]	8 [7, 10]	10 [7, 10]
		%Mod	4.20 [2.90, 5.30]	1.90 [1.00, 3.30]	2.50 [1.60, 5.20]	2.20 [1.10, 3.40]	4.40 [2.40, 5.50]	4.00 [2.70, 5.70]	2.50 [1.60, 5.20]	2.20 [1.10, 3.40]
Shenzhen (2024) PC-60NW	81	ITA (Forehead)	40.97 [36.65, 45.52]	19.24 [-2.77, 22.88]	-54.54 [-63.53, -32.14]	-63.53 [-73.37, -54.54]	41.53 [38.52, 45.82]	19.24 [-14.84, 22.88]	-50.02 [-63.53, -32.14]	-63.53 [-73.37, -54.54]
		ITA (DP)	28.17 [24.33, 36.70]	3.36 [-26.37, 12.68]	-50.20 [-68.77, -24.57]	-50.20 [-70.41, -39.31]	35.95 [28.10, 39.41]	3.36 [-11.47, 15.82]	-39.31 [-68.77, -24.57]	-50.20 [-70.41, -39.31]
		MST (Forehead)	4 [1, 4]	5 [5, 7]	8 [8, 8]	8 [8, 8]	4 [1, 4]	5 [3, 8]	8 [8, 8]	8 [8, 8]
		MST (DP)	2 [1, 5]	5 [5, 7]	8 [7, 9]	8 [8, 8]	2 [1, 4]	7 [3, 8]	8 [7, 9]	8 [8, 8]
		%Mod	2.80 [1.60, 4.70]	3.30 [2.40, 5.10]	1.00 [0.60, 2.50]	1.50 [0.60, 3.20]	3.30 [1.80, 5.60]	4.00 [2.40, 7.30]	1.30 [0.60, 2.80]	2.10 [0.70, 3.50]
Zug Medical Systems SAS SPLF-SE with SpotLife Reusable Probe	75	ITA (Forehead)	43.69 [40.97, 47.58]	15.16 [4.21, 21.28]	-57.20 [-58.96, -39.60]	-58.25 [-58.96, -57.20]	40.97 [36.65, 44.48]	15.16 [1.00, 21.04]	-54.54 [-58.25, -39.60]	-58.25 [-63.53, -54.54]
		ITA (DP)	33.85 [27.27, 39.41]	2.02 [-6.41, 10.64]	-61.09 [-67.15, -38.43]	-65.81 [-67.15, -61.09]	30.29 [24.67, 35.95]	6.07 [-11.47, 24.75]	-50.20 [-65.81, -38.85]	-50.20 [-67.15, -39.31]
		MST (Forehead)	2 [2, 4]	5 [4, 7]	8 [8, 9]	8 [8, 9]	4 [1, 4]	5 [4, 8]	8 [8, 9]	8 [8, 9]
		MST (DP)	2 [1, 4]	7 [4, 7]	8 [7, 9]	8 [8, 8]	2 [1, 5]	5 [2, 8]	8 [7, 9]	8 [8, 8]
		%Mod	2.00 [1.40, 3.80]	4.20 [1.80, 6.20]	1.50 [1.10, 2.50]	1.50 [1.10, 2.42]	2.30 [1.50, 4.40]	3.90 [2.20, 6.00]	1.50 [1.10, 2.80]	1.50 [1.10, 2.90]
Compass Health POX-ROS	12	ITA (Forehead)	32.79 [31.54, 40.24]	12.09 [3.35, 18.03]	-60.45 [-70.27, -51.44]	-66.82 [-70.27, -53.32]	38.90 [31.95, 42.16]	21.34 [7.62, 23.02]	-60.45 [-66.82, -51.44]	-64.00 [-66.82, -54.72]
		ITA (DP)	28.64 [11.48, 33.29]	8.24 [-11.29, 19.25]	-57.66 [-72.04, -43.71]	-57.66 [-72.04, -45.82]	28.64 [11.48, 30.29]	8.24 [-1.84, 27.24]	-61.80 [-68.65, -45.82]	-61.80 [-68.65, -48.78]
		MST (Forehead)	5 [2, 5]	6 [4, 7]	8 [8, 10]	9 [8, 10]	5 [2, 5]	5 [4, 8]	8 [8, 10]	8 [8, 10]
		MST (DP)	5 [1, 6]	7 [4, 7]	9 [8, 10]	9 [8, 10]	2 [1, 6]	5 [4, 8]	8 [8, 10]	8 [8, 10]
		%Mod	0.50 [0.40, 1.10]	2.00 [0.90, 5.00]	1.50 [0.70, 5.20]	1.80 [0.80, 5.80]	0.70 [0.40, 1.30]	3.00 [1.00, 5.40]	2.20 [1.10, 6.10]	2.40 [1.30, 5.60]

Beijing Choice MD300M	40	ITA (Forehead)	35.61 [32.78, 39.93]	17.97 [9.18, 25.65]	-51.72 [-58.69, -45.26]	-58.69 [-71.89, -51.72]	38.90 [34.88, 40.55]	17.97 [9.18, 24.39]	-51.72 [-58.69, -46.24]	-58.69 [-71.89, -51.72]
		ITA (DP)	23.42 [15.00, 26.31]	10.44 [6.69, 23.49]	-63.93 [-73.00, -44.50]	-63.93 [-84.81, -63.93]	21.98 [14.82, 26.82]	17.06 [-4.70, 26.63]	-56.96 [-63.93, -31.01]	-63.93 [-73.00, -44.50]
		MST (Forehead)	4 [1, 5]	5 [5, 8]	9 [8, 10]	9 [9, 10]	2 [1, 5]	5 [4, 8]	9 [8, 10]	9 [8, 10]
		MST (DP)	4 [3, 5]	5 [2, 8]	8 [8, 10]	8 [8, 10]	4 [2, 5]	5 [2, 8]	8 [7, 10]	9 [7, 10]
