Pulse Oximeters: Technology, Accuracy Limitations, and Regulation

James J. Lee Ph.D., Division Director
Neel Patel M.Eng. Senior Reviewer

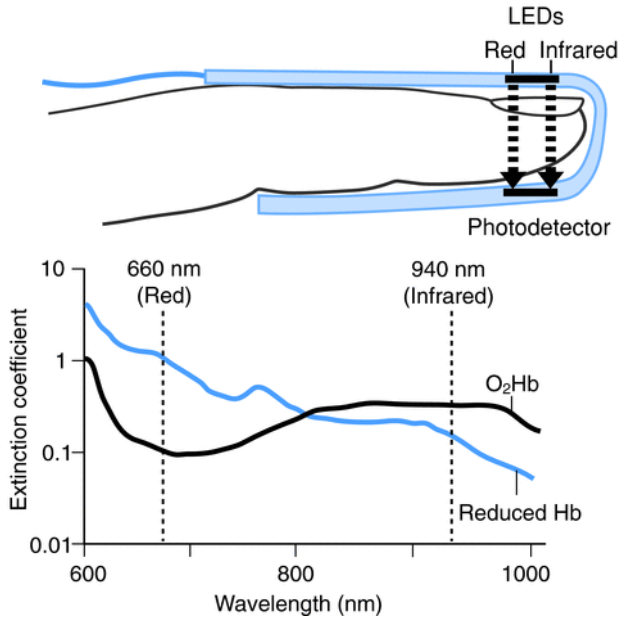
DHT1C: Division of Sleep Disordered Breathing, Respiratory and Anesthesia
OHT1: Ophthalmic, Anesthesia, Respiratory, ENT and Dental Devices

Fingertip Pulse Oximetry

- Wide use by healthcare providers, healthcare facilities, and patients to obtain an indirect measure (SpO_2) of arterial blood oxygen saturation (SaO_2)
  - Surrogate measurement of oxygen level in hemoglobin
  - SaO_2 measurement, obtained via arterial puncture, is considered the gold standard for assessment of blood oxygen saturation levels.
  - SpO_2 is an estimate of how much oxygen the hemoglobin contains compared to how much it could contain, expressed as a percentage
  - Pulse oximetry is a non-invasive and quick alternative to arterial puncture for estimating oxygen saturation

Principles of Operation

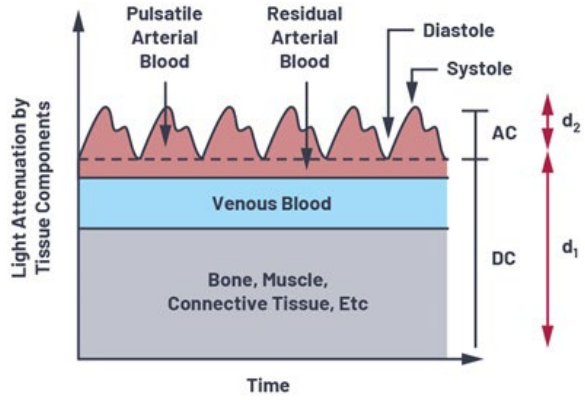
- Oxygenated and deoxygenated hemoglobin have different absorption spectra
- Photodetector used to measure the differential absorption of 2 or more wavelengths of light (typically red 660 nm and infrared 940 nm)
- SpO2 estimated as percentage of oxygenated hemoglobin to oxygenated + deoxygenated hemoglobin
  - Does not account for presence of dysfunctional hemoglobin



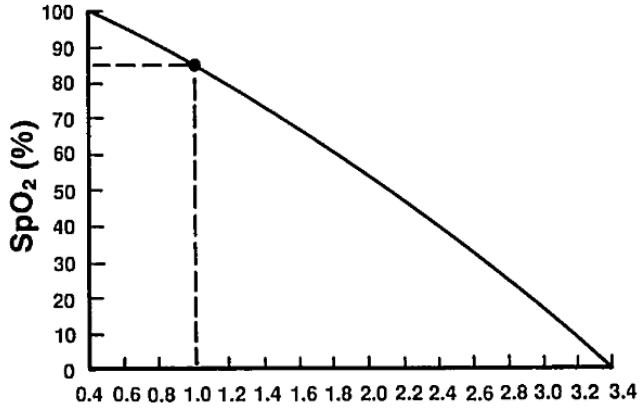
Xiong, Z., & Kodali, B. (2011). Pulse oximetry and capnography. In C. Vacanti, S. Segal, P. Sikka, & R. Urman (Eds.), *Essential Clinical Anesthesia* (pp. 186-190). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511842306.028)

Principles of Operation

- Absorption ratios change with pulse with arterial blood
- Ratio of absorption ratios converted to SpO2 using algorithm and lookup table developed with calibration data



https://www.analog.com/ru/technical-articles/how-to-design-a-better-pulse-oximeter.htmld text



$$R = \frac{AC_{660}/DC_{660}}{AC_{940}/DC_{940}}$$

Deshmane, Anagha. (2009). False arrhythmia alarm suppression using ECG, ABP, and photoplethysmogram.

- Pulse oximeters can be categorized as:

  - Prescription Use Pulse Oximeters: Regulated under product codes DQA (Oximeter), DPZ (Oximeter, Ear) and NLF (Reprocessed Oximeter)

    - Reviewed by the FDA through 510(k) pathway; currently prescription use only
    - Undergo clinical testing to confirm their accuracy
    - Used to monitor (i.e., trending or spot checking) oxygen saturation levels of patients, most often in hospitals and doctors' offices, although they may sometimes be prescribed for home use.
    - Recognized Consensus Standards 1-139 ISO 80601-2-61 Particular requirements for basic safety and essential performance of pulse oximeter equipment
    - Pulse Oximeters - Premarket Notification Submissions [510(k)s]: Guidance for Industry and Food and Drug Administration Staff (2013)

  - Over-the-Counter (OTC) Pulse Oximeters: Regulated under product codes PGJ and OCH

    - Most commonly intended for general wellness or sporting/aviation uses and are not intended for medical purposes.
    - They are often sold directly to consumers in stores or online and may utilize mobile medical apps intended for estimating oxygen saturation for non-medical purposes.
    - See FDA's guidance document *General Wellness: Policy for Low-Risk Devices* for additional information.

Prescription use pulse oximeters are Class II devices intended to measure blood oxygen saturation levels and are regulated under:

- **21 CFR 870.2700:** Oximeter (product codes: DQA and NLF). An oximeter is a device used to transmit radiation at a known wavelength(s) through blood and to measure the blood oxygen saturation based on the amount of reflected or scattered radiation. It may be used alone or in conjunction with a fiberoptic oximeter catheter.

- **21 CFR 870.2710:** Ear Oximeter, product code DPZ. An ear oximeter is an extravascular device used to transmit light at a known wavelength(s) through blood in the ear. The amount of reflected or scattered light as indicated by this device is used to measure the blood oxygen saturation level.

https://www.fda.gov/medical-devices/overview-device-regulation/classify-your-medical-device

The regulations 21 CFR 870.2700 and 870.2710 include devices using reflectance, transmittance, and fiber optic technologies. Prescription use pulse oximeters measure the amount of transmitted, reflected, and scattered light through various application sites (e.g., finger, ear, foot, hand, forehead, back, and nose).

- Some of the factors that can impact the accuracy of pulse oximeters include (but not limited to):
  - Skin pigmentation
  - Dyshemoglobinemias: disorders in which the hemoglobin molecule is functionally altered and prevented from carrying oxygen.
  - Severe anemia: disorder in which the blood has reduced ability to carry oxygen. Anemia occurs when there are not enough healthy red blood cells to carry oxygen to the body organs.
  - Low perfusion: reduced peripheral blood flow and subsequent reduction in the detectable signal at pulse oximeter sensor site.
  - Dyes
  - Nail polish
  - Ambient light

Medical Device Reporting (MDRs)

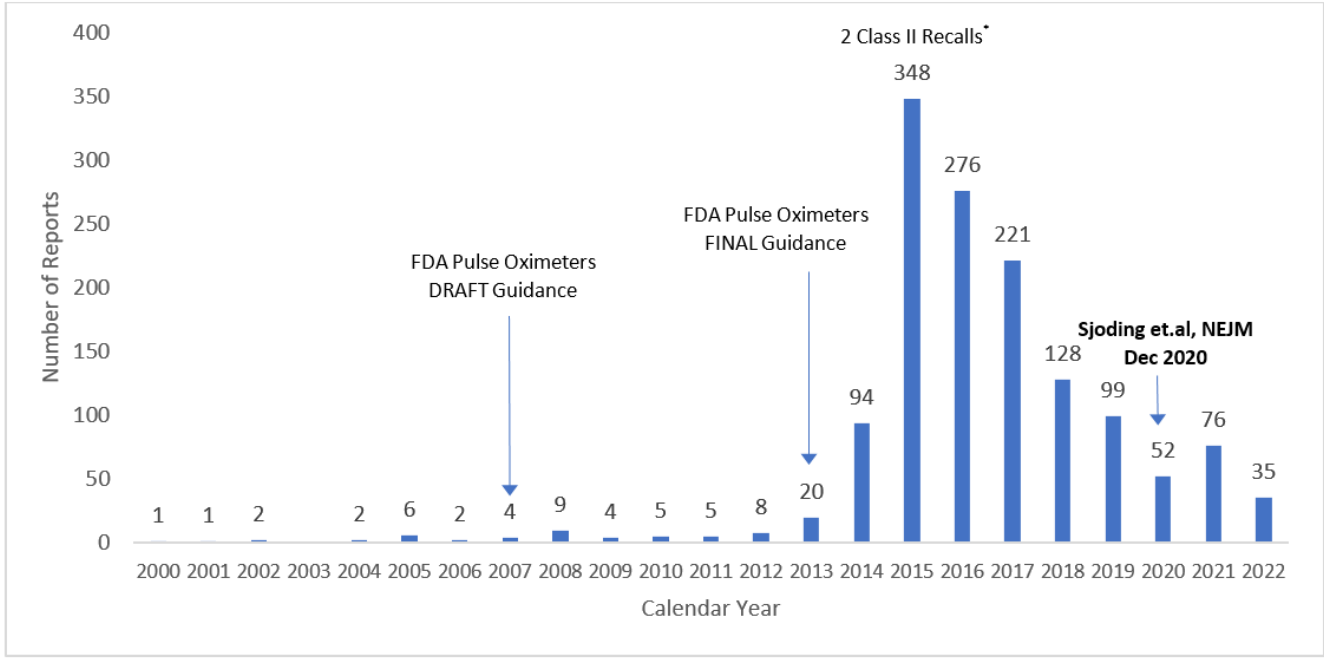
- Search for product codes DQA (Oximeter), DPZ (Oximeter, Ear) and NLF (Reprocessed Oximeter)
- January 1, 2000 to August 29, 2022
- Searched for term “skin”
- Manual review of death reports and those with term “skin” to assess:
  - If potentially related to inaccurate SpO2 reading
  - Any information that could inform potential source of problem
- Total 635 adverse events reports, including 296 deaths

	count
Potentially associated with SpO2 inaccurate reading	**99**
Mentions African American Race	3
Mentions Skin Pigmentation	3
Mentions Low Perfusion	2
Reports Device continued to provide readings after patient death	15
Insufficient information to assess reason for inaccuracy	83

Lacked sufficient information to assess association between use of device and adverse event, and the source for inaccurate reading

Medical Device Reporting (MDRs)

- Additional search for terms “inaccurate”, “inaccuracy”, and “incorrect”
- Number of reports by search term
  - Inaccurate: 626
  - Inaccuracy: 12
  - Incorrect: 2, 244
- Manual review to assess
  - If potentially related to inaccurate SpO2 reading
  - Any information that could inform potential source of problem



*Jul 27, 2015: Coviden Class II recall due to potential missing segments on the display that could result in misinterpretation of data, impacted 317,257 distributed units; and Nov 25, 2015: Massimo Class II recall due to sensors manufactured with incompatible configurations that could result in sensors that will provide either no readings or inaccurate readings, impacted 3,476 distributed units worldwide.

Figure 1 Number of Adverse Event Reports Related to Inaccurate SpO_2 Readings Submitted to the Agency by Year of Receipt, January 1, 2000 through August 29th, 2022, n=1,398

Medical Device Reporting (MDRs)

	count
Potentially associated with SpO2 inaccurate reading	**1,398**
Low perfusion	10
High perfusion	11
African American	5
White	7
Hispanic	1
Non-Hispanic	4
Problem with device component	196
Insufficient information to assess reason for inaccuracy	1, 198

- Lacked sufficient information to assess association between use of device and adverse event, and the source for inaccurate reading
- Top 3 reported health effects relate to no consequence to patient
- Top 3 device problems relate to inaccurate readings.

Top 10 Patient and Device Problems Reported in Adverse Event Reports Related to Inaccurate Pulse Oximeter Readings, n=1,398

Health Effect Clinical Code (Problem Code)	Count*	Device Problem Code (Problem Code)	Count*
No consequence or impact to patient (2199)	666	Incorrect measurement (1383)	1,187
No known impact or consequence to patient (2692)	462	Incorrect, Inadequate, or imprecise result or reading (1535)	72
No clinical signs, symptoms, or conditions (4582)	92	Low readings (2460)	49
No patient involvement (2645)	47	Device operates differently than expected (2913)	37
Low oxygen saturation (2477)	36	Incorrect or inadequate test results (2456)	34
No information (3190)	33	Device stops intermittently (1599)	32
Loss of pulse (2562)	16	Device displays incorrect message (2591)	30
Cyanosis (1798)	8	Device sensing problem (2917)	29
Hypoxia (1918)	7	Unable to obtain readings (1516)	21
Respiratory distress (2045)	7	High Readings (2459)	19

*Categories are not exclusive of each other, reports can include more than one problem code, therefore the numbers add up to more than 1,398

Summary

- Pulse oximeter provide immediate noninvasive estimates on oxygen saturation
- Limits of utility exist in measurements and clinical implications


FDA
U.S. FOOD & DRUG
ADMINISTRATION

U.S. FOOD & DRUG ADMINISTRATION

Standards for Pulse Oximeters: ISO 80601-261: 2017

Sandy Weininger, PhD
Senior Electrical/Biomedical Engineer – electrical safety gatekeeper
Co-chair ISO TC 121/SC3 JWG10 (with IEC 62D JWG5) - Oximeters
Co-chair AAMI Interoperability WG

Division of Biomedical Physics (DBP)
Office of Science and Engineering Laboratories (OSEL)
Center for Devices and Radiological Health (CDRH)

Pulse Oximeter clinical and engineering oversight

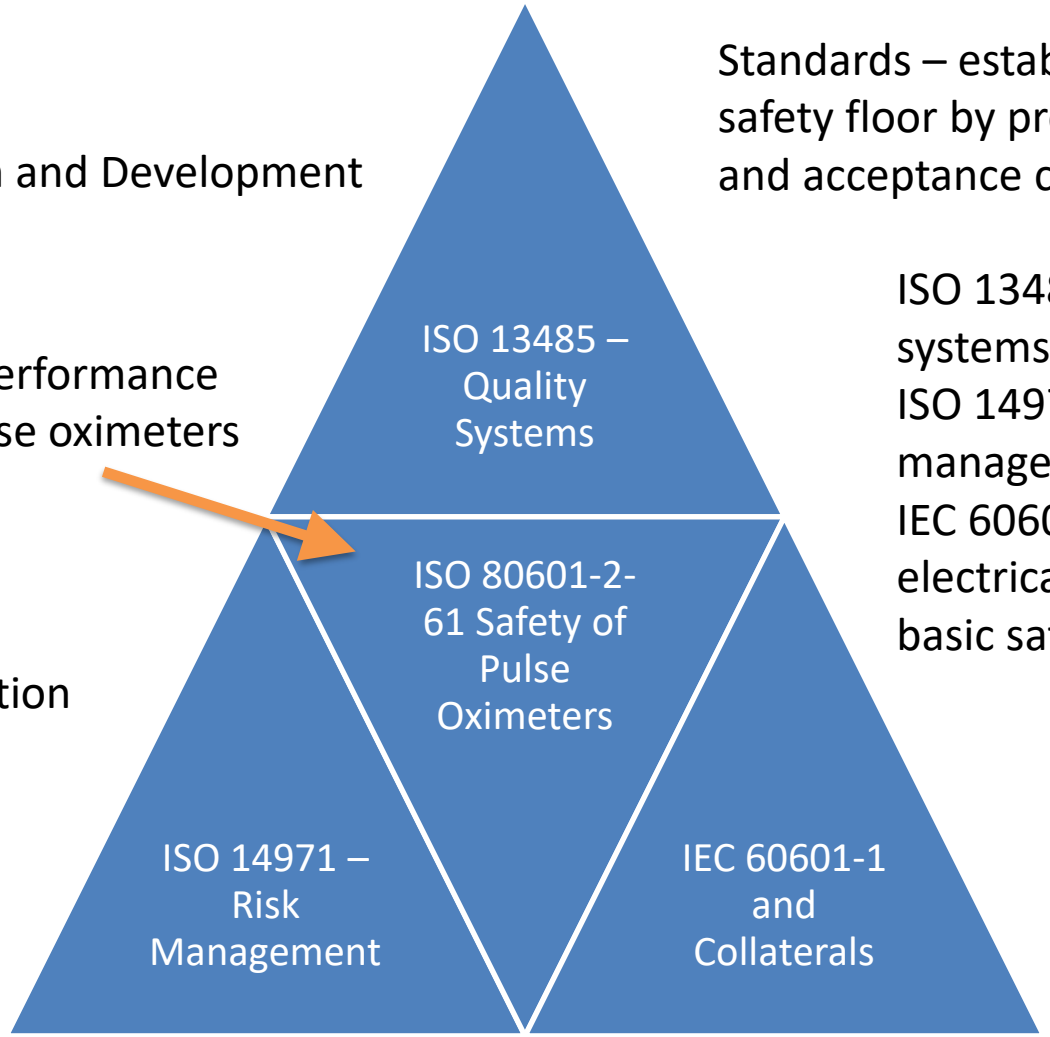
- 1982 – first marketing of pulse oximeter
- 1986 – American Society of Anesthesiologists: Standards for Basic Anesthesia Monitoring – “Blood oxygenation: During all anesthetics, a quantitative method of assessing oxygenation such as pulse oximetry shall be employed”
- 1992 – ASTM F1415-92 (and 2000) Standard Specification for Pulse Oximeters
- 1992 – FDA guidance - Noninvasive pulse oximeters and ear oximeters
- 1998 – ASA: amended standards to include oximetry
- 2000 - ASTM F1415-2000
- 2005 – Clinical and Laboratory Standards Institute (CLSI) – HS3-A-Pulse Oximetry; Approved Guideline
- 2005 – ISO 9919-2005 (IEC 60601-2-54)
- 2007 – FDA draft guidance - 7/19/2007
- 2011 – CLSI POCT11-A2 - Pulse Oximetry; Approved Guideline - Second Edition
- 2013 – FDA guidance
- 2010, 2020 – ASA: amended
- 2011, 2017, 202x – ISO/IEC 80601-2-61

Medical Device (Pulse Oximeter) Safety Standards

Process for Design and Development

Specific Safety and performance requirements for pulse oximeters

Process for Identification of unsafe situations



Standards – establish a minimum internationally accepted safety floor by providing consensus objective test methods and acceptance criteria

ISO 13485:2016 Medical devices — Quality management systems — Requirements for regulatory purposes
ISO 14971:2007 Medical devices — Application of risk management to medical devices
IEC 60601-1:2005+AMD1:2012+AMD2:2020 — Medical electrical equipment - Part 1: General requirements for basic safety and essential performance

Safety requirements – e.g. electrical, thermal, mechanical, biocompatibility

Scope of Pulse Oximeter standard

201.1.1 * Scope

Replacement:

This document applies to the *basic safety* and *essential performance* of *pulse oximeter equipment* intended for use on humans, hereafter referred to as *ME equipment*. This includes any part necessary for *normal use*, including the *pulse oximeter monitor*, *pulse oximeter probe*, and *probe cable extender*.

