UNITED STATES OF AMERICA

FOOD AND DRUG ADMINISTRATION

+++

CENTER FOR DEVICES AND RADIOLOGICAL HEALTH

+++

MEDICAL DEVICES ADVISORY COMMITTEE

+++

ANESTHESIOLOGY DEVICES PANEL

+++

November 1, 2022

9:00 a.m. EST

PANEL MEMBERS:

Steven Nathan, M.D.	Chair
Hugh A. Cassiere, M.D, FCCP, FACP	Voting Member
Richard D. Branson, M.Sc., R.R.T.	Voting Member
Lonny B. Yarmus, D.O., M.B.A.	Voting Member
Jeffrey R. Kirsch, M.D.	Temporary Nonvoting Member
Arlene J. Hudson, M.D.	Temporary Nonvoting Member
Robert G. Loeb, M.D.	Temporary Nonvoting Member
Andrea M. Kline, PhD, CPNPAC/PC, FCCM, FAAN	Temporary Nonvoting Member
Sean Hennessy, PharmD, Ph.D.	Temporary Nonvoting Member
Jason Connor, Ph.D.	Temporary Nonvoting Member
Eliot Katz, M.D.	Temporary Nonvoting Member
Nancy Collop, M.D.	Temporary Nonvoting Member
Jennifer Lynch, M.D.	Temporary Nonvoting Member
Murad Alam, M.D.	Temporary Nonvoting Member
Michael F. O'Connor, M.D.	Temporary Nonvoting Member

Veverly M. Edwards, M.F.A.	Consumer Representative
William C. Wilson, M.D., M.A.	Industry Representative
Joseph P. O'Brien, M.B.A.	Patient Representative
Akinola A. Awojope, MPH, Dr.PH	Designated Federal Officer

FDA STAFF:

Jeff Shuren, M.D., J.D.	Director, CDRH
RDML Richardae Araojo, PharmD, M.S.	Associate Commissioner, OMHHE
Malvina B. Eydelman, M.D.	Director, OHT1
Nilsa Loyo-Berrios, Ph.D., M.Sc.	Acting Associate Director, OHT1

FDA PRESENTERS:

James Lee, Ph.D.
CDRH, OHT1
Pulse Oximeters: Technology, Accuracy Limitations, and Regulation

Sandy Weininger, Ph.D.
CDRH, OSEL
Standards for Pulse Oximeters: ISO 80601-2-61: 2017

Allison O'Neill, Ph.D.
CDRH, OHT1
A Systematic Literature Review of the Real-World Performance of Pulse Oximeters

Mary Jung, Ph.D.
CDRH, OCEA
Overview of Desaturation Studies in Pulse Oximeter 510(k) Submissions

Kumudhini Hendrix, M.D.
CDRH, OHT1
Premarket Desaturation Studies for Pulse Oximeters

Gene Pennello, Ph.D.
CDRH, OSEL
Statistical Considerations in the Evaluation of Pulse Oximeters

Josh Pfefer, Ph.D.
CDRH, OSEL
Methods for Assessing Skin Pigmentation in Pulse Oximetry Studies

INVITED SPEAKERS:

Rekha Hagen

Jessica Cocolin, CRNA

Bob Kopotic, RN, RRT

Paul Batchelder, LRCP, RRT, Clinimark

Amal Jubran, M.D., American Academy of Sleep Medicine

Eric Gartman, M.D., American College of Chest Physicians

Ann G. Rizzo, M.D., FACS, DABS, American College of Surgeons

Jesse Ehrenfeld, M.D., American Medical Association

Steven Gay, M.D., M.S., American Thoracic Society

Julian Goldman, M.D., Anesthesia Patient Safety Foundation,

Garrett Burnett, M.D., Society of Technology in Anesthesia

Elizabeth Bridges, Ph.D., RN, CCNS, FCCM, American Association of Critical Care Nurses

Michael W. Sjoding, M.D., University of Michigan Medical School

An-Kwok Ian Wong, M.D., Ph.D., Duke University

Ashraf Fawzy, M.D., MPH, Johns Hopkins University

Eric Raphael Gottlieb, M.D., M.S., Brigham and Women's Hospital

Phil Bickler, M.D., Ph.D., UCSF-CERSI

Christopher Almond, M.D., Stanford-CERSI

Michael Lipnick, M.D., University of California UCSF

OPEN PUBLIC HEARING SPEAKERS:

Dr. Veronica Hickson

Sam Ajizian

David Stricken

Grace Berson

Jeff Matthews

Dr. Steven Barker

Dr. Eve Calender

Dr. Michael Abrams

Renee Kohi

INDEX

PAGE

CALL TO ORDER

Dr. Nathan: Good morning, everyone. I would like to call this Anesthesiology and Respiratory Therapy Devices Panel to order. I am Dr. Steve Nathan, the Chairperson of this panel. I am the Medical Director of the Advanced Lung Disease and Lung Transplant Program at Anno Fairfax Hospital, which is in Falls Church, Virginia. I note, for the record, that the voting members present constitute a quorum as required by 21 C.F.R. Part 14. I would also like to add that the Panel members participating in today's meeting have received training in FDA device law and regulations.

For today's agenda, the Panel will discuss the ongoing concerns that pulse oximeters may be less accurate in individuals with darker skin pigmentation. The committee will also discuss factors that may affect pulse oximeter accuracy and performance, the available evidence about the accuracy of pulse oximeters, recommendations for patients' health care providers, and amounts and type of data that should be provided by the manufacturer to assess pulse oximeter accuracy and to guide other regulatory actions as needed. I wanted to mention that the FDA has received 25 comments that were submitted to the docket as of October 31st, 2022. These documents are available to the public at the following website. www.regulations.gov.docket/FDA/2022-N-210 and these have been sent to the panelists and FDA for review.