		%Mod	1.10 [0.80, 2.30]	2.30 [0.90, 4.20]	0.80 [0.50, 3.10]	0.70 [0.30, 4.30]	0.90 [0.70, 1.60]	2.60 [1.20, 4.70]	0.90 [0.60, 2.80]	0.80 [0.40, 2.80]
Nonin PalmSAT 2500	36	ITA (Forehead)	38.03 [34.88, 42.67]	11.06 [8.44, 12.82]	-51.23 [-51.72, -38.43]	-51.72 [-58.69, -51.23]	39.76 [35.61, 43.69]	15.79 [3.37, 25.65]	-51.72 [-58.96, -39.52]	-58.96 [-71.89, -51.72]
		ITA (DP)	25.48 [11.95, 36.13]	8.17 [1.39, 22.26]	-63.93 [-68.35, -56.96]	-63.93 [-73.00, -44.50]	25.48 [16.52, 32.46]	5.56 [-18.99, 23.49]	-63.93 [-68.35, -44.50]	-63.93 [-73.00, -44.50]
		MST (Forehead)	2 [1, 5]	5 [4, 8]	9 [8, 9]	9 [9, 9]	2 [1, 5]	6 [4, 8]	9 [8, 10]	9 [8, 10]
		MST (DP)	3 [2, 5]	6 [2, 6]	8 [7, 9]	8 [8, 9]	3 [2, 5]	5 [4, 8]	8 [7, 10]	8 [8, 10]
		%Mod	1.00 [0.70, 1.50]	1.80 [1.10, 2.80]	0.60 [0.30, 1.60]	0.50 [0.30, 0.70]	1.00 [0.70, 1.50]	2.40 [1.30, 4.00]	1.00 [0.50, 2.50]	1.20 [0.60, 3.60]
Covidien PM1000N	59	ITA (Forehead)	38.03 [35.67, 42.67]	10.74 [8.44, 21.49]	-39.52 [-51.72, -38.16]	-51.72 [-58.69, -51.23]	39.37 [34.53, 43.50]	14.29 [3.56, 21.04]	-51.46 [-59.99, -40.73]	-58.69 [-66.82, -53.55]
		ITA (DP)	28.92 [21.63, 41.76]	14.12 [1.33, 22.26]	-63.93 [-68.35, -53.68]	-63.93 [-73.00, -44.50]	31.16 [24.20, 36.15]	16.20 [-3.60, 23.49]	-57.66 [-66.69, -43.71]	-61.80 [-67.37, -44.20]
		MST (Forehead)	4 [1, 5]	5 [4, 8]	8 [8, 9]	9 [9, 9]	4 [1, 5]	5 [4, 8]	8 [7, 10]	8 [8, 10]
		MST (DP)	3 [2, 5]	4 [4, 6]	8 [7, 9]	8 [8, 9]	4 [1, 6]	5 [2, 7]	8 [7, 10]	8 [7, 10]
		%Mod	1.00 [0.80, 1.50]	2.20 [1.40, 4.60]	0.60 [0.30, 1.50]	0.50 [0.30, 0.70]	2.00 [0.90, 3.80]	3.10 [1.40, 5.30]	1.60 [0.80, 4.70]	1.80 [0.90, 5.30]
Shenzhen AM801	10	ITA (Forehead)	40.24 [31.95, 43.50]	12.25 [0.72, 22.77]	-56.14 [-66.82, -43.04]	-66.82 [-71.89, -56.14]	40.24 [31.95, 43.50]	20.71 [3.56, 25.75]	-56.14 [-66.82, -43.04]	-65.41 [-70.62, -64.00]

		ITA (DP)	28.64 [15.00, 30.00]	8.24 [-11.35, 27.24]	-66.69 [-68.79, -57.66]	-57.66 [-84.81, -54.85]	30.00 [18.64, 33.29]	16.77 [-1.84, 27.42]	-66.28 [-68.79, -57.66]	-61.80 [-79.06, -57.66]
		MST (Forehead)	2 [1, 5]	5 [2, 8]	8 [8, 10]	10 [8, 10]	2 [1, 5]	6 [2, 7]	8 [8, 10]	8 [8, 10]
		MST (DP)	4 [1, 5]	7 [4, 8]	9 [8, 9]	9 [8, 9]	1 [1, 5]	5 [3, 7]	9 [8, 9]	8 [8, 9]
		%Mod	1.30 [0.80, 3.10]	3.00 [1.00, 5.40]	5.00 [2.80, 8.00]	3.90 [2.40, 7.40]	1.90 [0.80, 4.50]	4.00 [1.40, 6.10]	4.60 [2.50, 7.60]	3.40 [2.20, 6.60]
Nonin Onyx Vantage 9590	35	ITA (Forehead)	35.61 [31.54, 40.55]	20.78 [-10.06, 22.77]	-53.32 [-60.55, -43.66]	-60.55 [-71.89, -53.32]	35.61 [31.95, 39.72]	12.25 [-10.06, 21.85]	-53.32 [-64.00, -48.46]	-60.55 [-66.82, -53.32]
		ITA (DP)	25.48 [15.00, 30.22]	19.25 [-20.26, 27.24]	-51.76 [-68.79, -45.82]	-51.76 [-84.81, -45.82]	23.42 [11.48, 28.92]	4.95 [-16.34, 19.25]	-57.66 [-62.80, -45.82]	-54.85 [-61.80, -45.82]
		MST (Forehead)	3 [1, 5]	5 [4, 8]	9 [8, 10]	9 [9, 10]	3 [1, 5]	5 [4, 8]	8 [8, 10]	8 [8, 10]
		MST (DP)	2 [2, 5]	5 [4, 8]	8 [7, 10]	7 [7, 10]	4 [2, 5]	5 [4, 8]	8 [7, 10]	8 [7, 10]