These requirements also apply to pulse oximeter equipment, including pulse oximeter monitors, pulse oximeter probes and probe cable extenders, which have been reprocessed.

The intended use of *pulse oximeter equipment* includes, but is not limited to, the **estimation of arterial oxygen haemoglobin saturation** and **pulse rate** of *patients* in professional healthcare institutions as well as *patients* in the *home healthcare environment* and the *emergency medical services environment*.

This document is not applicable to *pulse oximeter equipment* intended for use in laboratory research applications nor to oximeters that require a blood sample from the *patient*.

Hazards inherent in the intended physiological function of *ME equipment* or *ME systems* within the scope of this document are not covered by specific requirements in this document except in 201.11 and in 7.2.13 and IEC 60601-1:2005+AMD1:2012+AMD2:2020, 8.4.1.

The FDA has a process to recognize standards.

FDA Supplemental Information Sheet which details the extent of recognition

INTERNATIONAL STANDARD **ISO 80601-2-61**

Second edition
2017-12

Corrected version
2018-02

Medical electrical equipment —

Part 2-61:
Particular requirements for basic safety and essential performance of pulse oximeter equipment

Appareils électromédicaux —

Partie 2-61: Exigences particulières pour la sécurité de base et les performances essentielles pour les oxymètres de pouls

Part B: Supplementary Information Sheet (SIS)

FR Recognition List Number 050 **Date of Entry** 09/17/2018

FR Recognition Number 1-139

Standard (Included in ASCA pilot)

ISO 80601-2-61 Second edition 2017-12 (Corrected version 2018-02)
Medical electrical equipment - Part 2-61: Particular requirements for basic safety and essential performance of pulse oximeter equipment

Scope/Abstract

ISO 80601-2-61:2017 applies to the basic safety and essential performance of pulse oximeter equipment intended for use on humans, hereafter referred to as me equipment. This includes any part necessary for normal use, including the pulse oximeter monitor, pulse oximeter probe, and probe cable extender.

These requirements also apply to pulse oximeter equipment, including pulse oximeter monitors, pulse oximeter probes and probe cable extenders, which have been reprocessed.

The intended use of pulse oximeter equipment includes, but is not limited to, the estimation of arterial oxygen haemoglobin saturation and pulse rate of patients in professional healthcare institutions as well as patients in the home healthcare environment and the emergency medical services environment.

ISO 80601-2-61:2017 is not applicable to pulse oximeter equipment intended for use in laboratory research applications nor to oximeters that require a blood sample from the patient.

If a clause or subclause is specifically intended to be applicable to me equipment only, or to me systems only, the title and content of that clause or subclause will say so. If that is not the case, the clause or subclause applies both to me equipment and to me systems, as relevant.

Hazards inherent in the intended physiological function of me equipment or me systems within the scope of this document are not covered by specific requirements in this document except in 201.11 and in 7.2.13 and 8.4.1 of the general standard.

Extent of Recognition

Partial recognition. The following part(s) of the standard is (are) not recognized:

Clause 201.12.1.101.1 SpO2 ACCURACY of PULSE OXIMETER EQUIPMENT, Specification is not recognized.

Rationale for Recognition

This standard is relevant to medical devices and is recognized on its scientific and technical merit and/or because it supports existing regulatory policies.
This standard is recognized in part because "Clause 201.12.1.101.1 SpO2 ACCURACY of PULSE OXIMETER EQUIPMENT, Specification" is in conflict with an existing published final guidance. See Section 4.1 Accuracy of Pulse Oximeters, Table 3. Typical Arms Specification by Sensor Type, of the Guidance document cited below.

Accuracy spec not recognized

201.12.1.101 * *Sp*O2 ACCURACY of PULSE OXIMETER EQUIPMENT

201.12.1.101.1 * Specification

The SpO_2 ACCURACY of PULSE OXIMETER EQUIPMENT shall be a root-mean-square difference of less than or equal to 4,0 % SpO_2 over the range of 70 % to 100 % SaO_2. The SpO_2 shall be indicated as FUNCTIONAL OXYGEN SATURATION and shall not be indicated as FRACTIONAL OXYHAEMOGLOBIN.

The DECLARED RANGES of SpO_2 and SpO_2 ACCURACY over those ranges shall be disclosed in the instructions for use. The SpO_2 ACCURACY shall be stated over the range 70 % to 100 % (additional information is found in 201.12.1.101.2). SpO_2 ACCURACY information shall be accompanied by a note reminding the reader that, because PULSE OXIMETER EQUIPMENT measurements are statistically distributed, only about two-thirds of PULSE OXIMETER EQUIPMENT measurements can be expected to fall within $\pm A_{rms}$ of the value measured by a CO-OXIMETER. When a PULSE OXIMETER MONITOR is suitable for use with a variety of PULSE OXIMETER PROBES, SpO_2 ACCURACY information shall be made available for each type of PULSE OXIMETER PROBE.

The modified Bland and Altman plot (i.e., (SpO_2- SaO_2) versus SaO_2)[10][11] for each combination of PULSE OXIMETER PROBE and PULSE OXIMETER MONITOR listed in the instructions for use for all subjects pooled, including upper 95 % and lower 95 % limits of agreement shall at a minimum be provided to the RESPONSIBLE ORGANIZATION upon request. See Annex CC for an example:

Accuracy formula

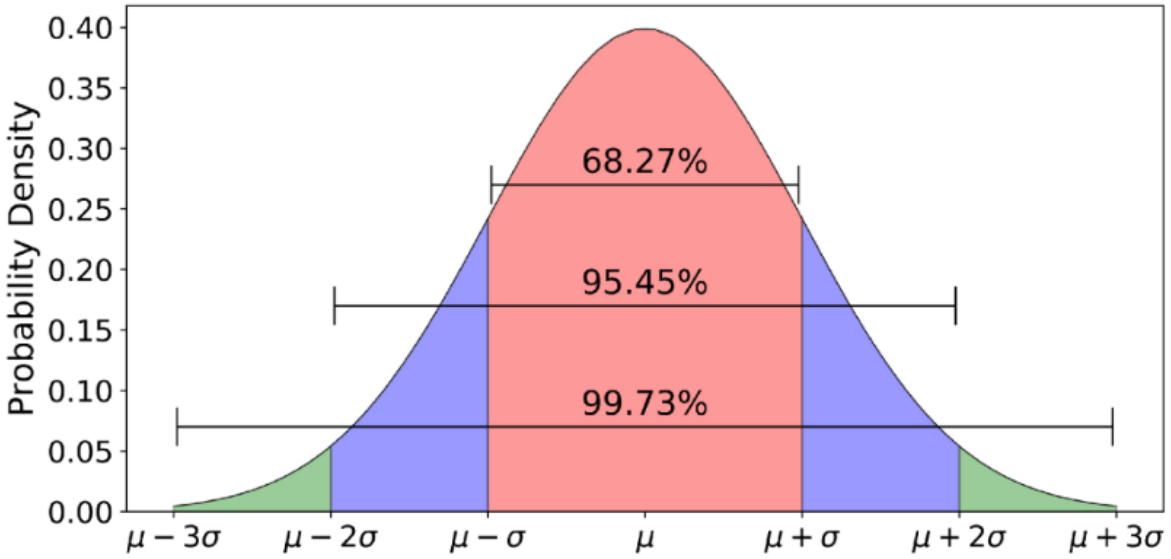
201.12.1.101.3 * Data analysis for determination of SpO_2 ACCURACY

For each range specified, SpO_2 ACCURACY of the PULSE OXIMETER EQUIPMENT shall be stated in terms of the root-mean-square (rms) difference between measured values (SpO_{2i}) and reference values (S_{Ri}), as given by Formula (1).

$$A_{rms} = \sqrt{\frac{\sum_{i=1}^{n} \left(SpO_{2i} - S_{Ri} \right)^2}{n}} \qquad (1)$$

SpO2 is an estimate of SaO2

- Arms behaves like a normal statistical distribution; one expects uncertainty:
  - For a Arms of 3%, if the true value, SaO2, is 92% then 95% of the SpO_2 estimates can be expected to fall within the interval [86% - 98%].
  - 5 of 100 patients [green tails] are expected to be OUTSIDE of this interval



Normal Distribution Definition (isixsigma.com)

Controlling variability in collection of verification data

Annex EE
(informative)

Guideline for evaluating and documenting *Sp*O2 ACCURACY in human subjects

EE.1 General

This annex is provided as a guideline for evaluating and documenting the SpO_2 ACCURACY of PULSE OXIMETER EQUIPMENT. The methods described in this annex are applicable to both new PULSE OXIMETER EQUIPMENT and modified PULSE OXIMETER EQUIPMENT or parts whenever human testing is required.

NOTE Subclause 201.12.1.101.2 requires that a study conducted to evaluate the SpO_2 ACCURACY of PULSE OXIMETER EQUIPMENT complies with ISO 14155:2011.

This annex describes testing methods for assessing the SpO_2 ACCURACY of PULSE OXIMETER EQUIPMENT. It does not prescribe medical practice, proper safety PROCEDURES or institutional review board (IRB) or ethics committee (EC) PROCESSES.

Inclusion criteria

201.12.1.101.4 Characteristics of the clinical study population for determination of SpO_2 ACCURACY

The summary of the clinical study report used to assess SpO_2 ACCURACY shall state whether the test subjects were sick or healthy and shall describe their skin colour, age and gender. This information shall be disclosed in the ACCOMPANYING DOCUMENT.

Check compliance by inspection of the ACCOMPANYING DOCUMENT.

Factors Influencing the Performance of Oximeters

201.12.1.101.2 * Data collection for determination of SpO_2 ACCURACY

The claims of *Sp*O2 accuracy shall be supported by controlled desaturation study measurements taken over the full range of *Sa*O2 values +3 % of the lower value and −3 % of the upper value for which *Sp*O2 accuracy is claimed.

EXAMPLE 1 A controlled desaturation study supporting a claimed range of *Sp*O2 accuracy from 70 % *Sa*O2 to 100 % *Sa*O2 can be supported with *Sa*O2 data collected over the range of 73 % *Sa*O2 to 97 % *Sa*O2.

The controlled desaturation study shall comply with the requirements of ISO 14155:2011.

The residual risk inherent in a controlled hypoxia study on healthy adult volunteers, can be reduced to a non-significant level by following recommended additional procedures[12].

The accuracy of pulse oximeter equipment for paediatric patients shall be supported via controlled desaturation study measurements on adult subjects. Paediatric subjects are a vulnerable population. Data points should be recorded with comparable density over the full range claimed.

NOTE 1 Additional information is found in Annex EE.

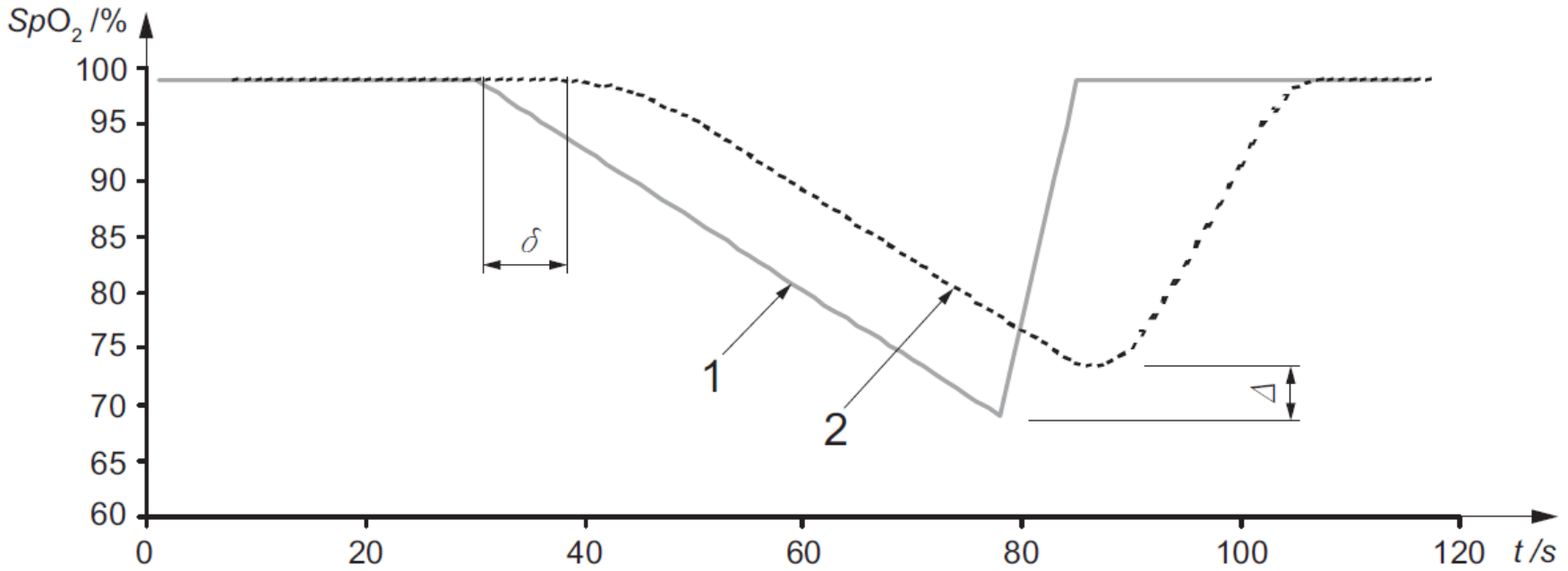
Any types of interference known to influence or affect the *Sp*O2 accuracy need not be stated as part of the *Sp*O2 accuracy specification, but shall be disclosed in the instructions for use.

EXAMPLE 2 Ambient light (including photodynamic therapy); physical movement (patient and imposed motion); diagnostic testing; low perfusion; electromagnetic interference; hf surgical equipment; dysfunctional haemoglobin; presence of certain dyes; inappropriate positioning of the pulse oximeter probe.

A summary of the test methods used to establish the *Sp*O2 accuracy claims shall be disclosed in the technical description.

Oximeter fidelity

Device performance and user settings may influence the fidelity of the displayed saturation value – especially when the interpretation when the saturation is changing



Key

1	SaO_2	Δ	is saturation deviation
2	displayed SpO_2	δ	is time delay
		SpO_2	is saturation
		t	is time

Figure GG.1 — Illustration of fidelity of PULSE OXIMETER EQUIPMENT performance in tracking saturation changes

Panel Question

Current labeling for prescription use pulse oximeters is intended for clinicians and generally it does not address inaccuracies that may be associated with skin pigmentation. In your deliberations please discuss:

a. Labeling modifications to address inaccuracies that may be associated with skin pigmentation.

b. Recommendations for the content of labeling for lay users who may use pulse oximeters at home.

Summary

- The Standard provides definitions and requirements that address hazardous situations found in pulse oximeters and
- Establishes test methods, acceptance criteria, and rationale to verify risk control measures are in place and effective and acceptable
- The Standard harmonizes with FDA's guidance document to support the regulation of pulse oximeters and assure reasonable safety
- Thank you all and especially distinguished panel members


FDA
U.S. FOOD & DRUG
ADMINISTRATION

A Systematic Literature Review of the Real-World Performance of Pulse Oximeters

Allison O'Neill, PhD, MA
Safety Signal Coordinator
Office of Health Technology 1
Center for Devices and Radiological Health

PURPOSE OF REVIEW

2013 | … | 2020 | 2021 | 2022

Pulse Oximeters - Premarket Notification Submissions [510(k)s]

Guidance for Industry and Food and Drug Administration Staff

Document issued on: March 4, 2013

This document supersedes Non-invasive Pulse Oximeter General Guidance Document, September 7, 1992
The draft of this document was issued on July 19, 2007.