Before we begin, I'd like to remind the public and panelists that this is a non-voting meeting and ask our distinguished committee members and FDA participants attending virtually to introduce themselves. Committee members, please turn on your video monitors if you have not done so already and unmute your line so that we can hear you when you speak. When I call

your name, please state your area of expertise, your position, and affiliation. Let's start off with Dr. Cassiere, please.

PANEL INTRODUCTIONS

Dr. Cassiere: Good morning everyone, I am Hugh Cassiere. I'm a Critical Care Pulmonary. My specialty is I'm the Medical Director for Respiratory Therapy Services at a large transplant center. I work for Northwell Health, and I'm very excited to join the meeting.

Dr. Nathan: Thank you, Dr. Cassiere. Sorry, I butchered your name a little bit. Next, Dr. Kirsch.

Dr. Kirsch: Good morning. I'm Jeff Kirsch. I am the Professor of Anesthesiology and Pain Management at the University of Washington. I'm the Chief of Anesthesiology at Harvard View Medical Center.

Dr. Nathan: Thank you, Dr. Kirsch. Next is Mr. Branson.

Mr. Branson: Hi, this is Rich Branson. I'm a Respiratory Therapist. I'm a Professor in the Division of Trauma and Critical Care at the University of Cincinnati, where I've been Director of Clinical Research for 30 years. I'm also the Editor in Chief of Respiratory Care.

Dr. Nathan: Thank you. Mr. Branson, I'm not sure if Dr. Yarmus is on. Dr. Yarmus, if you are on, if you could please introduce yourself.

Dr. Yarmus: Hi, I'm Lonny Yarmus, I am the Clinical Chief of the Division of Pulmonary Critical Care at Johns Hopkins Medicine.

Dr. Nathan: Next we have Dr. Wilson.

Dr. Wilson: Hi, I'm Bill Wilson, I'm a Cardiac Anesthesiologist and Critical Care Physician, I was the CMO at the University of California Irvine Medical Center and I am now the CMO at Masimo in charge of clinical research.

Dr. Nathan: Thank you. Next, Mr. O'Brien.

Mr. O'Brien: Hi I am Joe O'Brien, and I am the President and Chief Executive Officer of the National Scoliosis Foundation based in Stoughton, Massachusetts. I am also a patient with 21 surgeries.

Dr. Nathan: Next is Ms. Edwards.

Ms. Edwards: Hi. I am Veverly Edwards, I am the Consumer Representative on the panel, I teach at the University of Memphis in Tennessee.

Dr. Nathan: Dr. Hudson is next.

Dr. Hudson: I am Arlene Hudson, I am a Cardiac Anesthesiologist, and I am Chair of the Department of Anesthesiology Uniformed Services University, University of Hudson.

Dr. Nathan: Thank you Dr. Hudson. Next, please Dr. Loeb.

Dr. Loeb: I'm Robert Loeb, I go by Butch. Robert is fine for this meeting. I am Professor Emeritus of Anesthesiology, University of Florida, now living in the San Diego area.

Dr. Nathan: Thank you, Next, Dr. Klein.

Dr. Klein: Good morning, I'm Andrea Kline, I am a Pediatric Acute Care Chief Nurse Practitioner, and Nurse Practitioner Director at University of Michigan Health, Ann Arbor, Michigan.

Dr. Nathan: Thank you very much, next please, Dr. Hennessy.

Dr. Hennessy: Good morning, my name is Sean Hennessy, I'm a Pharmacist Epidemiologist at the University of Pennsylvania, where I lead the Division of Epidemiology.

Dr. Nathan: Thank you very much, and next please, Dr. Connor.

Dr. Connor: I'm Jason Connor. I'm a Biostatistical consultant, owner of Consulate and also Assistant Professor of Medical Education at University of Central Florida. I usually don't introduce myself further back, but I also before becoming a PhD Biostatistician did PhD work in biomedical optics specifically involving using lasers for diagnosis. That seemed relevant today.

Dr. Nathan: Thank you, Dr. Connor. Next, please, Dr. Katz.

Dr. Katz: Hi, I am Eliot Katz, from Boston Children's Hospital. I'm a Pediatric Pulmonologist and sleep Physician at Harvard Medical School.

Dr. Nathan: Thank you, next please, Dr. Collop.

Dr. Collop: Hi, I am Nancy Collop, Professor of Medical and Neurology, at Emory University, my expertise is in sleep medicine and I run the Emory Sleep Center.

Dr. Nathan: Thank you, Next please Dr. Lynch.

Dr. Lynch: Hi, I'm Jennifer Lynch I am a pediatric Cardiac Anesthesiologist, at the Children’s Hospital in Philadelphia, I also have a PhD in physics and also do research in the field of Biomedical Optics.

Dr. Nathan: Thank you very much, Dr. Lynch.

Dr. Nathan: Dr. Alam.

Dr. Alam: Good morning, my name is Murad Alam, I’m a dermatologist and clinical researcher at Northwestern University, where I am a Professor and Vice chair of the Department of Dermatology.

Dr. Nathan: Thank you very much, next please, Dr O’Connor.

Dr. O’Connor: Hi I am Michael O’Connor. I’m a professor of Anesthesiology and Critical Care at the University of Chicago. I'm a practicing Anesthesiologist and Intensivist.