		%Mod	1.30 [0.60, 4.00]	1.40 [0.60, 3.40]	2.00 [0.70, 3.90]	2.20 [0.30, 4.10]	1.70 [0.70, 4.50]	2.50 [0.90, 4.20]	2.60 [1.10, 5.60]	2.60 [1.00, 5.72]
Choicemed MD300CN350R	14	ITA (Forehead)	35.61 [31.22, 40.24]	12.09 [-10.06, 20.71]	-53.32 [-70.27, -51.44]	-66.82 [-70.27, -53.32]	35.61 [31.54, 42.16]	12.25 [2.69, 23.02]	-58.96 [-66.82, -51.44]	-64.00 [-70.27, -58.94]
		ITA (DP)	28.64 [18.64, 31.16]	18.51 [-20.26, 21.30]	-57.66 [-72.04, -45.82]	-57.66 [-72.04, -45.82]	28.92 [18.64, 31.16]	2.39 [-11.30, 18.51]	-60.10 [-66.28, -43.71]	-61.80 [-68.65, -45.82]
		MST (Forehead)	5 [1, 5]	5 [2, 7]	9 [8, 10]	9 [8, 10]	2 [1, 5]	5 [2, 8]	8 [8, 10]	8 [8, 10]
		MST (DP)	4 [1, 5]	5 [3, 7]	9 [8, 10]	9 [8, 10]	4 [1, 5]	5 [3, 7]	8 [7, 10]	8 [8, 10]
		%Mod	1.10 [0.60, 2.90]	2.10 [0.60, 4.60]	3.40 [0.90, 6.60]	2.50 [0.80, 5.80]	1.30 [0.70, 3.40]	2.10 [0.60, 4.30]	2.50 [1.30, 5.10]	2.10 [1.00, 4.50]
Masimo (2024) Mightysat	78	ITA (Forehead)	38.40 [35.35, 43.69]	15.16 [8.70, 21.28]	-59.99 [-70.08, -43.61]	-70.08 [-70.08, -59.99]	40.97 [38.40, 43.69]	15.16 [4.21, 21.28]	-55.36 [-59.99, -39.60]	-59.99 [-72.28, -58.96]
		ITA (DP)	35.24 [27.27, 47.60]	2.02 [-6.41, 14.96]	-60.16 [-66.83, -38.43]	-66.83 [-69.20, -60.16]	35.24 [27.27, 37.28]	2.02 [-6.41, 14.96]	-57.14 [-61.09, -28.63]	-61.09 [-65.81, -60.16]

		MST (Forehead)	2 [2, 4]	6 [4, 8]	8 [8, 9]	9 [8, 9]	2 [2, 4]	6 [4, 7]	8 [8, 9]	8 [8, 9]
		MST (DP)	2 [2, 4]	5 [4, 7]	8 [8, 9]	8 [8, 9]	2 [1, 4]	5 [4, 7]	8 [7, 9]	8 [8, 9]
		%Mod	2.50 [1.80, 3.80]	3.90 [2.48, 5.30]	1.80 [1.20, 3.50]	1.80 [1.20, 3.80]	2.50 [1.70, 3.80]	3.80 [2.40, 5.20]	1.80 [1.20, 3.30]	1.80 [1.20, 3.60]
Acare (2023) Ab-F1/Ah-S1/Ah-M1	56	ITA (Forehead)	37.54 [34.88, 42.67]	9.18 [-5.29, 16.59]	-53.32 [-60.55, -42.00]	-60.55 [-71.89, -53.32]	35.85 [32.95, 42.67]	9.18 [-7.51, 23.74]	-53.32 [-60.55, -42.57]	-59.99 [-71.89, -53.32]
		ITA (DP)	28.92 [18.29, 30.22]	6.69 [-11.27, 9.73]	-59.38 [-68.35, -51.76]	-51.76 [-84.81, -45.82]	26.31 [18.64, 30.22]	-4.70 [-20.26, 12.84]	-59.38 [-66.73, -45.30]	-51.76 [-66.83, -43.71]
		MST (Forehead)	4 [2, 5]	5 [5, 7]	9 [8, 10]	9 [9, 10]	4 [1, 5]	6 [4, 8]	8 [8, 10]	9 [8, 10]
		MST (DP)	5 [2, 5]	6 [2, 8]	8 [7, 10]	7 [7, 10]	3 [2, 5]	6 [2, 8]	8 [7, 10]	8 [7, 10]
		%Mod	1.20 [0.80, 3.00]	2.80 [0.90, 4.40]	2.70 [1.10, 5.70]	2.10 [0.20, 5.10]	1.70 [0.90, 3.60]	3.20 [1.50, 5.00]	1.40 [0.50, 4.10]	1.70 [0.50, 3.80]
Philips M1020B IntelliVue with M1191BL Probe	71	ITA (Forehead)	38.91 [34.33, 41.15]	18.28 [-18.79, 22.91]	-47.91 [-60.45, -42.57]	-60.45 [-71.89, -51.44]	41.00 [39.02, 44.48]	21.01 [14.05, 23.62]	-51.44 [-67.50, -42.71]	-67.50 [-68.06, -60.45]
		ITA (DP)	25.98 [19.90, 40.20]	25.19 [-9.08, 27.26]	-66.69 [-68.79, -43.71]	-68.65 [-84.81, -43.71]	30.85 [21.04, 39.68]	19.13 [2.21, 26.56]	-66.69 [-70.41, -42.89]	-70.41 [-77.28, -68.65]
		MST (Forehead)	4 [2, 5]	5 [4, 8]	8 [8, 10]	8 [8, 10]	2 [1, 5]	4 [4, 8]	8 [7, 9]	8 [8, 9]
		MST (DP)	2 [2, 6]	5 [2, 8]	8 [8, 9]	9 [8, 9]	2 [1, 6]	4 [2, 8]	8 [7, 10]	8 [8, 10]