For questions regarding this document contact Neel Patel at 301-796-5580 or neel.patel@fda.hhs.gov.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Devices and Radiological Health
Office of Device Evaluation
Division of Anesthesiology, General Hospital, Infection Control, and Dental Devices
Anesthesiology and Respiratory Devices Branch

COVID-19 Pandemic

Sjoding et al (2020)
Media attention

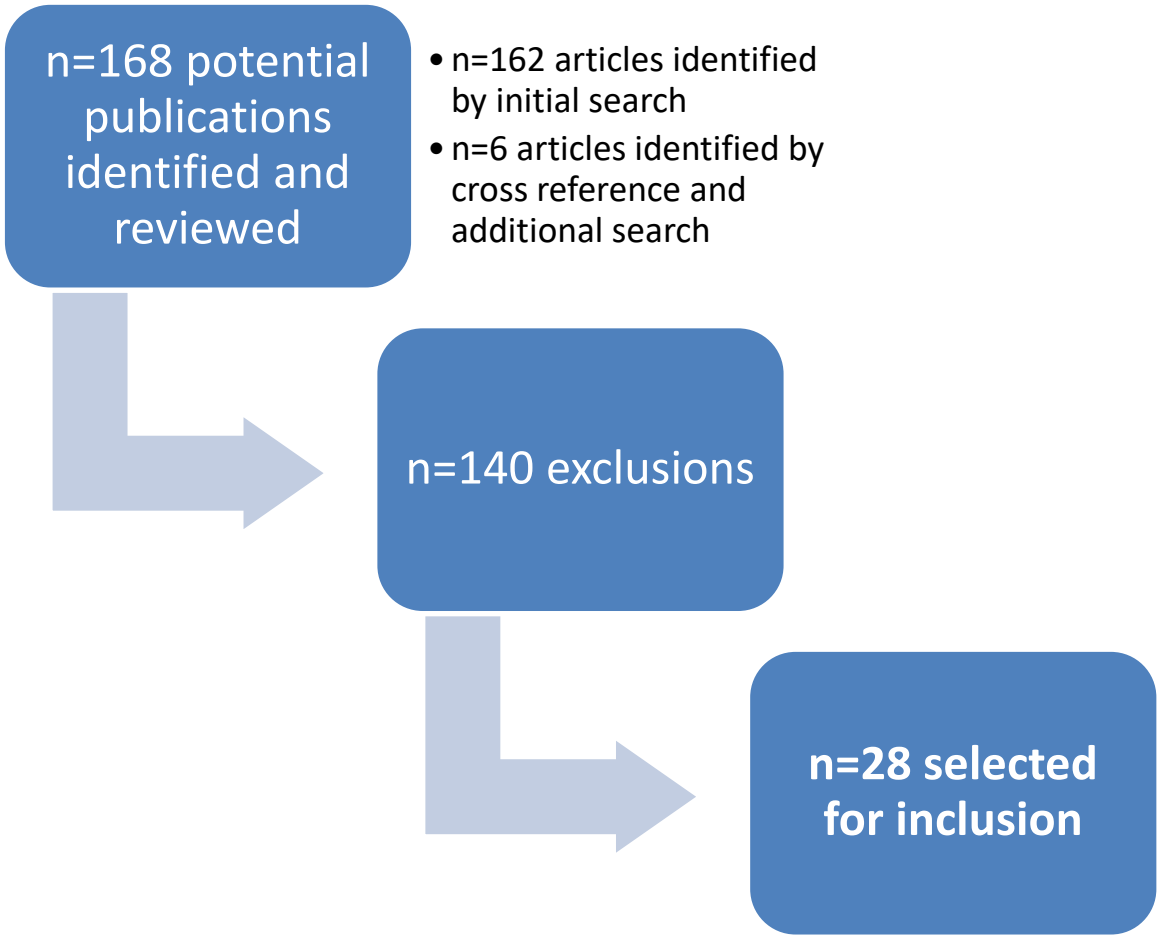
Pulse Oximeter Accuracy and Limitations: FDA Safety Communication

Nov 2022 Advisory Committee Meeting

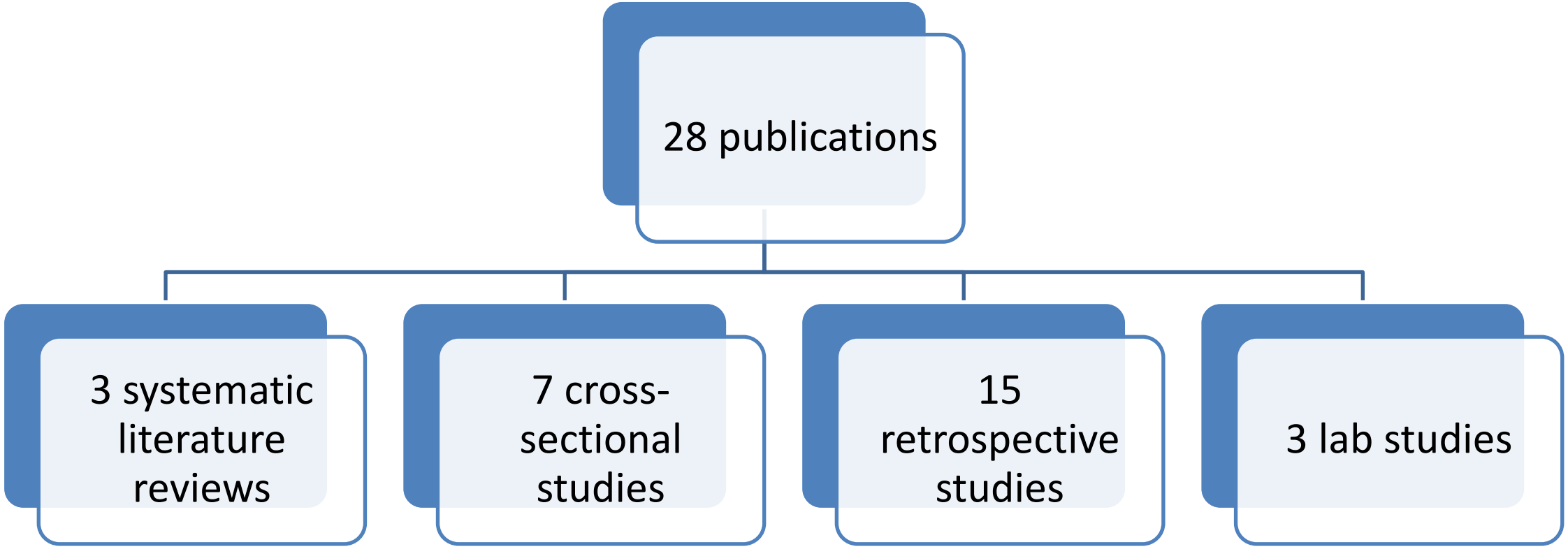
SYSTEMATIC LITERATURE REVIEW

Methodology

- PubMed search strategy: pulse oxim* AND (race OR racial OR pigment*)
- Inclusion criteria:
  - Publication date after FDA guidance document (2013)
  - Must include clinical data, with at least 9 participants
  - Must include measurement of either skin pigmentation or race/ethnicity



Results


28 publications
3 systematic literature reviews
7 cross-sectional studies
15 retrospective studies
3 lab studies

Systematic Reviews

Reference	Description
Cabanas 2022	• 41 articles published bet. Jan 1976 - Feb 2022 related to skin pigmentation • Of 11 studies at low risk of bias, 8 studies found inaccuracies due to skin pigmentation, and 3 reported no loss of accuracy • Reported overestimations in populations with darker skin pigmentation, especially at low saturation or hypoxemia conditions
Shi 2022	• 32 articles published before Dec 2021 related to skin pigmentation • Meta-analysis: Reported overestimations in people with high level of skin pigmentation (pooled mean bias 1.11%; 95% confidence interval 0.29 to 1.93%) and people described as Black/African American (1.52%; 0.95 to 2.09%) (moderate- and low-certainty evidence)
Poorzargar 2022	• 22 articles published before Dec 2020 related to poor perfusion • Only 1 study controlled for skin pigmentation, and none strictly followed FDA recommendations

Real World Studies: Cross Sectional

- Statistically significant association between skin pigmentation/race and occult hypoxemia in 5 of 7 studies
  - Henry 2022:
    - Adjusted OR (95%CI) of OH (SaO_2 < 88% despite concurrent SpO_2 > 92%)
      - White: Reference
      - Black: **1.65 (95%CI: 1.28, 2.14)**
      - Asian: 1.53 (95%CI: 0.95, 2,47)
      - American Indian: 1.31 (95%CI: 0.80, 2.16)
    - OH associated with in-hospital mortality: OR=2.96 (95%CI: 1.20, 7.28)
- No significant association in 2 of 7 studies

Source	Study Population (location, year(s) of data collection, age, health status)	Sample Size (patients, paired measures)	Skin Pigmentation Measurement or Proxy	Outcome Variable Definition	Reported Measure
Ebmeier 2018[46]	Aus/NZ, 2015, ≥16yo, ICU			Bias: Bland-Altman	Unadjusted regression coefficient (95% CI) Light: Reference Medium: 0.9 (0.4, 1.3) Dark: 2.4 (1.2, 3.6)
Foglia 2017[37]	US, 2013-2015, infants, Cyanotic Congenital Heart Disease	35 35	Munsell System Soil Color Chart	Bias: SpO_2-SaO_2	Device #1: mean bias (SD) Light: 0.2% (3.8%) Dark: 1.6% (4.8%) Device #2: mean bias (SD) Light: 3.0% (5.0%) Dark: 5.4% (5.1%)
Harskamp 2021[47]		35 234	Fitzpatrick	Mean bias in SpO_2 Accuracy measured by Arms, and Mean Absolute Error (MAE)	Mean bias range: -0.6 to -4.8 None of the pulse oximeters met Arms < 3% in SaO_2 range of 70-100%. MAE range: 2.3 to 5.1 and 5 of the pulse oximeters met < 3% Darker skin complexion associated with poorer SpO_2 performance
Henry 2022[6]		26,603 128,285	Race	OH: SaO_2 < 88% despite concurrent SpO_2 > 92%	Adjusted OR (95%CI) White: Reference Black: 1.65 (1.28, 2.14) Asian: 1.53 (0.95, 2,47) American Indian: 1.31 (0.80, 2.16)
Seitz 2022[71]					
Smith 2019[32]	South Africa, years NR, ≥18yo, surgical	220 220	Fitzpatrick	Bias: Bland-Altman	No significant differences
Stell 2022[48]					

Real World Studies: Retrospective

- Statistically significant association between race and OH/bias in 14/15 studies
  - Wong 2022:
    - Incidence of OH ($SaO_2 < 88\%$ despite $SpO_2 \geq 88\%$)
      - Asian: 4.9%
      - Black: 6.9%
      - Hispanic: 6.0%
      - White: 4.9%
    - Between group differences: $p<.001$
    - OH was associated with greater organ dysfunction and higher in-hospital mortality.

Source	Study Population (location, year(s) of data collection, age, health status)	Sample Size (patients, paired measures)	Skin Pigmentation Measurement or Proxy (categories, n, %)	Outcome Variable Definition	Association?
Andrist 2022[2]		1,061 9,023	Race	OH: SaO_2 < 88% despite SpO_2 > 92%	Yes
Bangash 2022[72]	UK, 2017-2021, adults, inpatient	16,818 20,231	Race	OH: SaO_2 < 94% despite SpO_2 ≥ 94%	Yes
Burnett 2022[3]	US, 2008-2019, ≥18yo, patients receiving anesthetic	46,253 151,070	Race/ethnicity	OH: SaO_2 < 88% despite SpO_2 ≥ 92%	Yes
Chesley 2022[35]	US, 2019-2021, adults (no age cutoff reported), ICU	7,693 105,467	Race/ethnicity	OH: SaO_2 < 88% despite SpO_2 between 92%-96%	Yes
Crooks 2022[73]	UK, 2020-2021, no age limit reported, inpatients with COVID-19	2,997 5,374	Race	Mean difference: SpO_2-SaO_2	Yes
Fawzy 2022[5]	US, 2020-2021, no age limit reported, Emergency department visit or hospitalized for COVID-19	1,216 32,282	Race/ethnicity	OH: SaO_2 < 88% despite concurrent SpO_2 of 92% to 96% Treatment Initiation: SpO2 ≤94% or use of supplemental oxygen	Yes
Gadrey 2022[36] *preprint	US, 2020-2021, ≥18yo, Emergency department visit or hospitalized for COVID-19	5,319 1,909,867	Race/ethnicity	Clinical deterioration (either transfer to ICU or in-hospital mortality)	Yes
Gottlieb 2022[74]	US, 2008-2019, no age limit reported, ICU	3,069 n/a	Race/ethnicity	Time-weighted average supplemental oxygen rate	Yes
Sjoding 2020[1]	US, 2014-2015 and 2020, no age limit reported, ICU	10,001 48,097	Race	OH: SaO_2 < 88% despite 92% ≥ SpO_2 ≥ 96%	Yes
Sudat 2022[75]	US, 2020-2021, adult (no age limit reported), Cohort 1 (hospital visits with ABG), Cohort 2 (emergency visits with COVID-19)	Cohort 1: 43,753 paired measures Cohort 2: 8,735 paired measures	Race/ethnicity	Cohort 1: SpO_2 SaO_2 pairs, OH Cohort 2: clinical and treatment characteristics	Yes
Valbuena 2022a[7]	US, 2019-2020, ≥18yo, patients on ECMO due to ARDS or COVID-19	372 372	Race/ethnicity	Pre-ECMO OH: SpO2 between 92 to 96% despite SaO_2 < 88%	Yes
Valbuena 2022b[76]	US, 2013-2019, US veterans (no age limit reported), inpatient excluding ICU	30,039 30,039	Race	OH: SpO_2 ≥ 92% despite SaO_2 < 88%	Yes
Vesoulis 2022[8]	US, 2012-2019, infants <32 weeks gestation, NICU	294	Race	OH: SaO_2 < 85% despite concurrent SpO_2 > 90%	Yes
Wiles 2022[77]	UK, 2020-2020, ≥16yo, COVID pneumonitis	194 6,216	Race	Bias: SpO_2-SaO_2	No
Wong 2021[9]	US, 2014-2021, no age limit reported, ICU	79,044 87,971	Race	OH: SaO_2 < 88% despite SpO_2 ≥ 88% Organ dysfunction, Length of hospital stay, In-hospital mortality	Yes

Lab Studies

- Baek 2018:
  - Large error in dark skin pigmentation subjects
- Mantri 2022:
  - No significant differences in SpO_2 by different skin types
  - Subjects with darker skin tones exhibit significantly higher photoacoustic signal
- Okunlola 2022:
  - Small positive bias in dark skin pigmentation group

Limitations of Literature

- Variable definitions
  - Occult Hypoxemia
  - Skin pigmentation vs. race/ethnicity
- Real world and mostly retrospective data
  - Time between paired measurements
  - Rely on self reported race/ethnicity
  - Residual confounding
- Heterogeneity of population
  - Sick patients vs. healthy volunteers
  - Prevalence of hypoxemia impacts the positive and negative predictive values (e.g., see Pennello 2022)
- Heterogeneity of technology
  - Brand of oximeter not always reported
  - Technology changes over time
- Publication bias

Literature Summary

- Mounting real-world evidence from literature that suggests that pulse oximeter accuracy may vary by self-reported race, and skin pigmentation
- Need for prospective studies that:
  - utilize standardized measurement of skin pigmentation
  - capture simultaneous measurement of SaO_2 and SpO_2 paired data
  - systematically collect data on important confounders

Panel Question

Please discuss the clinical evidence from the scientific literature about the accuracy of pulse oximetry among patients with darker skin pigmentation. In your deliberations consider the strengths and limitations of the studies, including study design, outcome definitions, and potential confounding factors that can impact interpretation of the evidence. Specifically, please address:

a. Does the currently available clinical evidence demonstrate disparate performance in patients with darker skin pigmentation? If so, do you believe such disparate performance may lead to increased risks? Please include prescription use and OTC pulse oximeters (when used for medical purposes) in your deliberations.

b. Do you believe the reported disparate performance or increased risks may be explained by factors other than darker skin pigmentation such as perfusion index, motion artifacts?