Dr. Nathan: Thank you very much, Dr. Lee.

Dr. Lee: Hi, James Lee, Division Director for Sleep Disorder Breathing, Respiratory and Anesthesia Devices here at FDA.

Dr. Abrams: Hi, I'm Michael Abrams at the University of Chicago. I'm a practicing anesthesiologist.

Dr. Nathan: Thank you very much, Dr. Loyo-Berrios.

Dr. Loyo-Berrios: Hi my name is Nilsa Loyo-Berrios, I'm an Epidemiologist and Acting Associate Director in the Office of Health Technology 1 at CDRH.

Dr. Nathan: And last but by no means least, Dr. Eydelman.

Dr. Eydelman: Good morning, everyone. My name is Melvin Eydelman. I'm Director of the Office of Health Technology One, which oversees all of the anesthesia devices. Thank you all for joining us today, and we'll truly look forward to hearing your thoughts and opinions on this very important matter.

Dr. Nathan: Thank you for introducing yourselves, as you can tell from introductions, a very well-rounded panel. Thank you very much. Next, Dr. Akinola Awojope, our Designated Federal Officer for today's Anesthesiology Devices Panel will make some introductory remarks.

CONFLICT OF INTEREST STATEMENT

Dr. Awojope: Good morning. I am Dr. Akinola Awojope, I'm the DFO, the designated officer for this advisory meeting. I will now take this interest statement to everyone.

The Food and Drug Administration, FDA, is convening today's meeting of the Anesthesiology Devices Panel of Medical Devices Advisory Committee under the authority of the Federal Advisory Committee Act FACA of 1972. With the exception of the industry

representative, all members and consultants of the panel are special government employees or regular Federal employees from other agencies and are subject to Federal conflict of interest laws and regulations.

The following information on the status of these panels, compliance with the federal ethics and conflict of interest laws covered by but not limited to, those found at 18 USC Section 208 are being provided to the participants in today's meeting and to the public.

FDA has determined that members and consultant of this panel are in compliance with the federal ethics and conflict of interest laws under 18 USC Section 208, Congress has authorized FDA to grant waivers to special government employees and regular federal employees who have a financial conflict when it is determined that the agencies need for the particular individuals services, outweigh is our potential financial conflict of interest.

Related to the discussions of today's meeting, members and consultants of this Panel who are special Government employees or regular Federal employees have been screened for potential financial conflicts of interest of their own as well as those imputed to them, including those of their spouses or minor children and, for purposes of 18 U.S.C 208 and their employers.

These interests may include investments, consulting, expert witness testimony, contracts, grants, CRADAS, teaching, speaking writing, patents and royalties, and primary employment.

For today's agenda, the Panel will discuss the ongoing concerns that pulse oximeters may be less accurate in individuals with darker skin pigmentations. The advisory panel will discuss factors that may affect pulse oximeter accuracy and performance, the available evidence about the accuracy of pulse oximeters, recommendations for patients and health care providers, the amount, and type of data that should be provided by manufacturers to assess pulse oximeter accuracy and to guide other regulatory actions as needed.

Based on the agenda for today's meeting, and all financial interests reported by panel members and consultants, a conflict of interest waiver has been issued in accordance with 18 U.S.C. 208(b)(3) to Dr. Steve Nathan.

Dr. Nathan's waiver addresses his personal financial interests in health sector mutual funds that contain underlying assets shares potentially affected competing firms. The aggregate value of his holdings in the funds is between $200,000 and $300,000. The waiver allows this individual to participate fully in the panel deliberations. FDA's reasons for issuing the waivers are described in the waiver documents which are posted FDA's website, copies of this waiver may also be obtained by submitting the written request to the mailing address is Agency's Division of Freedom of Information, 5630 Fishers Lane, Room-1035, Rockville, MD 20857.

Dr. William Wilson is serving as the industry representative, acting on behalf of all related industry. He's employed by Masimo Corporation.

For the record, the agency notes that Dr. Christopher Almond who is an invited guest speaker has acknowledged his employer's interest in the form of a research grant that is federally funded.

Mr. Paul Batchelder, who is also an invited guest speaker with us today, has acknowledged his employer's interest with numerous affected competing firms in the form of research contracts.

Dr. Philip Bickler, another invited guest speaker with us today, has acknowledged his employer's interest with multiple affected competing firms in the form of research studies.

We would like to remind members and consultants that if the discussions involve any other products or firms are not already on the agenda from which an FDA participant has a

personal or imputed financial interest the participants need to exclude themselves from such involvement, and their exclusion will be noted for the record.

The FDA encourages all participants to advise the panel of any financial relationships they have with any firms at issues.

I will now read the Temporary Borrow Memo.

For the duration of Anesthesiology and Respiratory Therapy Devices Panel on November 1, 2022, Sean Hennessy, PharmD, PhD has been appointed to serve as Temporary Non-Voting Member, and for the record Dr. Hennessy, served as a consultant to the Drug Safety and Risk Advisory Committee, at the Center for Drug Evaluation and Research. This individual is a special government employee who has undergone the customary conflict of interest review and has reviewed materials to be considered for this meeting. The appointment was authorized by Russell Forney, The Director Advisory Committee Oversight, our management staff on October 24th, 2022. A copy of this statement will be available for review and will be included as a part of the official transcript.

Before I turn the meeting back to Dr. Nathan, I would like to make a few general announcements. In order to help the transcribers identify who is speaking, please be sure to identify yourself each and every time that you speak. The press contact for today's meeting is Audra Harrison. Thank you very much. Thank you, all the panelists. I'll hand it back to Dr. Nathan again.