		%Mod	2.70 [2.00, 4.08]	2.40 [1.40, 6.50]	1.20 [0.60, 2.32]	0.60 [0.40, 1.50]	3.00 [1.60, 5.70]	5.00 [3.70, 6.90]	0.90 [0.50, 2.50]	0.60 [0.50, 1.50]
CMS Zacurate 500C	15	ITA (Forehead)	38.90 [31.95, 41.13]	16.59 [-10.06, 22.77]	-66.82 [-70.27, -43.66]	-66.82 [-71.49, -51.44]	35.61 [31.95, 41.13]	14.45 [3.56, 23.18]	-56.14 [-66.82, -43.66]	-64.00 [-70.27, -56.14]
		ITA (DP)	28.92 [15.00, 32.22]	4.95 [-20.26, 27.24]	-57.66 [-72.04, -43.71]	-57.66 [-81.62, -43.71]	28.92 [11.48, 32.22]	12.50 [-1.84, 27.24]	-66.28 [-68.79, -54.85]	-61.80 [-72.04, -54.85]
		MST (Forehead)	3 [2, 5]	5 [4, 8]	8 [8, 10]	9 [8, 10]	5 [2, 5]	6 [4, 8]	8 [8, 10]	8 [8, 10]
		MST (DP)	2 [2, 5]	5 [3, 8]	8 [8, 10]	8 [8, 10]	2 [2, 6]	6 [3, 8]	8 [8, 10]	8 [8, 10]

		%Mod	0.80 [0.50, 1.40]	2.85 [1.00, 4.60]	1.50 [0.70, 4.50]	2.60 [0.60, 5.70]	1.00 [0.50, 1.90]	4.20 [2.80, 6.50]	2.40 [1.20, 5.60]	2.20 [1.20, 5.30]
Covidien (2022) Nellcor PM10N	14	ITA (Forehead)	38.03 [35.32, 40.55]	9.18 [-18.79, 24.39]	-58.69 [-58.94, -40.55]	-58.94 [-71.89, -58.69]	40.55 [38.03, 45.92]	9.18 [8.70, 23.88]	-51.48 [-58.69, -40.55]	-58.69 [-58.94, -51.48]
		ITA (DP)	25.48 [21.63, 33.04]	6.69 [-35.93, 17.06]	-63.93 [-68.35, -59.38]	-63.93 [-84.81, -60.10]	28.92 [21.98, 36.05]	10.44 [-5.93, 17.13]	-59.38 [-63.93, -30.35]	-60.10 [-63.93, -44.20]
		MST (Forehead)	4 [1, 5]	5 [5, 8]	8 [8, 10]	9 [8, 10]	2 [1, 5]	6 [4, 8]	8 [8, 10]	8 [8, 10]
		MST (DP)	4 [2, 5]	8 [2, 8]	8 [8, 10]	9 [8, 10]	2 [1, 5]	6 [4, 8]	8 [7, 10]	8 [8, 10]
		%Mod	1.00 [0.70, 1.40]	1.00 [0.70, 2.60]	2.20 [0.70, 4.70]	2.80 [0.30, 5.80]	1.30 [0.70, 2.60]	3.80 [2.50, 5.90]	2.30 [0.70, 6.30]	5.30 [1.00, 7.40]
Utech UT100	39	ITA (Forehead)	34.88 [33.85, 45.61]	9.82 [3.37, 24.93]	-45.26 [-51.72, -39.52]	-51.72 [-60.55, -51.23]	34.88 [32.81, 38.90]	23.74 [5.59, 25.65]	-51.72 [-60.55, -45.26]	-60.55 [-71.89, -51.72]
		ITA (DP)	31.36 [16.63, 33.04]	5.56 [-9.95, 14.79]	-51.76 [-68.35, -44.50]	-51.76 [-73.00, -44.50]	31.36 [16.63, 33.04]	12.84 [-4.70, 22.26]	-56.96 [-68.35, -44.50]	-51.76 [-73.00, -44.50]
		MST (Forehead)	4 [1, 4]	5 [4, 8]	9 [8, 9]	9 [9, 9]	2 [1, 4]	5 [4, 8]	9 [8, 10]	9 [8, 10]
		MST (DP)	5 [2, 5]	5 [4, 8]	8 [8, 9]	8 [8, 8]	4 [2, 5]	5 [2, 8]	8 [8, 10]	8 [8, 10]
		%Mod	0.90 [0.70, 1.10]	2.20 [0.80, 4.20]	0.70 [0.50, 1.60]	0.50 [0.30, 0.70]	1.10 [0.80, 3.10]	2.50 [1.20, 4.20]	1.20 [0.70, 3.30]	1.20 [0.60, 3.70]
Nihon Kohden OLV-4202 with NK-SPO2-ADVT Probe	64.03	ITA (Forehead)	35.61 [32.01, 41.15]	24.34 [14.05, 26.72]	-45.26 [-67.50, -39.83]	-67.50 [-71.89, -53.32]	39.55 [35.61, 44.83]	22.45 [22.06, 25.88]	-42.71 [-67.97, -39.83]	-67.97 [-67.97, -53.32]
		ITA (DP)	28.64 [23.42, 30.22]	19.13 [9.19, 26.58]	-45.82 [-77.28, -29.87]	-77.28 [-84.81, -45.82]	25.98 [23.42, 31.50]	18.38 [10.64, 27.24]	-68.77 [-69.25, -20.30]	-69.25 [-77.28, -45.82]
		MST (Forehead)	3 [2, 5]	5 [5, 6]	8 [8, 10]	8 [8, 10]	2 [1, 4]	5 [4, 8]	8 [7, 10]	9 [8, 10]
		MST (DP)	4 [1, 6]	5 [4, 6]	8 [8, 10]	8 [8, 10]	4 [1, 6]	4 [4, 8]	9 [7, 10]	9 [8, 10]
		%Mod	1.00 [0.80, 1.80]	5.60 [4.07, 7.60]	1.50 [0.80, 5.00]	5.40 [4.40, 6.00]	1.30 [0.80, 2.70]	5.10 [3.30, 7.00]	2.50 [1.30, 4.40]	4.00 [1.00, 5.80]