FDA
U.S. FOOD & DRUG
ADMINISTRATION

Overview of Desaturation Studies in Pulse Oximeter 510(k) Submissions

Mary Jung, PhD, MPH, Epidemiologist
Office of Clinical Evidence and Analysis (OCEA)
Center for Devices and Radiological Health (CDRH)

Study Background & Objectives

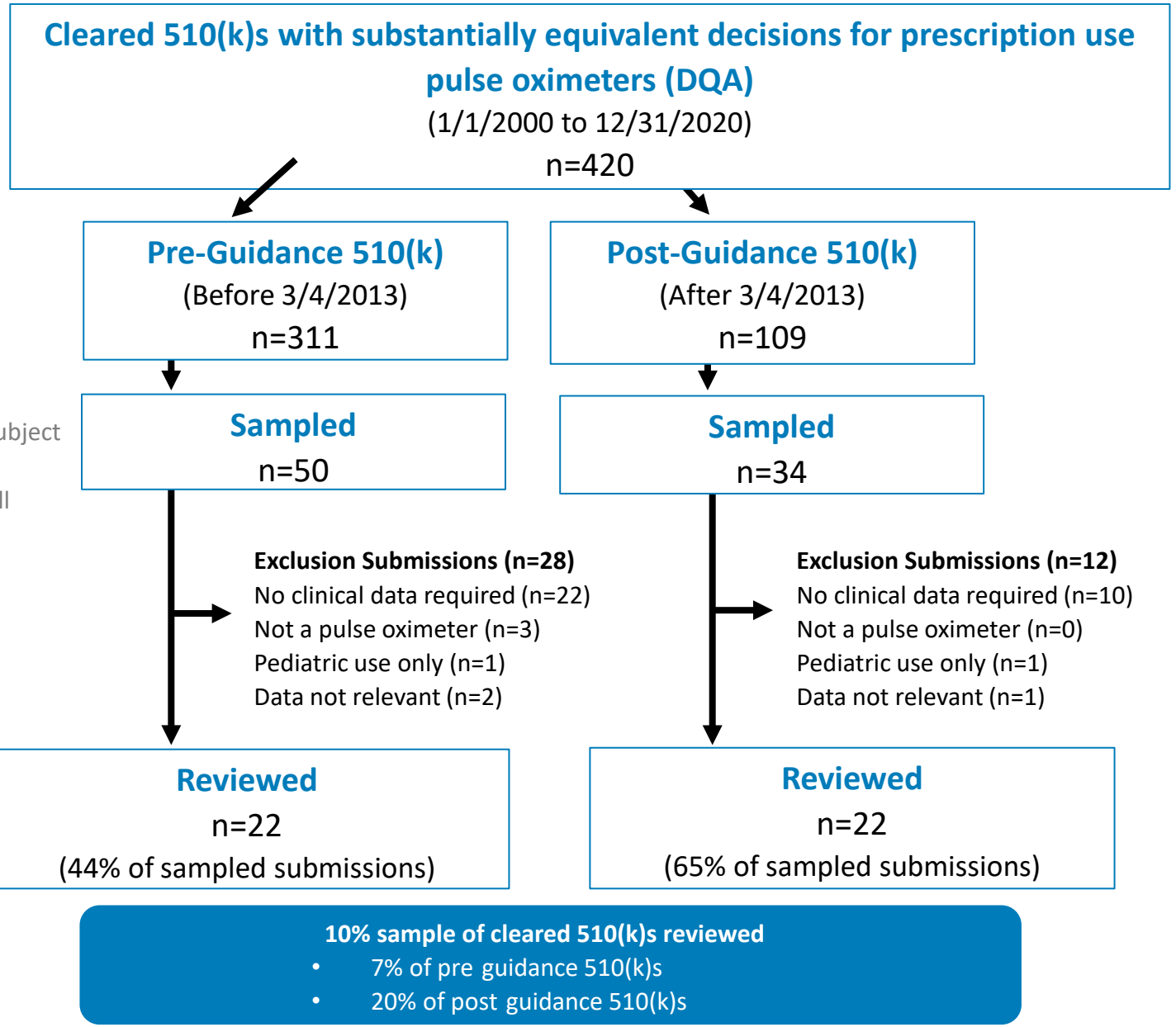
- **Background:** Need to better understand skin pigmentation information in cleared 510(k) submissions given concerns about the potential impact on pulse oximeter accuracy
- Study performed to evaluate cleared 510(k) submissions for prescription use pulse oximeters
- Findings to provide an initial examination of available skin pigmentation information within cleared submissions

Study Objectives

Among cleared 510(k) submissions for prescription use pulse oximeters (DQA):

- Objective 1: To examine premarket clinical study characteristics and reporting of skin pigmentation classification
- Objective 2: To assess reporting of factors potentially impacting device accuracy, including skin pigmentation, in device labeling

Methods – Sample Ascertainment



1. Random selection
2. Review for eligibility by Subject Matter Experts
3. Random selection to fulfill ~10% target, excluding ineligible 510(k)s
4. Assessment of eligibility

Methods – Data Elements

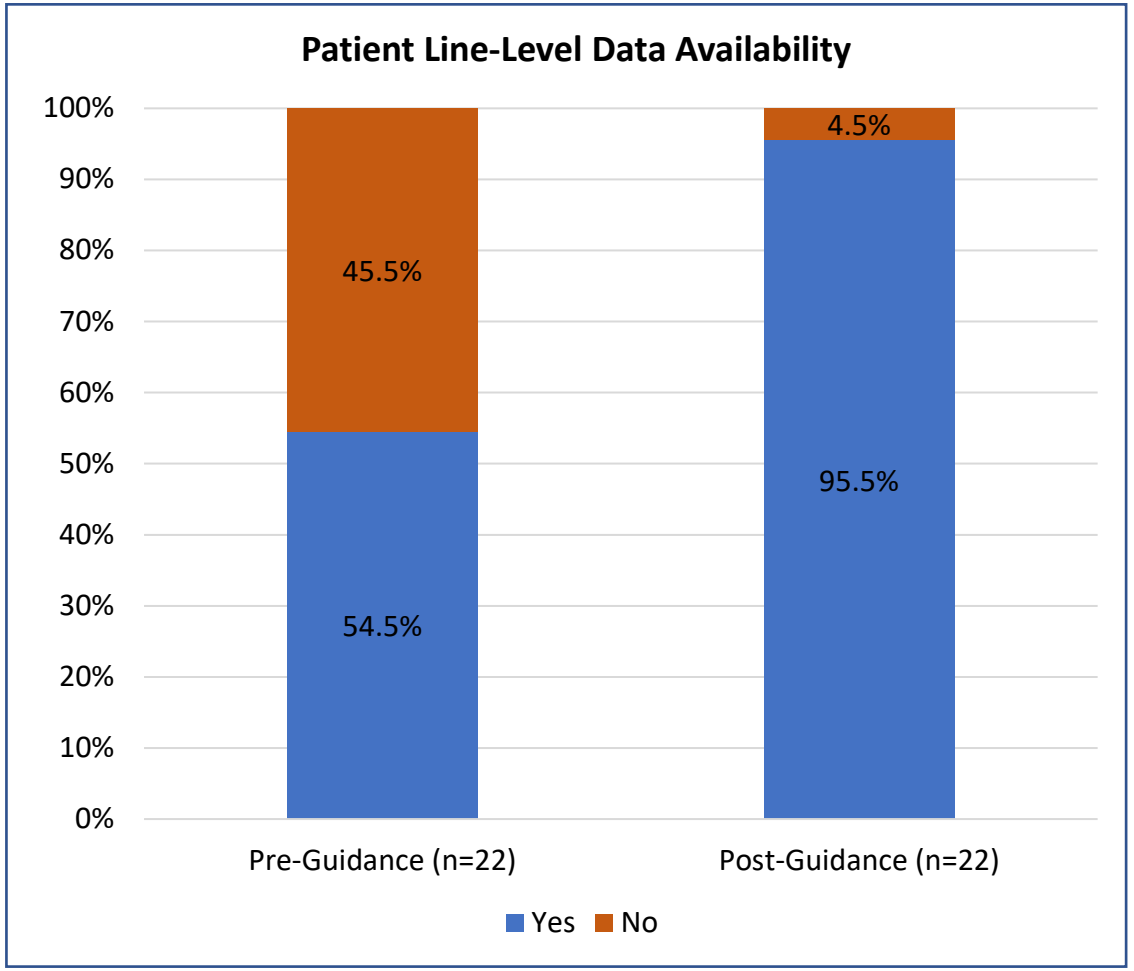
Extracted Data Elements

	Source	Extracted Data Elements	
Objective 1	Clinical studies	• Number of subjects • Number of paired oxygen measures per subject • Availability of patient line-level data • Participant inclusion criteria	• Reporting of participant characteristics: age, sex, race/ethnicity, nationality, weight, height, finger size • Use of Bland-Altman plots • Skin pigmentation assessment and categories
Obj 2	Device labeling	• Factors listed in the labeling that may impact device accuracy	

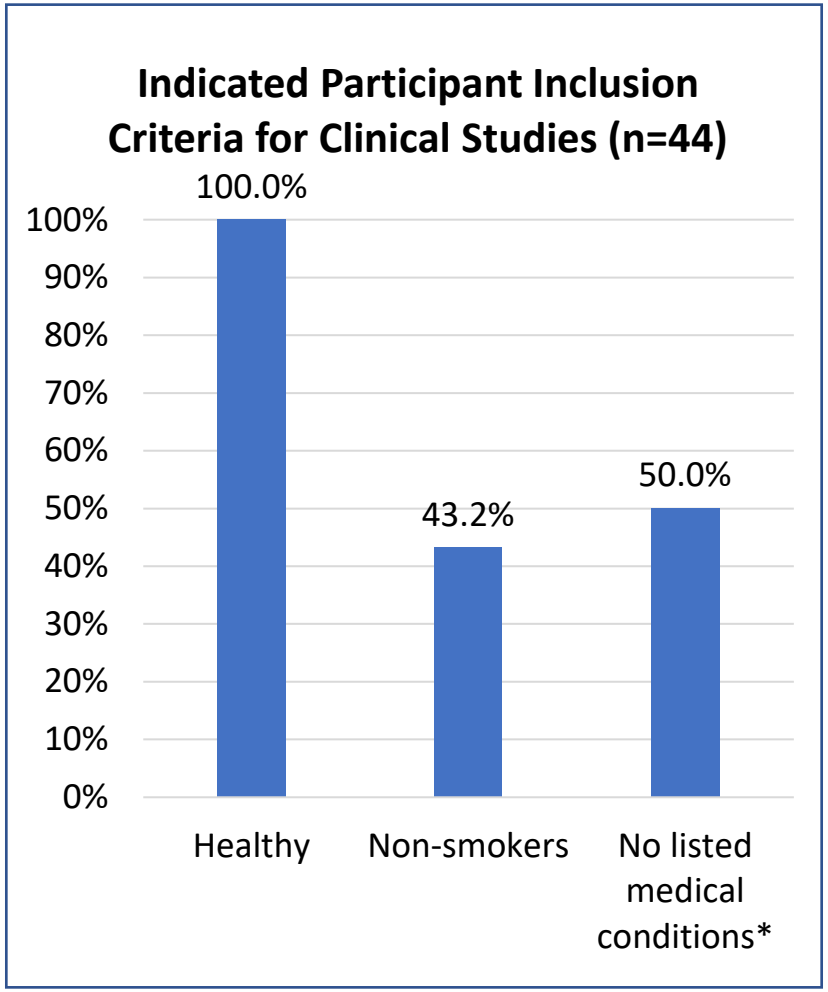
- Descriptive analysis performed to examine distributions of data elements, including overall and for pre- and post-guidance

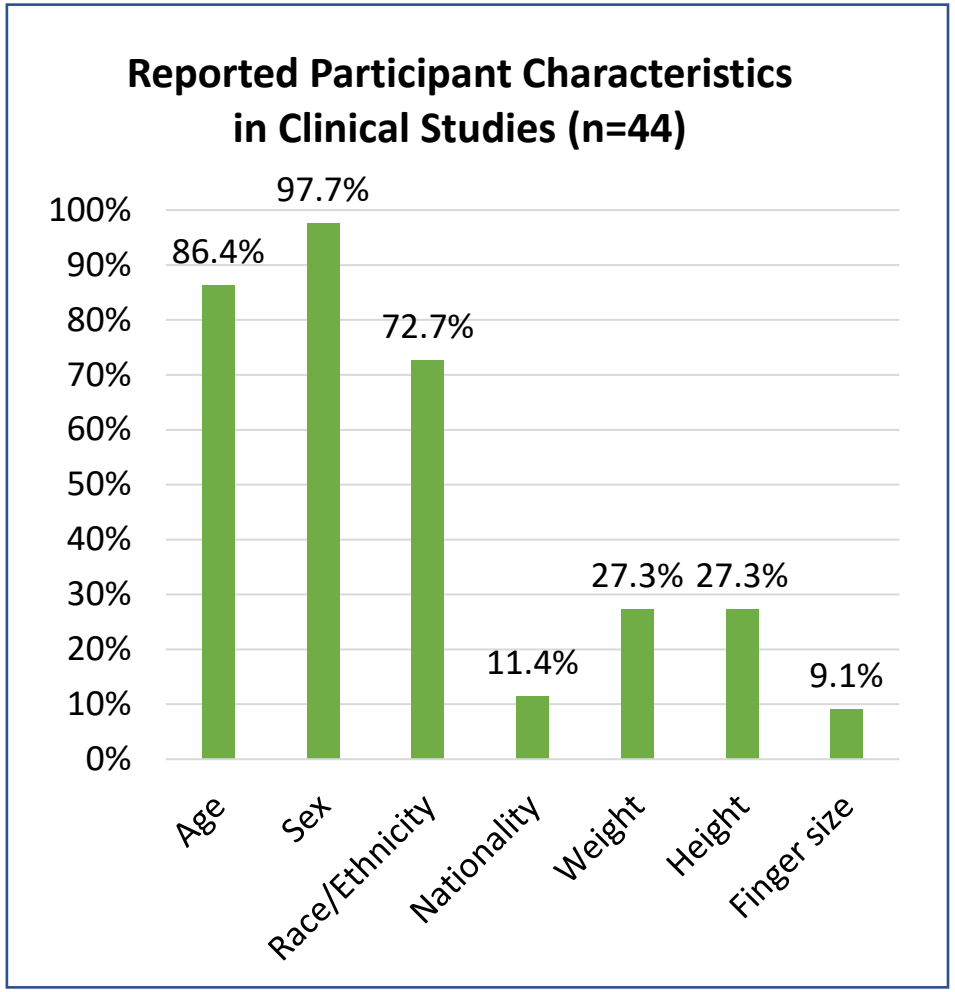
Objective 1 Results – Clinical Study Sample Sizes and Patient Line-Level Data Availability

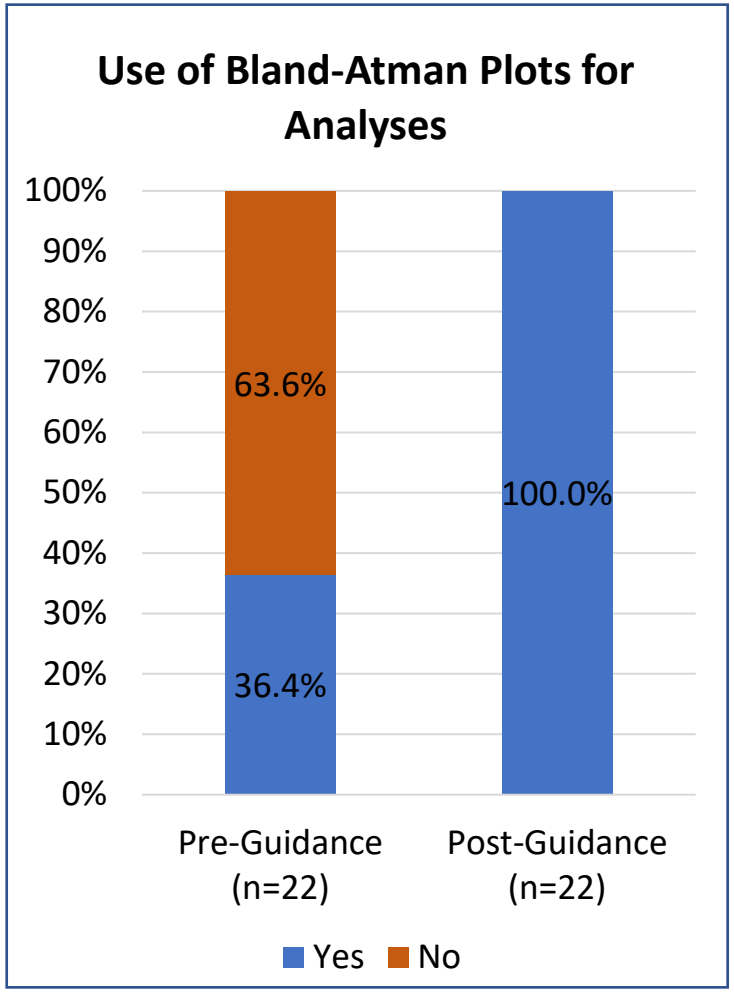
Variable	Median (min, max)
Number of study subjects	12 (6, 33)
Number of paired oxygen measurements per subject	24 (13, 35)



Objective 1 Results – Inclusion Criteria, Participant Characteristics, and Analyses





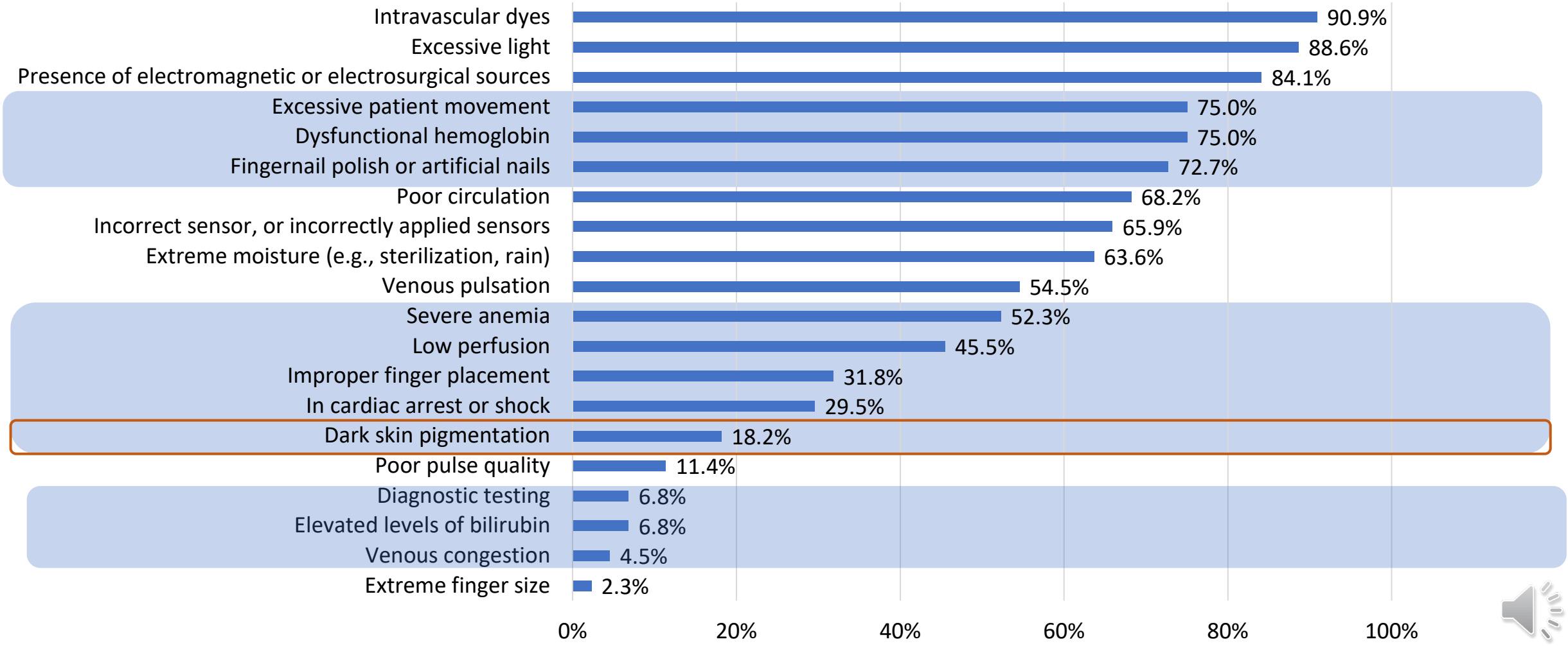


* Examples of indicated medical conditions: Hypertension, cardiovascular disease, respiratory disease

Objective 1 Results – Skin Pigmentation Assessment and Categories

	All Submissions (n=44)	Pre-Guidance (n=22)		Post-Guidance (n=22)	
Singular	2 (4.5%)	2 (9.1%)			
Binary	8 (18.2%)	3 (13.6%)			
Ternary					
Quaternary		2 (9.1%)	50.0%	4 (18.2%)	95.5%
Quinary	1 (2.3%)	.			
Von Luschan scale	1 (2.3%)	1 (4.5%)		.	
Fitzpatrick scale	1 (2.3%)	.		1 (4.5%)	
Not specified	12 (27.3%)				

Objective 2 Results – Factors Listed in the Labeling that May Impact Device Accuracy



*Blue shading indicates increased reporting of factor from pre- to post-guidance

Summary

Objective 1 - Clinical Studies

- Comparing pre- and post- guidance submissions, there was greater:
  - Indication of skin pigmentation classifications
  - Availability of patient line-level data
  - Use of Bland-Altman plots
- A wide variety of skin pigmentation categories was observed

Objective 2 – Device Labeling

- Increases in reporting of factors that may impact accuracy observed from pre- to post-guidance
- Indication that skin pigmentation may impact device accuracy was not included in 73% of post-guidance 510(k) submissions included in the analysis

Panel Questions

There are several tools to assess skin pigmentation, including but not limited to, colorimetry, spectrophotometer, melanosome volume fraction, and skin color scales (e.g., Fitzpatrick and von Luschan color scale). Please provide recommendations for studies evaluating pulse oximeters, for the following:

a. Standardization of skin pigmentation assessment.

b. Categorization and reporting of skin pigmentation data.