FDA OPENING REMARKS

Dr. Nathan: Thank you very much. We're going to now move on to some FDA opening remarks, and I'd like to now invite Dr. Jeff Shuren, Center Director for the Center for

Devices and Radiological Health, as well as Rear Admiral Richardae Araojo, Associate Commissioner for Minority Health, to provide introductory remarks, starting with Dr. Shuren.

JEFF SHUREN, CDRH

Dr. Shuren: Thank you. Good morning and thank you to our panel members and other participants in today's meeting.

Pulse oximeters are non-invasive tools used to estimate Blood oxygen saturation, and associated with inaccurate readings must be well understood. Although pulse oximetry presents benefits for patients, the limitations and risks associated with inaccurate pulse oximeter readings must be well understood. During the COVID 19 pandemic, the use of pulse oximeters significantly increased for both clinical and at home settings.

In 2020, FDA became aware of a post-market signal regarding pulse oximeter performance. In a December, 2020, letter to the editor in the New England Journal of Medicine, the authors of a retrospective cohort study found that patients who self-identified as blacks were almost three times more likely to have hypoxemia not detected by pulse oximetry when compared to white patient counterparts.

These results suggest that black patients are at increased risk for failing to detect occult hypoxemia when pulse oximetry is used, or at least some pulse oximetry devices are used to triage patients and adjust supplemental oxygen levels. Numerous subsequent publications have also raised concerns that there is a potential bias in pulse oximetry performance due to skin pigmentation, potentially putting patients with darker skin pigmentation and delays in treatment and worst patient outcomes.

Therefore, we are holding today's panel meeting for the following purposes and objectives, specifically: to discuss the impact of skin pigmentation on clinical performance of

pulse oximetry technology, which could be indicative of racial disparities in the performance of these devices; discuss and make recommendations regarding the design, conduct, and reporting of data for studies, assessing the accuracy of pulse oximeters and potential bias due to skin pigmentation; promote transparency on this important public health issue and on the agency's activities to date; and provide a public forum for the many stakeholders impacted by this issue to express their views, patients, healthcare providers, professional societies, researchers, and industry.

Our agenda for today's meeting includes the following, an overview of the current regulatory framework for pulse oximeters; including relevant standards, guidance documents, and pre-market study requirements; a summary of the currently available real-world evidence regarding the potential bias in pulse oximetry due to skin pigmentation; invited speaker sessions with the perspectives from adult and pediatric patients researchers, 10 professional societies, industry, and numerous authors of real-world evidence. Papers will be shared, an open public hearing and a panel discussion of FDA's questions, as well as the panel's interpretation of the currently available real-world evidence on this issue.

Recommendations regarding tools to assess skin pigmentation for future studies. Recommendations regarding expectations of pulse oximetry accuracy across various clinical settings, arterial oxygen saturation ranges, and patient subpopulations of varying skin pigmentation. Also, whether ARMS is the best metric to assess device accuracy and recommendations regarding device labeling to convey the potential inaccuracies due to skin pigmentation, as well as the content of labeling for lay users who may use pulse oximeters at home.

FDA has taken several actions to inform today's discussion and to address this issue. They include conducting a comprehensive literature review, reviewing data from medical device reports, conducting a market assessment of pulse oximeter devices, performing a systematic review of desaturation studies submitted in pulse oximeter 510(k)s since 2000.

Issuing a safety communication at February, 2021 to inform patients and healthcare providers regarding the limitations and risks of inaccuracy for pulse oximeters, and to consider these when using such devices to make diagnostic and treatment decisions. And finally, awarding grants to the USCF Stanford Cersi to conduct two real world prospective studies that will assess the performance of pulse oximeters in the adult and pediatric patient populations.

The issue of racial disparities in pulse oximetry is of great public health importance, and it further highlights the need to appropriately assure that all medical devices is safe and effective in all the populations in which they're intended to be used, which is one of the goals of CDRH's 2022 to 2025, strategic priority to advance health equity.

We have invited Rear Admiral Richardae Araojo, Associate Commissioner for Minority Health, the Office Commissioner, and the Director of the FDA's Office of Minority Health and Health Equity to join us today for discussion. She'll deliver some welcoming remarks, so let me now turn it over to Read Admiral Richardae Araojo. Welcome.

RICHARDAE ARAOJO, OMHHE

Dr. Araojo: Thank you, Dr. Shuren. I am glad to be able to join you for this advisory committee meeting on such an important topic. The ongoing effort to better understand and evaluate the relationship between race, ethnicity, skin pigmentation, and oximeter accuracy is crucial and related to the priority across FDA, including CDRH and the Office of Minority

Health and Health Equity to advance health equity and address disparities among racial and ethnic minority populations.

And of course, the FDA's Office of Minority Health and Health Equity engages broadly across our agency as well as with external stakeholders. And we have a dedicated mission to protect and promote the health of racial and ethnic minority and other diverse populations through research and communication that addresses health disparities and advances health equity.

As we have all witnessed over the past two and a half plus years, the COVID-19 pandemic brought health equity to the forefront of our national dialogue and highlighted racial inequities and disparities that have long, adversely impacted diverse communities. So throughout the FDA, information on race and ethnicity has and continues to be integral to our understanding of health issues affecting the US population and support of improving population health outcomes.

Central to our work at FDA is also to support the consistent collection and reporting of race and ethnicity data because we know that variations associated with race and ethnicity have been correlated with risks for certain diseases and conditions and responses to regulated products. Understanding these factors remains important for identifying and addressing disparities in health, which is vital to achieving health equity.