EDAN (2020) H100b	44	ITA (Forehead)	34.88 [32.81, 40.55]	11.06 [-7.51, 16.79]	-51.23 [-51.72, -38.43]	-51.72 [-58.69, -51.23]	43.48 [38.40, 45.93]	12.82 [8.89, 21.49]	-47.91 [-58.69, -31.93]	-58.69 [-61.99, -51.72]
		ITA (DP)	24.45 [21.98, 33.04]	-7.35 [-31.34, 22.26]	-63.93 [-68.35, -56.96]	-63.93 [-73.00, -44.50]	34.54 [21.98, 36.20]	10.44 [-13.62, 18.58]	-63.93 [-75.10, -43.61]	-63.93 [-75.95, -59.07]
		MST (Forehead)	4 [2, 4]	8 [4, 8]	9 [8, 9]	9 [9, 9]	2 [1, 5]	5 [4, 8]	9 [7, 10]	9 [9, 10]
		MST (DP)	4 [2, 5]	4 [2, 8]	8 [7, 9]	8 [8, 9]	2 [1, 4]	5 [4, 8]	9 [7, 10]	8 [8, 10]
		%Mod	1.00 [0.80, 1.80]	1.70 [0.80, 3.60]	0.60 [0.30, 1.60]	0.50 [0.30, 0.70]	1.40 [0.80, 3.70]	3.00 [1.80, 4.20]	1.30 [0.50, 4.50]	0.60 [0.30, 2.20]
Shenzhen WIT-S300C	62	ITA (Forehead)	43.48 [42.16, 43.83]	20.78 [-18.79, 25.95]	-54.63 [-58.25, -47.91]	-57.20 [-66.82, -54.63]	43.48 [38.91, 45.29]	17.97 [3.35, 23.02]	-54.63 [-58.25, -47.91]	-57.20 [-66.82, -54.63]
		ITA (DP)	30.22 [20.94, 30.29]	11.33 [-49.67, 32.46]	-62.80 [-67.15, -45.82]	-57.66 [-67.37, -45.82]	30.29 [26.36, 38.64]	9.73 [-11.27, 17.06]	-62.80 [-67.15, -52.59]	-57.66 [-67.37, -45.82]
		MST (Forehead)	2 [1, 3]	5 [2, 8]	8 [8, 10]	8 [8, 10]	2 [1, 4]	5 [2, 8]	8 [8, 10]	8 [8, 10]
		MST (DP)	2 [2, 4]	5 [3, 8]	8 [8, 10]	8 [8, 10]	2 [2, 4]	5 [2, 8]	8 [8, 10]	8 [8, 10]
		%Mod	1.40 [0.50, 3.90]	4.40 [3.20, 6.30]	3.30 [1.30, 6.30]	2.40 [1.10, 6.00]	2.90 [1.20, 4.50]	1.60 [0.80, 4.80]	2.80 [1.20, 6.00]	2.40 [1.10, 6.00]
CONTEC CMS50M	17	ITA (Forehead)	35.61 [31.22, 43.48]	10.74 [7.42, 20.78]	-51.44 [-70.27, -42.57]	-66.82 [-71.49, -51.44]	43.48 [35.39, 45.29]	17.78 [8.44, 20.08]	-51.44 [-70.27, -42.57]	-66.82 [-71.49, -51.44]
		ITA (DP)	30.00 [16.47, 30.95]	11.33 [-4.47, 14.12]	-57.66 [-72.04, -53.68]	-57.66 [-81.62, -43.71]	30.00 [16.47, 32.22]	11.33 [-0.58, 14.12]	-66.28 [-72.04, -53.68]	-57.66 [-81.62, -43.71]
		MST (Forehead)	4 [2, 4]	5 [4, 7]	8 [8, 10]	9 [8, 10]	2 [2, 4]	6 [4, 7]	8 [8, 10]	9 [8, 10]
		MST (DP)	4 [1, 4]	4 [4, 7]	8 [8, 10]	8 [8, 10]	2 [1, 4]	4 [4, 7]	8 [8, 10]	8 [8, 10]
		%Mod	1.60 [0.50, 5.60]	1.60 [0.90, 4.40]	2.60 [1.20, 5.30]	2.60 [0.60, 5.70]	2.20 [0.70, 5.90]	2.10 [0.90, 5.20]	3.20 [1.40, 5.70]	2.60 [0.60, 5.70]
Shenzhen A310L	18	ITA (Forehead)	32.79 [31.95, 40.24]	16.59 [3.56, 22.77]	-56.88 [-66.82, -43.66]	-66.82 [-66.82, -53.32]	35.96 [32.16, 43.50]	18.03 [3.56, 22.77]	-53.32 [-64.00, -43.66]	-64.00 [-66.82, -58.94]

		ITA (DP)	18.64 [11.48, 28.64]	18.51 [-1.84, 27.42]	-63.15 [-68.79, -45.82]	-57.66 [-68.65, -45.82]	28.64 [15.00, 37.31]	18.51 [-1.84, 27.42]	-61.80 [-68.65, -45.82]	-61.80 [-68.65, -57.66]
		MST (Forehead)	3 [1, 5]	5 [2, 7]	8 [8, 10]	9 [8, 10]	2 [1, 5]	5 [2, 8]	8 [8, 10]	8 [8, 10]
		MST (DP)	2 [2, 5]	5 [3, 7]	9 [8, 10]	9 [9, 10]	2 [2, 5]	5 [3, 8]	8 [8, 10]	9 [8, 10]
		%Mod	2.70 [0.80, 4.50]	2.60 [0.90, 5.40]	3.20 [1.20, 6.00]	4.00 [1.50, 7.40]	3.70 [1.20, 5.80]	2.00 [0.80, 5.20]	2.40 [1.20, 5.80]	3.40 [1.80, 6.10]