Current labeling for prescription use pulse oximeters is intended for clinicians and generally it does not address inaccuracies that may be associated with skin pigmentation. In your deliberations please discuss:

a. Labeling modifications to address inaccuracies that may be associated with skin pigmentation.

b. Recommendations for the content of labeling for lay users who may use pulse oximeters at home.

Thank You


FDA
U.S. FOOD & DRUG
ADMINISTRATION

U.S. FOOD & DRUG ADMINISTRATION

Premarket Desaturation Studies for Pulse Oximeters

Kumudhini Hendrix, MD

Medical Officer

Anesthesia Devices Team

OHT1/CDRH

Outline

- Premarket desaturation testing
  - Submissions requiring pre-market desaturation study
  - Convenience sample verification in neonatal populations
  - Data submission for FDA review
  - Limitations of pre-market desaturation study for clinical applications of pulse oximetry

U.S. FOOD & DRUG ADMINISTRATION

Premarket Desaturation Study

Premarket Desaturation Studies

- Recommendation of *in vivo* testing for SpO2 accuracy under laboratory conditions
  - All new pulse oximeters
  - All prior cleared pulse oximeters with following modifications
    - Significant electro-optical sensor modifications
    - SpO2 algorithm modifications

Purpose

- To verify the SpO2 accuracy in comparison to the gold-standard measurements of blood SaO2 by a co-oximeter over the specified range (SaO2 70-100%)

- Additional clinical safety and effectiveness data should be submitted for devices with *specific clinical indication(s)*

Subjects

- Sufficient number of subjects to attain statistical significance to demonstrate a specified SpO2 accuracy
  - Minimum of pooled 200 data pairs (SpO2, SaO2) evenly distributed over the tested range (70-100%)
- Healthy adult volunteers who vary in physical characteristics to the greatest extent possible
  - 10 or more healthy subjects with a range of skin pigmentation, age and gender
  - At least 2 darkly pigmented subjects or 15% of subject pool, whichever is greater
  - COHb <3%, MetHb <2%, ctHb> 10g/dL

Allowable Testing Conditions

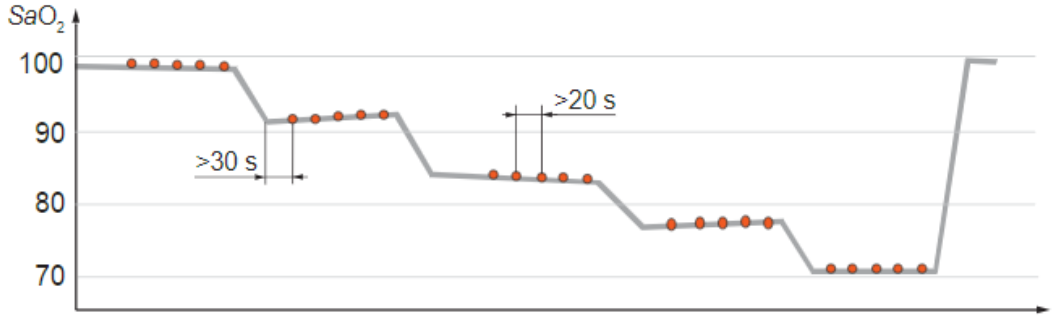
- Application of warming techniques to improve circulation and pulse amplitude at a pulse oximeter probe site
- Pulse oximeter probes can be covered with opaque material to prevent optical interference
- CO_2 gas can be added to inspired gas mixture to maintain normal carbon dioxide levels to prevent respiratory alkalosis secondary to hyperventilation caused by hypoxia

Testing Procedure

- Methodology
  - Fraction of inspired oxygen [FiO2] delivered to test subjects is varied in a stepwise manner to achieve a series of targeted steady-state saturation periods
  - When reference oximetry blood saturation stabilizes (≥30s) to an acceptable plateau, arterial blood samples from an indwelling arterial catheter is sampled for comparison of simultaneous data pairs (premarket device SpO2, SaO2)

Table EE.1 — Example of target plateaus and ranges

SaO_2 plateau range %	Target number of samples
100 to 97	5
97 to 92	5
92 to 85	5
84 to 78	5
77 to 70	5
Total	25



Points are SaO_2 values at the time of the blood draws.

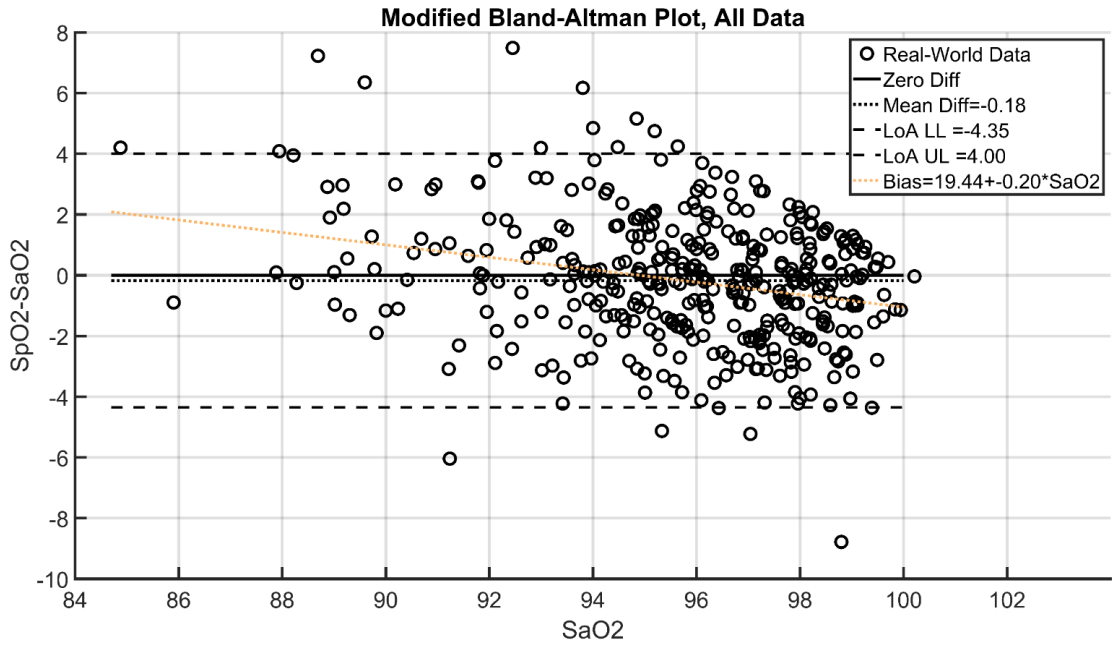
ISO 80601-2-61:2011

Data submission for FDA Review

- Pertinent test apparatus used
- Inclusion/exclusion criteria
- Number of samples taken per subject
- Specific conditions of testing [laboratory conditions, subject motion, low amplitude]
- Type and frequency of motion for testing, if applicable
- Criteria and methods for determining stability of reference SaO2 at the pulse oximeter sensor site
- Desaturation profile [target plateaus, and ranges]
- Formula used for determination of root mean square difference (A_{rms})

Recommended Analyses for FDA review

- Individual and pooled modified Bland-Altman Plots [SaO2, SpO2-SaO2]
- Population mean bias (μo)
- Between-subject variance $(\sigma\mu^2)$
- Within-subject variance (σ^2)
- Upper 95% and lower 95% limits of agreement
- Rationale for any points excluded from analysis



Ebmeier et al, 2018

Outliers

- Discussion of factors that may have affected these data points
- Discussion of how outliers do not raise safety and performance concerns regarding the accuracy of the device

U.S. FOOD & DRUG ADMINISTRATION

Neonatal populations

Recommended *in vivo* testing for neonates

- Any new or significantly modified sensor
- Additional convenience sampling of data pairs (SpO2, SaO2) drawn for clinical decision-making in a clinical environment to verify form, fit, function
- Pulse oximeter probe placed in same circulatory stream as the sampled artery
- Number of data pairs to demonstrate *confidence* the device will assure form, fit and function
- Justification of specific number of samples, subjects and analysis technique

Limitations of Premarket Desaturation Studies

Subjects

- Healthy adult volunteers
  - Limited number of subjects
  - Subjective skin pigmentation values
- No verification for non-neonatal pediatric populations
  - Extrapolated from smaller finger sizes of adult volunteers
  - Unverified form, fit, function for non-neonatal, pediatric populations

Conditions

- Ideal conditions
  - Hands are warmed for optimal performance
  - Shielding of optical sensors
  - Simultaneous data pairs sampled only at stable plateaus (≥ 30s)
  - Normocarbic hypoxic conditions (CO2 in breathing gas mixtures)
  - Desaturated down to tolerable limits only

Analyses

- To verify the SpO2 accuracy in comparison to the gold-standard measurements of blood SaO2 by a co-oximeter over the specified range (SaO2 70-100%)
  - Pooled accuracy (A_{rms}) across the entire tested range of SpO2
  - No threshold accuracy analyses
  - Not powered to determine significant difference between cohorts (age, sex, specific pigmentation values)

Reconsidering Premarket Clinical Desaturation Studies

- Indications
  - Is currently cleared indications for use for monitoring trends, spot checking adequate?
    - Should certain target SaO2 range[s] require a higher degree of accuracy?
      - » If so, what is the needed accuracy?
- Clinically relevant pulse oximetry performance
  - Can premarket desaturation studies be better designed to answer reported disparate performance in the real world?


FDA
U.S. FOOD & DRUG
ADMINISTRATION

Statistical Considerations in the Evaluation of Pulse Oximeters

GENE PENNELLO, PhD
Division of Imaging Diagnostics & Software Reliability
U.S. Food and Drug Administration
Center for Devices and Radiological Health
Office of Science and Engineering Laboratories

Performance Metrics

Performance Metrics

- **Mean bias** = **mean** of $SpO_2 - SaO_2$
- **Precision** = standard deviation (SD) of $SpO_2 - SaO_2$.
- **Mean absolute deviation (MAD)** = **mean** of $|SpO_2 - SaO_2|$
- $\boldsymbol{A_{RMS}}$ = root **mean** square of $SpO_2 - SaO_2$

- These performance metrics are averages across all paired measurements of (SaO_2,SpO_2) from all subjects.

If groups differ in SaO_2 distribution, then group comparisons in these performance metrics may be difficult to interpret.

Statistical Analysis Considerations

- Skin color groups may be compared on **SpO_2 bias** using ***analysis of covariance (ANCOVA)***. That is, regress **SpO_2** on
  - skin color effects, *and*
  - **SaO_2**.
- The skin color effects are differences in **SpO_2** bias between the skin color groups *at the same **SaO_2** value*.
- **ANCOVA**
  - facilitates comparing "like-with-like".
  - can adjust for other covariates (e.g., age, sex, perfusion index, finger size)
  - adjusts group comparisons for *regression-to-the mean (*Barnett et al, 2005[1]*).*

[1]Barnett et al. Regression to the mean: what it is and how to deal with it. *Int J Epidemiol*. 2005 Feb;34(1): 215-20.