And advancing health equity is a top priority for the FDA, including CDRH and the Office of Minority Health and Health Equity. The performance of medical devices needs to be well understood to mitigate any negative unintended consequences for patients in different populations. We are committed to the continued evaluation of the safety and effectiveness and

availability of regulated products, including how medical devices perform across racial and ethnic groups.

So today's discussion on the real world performance of pulse oximeters and factors that impact their accuracy, we'll inform FDA as we consider the regulation of these medical devices. So with that, again, I'm glad to be able to join you today and I look forward to this important discussion. Thank you.

Dr. Nathan: Thank you very much, both Dr. Shuren and Rear Admiral Richardae Araojo. I'd now like to invite the FDA review team to start their presentations. The first session will have presentations covering the regulation of pulse oximeters, pulse oximeter technology standards for pulse oximeters, and the real-world performance of pulse oximeters.

The second FDA session will cover various topics related to clinical study design methods and analyses for studies evaluating pulse oximeters as well. I would like to remind public observers at this meeting that while this meeting is open for public observation, public attendees may not participate except at the specific request of the panel chair, that'd be me.

FDA. You may now begin your presentation.

FDA PRESENTATIONS

PULSE OXIMETERS: TECHNOLOGY, ACCURACY LIMITATIONS, AND REGULATION — DR. LEE

Dr. Lee: Thank you. Good morning. I'm here to present regarding the agency's regulations of pulse oximeters, which includes reviewing the technology, its limitations on accuracy, and how the agency regulates the device type. My name is James Lee and I am the

division director here at OPEC OT-1 Division for Sleep Disorder, Breathing, Respiratory and Anesthesia devices.

My co-author today is Mr. Neil Patel, our senior reviewer in the anesthesia devices team. Pulse oximeters are widely used in various situations and clinical settings in particular, and in recent years, there has been an increase utilization and public familiarization with Pulse oximeters inclusive of the COVID pandemic.

In general, the FDA considers the environment of use and intended populations when evaluating the safety and effectiveness of devices like Pulse oximeters factors for consideration in our review include the fields of use and populations like healthcare providers and facilities, and also patients who might require indirect measurements of blood saturation or SpO2.

It is important to keep in mind that these measurements of oxygenation levels in the blood are ratios of oxygen bound Hemoglobin are surrogate measurements of physiological parameters, meaning that they are not direct measurements of oxygenation levels, but indirect measurements of oxygenation via light wave absorption of the tissue.

True oxygen levels in the hemoglobin are from SaO2 measurements obtained via arterial puncture, which is considered the gold standard for assessment of blood oxygenation, saturation levels, blood gases, SaO2s while the most accurate and assessment of oxygenation levels involves an invasive procedure.

SpO2 is an estimate, therefore has an error, and the measured values are expressed as a percentage with a labeled error range, as mentioned on the prior slide SpO2is estimated as a percentage of oxygenated hemoglobin over the sum of oxygenated and deoxygenated hemoglobin. An important factor is that oxygenated and deoxygenated hemoglobin have different absorption spectra, which allows the use of optical techniques that utilize two or more

different wavelengths of light to measure differences in absorption of the tissues within the targeted field

The lower figure shows two typical wavelengths of light used to measure SpO2, we have red light at 660 nanometers and infrared light at 940 nanometers. The total absorption is relatively low, allowing enough light to pass through the light emitting LEDs to the photo detector sitting opposite as shown in the figure above oxygenating hemoglobin defined as O2HB in the lower figure absorbs or attenuates the amount of infrared light more than compared to the red light.

The opposite is true for deoxygenated or reduced hemoglobin, which absorbs more red light and allows relatively more infrared light to pass. A second basic principle of pulse oximetry is the presence of a pulsatile arterial signal, which allows changes in light absorption to be measured. The upper figure to the right shows a changing absorbance of light during an arterial pulse.

The portion that changes is often referred to as the AC portion of the signal. The portion that remains constant due to the presence of residual arterial blood, venous blood and other tissues is referred to as the DC portion of the signal with a ratio of the AC to DC signal, both red and infrared light, a ratio of ratios referred to as the R value, which again is a ratio of AC to DC for the red signal over the ratio of AC/DC for the infrared signal can be calculated with the calculated R value. You can estimate the SpO2 value using empirical calibration data as shown in the lower figure.

Regulation of prescription Pulse oximeters are based on a moderate device risk level, meaning these devices are reviewed under the FDA CDRH 510(k) program and cleared under the basis of substantial equivalents as described in our 2013 FDA guidance document for pulse

oximeters. These devices undergo clinical testing, bench testing, and other standardized tests like biocompatibility to review the devices relative safety and effectiveness against predicates with the same intended use.

These devices are considered for spot checking and trending tools and not for diagnosis of a disease. A cornerstone of the FDA guidance document in review is the use of internationally recognized standard for the evaluation of pulse oximeters. This standard establishes the basic safety and testing needed and defines protocols and tests that are utilized in the evaluation of a pulse oximeters defined performance. Specifics on these tests and framework of the standards will be covered by our next presentation. These pulse oximeters are labeled for trending or spot checking for oxygen saturation levels of patients who are hospitalized in doctor's offices, and also for prescribed use in homes by patients under the care of a physician. Regarding what is considered to be over the counter regulation of these devices are under enforcement discretion and fall under product codes like PGJ, or OCH. These are specifically for general wellness, which apply to sports and aviation use, but are not intended for medical purposes. For this reason, they do not undergo FDA pre-market review and can be sold directly to customers and patients. Please see our reference here at the link for our general wellness guidance on low-risk devices.