Masimo (2020) Mightysat Rx	7	ITA (Forehead)	41.13 [35.61, 43.50]	10.74 [6.40, 21.96]	-53.32 [-66.82, -48.46]	-54.63 [-66.82, -53.32]	41.96 [37.61, 43.50]	18.20 [10.74, 22.06]	-53.32 [-66.35, -42.71]	-56.14 [-66.82, -53.32]
		ITA (DP)	28.92 [18.64, 30.95]	5.06 [-4.47, 14.12]	-53.68 [-62.80, -43.71]	-45.82 [-57.66, -43.71]	28.92 [25.06, 30.95]	10.26 [1.33, 23.54]	-53.68 [-62.80, -43.61]	-54.85 [-74.08, -43.71]
		MST (Forehead)	2 [1, 5]	6 [2, 8]	9 [8, 10]	9 [8, 10]	2 [1, 5]	5 [2, 8]	8 [7, 10]	8 [8, 10]
		MST (DP)	2 [1, 4]	4 [3, 8]	8 [8, 10]	8 [8, 10]	2 [1, 4]	4 [3, 8]	8 [7, 10]	8 [8, 10]
		%Mod	1.20 [0.50, 2.73]	0.90 [0.60, 3.20]	2.00 [0.90, 6.60]	2.10 [0.70, 6.70]	1.80 [0.80, 4.20]	2.00 [0.70, 6.00]	2.30 [1.30, 5.50]	2.40 [0.80, 6.30]
Masimo Rad-G with Rad-G Finger Clip Sensor	55.03	ITA (Forehead)	41.53 [40.97, 47.58]	15.16 [-9.75, 21.28]	-55.36 [-63.53, -43.61]	-63.53 [-73.37, -55.36]	42.16 [39.55, 46.26]	15.16 [-7.62, 19.62]	-55.36 [-66.35, -33.46]	-66.35 [-73.37, -63.53]
		ITA (DP)	35.24 [28.10, 36.70]	2.02 [-6.41, 15.82]	-57.14 [-61.09, -24.57]	-61.09 [-70.41, -39.31]	34.58 [24.20, 37.28]	2.02 [-6.17, 10.26]	-43.61 [-70.41, -28.63]	-70.41 [-74.08, -61.09]
		MST (Forehead)	4 [2, 4]	5 [4, 7]	8 [8, 8]	8 [8, 8]	2 [1, 4]	5 [3, 7]	8 [8, 8]	8 [8, 8]
		MST (DP)	2 [1, 5]	6 [4, 7]	8 [7, 9]	8 [8, 8]	2 [1, 5]	7 [3, 7]	8 [7, 9]	8 [8, 9]
		%Mod	2.80 [1.70, 4.30]	4.40 [2.60, 6.20]	1.30 [0.50, 2.80]	1.50 [0.60, 3.80]	3.00 [1.70, 4.50]	4.60 [2.30, 7.50]	1.60 [0.70, 3.10]	1.60 [0.70, 3.50]
Bistos Co., Ltd. BT-710	61	ITA (Forehead)	40.55 [38.90, 43.48]	9.18 [-11.97, 9.22]	-45.26 [-60.55, -42.57]	-60.55 [-71.89, -53.32]	39.30 [35.61, 42.67]	10.74 [-5.29, 23.02]	-51.44 [-58.25, -43.66]	-57.20 [-60.55, -53.32]
		ITA (DP)	21.98 [16.63, 30.00]	9.73 [-28.79, 10.44]	-51.76 [-68.79, -45.82]	-51.76 [-84.81, -45.82]	24.39 [19.90, 30.22]	8.17 [-11.27, 14.12]	-43.71 [-67.15, -31.01]	-51.76 [-67.15, -43.71]

		MST (Forehead)	2 [1, 4]	6 [2, 8]	8 [8, 10]	9 [9, 10]	2 [1, 5]	6 [2, 8]	8 [8, 10]	8 [8, 10]
		MST (DP)	2 [2, 5]	8 [3, 8]	8 [7, 10]	7 [7, 10]	4 [2, 5]	5 [2, 8]	8 [7, 10]	8 [7, 10]
		%Mod	1.00 [0.80, 1.70]	5.00 [2.90, 7.80]	1.10 [0.70, 2.30]	2.10 [0.20, 5.10]	1.00 [0.70, 1.90]	2.00 [0.90, 5.20]	1.10 [0.60, 2.10]	1.20 [0.50, 2.40]
BodyMed BDMOXMTRBLK	2	ITA (Forehead)	35.61 [31.54, 43.48]	17.98 [6.40, 21.96]	-53.32 [-54.63, -42.57]	-54.63 [-66.82, -53.32]	40.24 [32.61, 42.16]	12.31 [8.44, 20.78]	-53.32 [-58.94, -43.04]	-54.63 [-66.82, -53.32]
		ITA (DP)	30.22 [28.64, 31.16]	5.06 [-4.47, 19.25]	-53.68 [-62.80, -45.82]	-45.82 [-57.66, -10.91]	30.22 [28.64, 30.95]	11.33 [-0.58, 13.71]	-57.66 [-66.28, -45.82]	-45.82 [-60.10, -43.71]
		MST (Forehead)	5 [2, 5]	6 [2, 8]	9 [8, 9]	9 [8, 9]	2 [2, 5]	6 [2, 8]	8 [8, 10]	8 [8, 10]
		MST (DP)	4 [1, 4]	5 [4, 8]	9 [8, 9]	9 [8, 9]	2 [1, 4]	5 [4, 8]	9 [8, 9]	9 [8, 9]
		%Mod	0.80 [0.50, 3.40]	1.00 [0.60, 3.40]	1.90 [0.90, 5.90]	1.70 [0.30, 6.90]	0.90 [0.50, 3.42]	1.10 [0.60, 4.80]	2.20 [1.10, 5.10]	1.80 [0.80, 4.70]