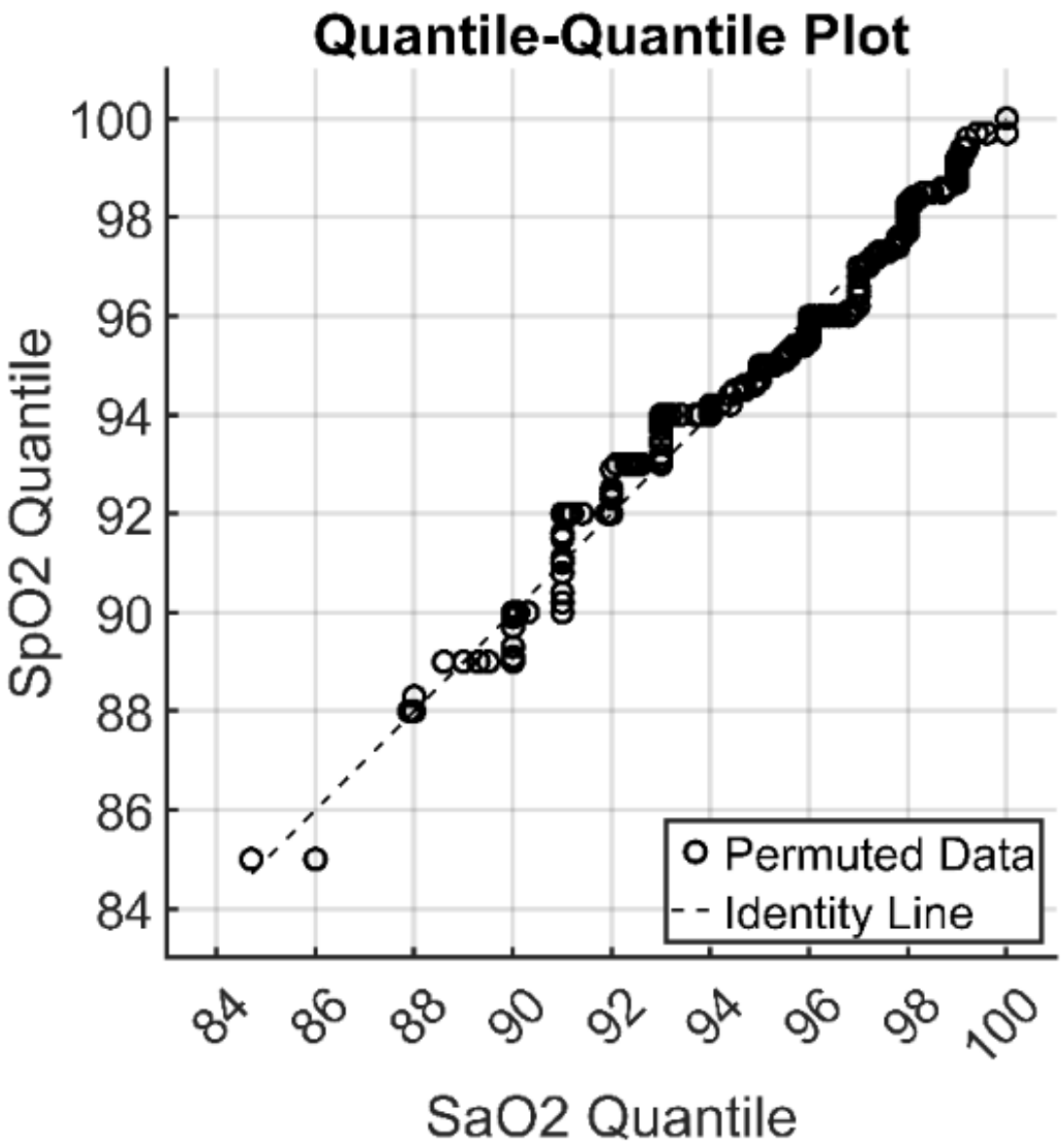
Regression towards the Mean

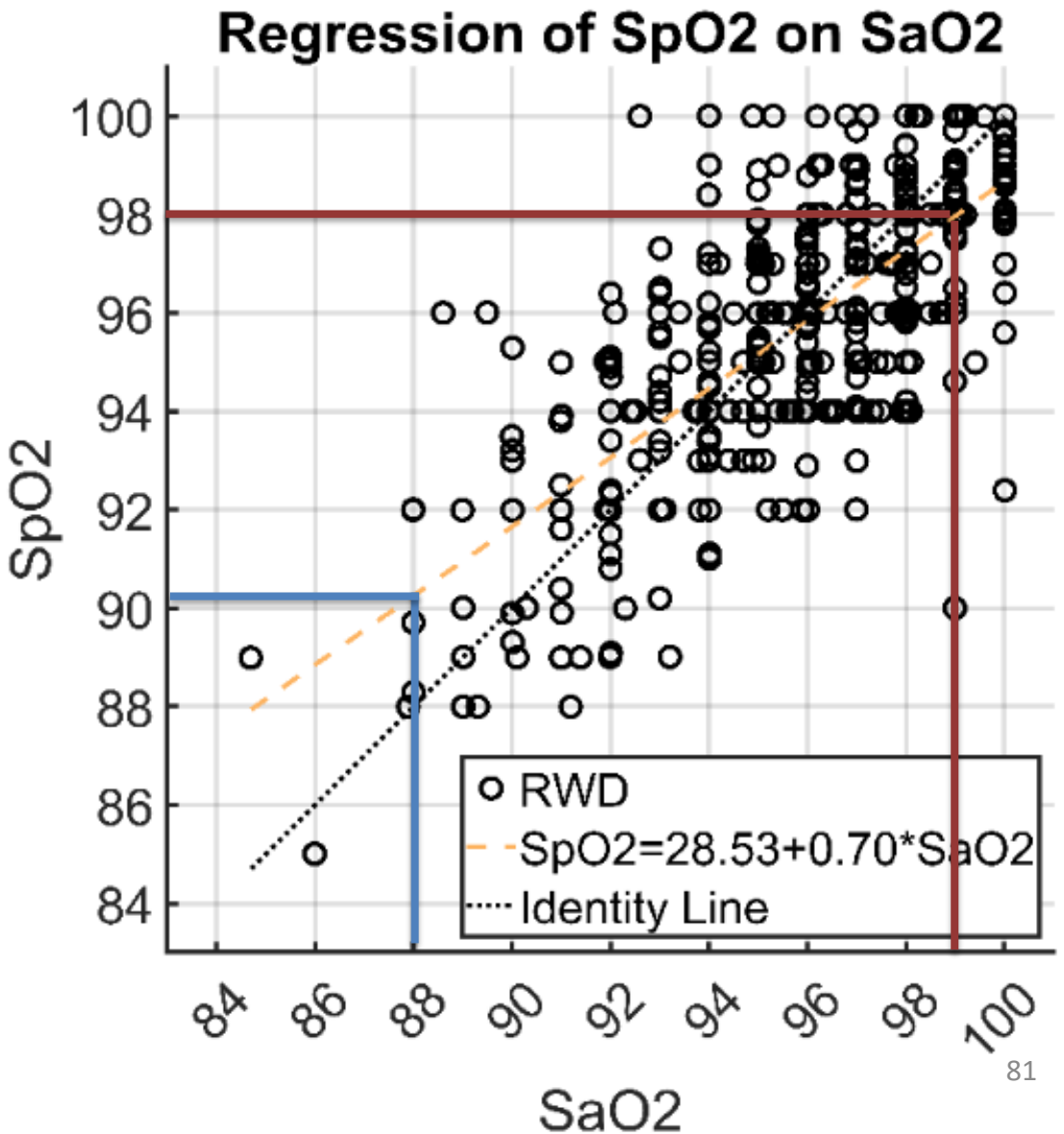
Regression-to-the-Mean

- **EX 1. Heights of parents and children**[2]
  - **Tall** parents tend to have **shorter** children.
  - **Short** parents tend to have **taller** children.
- **EX 2. Comparing two methods of measurement**[3]
  - A **low *SaO_2*** value of oxygen saturation will tend to be paired with a **higher *SpO_2*** value of oxygen saturation – even when the two methods are unbiased for each other.
- **RTM in Two Methods of Measurement (e.g., SaO_2, SpO_2)** occurs because of
  - variation between the measurements,
  - measurement imprecision (variation in repeated measurements),
  - biological variation (if paired measurements are not simultaneous)

[2]Galton F. Regression towards mediocrity in hereditary stature. *J Anthropological Institute* 1886; 15:246-63.
[3]Bland JM, Altman DG. Some examples of regression towards the mean. *BMJ* 1994 Sep 24; 309(6957): 780.

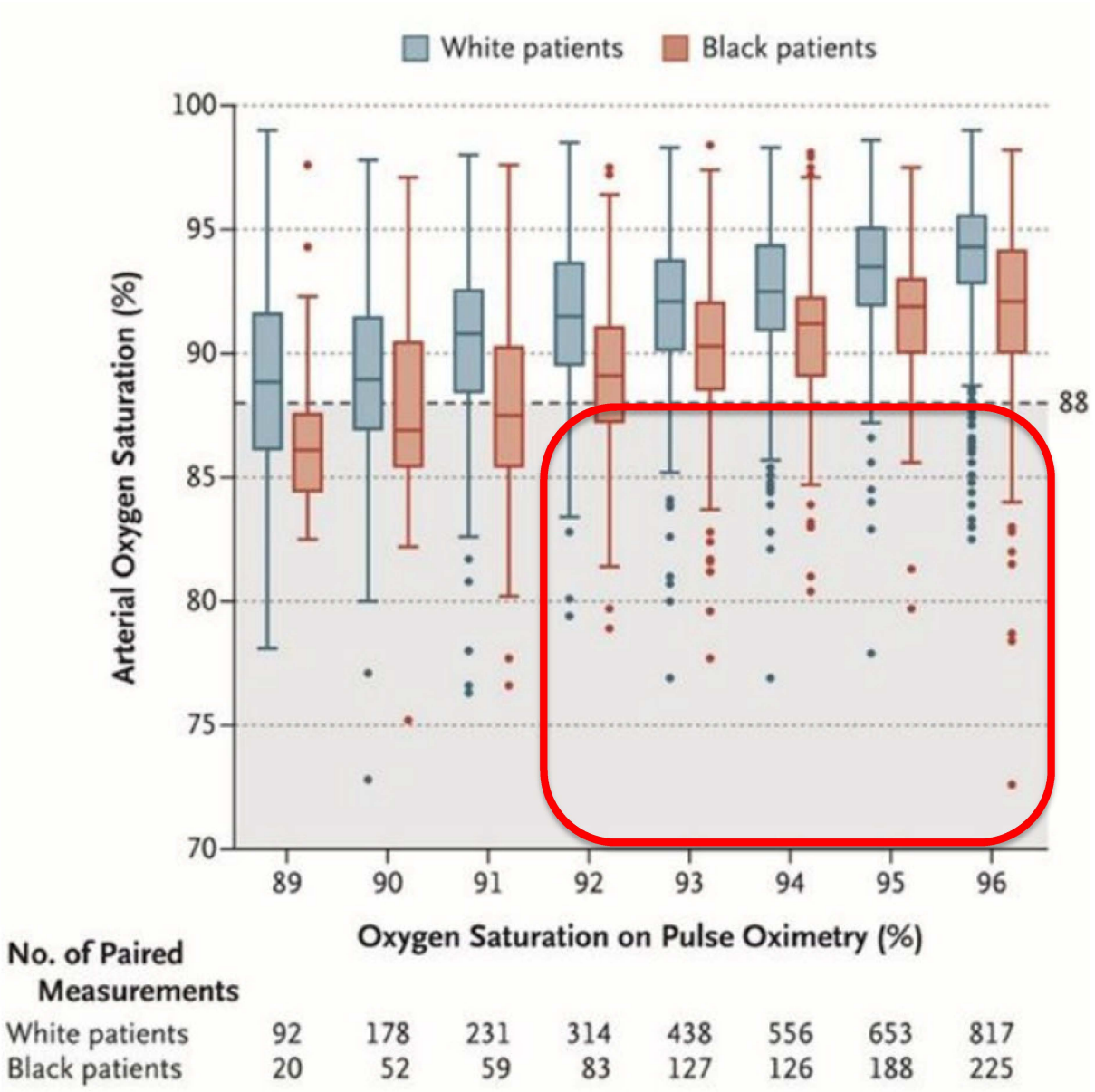
Illustration of RTM in a Dataset in Which SaO_2 and SpO_2 Labels Were Randomly Permuted





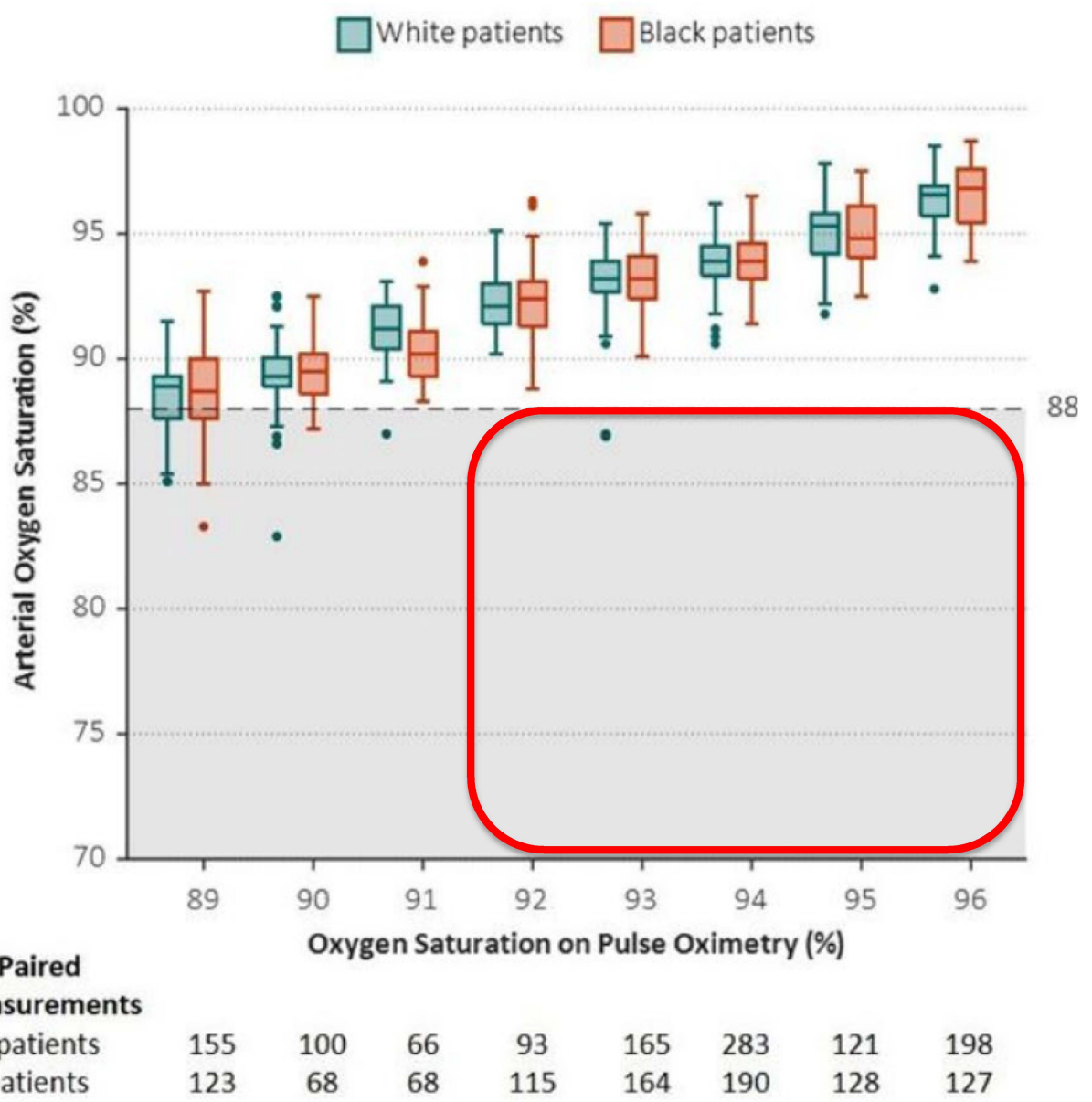
Box Plots

of SaO_2 at each SpO_2 value
Stratified by Skin Color



No. of Paired Measurements								
White patients	92	178	231	314	438	556	653	817
Black patients	20	52	59	83	127	126	188	225

RW Study: Sjoding et al, *NEJM*, 2020



No. of Paired Measurements								
White patients	155	100	66	93	165	283	121	198
Black patients	123	68	68	115	164	190	128	127

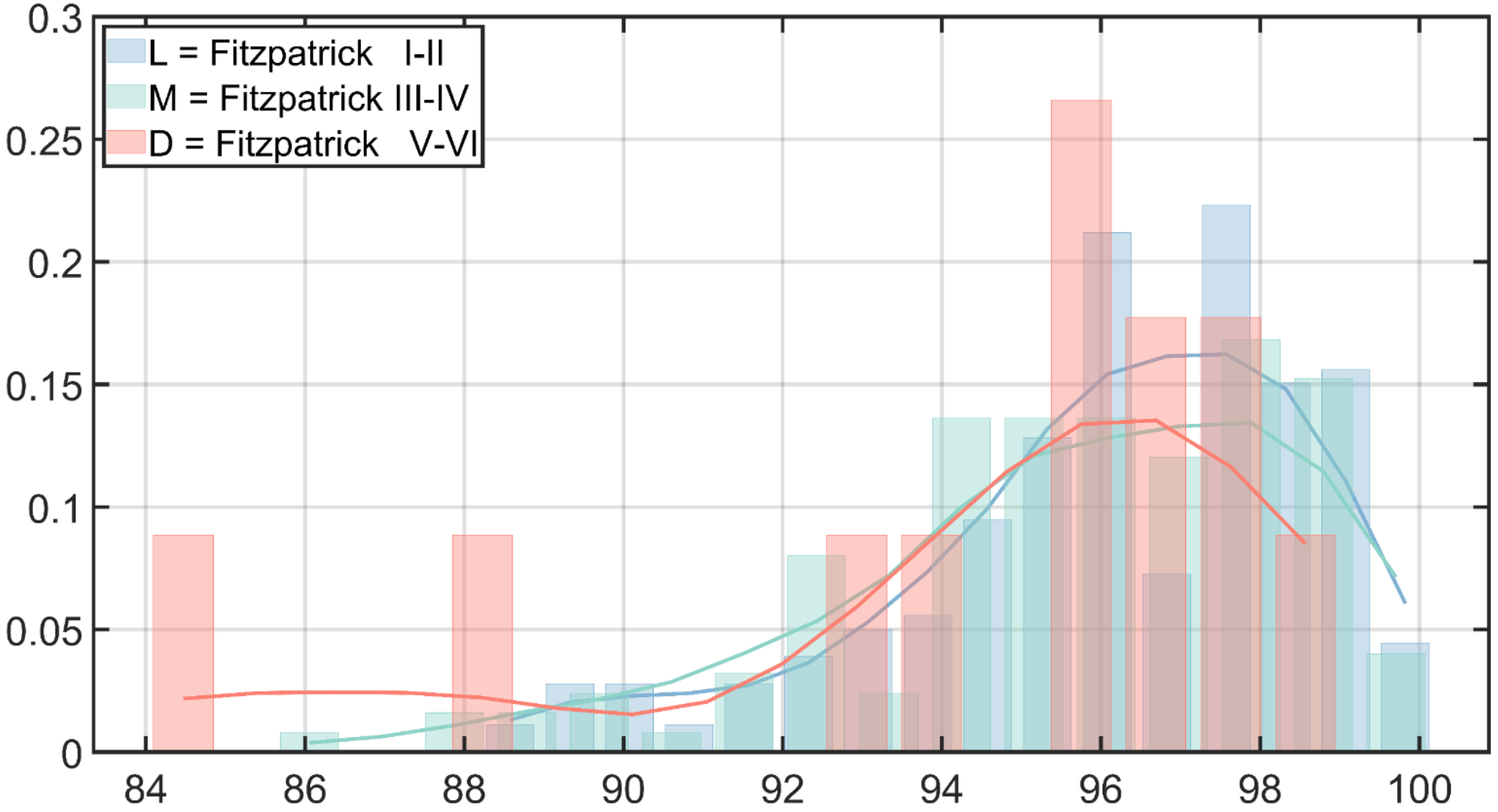
De-Sat Study: Barker, Wilson, *Res Square* 2022

Box Plot Comparisons

- Statistically, the distribution of SaO_2 given SpO_2 (i.e., the box plot) depends on the marginal distribution of SaO_2.
  - The lower the marginal distribution of SaO_2,
  - the lower the distribution of SaO_2 given SpO_2 (all else being equal).
- A box plot comparison of two groups does not consider that
  - the marginal distribution of SaO_2 could be lower (stochastically) for one group than the other group.

If groups differ in the SaO_2 marginal distribution, then differences between groups in the box plot may be difficult to interpret.

Real-World Data, Adults in ICU[4]
SaO_2 Distribution by Skin Color Group



[4]**Data Source**: Ebmeier et al, *Anaesth Intensive Care*. 2018 May; 46(3):297-303.

Statistical Considerations

- If **Box Plot** comparisons are desired, then standardize the groups to the same SaO_2 distribution before drawing the box plot:
  - **Method 1**. Weight the observations in group 1 so that they have the same SaO_2 distribution as observed in group 2.
  - **Method 2**. Weight the observations in groups 1 and 2 using an identical set of weights corresponding to a *standard* SaO_2 distribution.

Diagnostic Accuracy

of Pulse Oximetry to Detect Hypoxemia

Real-World Data, Infants in NICU[5]

White Infants		Reference Standard		
		$SaO_2 < 85\%$	$SaO_2 \geq 85\%$	
Test Result		+	−	Total
$SpO_2 < 90\%$	+		451	
$SpO_2 \geq 90\%$	−			
Total		293	2049	2342

Occult hypoxemia rate = 181/1779 = 10.2%

Black Infants		Reference Standard		
		$SaO_2 < 85\%$	$SaO_2 \geq 85\%$	
Test Result		+	−	Total
$SpO_2 < 90\%$	+	124	329	453
$SpO_2 \geq 90\%$	−			
Total		312	1732	2044

Occult hypoxemia rate = 188/1591 = 11.8%

[5]Vesoulis et al, Racial discrepancy in pulse oximeter accuracy in preterm infants. *J Perinatol.* 2022 Jan; 42(1):79-85.