The FDA reviews medical technology like oximetry under established regulations under the Code of Federal Regulations or CFR. Here we would like to provide the established definitions of these regulations.

Furthermore, the FDA buckets or groups within these regulations by product codes, which allow for a more granular definition of products and allow for detailed assignment according to both established regulations and sorting devices by product code.

What should be noted is that oximetry devices utilize both visible and invisible wavelengths of light and rely on absorption by tissues where the remaining signal can be transmitted to an opposing sensor. Innate configurations like this, while effective in picking up physiological signals, have limitations due to technology and configuration of the sensor system highlighted here are several confounding factors that may contribute to either inaccuracies of the signal or increase the variability in relation to ground truth or absolute oxygen saturation. In hemoglobin, one of the main considerations is a level of skin pigmentation, which is a major part of our discussions today.

In addition, there are patient dispositions that may arise or challenge the accuracy of the SpO2 signal due to the ratio of oxygenated and deoxygenated and hemoglobin, which does not include considerations for other types of hemoglobin. Therefore, disorders of the hemoglobin molecule or anemia may cause a reduction in pulse oximeter accuracy.

In addition, other factors like intravascular dyes, tattoos, nail polish, and ambient light may interfere with the sensor and contribute to lower levels of accuracy. Regarding MDRs. Please note that this information was included in our executive summary. We conducted a keyword search using skin for adverse events submitted to our Medical Device Reporting or MDR system dating back some 22 years.

Reports of death were reviewed for inaccurate SpO2 readings as a factor of which there were 99 such reports, but a detailed review found that there was a lack of sufficient information to evaluate and association between these adverse events that were reported in inaccurate readings.

Please note that this was not included in our executive summary. However, we are including this slide here to provide a full context of information for consideration. The chart on

the right shows the total number of MDRs spiked after the 2013 guidance was published and have trended downward. An additional search was conducted with keywords such as inaccurate, inaccuracy and incorrect to see if those helped in informing a potential source of the problem.

Continuing with the agency's analysis, we took a granular look at the information contain in the individual reports. Again, this was not included in the executive summary. A further detailed review found that there was a lack of sufficient information to evaluate the association between adverse events and inaccurate readings. We found that the top three reported health effects relate to no consequence to the patient, and the top three reported device problems related to inaccurate measurements by pulse oximeter devices.

Importantly, although MDRs are a valuable source of information, this passive surveillance system has limitations. The incidents, prevalence, or cause of an event typically cannot be determined from this reporting system alone due to under-reporting of events, inaccuracy in reports, lack of verification that the device caused a reported event, and lack of information about the details such as frequency of device use.

Because of these limitations, MDRs comprise only one of the FDA's several important post-market surveillance system data sources. These reports along with data from other sources can contribute important information to a medical device's benefit risk assessment. Reports like these help the agency look at a class of devices and consider in totality of post-market performance.

In total, while oximetry has provided an excellent tool in evaluating noninvasively important vital signs like oxygenation levels, our conversation will continue with our next speaker, Dr. Sandy Weininger from our Office of Science and Engineering Laboratory, who will

discuss how the FDA uses international standards to address issues of effectiveness and basic safety.

STANDARDS FOR PULSE OXIMETERS: ISO 80601-2-61: 2017 — DR. WEININGER

Dr. Weininger: Welcome distinguished panel members and thank you for participating in this important activity to understand the performance of Pulse oximeters. I am Sandy Weininger, an Electrical Biomedical Engineer in CDRH Office of Science and Engineering Labs, where I perform regulatory science research on the safety of interoperable systems and consults for the safety of electromedical devices.

I am the current chair of the ISO IEC Pulse Oximeter Committee and have been involved in oximeter standards and performance assessment of pulse oximeters since the early 1990s. Pulse oximeters were introduced in the early eighties, and clinical societies very rapidly adopted them to support safe care. It was quickly recognized that engineering safety needed to be addressed as well as technology improved over the years, devices have evolved in response to clinical observations of underperformance, including darkly, pigmented and lightly pigmented individuals. Motion tolerance and low perfusion.

Let's first look generally at how safety is supported by standards. Then specifically at how the pulse oximeter standard addresses these issues. What does it mean to be safe? It means that the relevant hazardous situations have been identified and risk control measures built into the device. This happens most effectively. When one has a structured design development process in place and then makes good decisions within that process. ISO 13485 establishes the process for design development of medical devices. It works in concert with ISO 14971, which provides the process for identifying and assigning risk to hazardous situations. These standards establish the how to for design, but not the how good should the design be. The how good comes

from IEC 60601, the general standard for the safety of Electromedical devices, which provides answers in terms of safety requirements, test methods, and acceptance criteria. ISO 80601-2-61, Our Pulse oximeter standard, inherits these processes and requirements and customizes them for pulse oximeters. In general, international standards are intended to promote a minimum level of safety and performance that is accepted by regulators around the world. The standard is developed with the intent of harmonizing with these local requirements as has been done with FDA's guidance document.

Let's look at what the scope of the standard is for devices that provide an estimate of functional arterial hemoglobin saturation in all use environments. The standard provides requirements, test methods, and acceptance criteria for engineering basic safety and essential performance but does not address what performance is necessary for our particular type of clinical application. In addition to the safety requirements, the standard provides significant supplemental and educational material to help properly understand and interpret the results from an oximeter. The standard recognizes and addresses sources of error and attempts to address these to facilitate testing and characterization of the device.