Masimo Rad-97 with Adult Probe, Product Code 4050	60	ITA (Forehead)	40.55 [38.91, 43.48]	8.36 [-11.97, 21.34]	-47.91 [-58.69, -43.04]	-58.69 [-71.89, -51.72]	39.55 [32.95, 43.48]	8.70 [-2.97, 18.20]	-51.44 [-58.69, -39.60]	-58.69 [-68.03, -51.72]
		ITA (DP)	21.98 [11.95, 30.22]	1.76 [-28.79, 29.93]	-66.28 [-68.35, -63.93]	-63.93 [-84.81, -44.50]	28.17 [19.19, 35.95]	4.95 [-11.27, 20.73]	-60.53 [-66.28, -44.20]	-57.66 [-63.93, -43.83]
		MST (Forehead)	3 [1, 5]	5 [5, 8]	9 [8, 10]	9 [9, 10]	2 [1, 5]	5 [4, 7]	8 [7, 9]	9 [8, 9]
		MST (DP)	2 [2, 5]	5 [3, 7]	9 [8, 10]	9 [8, 10]	2 [1, 5]	5 [2, 7]	9 [7, 10]	8 [8, 10]
		%Mod	1.30 [0.70, 2.20]	2.20 [0.70, 3.20]	1.40 [0.40, 4.60]	0.80 [0.40, 4.60]	2.00 [0.80, 3.80]	3.60 [1.80, 6.90]	2.30 [1.00, 5.30]	2.60 [0.60, 5.30]
Guangdong Biolight M800	43	ITA (Forehead)	35.85 [32.81, 39.30]	8.44 [-18.79, 21.49]	-51.72 [-58.69, -43.66]	-58.69 [-71.89, -51.72]	38.90 [32.81, 43.50]	15.79 [3.37, 23.74]	-51.48 [-58.25, -43.66]	-58.25 [-58.69, -51.72]
		ITA (DP)	25.36 [16.63, 36.13]	5.56 [-49.67, 23.49]	-44.50 [-63.93, -37.96]	-63.93 [-84.81, -44.50]	30.22 [18.64, 36.05]	5.56 [-9.95, 23.49]	-44.20 [-63.94, -31.01]	-63.93 [-67.15, -44.20]
		MST (Forehead)	4 [2, 5]	8 [5, 8]	9 [8, 10]	9 [9, 10]	2 [1, 5]	6 [4, 8]	8 [8, 10]	8 [8, 10]
		MST (DP)	3 [2, 5]	5 [5, 8]	8 [8, 10]	9 [8, 10]	2 [2, 5]	6 [2, 8]	8 [7, 10]	8 [7, 10]

		%Mod	1.00 [0.70, 1.50]	2.70 [1.40, 4.60]	0.80 [0.60, 2.50]	0.70 [0.30, 4.00]	1.20 [0.80, 3.50]	2.70 [1.30, 4.20]	1.20 [0.60, 3.40]	1.90 [0.80, 6.00]
Masimo Rad-G with Adult Probe, Product Code 4050	55	ITA (Forehead)	35.61 [34.22, 39.30]	25.65 [9.18, 25.95]	-58.69 [-66.82, -45.26]	-66.82 [-71.89, -58.69]	37.54 [35.32, 45.92]	16.79 [8.89, 25.65]	-51.44 [-66.82, -40.73]	-58.96 [-71.89, -53.25]
		ITA (DP)	24.45 [11.95, 29.97]	22.66 [10.44, 31.81]	-60.79 [-73.00, -56.96]	-63.93 [-84.81, -57.66]	29.44 [17.18, 36.20]	17.13 [-5.02, 23.49]	-57.66 [-73.00, -38.43]	-63.93 [-73.00, -43.71]
		MST (Forehead)	4 [1, 5]	5 [2, 8]	9 [8, 10]	9 [9, 10]	2 [1, 5]	6 [2, 8]	9 [7, 10]	9 [8, 10]
		MST (DP)	4 [2, 5]	5 [3, 6]	8 [8, 10]	8 [8, 10]	4 [2, 5]	5 [2, 7]	8 [8, 10]	8 [8, 10]
		%Mod	1.20 [0.90, 2.00]	3.80 [1.90, 5.90]	1.10 [0.50, 6.20]	4.70 [0.40, 6.90]	1.50 [0.90, 2.60]	2.50 [1.30, 4.60]	1.30 [0.60, 4.90]	1.50 [0.50, 5.30]
Guangdong Biolight M70	37	ITA (Forehead)	35.61 [35.61, 43.48]	16.59 [3.35, 20.08]	-53.32 [-60.55, -43.66]	-60.55 [-71.89, -53.32]	39.37 [35.61, 43.50]	12.31 [-7.51, 16.85]	-53.32 [-58.25, -48.46]	-57.20 [-60.55, -51.94]
		ITA (DP)	30.00 [23.42, 30.95]	1.33 [-11.29, 11.33]	-51.76 [-62.80, -45.82]	-51.76 [-84.81, -45.82]	30.22 [23.42, 33.29]	-0.58 [-34.75, 11.33]	-54.85 [-62.80, -44.20]	-54.85 [-67.15, -44.60]
		MST (Forehead)	2 [2, 5]	5 [4, 7]	9 [8, 10]	9 [9, 10]	2 [1, 5]	6 [4, 8]	8 [8, 10]	8 [8, 10]
		MST (DP)	4 [1, 5]	6 [4, 8]	8 [8, 10]	8 [8, 10]	2 [1, 5]	6 [4, 8]	8 [8, 10]	8 [8, 10]
		%Mod	0.80 [0.60, 1.30]	1.50 [0.70, 3.60]	3.20 [0.80, 5.40]	3.20 [0.30, 4.50]	1.60 [0.80, 2.70]	1.80 [0.90, 4.40]	2.40 [1.20, 5.90]	2.60 [1.10, 6.30]