Real-World Data, Infants in NICU[5]

White Infants		Reference Standard		
		$SaO_2 < 85\%$	$SaO_2 \geq 85\%$	
Test Result		+	−	Total
$SpO_2 < 90\%$	+		451	563
$SpO_2 \geq 90\%$	−			
Total		293	2049	2342

Occult hypoxemia rate = 181/1779 = 10.2%

Sensitivity = 112/ 293 = 38.2%

Black Infants		Reference Standard		
		$SaO_2 < 85\%$	$SaO_2 \geq 85\%$	
Test Result		+	−	Total
$SpO_2 < 90\%$	+	124	329	453
$SpO_2 \geq 90\%$	−			
Total		312	1732	2044

Occult hypoxemia rate = 188/1591 = 11.8%

Sensitivity = 124/ 312 = 39.7%

[5]Vesoulis et al, Racial discrepancy in pulse oximeter accuracy in preterm infants. *J Perinatol.* 2022 Jan; 42(1):79-85.

Real-World Data, Infants in NICU[5]

White Infants		Reference Standard		
		$SaO_2 < 85\%$	$SaO_2 \geq 85\%$	
Test Result		+	−	Total
$SpO_2 < 90\%$	+	112	451	563
$SpO_2 \geq 90\%$	−			
Total		293	2049	2342

Occult hypoxemia rate = 181/1779 = 10.2%
Sensitivity = 112/ 293 = 38.2%
Specificity = 1598/2049 = 78.0%

Black Infants		Reference Standard		
		$SaO_2 < 85\%$	$SaO_2 \geq 85\%$	
Test Result		+	−	Total
$SpO_2 < 90\%$	+	124	329	453
$SpO_2 \geq 90\%$	−			
Total		312	1732	2044

Occult hypoxemia rate = 188/1591 = 11.8%
Sensitivity = 124/ 312 = 39.7%
Specificity = 1403/1732 = 81.0%

[5]Vesoulis et al, Racial discrepancy in pulse oximeter accuracy in preterm infants. *J Perinatol.* 2022 Jan; 42(1):79-85.

Real-World Data, Infants in NICU[5]

White Infants		Reference Standard		
		$SaO_2 < 85\%$	$SaO_2 \geq 85\%$	
Test Result		+	−	Total
$SpO_2 < 90\%$	+	112	451	563
$SpO_2 \geq 90\%$	−	181	1598	1779
Total		293	2049	2342

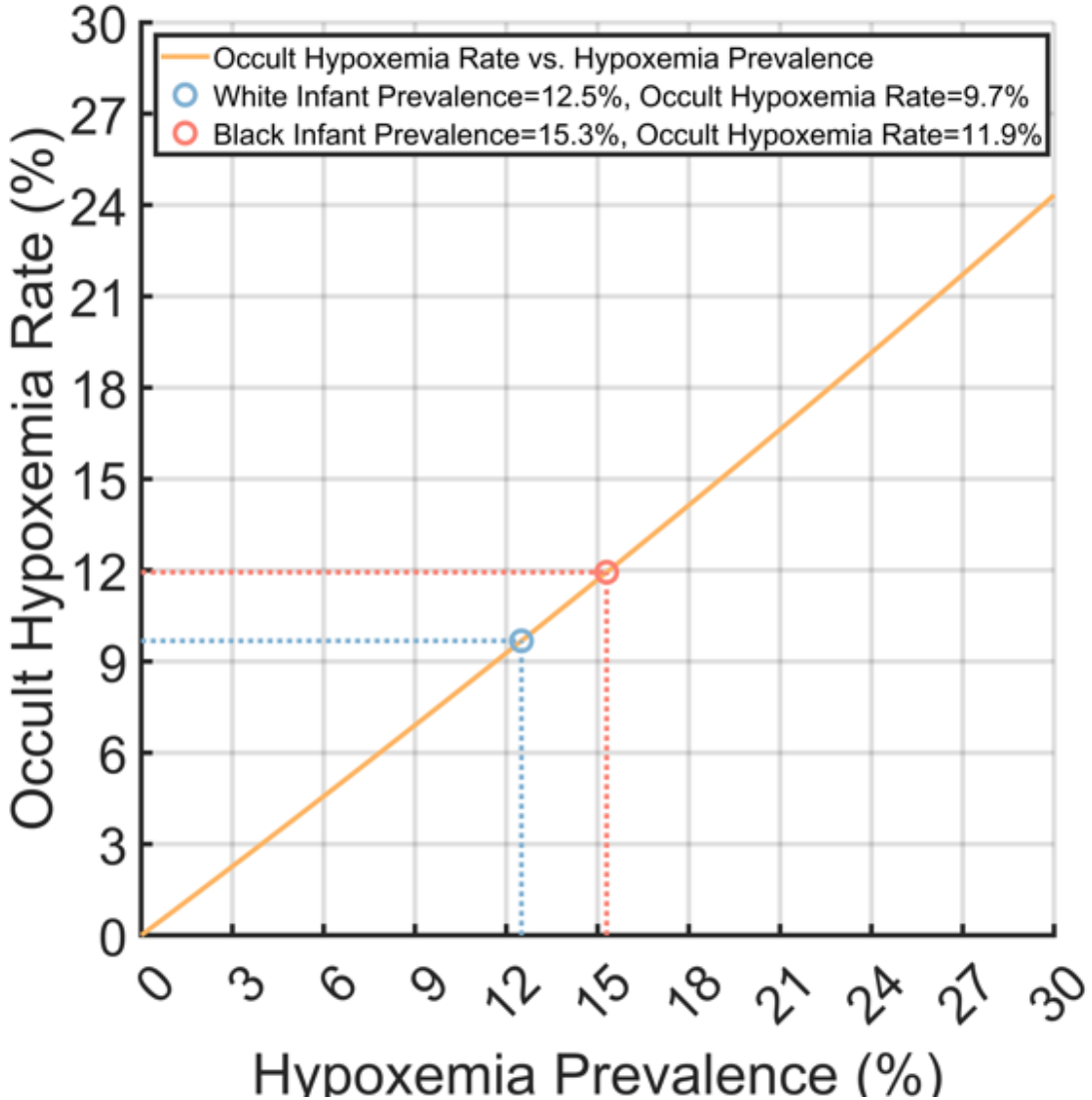
Occult hypoxemia rate = 181/1779 = 10.2%
Sensitivity = 112/ 293 = 38.2%
Specificity = 1598/2049 = 78.0%
Hypoxemia Prevalence = 293/2342 = 12.5%

Black Infants		Reference Standard		
		$SaO_2 < 85\%$	$SaO_2 \geq 85\%$	
Test Result		+	−	Total
$SpO_2 < 90\%$	+	124	329	453
$SpO_2 \geq 90\%$	−	188	1403	1591
Total		312	1732	2044

Occult hypoxemia rate = 188/1591 = 11.8%
Sensitivity = 124/ 312 = 39.7%
Specificity = 1403/1732 = 81.0%
Hypoxemia Prevalence = 312/2044 = 15.3%

[5]Vesoulis et al, Racial discrepancy in pulse oximeter accuracy in preterm infants. *J Perinatol.* 2022 Jan; 42(1):79-85.

Occult Hypoxemia Rate Depends on Hypoxemia Prevalence



- **Occult hypoxemia rate vs. hypoxemia prevalence (gold line in plot on left) is based on**
  - sensitivity=40%, specificity=80%, and
  - Bayes formula:

$$\frac{occult\ hypoxemia\ rate}{1-occult\ hypoxemia\ rate}=\frac{1-sensitivity}{specificity}\times\frac{prevalence}{1-prevalence}$$

Real-World Data, Infants in NICU[5]

White Infants		Reference Standard		
		$SaO_2 < 85\%$	$SaO_2 \geq 85\%$	
Test Result		+	−	Total
$SpO_2 < 90\%$	+	112	451	563
$SpO_2 \geq 90\%$	−	181	1598	1779
Total		293	2049	2342

Occult hypoxemia rate = 181/1779 = 10.2%

Sensitivity = 112/293 = 38.2%

Specificity = 1598/2049 = 78.0%

Hypoxemia Prevalence = 293/2342 = 12.5%

Black Infants		Reference Standard		
		$SaO_2 < 85\%$	$SaO_2 \geq 85\%$	
Test Result		+	−	Total
$SpO_2 < 90\%$	+			
$SpO_2 \geq 90\%$	−	188	1403	1591
Total		312	1732	2044

Occult hypoxemia rate = 188/1591 = 11.8%

Sensitivity = 124/312 = 39.7%

Specificity = 1403/1732 = 81.0%

Hypoxemia Prevalence = 312/2044 = 15.3%

[5]Vesoulis et al, Racial discrepancy in pulse oximeter accuracy in preterm infants. *J Perinatol.* 2022 Jan; 42(1):79-85.

Summary

Non-randomized comparisons of groups
may be difficult to interpret
without adjustment for potential confounders.

Non-randomized comparisons of groups on occult hypoxemia rate
are difficult to interpret
because of confounding by hypoxemia prevalence.

A pulse oximeter study may be difficult to interpret when

- paired measurements of (SaO_2,SpO_2) are not simultaneous,
- data were excluded from analysis or not reported at all,
- limitations exist in study design, conduct, analysis, or reporting.

Panel Question

FDA currently recommends assessment of the effectiveness of pulse oximeters using Arms [Root mean square of pooled data pairs] and adherence to the currently recognized ISO 80601-2-61:2017 standard. For this variable (Arms), currently pulse oximeters are expected to have accuracy within 1 standard deviation (SD) (66% of the time) and within 2 SD (95% of the time). Please address the following:

a. Please discuss how accurate pulse oximeters should be for clinical use. In your discussion, please address whether the accuracy varies based on: (i) the clinical setting or (ii) the levels of SaO_2.

b. Please discuss your recommendations for pulse oximeters performance across sub-groups of subjects with different skin pigmentations.

c. Please discuss if Arms is an appropriate measure of device effectiveness for clinicians and users. If you do not believe Arms is appropriate, please discuss alternative methods to assess the accuracy of a pulse oximeter.

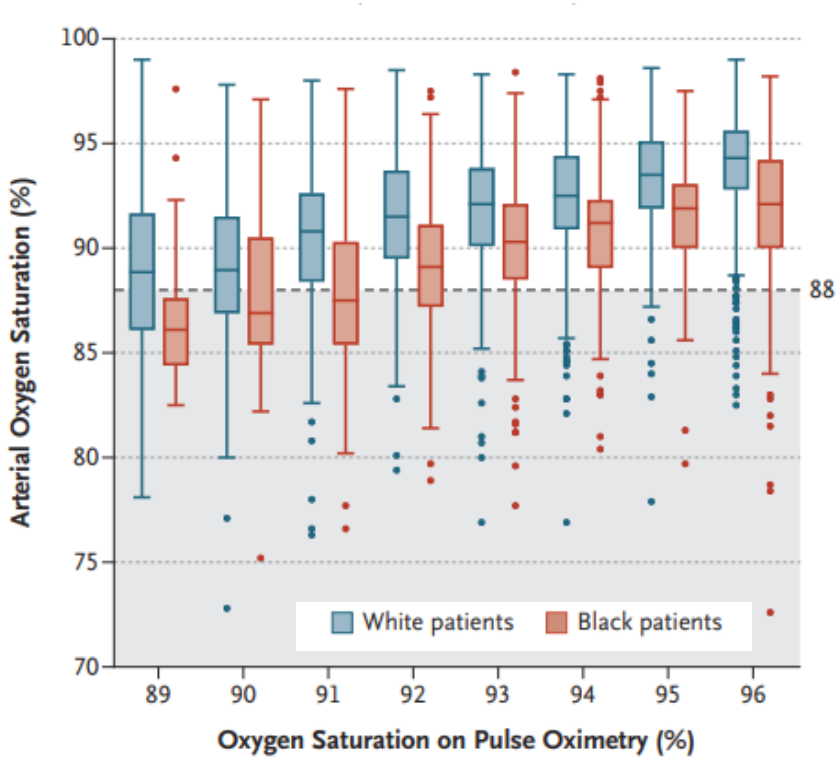
Methods of Assessing Skin Pigmentation In Pulse Oximetry Studies

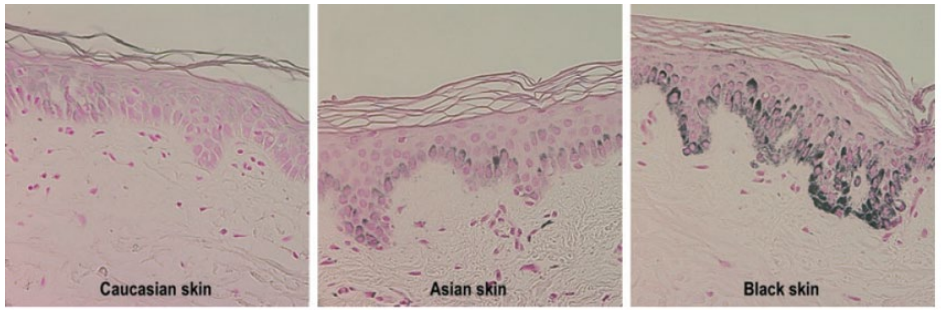
Sandhya Vasudevan, PhD, William Vogt, PhD, Sandy Weininger, PhD and Josh Pfefer, PhD

Office of Science and Engineering Laboratories

Center for Devices and Radiological Health

Ethnicity/Race and Skin Pigmentation in Pulse Oximetry





Epidermal Melanin Distribution By Histology

Sjoding, M.W., et al. NEJM 2020

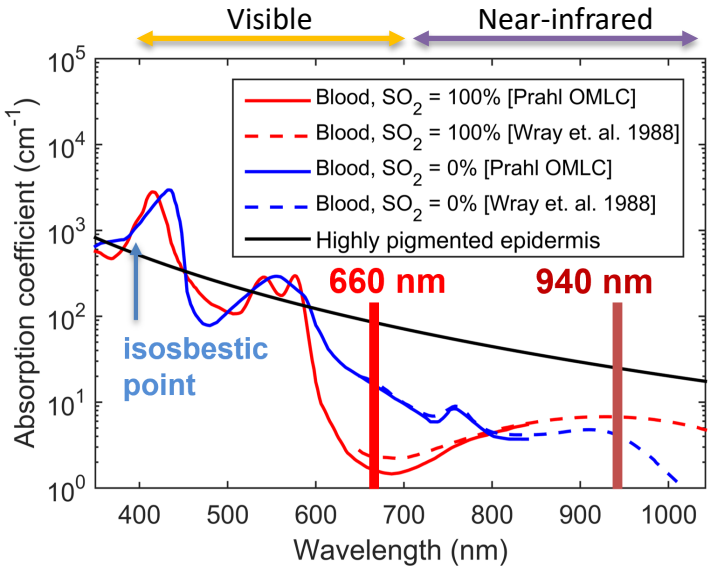
NY Times (R.C. Rabin), Dec 22, 2020 (BSIP SA/Alamy)
Brenner, M. and V.J. Hearing, Photochem. Photobiol. 2008

Pulse Oximetry Tissue Optics

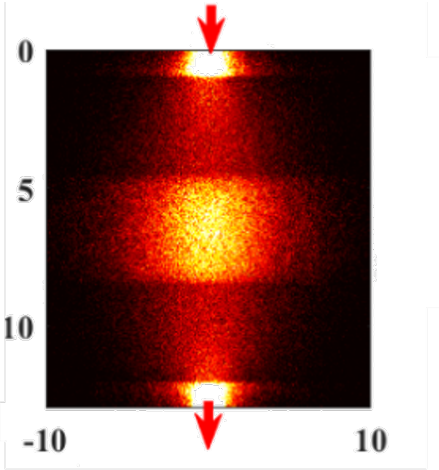
- **Light propagation mechanisms**
  - Scattering (e.g., collagen, cells)
  - Absorption (e.g., blood, melanin, water)
  - Exploit the "optical window" in red/near-IR

- **Epidermal melanin**
  - Significant red/NIR absorber
  - Can reduce detected light at 660 nm by ~40% per pass through epidermis
  - Detected signals processed to mitigate optical effects, may not be fully successful

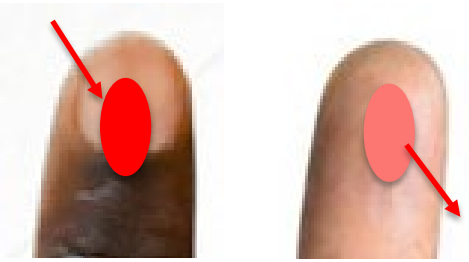
- **Anatomical site considerations**
  - Fingernail, palmar finger contain *low* levels of melanin
  - Other sites (e.g., forehead) and skin adjacent to fingernail may contain *high* levels of melanin