One difference from the FDA guidance document is the standard does not specify the type of use, spot testing versus continuous monitoring. The standard uses the construct of whether the device has a physiological alarm signal, as this was determined to be the primary difference in clinical use, there's no need for an alarm on spot checking.

FDA has a process to recognize standards for use and regulatory activities. While FDA encourages the use of recognized standards, it may not recognize the full extent of the standard. This is the case for Pulse oximeters. FDA does not recognize the clause that sets a limit on accuracy, in the U.S. the FDA guidance document takes president.

In the FDA guidance document, accuracy limit is set to 3% for transmittance oximeters, where in the standard the limit is 4%. These were found to be the worst performers at the time that the limit was set. There are times when health delivery organizations, clinicians, or consumers may want to go with a device with different performance or features.

Technology continues to advance, enabling better performing devices. The standard is updated to reflect these developments. Let's look into the standard more deeply. The primary measure of performance of a pulse oximeter is defined to be accuracy. The standard defines accuracy as the root means square difference between the estimated value of saturation as measured by the pulse oximeter and the reference value as measured by a well performing co.

This metric, ARMS, is presented as a single statistic to communicate the average bias and precision across the operating range, 70 to 100, additional subagents can be specified with a minimum width of 15% to prevent any false indication of performance. Accuracy at higher ranges is typically much better and should not be used to represent performance at lower saturation values and alternative currently being proposed is to add disclosure for the bias and precision for each decade.

The downside is this is more information to comprehend and understand the standard requires disclosure of performance curves, Bland-Altman plots to communicate. The actual performance over the operating range. Standard has focused on conveying clear and clinical useful information regarding pulse oximeter performance.

The fixed declared range is one of many constraints designed to provide a standardized reporting characteristic that enables comparison across devices and supports a determination of substantial equivalency. An oximeter must achieve an RMS less than 4% to be in compliance

with the standard. The standard recognizes that there is an inherent uncertainty in the estimate of SaO2 by pulse oximeters, some of which you're hearing about today.

These contribute to the calculated ARMS. The standard defines the formula. Note that it looks like a standard deviation, and in fact, it behaves like one. As a result. If we knew the true value of saturation, and of course we do not without an arterial stick, then there's an expected range of values estimated by the pulse oximeter.

For example, if the measured value from a CO oximeter is 92%, a pulse ox could be expected to read below 86 or above 98 in five of 100 patients. Given the number of times oximeters are used per day, a significant number of estimates from a properly operating pulse oximeter can be expected to appear to be inaccurate.

How does one collect the data? From a measurement that has a significant variability? Standard defines a protocol for controlling the variability in collection of verification data. It relies on healthy subjects that can safely tolerate repeated desaturations, can tolerate an arterial blood draw, can follow instructions to perform the maneuvers, are capable of providing informed consent and be physically located in a laboratory with the necessary testing equipment. Methods require a stable desaturation plateau to perform simultaneous measurements of SpO2 and SaO2. This is in contrast to the convenience samples that are collected during real-world studies.

The use of laboratory desaturation studies enables stable measurement conditions for simultaneous samples to be obtained. The goal is to minimize variability in testing so that device performance can be compared and a maximum performance can be communicated to the clinician. It is known that real-world performance is likely to be degraded when conditions deviate from the idealized testing environment in the lab.

The protocol recommends parameters to be included in the study. The standard recommends including ranges of age, gender, finger size, and skin color, and requires the population demographics used to verify the performance of devices be disclosed. In the U.S. these inclusion criteria are narrowed by recommending at least 10 subjects with a minimum of two or 15% with darkly pigmented skin.

Additional factors affect performance. The standard goes to great lengths to document external factors that can adversely affect pulse oximetry accuracy, and these are detailed in the standard for consideration during pulse oximetry laboratory testing. These factors include ambient light, low motion, low perfusion, electromagnetic infrared interference, certain dyes, and dysfunctional hemoglobins

Device fidelity is another issue that needs to be understood in order to collect high quality data. Device and user settings and algorithmic complexity can influence the fidelity of the estimate of the depth of the desaturation and duration of the event. The pictogram shows that the Nader of saturation is underestimated and the onset is delayed.

The causes of these errors may be inherent in the device design, in user settings, and from external factors such as inflation of a blood pressure cuff upstream from the oximeter probe on the same arm. Capturing context, for example, sensor location and bandage pressure may be challenging in real-world studies, the panel is asked to consider the influence of lay use with respect to pigmentation and lay use. The standard recognizes that there is increasing use by laypersons and is trying to standardize language directed at this population. Regarding best practices for using and interpreting the data from oximeters, the committee awaits the panel's deliberation regarding pigmentation and lay users as input to updating requirements in the standard.

In conclusion, 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.

A SYSTEMATIC LITERATURE REVIEW OF THE REAL-WORLD PERFORMANCE OF PULSE OXIMETERS — DR. O'NEILL

Dr. O'Neil: Good morning. My name is Allison O'Neill and I'm the Safety Signal Coordinator for OHT 1 in CDRH. I will be presenting a systematic literature review of the real-world performance of pulse oximeters. First, I will begin by explaining a bit about the history of the safety signal and the purpose of the lit review. The scientific community has been aware for some time that certain patient factors may impact the accuracy of pulse oximeters, and one of these factors may include skin pigmentation.

In 2013, FDA published a guidance document which makes recommendations for pulse ox manufacturers specifically that each pre-market study should have subjects with a range of skin pigmentations between 2013 and 2020 there was some limited literature on this subject. In 2020, the Covid 19 pandemic brought increased awareness of the importance of pulse oximetry, both in the hospital and home settings.