Acare (2023) AH-M1	57	ITA (Forehead)	35.85 [32.95, 39.30]	9.22 [-5.29, 20.76]	-51.72 [-60.55, -42.00]	-60.55 [-71.89, -51.72]	38.90 [35.61, 43.48]	9.18 [-5.29, 15.79]	-51.48 [-58.25, -40.73]	-58.25 [-60.55, -51.72]
		ITA (DP)	25.36 [17.18, 29.97]	1.39 [-11.27, 11.64]	-59.38 [-66.69, -51.76]	-51.76 [-84.81, -44.50]	28.92 [17.18, 33.04]	-7.35 [-27.34, 9.73]	-63.97 [-67.37, -44.50]	-51.76 [-67.37, -44.50]
		MST (Forehead)	4 [2, 5]	5 [5, 8]	9 [8, 10]	9 [9, 10]	2 [1, 5]	6 [4, 8]	8 [7, 10]	8 [8, 10]
		MST (DP)	5 [2, 5]	5 [5, 8]	8 [7, 10]	8 [8, 10]	4 [1, 5]	6 [4, 8]	8 [7, 10]	8 [8, 10]
		%Mod	1.50 [0.80, 3.20]	1.90 [0.50, 4.40]	1.10 [0.30, 5.40]	0.70 [0.30, 4.50]	2.30 [0.80, 4.50]	2.70 [0.70, 4.90]	1.60 [0.70, 6.00]	1.40 [0.80, 5.80]

Acare (2020) AH-MX	42	ITA (Forehead)	35.61 [32.81, 38.66]	11.06 [-2.37, 21.49]	-51.23 [-51.72, -38.43]	-51.72 [-58.69, -51.23]	35.61 [32.81, 45.61]	11.06 [-2.37, 21.49]	-46.24 [-57.20, -40.73]	-57.20 [-58.69, -52.64]
		ITA (DP)	25.48 [16.63, 36.13]	5.56 [-31.34, 31.71]	-63.93 [-68.35, -56.96]	-63.93 [-73.00, -44.50]	30.22 [21.72, 36.13]	5.56 [-10.86, 22.28]	-63.93 [-68.35, -36.92]	-67.15 [-70.18, -44.50]
		MST (Forehead)	4 [1, 5]	5 [4, 8]	9 [8, 9]	9 [9, 9]	4 [1, 5]	6 [4, 8]	8 [7, 10]	8 [8, 10]
		MST (DP)	4 [2, 5]	6 [2, 8]	8 [7, 9]	8 [8, 9]	4 [1, 5]	6 [2, 8]	8 [7, 10]	8 [7, 10]
		%Mod	1.00 [0.80, 1.60]	1.20 [0.60, 2.23]	0.60 [0.30, 1.60]	0.50 [0.30, 0.70]	1.00 [0.60, 2.60]	2.10 [1.00, 4.10]	1.10 [0.60, 2.40]	1.10 [0.60, 2.20]
American Oxygen LK-87	20	ITA (Forehead)	39.56 [32.79, 43.50]	16.59 [8.36, 21.85]	-61.48 [-70.27, -48.46]	-66.82 [-71.89, -61.48]	39.76 [32.79, 43.50]	13.52 [8.75, 22.77]	-64.00 [-66.82, -42.57]	-66.82 [-70.27, -64.00]
		ITA (DP)	19.90 [15.00, 30.00]	8.24 [-10.51, 21.30]	-68.79 [-72.04, -57.66]	-57.66 [-84.81, -56.25]	19.90 [15.00, 30.00]	1.76 [-13.62, 17.13]	-61.80 [-68.79, -57.66]	-61.80 [-72.04, -57.66]
		MST (Forehead)	2 [1, 5]	5 [5, 8]	9 [8, 10]	9 [8, 10]	2 [1, 5]	5 [4, 8]	8 [8, 10]	8 [8, 10]
		MST (DP)	4 [1, 5]	5 [5, 8]	9 [8, 10]	8 [8, 10]	2 [1, 5]	5 [4, 8]	8 [7, 10]	8 [8, 10]
		%Mod	2.50 [0.90, 3.70]	4.20 [1.40, 5.60]	3.30 [2.00, 6.00]	2.80 [1.50, 5.70]	2.70 [1.20, 4.60]	3.70 [1.40, 5.40]	2.80 [1.60, 5.90]	2.70 [1.70, 6.10]

Distribution of Individual Typology Angle (ITA) at the forehead and dorsal distal phalanx (DP), Monk Skin Tone (MST) at the forehead and DP, and percent modulation (%Mod), an indicator of participant perfusion —measured by the Nellcor PM1000N and scaled by 1/10—across four ITA-defined skin color bins: light (ITA > 30°), medium (-30° ≤ ITA ≤ 30°), dark (ITA < -30°), and very dark (ITA < -50°); note that “very dark” is a subset of the dark bin. Data for the 24 participant cohorts are shown on the left and data for the larger cohorts meeting anticipated 2025 U.S. Food and Drug Administration (FDA) and International Organization for Standardization (ISO) enrollment criteria are reported on the right. ITA and %Mod are reported as median [Q1, Q3], and MST as mode [minimum, maximum].