Light transport across finger



Illumination / Detection

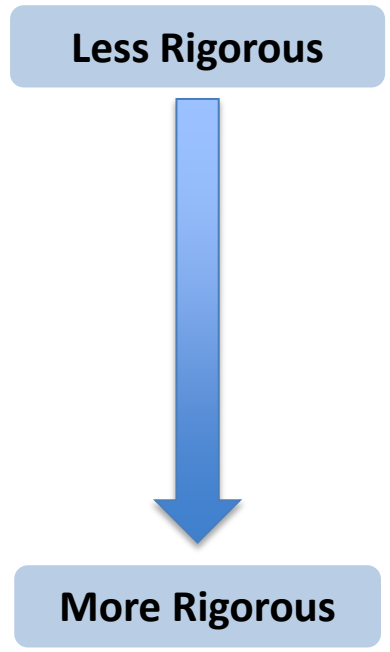


Dorsal Palmar/Ventral

Kyriacou et al., Academic Press 2021
Chatterjee et al., Sensors 2019

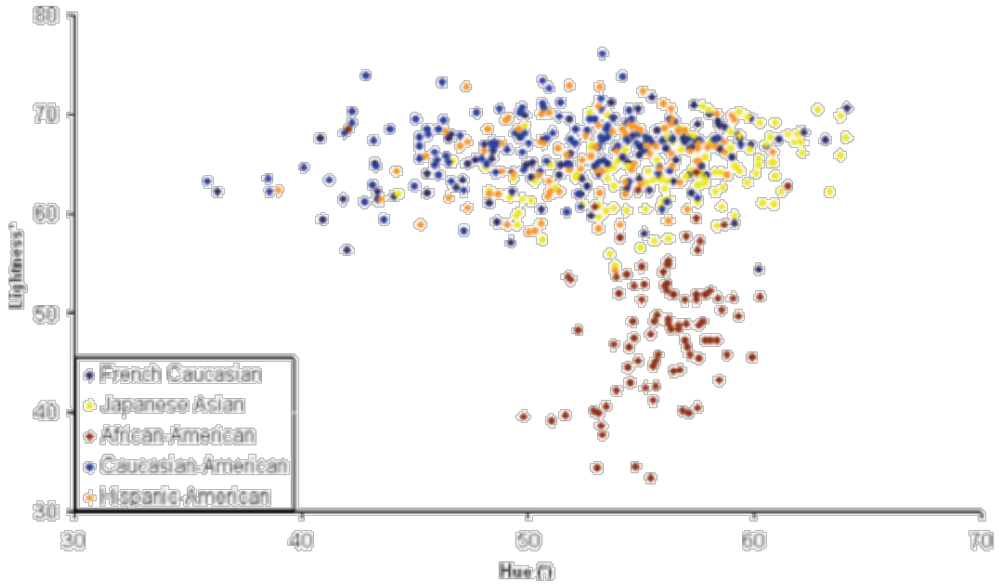
Skin Pigmentation Assessment Methods

Skin Pigmentation Assessment Method	Degree of Objectivity/Quantitation
Racial/Ethnic Self-Identification (e.g., Black, White, Hispanic, Asian)	Subjective with limited/no skin pigmentation information
Skin Color Descriptors (e.g., light, medium, medium dark, dark)	Subjective due to lack of a standardized scale. Large variance within skin color groupings
Sunburn susceptibility / color scale (Fitzpatrick skin phototypes, I-VI)	Subjective (questionnaire-based), quantitative (categorical); sometimes used as a non-standardized color scale
Color scales (Massey-Martin, Munsell)	Subjective but with lower variance due to the use of standardized color categories
Optical Methods (Spectroscopy, Colorimetry)	Quantifiable information that is not dependent on a subjective evaluation; but some metrics not standardized
***Biopsy** with histological/optical processing or high-performance liquid chromatography*	*Quantitative melanin content, but can involve reader-dependent steps (e.g., layer identification)*



Racial/Ethnic Self-Identification

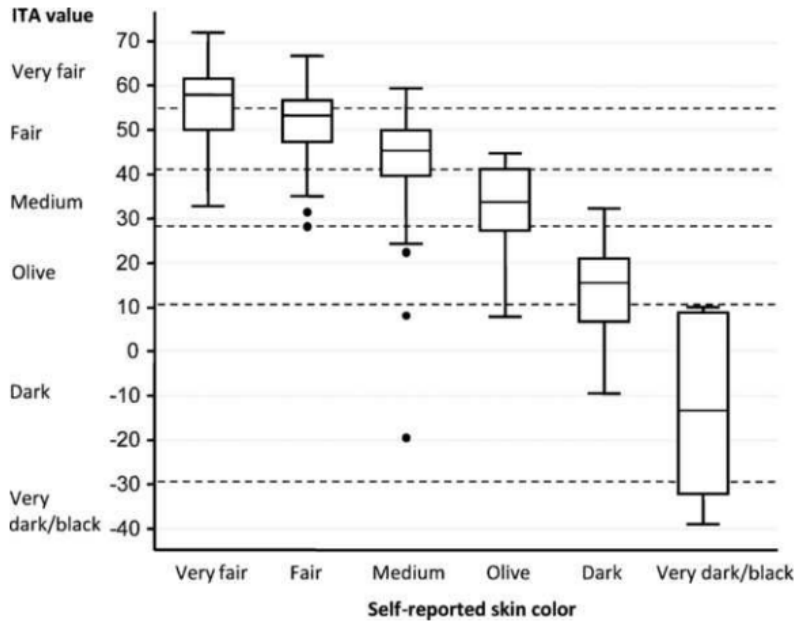
- Commonly used in retrospective pulse oximetry studies to classify subjects (self-identified)
- Subjective and qualitative, with limited skin pigmentation information
- People have mixed ethnicities and skin pigmentation levels can vary within any ethnic group
- Conflation of race/ethnicity with skin pigmentation may produce misleading results when investigating oximeter bias



Caisey, L., et al., Int. J Cosm Sci. 2006

Skin Color Descriptors

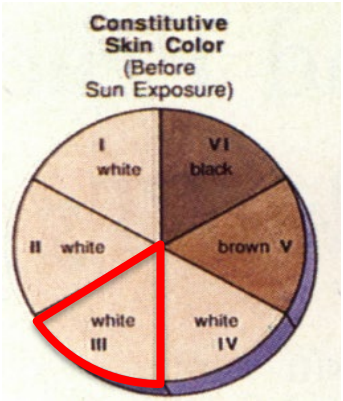
- Tend to be used in controlled pulse oximetry studies
  - e.g., light, medium, medium-dark, and dark
- Highly subjective and qualitative
  - Lack of any standardized definitions or scale
  - Large variance within each category
- Skin color differences in dark-pigmented skin are overlooked by visual assessment
- Not highly repeatable or reproducible

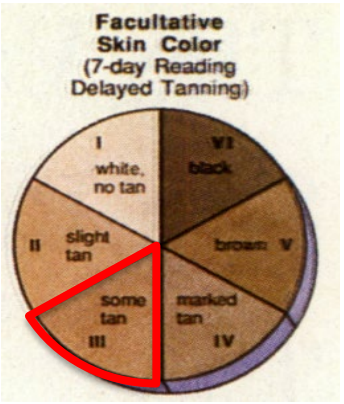


Campbell, M.E., et. al. *Race Soc Probl* 2020
Reeder, A.I. et al. *Cancer Epidemiology and Prevention Biomarkers* 2010

Fitzpatrick Skin Phototype (FSP) Scale

- Originally developed to assess the propensity for skin to burn/tan, later adapted to categorize skin color
- FSP classification (I-VI)
  - Questionnaire/interview on subject's response to sun exposure
  - Types V and VI based on color descriptors ("brown" and "black")
  - FSP color charts/descriptors not standardized
- Weak correlation with pigmentation, does not change with tanning
- Physicians typically assign individuals of color to FSP IV-VI based on their ethnicity, which is unreliable
- Relatively coarse, uneven categorization for continuum of skin tones
  - Can lead to loss of information (e.g., variations within Type VI)
- FST and objective methods compared, but results were highly variable (|R| = 0.23 – 0.90)



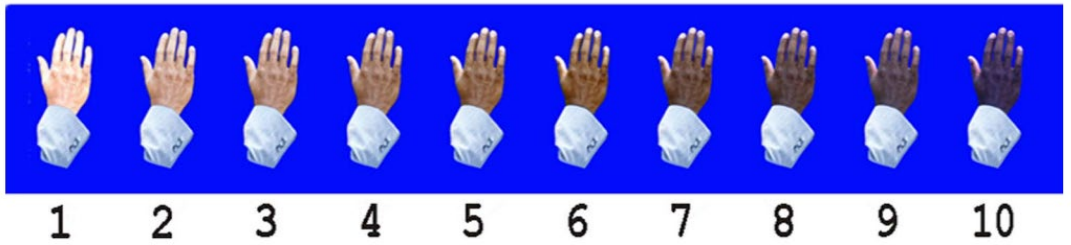


Fitzpatrick, Arch Dermatol 1988
Charlton et al., PLOS One 2020
Ash et al., Photon Lasers Med 2015
Varughese, P. M., et. al. *IJPD* 2018

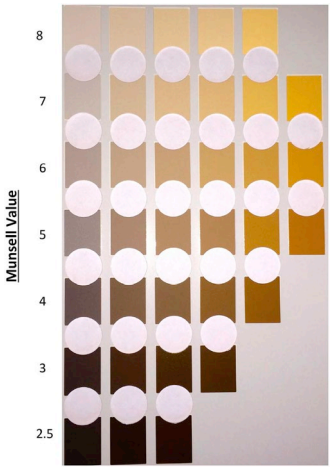
Skin Color Scales/Charts

- Infrequently used in pulse oximeter studies
- Several skin color scales implemented
  - Massey-Martin, Munsell, von Luschan
- Theoretically, should have lower variance due to standardized color charts
- Access to validated charts is a common limitation
- Prone to subjective errors, lighting conditions, inter-observer bias

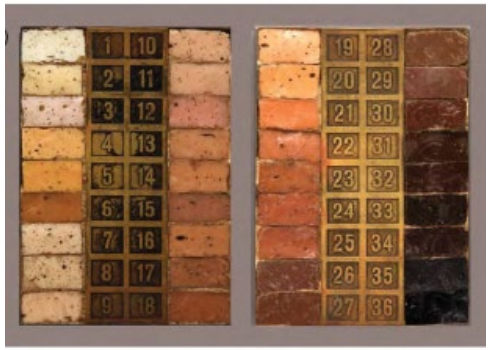
Massey-Martin Scale



Munsell color chart



von Luschan Scale

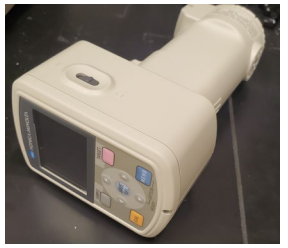


Massey, D. S., and J. A. Martin 2003
Charoenngam, N. et. al, J Nutrit Metabol, 2021
Foglia et al., J Pediatrics 2017

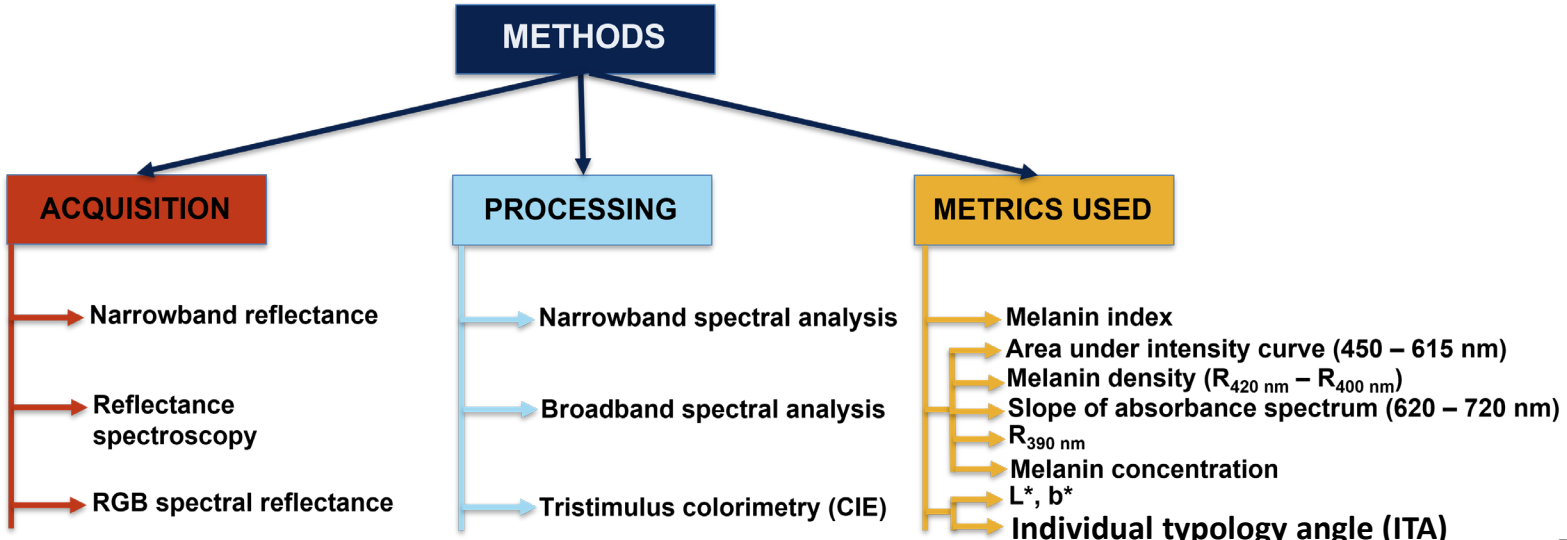
Objective (Optical) Approaches

- Vis-NearIR reflectance to quantify objective metrics for melanin (melanometry) has been studied extensively in the literature
- Commercial systems include spectroscopy, colorimetry devices
- Provides site-specific pigmentation data (e.g., palmar finger)
- Some outputs, metrics are not standardized
- Commercial systems show good inter-device agreement (|R| = 0.56 – 0.98)

Colorimeter



Spectrophotometer

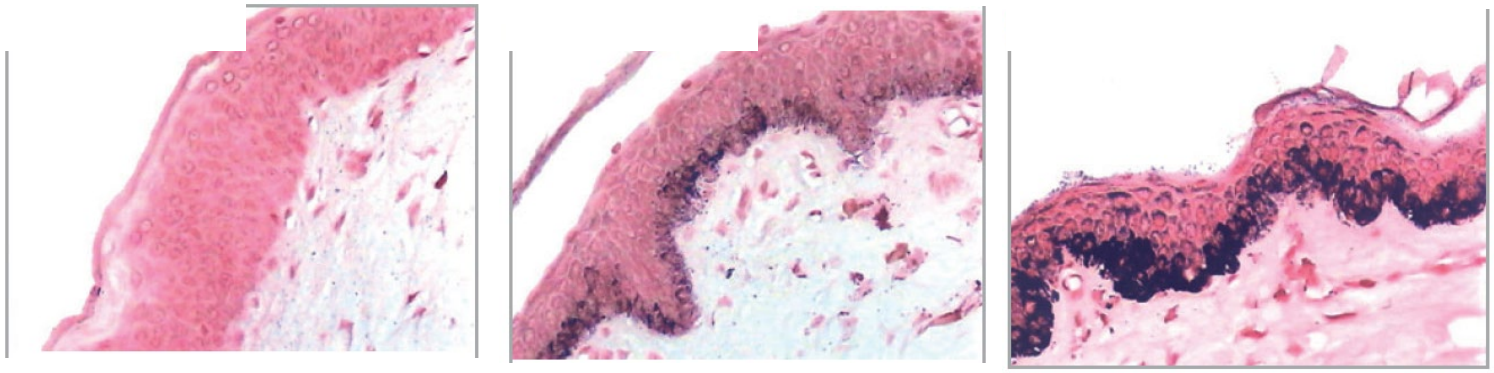


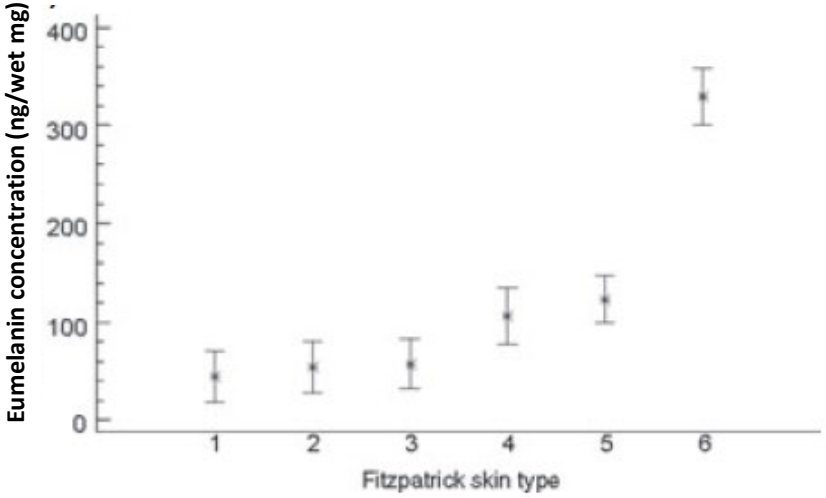
Individual Typology Angle ITA°	Skin classification
ITA° > 55°	Very light
41° < ITA° < 55°	Light
28° < ITA° < 41°	Intermediate
10° < ITA°< 28°	Tan
–30° < ITA°< 10°	Brown
ITA° < –30°	Dark

Del Bino S. and F. Bernerd, Brit J Derm 2013

“Gold Standard” Melanin Quantification

Visualization and quantification of melanin-stained skin sections





- Biopsy with histological/optical processing or high-performance liquid chromatography are considered high quality reference approaches for quantifying melanin content
- Invasive, so primary role is validating pigmentation assessment methods
- Objective pigmentation assessment methods show better agreement with gold standard approaches (|R| > 0.75 in 59% of compiled results)

Matts, P.J. et al., Brit J Derm, 2007

Rigor in Skin Pigmentation Assessment

- **Articles on retrospective studies have criticized subjective methods as:**
  - A "limiting factor" in clinical studies (Andrist et al. 2022, Burnett et al. 2022, Wiles et al. 2021)
  - "Surrogate" for skin pigmentation or melanin content (Burnett et al., 2022, Fawzy et al., 2022)
  - Having categories that are "too blunt" or "broad" (Shi et al. 2022, Henry et al. 2022)
  - "…stop the use of subjective skin tone scales altogether" (Colvonen et al. 2021)

Summary

- **Subjective Methods**
  - Commonly used, inexpensive, easy to implement
  - Less accurate, repeatable than objective methods
  - Variety of options with different levels of differentiation capability
  - May be sufficient for comparing scenarios involving large differences in pigmentation

- **Objective (Optical) Methods**
  - Alternatives to subjective methods with higher rigor/accuracy
  - Demonstrated strong correlation to biopsied melanin content
  - May be most useful for specific sites (e.g., finger) but not necessary for all studies
  - Likely more expensive to implement than subjective methods
  - Currently not a standard approach in dermatology

Panel Question

There are several tools to assess skin pigmentation, including but not limited to, colorimetry, spectrophotometer, melanosome volume fraction, and skin color scales (e.g., Fitzpatrick and von Luschan color scale). Please provide recommendations for studies evaluating pulse oximeters, for the following:

a. Standardization of skin pigmentation assessment.

b. Categorization and reporting of skin pigmentation data.


FDA
U.S. FOOD & DRUG
ADMINISTRATION