In December, 2020, Sjoding et al published an article which reported that black ICU patients had approximately three times the rate of occult hypoxemia compared to white patients. This article concluded "that reliance on pulse oximetry to triage patients and adjust supplemental oxygen levels may place Black patients at increased risk for hypoxemia" The article received

media attention and spurred additional interest in related research studies. The FDA wanted to provide the public with information on the interpretation and limitations of pulse oximetry.

In February, 2021, the FDA released a safety communication with recommendations for providers and patients. At that time, the FDA committed to the continued evaluation of the safety and effectiveness of pulse oximeters, including a review of relevant literature.

Today I will briefly present a high level overview of our systematic review of relevant literature published since 2013. The panel will be asked to incorporate this information in the discussion of the clinical evidence from the scientific literature about the accuracy of pulse oximetry among patients with darker skin pigmentation. We conducted a search of the publicly available PubMed database to identify published articles relevant to the real-world performance of pulse oximeters as it relates to skin pigmentation.

162 publications were identified from the initial search, and six additional publications were identified by cross reference and additional searching after the initial search date. These 168 publications were reviewed for possible inclusion. We excluded articles published before the 2013 FDA guidance document, which outlined more stringent recommendations for accuracy. We also excluded publications that were not relevant to the topic or did not include any clinical data such as nonsystematic reviews and editorials. After exclusions, 28 publications were selected for inclusion in the review. I will now briefly describe the overall body of literature, but in the interest of time, I will not cover each individual study for more detail please see Appendix A of the executive summary.

Of 28 selected articles, there were three systematic literature reviews and three lab studies. There were also 22 publications that used real world data from hospital patients, usually in patients from the ICU or surgical unit. Of the 22, 7 were considered cross-sectional studies

that prospectively collected data and 15 were considered retrospective studies that relied on electronic health record data. Most studies were conducted using U.S. patients, although there were some from Europe, Asia, Africa, and Australia.

The retrospective studies all relied on self-reported race and or ethnicity. Whereas five of the cross-sectional studies were able to prospectively collect pigmentation data via either the Fitzpatrick scale or in one case the Munsell System. Many use the endpoint of bias between paired measurements or occult hypoxemia with varying definitions. Others use an endpoint such as less than or equal to 94% SpO2 to qualify for treatment. These variable definitions are listed by study in the appendix.

Three systematic reviews pertaining to the review topic were published in 2022. There was some overlap in the articles that were included in each review. Two of the three reviews concluded that there were overestimations in people with darker skin pigmentation based on the evidence reviewed. The systematic review by Cabanas et al found the eight of 11 studies at low risk for bias found inaccuracies due to skin pigmentation.

The meta-analysis by ~~~she et al, reported about a 1% bias in both people with high level of skin pigmentation and people described as Black or African-American based on moderate to low certainty evidence. The third review was focused on poor perfusion, but it did also comment on skin pigmentation. The authors noted that only 1 of 22 studies in their review controlled for skin pigmentation and none strictly followed FDA recommendations for including at least 15% subjects with dark skin pigmentation.

Systematic reviews typically represent a higher level of evidence than one study alone. Of course, it should be noted that lit reviews inherently have the same limitations as the studies that are included in the review. Also, all three of these reviews contain articles going back to the

1970s and eighties, which is a larger timeframe than considered for the FDA's review, which included articles published since 2013.

Regarding the seven cross-sectional studies, each study reported either the bias between SpO2 and SaO2 or an odds ratio for occult hypoxemia. Five of seven papers reported a positive bias or statistically significant odds ratio for subjects with darker skin pigmentation or African American race compared to the reference group.

The largest study, Henry 2022, collected approximately 130,000 simultaneous paired measurements of SaO2 and SpO2o in more than 26,000 ICU or surgical patients at three medical centers in the U.S. and they reported that black patients were 1.65 times more likely to experience occult hypoxemia compared to white patients. The odds ratios for Asian or American Indian patients were not statistically significant. Additionally, occult hypoxemia was significantly associated within hospital mortality.

Of note, one study also reported that black race was also associated with hyper-toxemia. The details of the cross-sectional studies that were identified by the review are listed in table 12 dash one. In the executive summary, the 15 retrospective studies were more varied in design and chosen endpoints. Overall, most studies reported a statistically significant association between Black race and either occult hypoxemia or SpO2. Additionally, some individual studies also have reported an association with certain clinical endpoints including treatment probability, organ dysfunction, and in-hospital mortality.

It's important to note that real world retrospective studies have certain limitations and nuance differences in study design. I will discuss this more in depth on an upcoming slide. The largest study was published by Wong et al in 2022 with electronic health record data from more than 79,000 patients admitted to the ICU.

Paired measurements occurred within five minutes of each other. This paper reported a higher incidence of occult hypoxemia in Black and Hispanic patients compared to White. occult hypoxemia was associated with greater organ dysfunction and higher in hospital mortality. Please refer to table 12 dash two in the executive summary for a table of results by study.

You will hear presentations later today by some of the authors of these studies, including Doctors Sjoding, Wong, Fawzy, and Gottlieb. The lit search identified three laboratory studies using healthy volunteers. The first was a desaturation study performed in Korea to compare pulse ox sensors with or without optical crosstalk.

The authors reported a measurement error in African American subjects compared to Caucasian or Asian in the sensor that did not prevent crosstalk. The second was a case series performed in the U.S. designed to compare photo acoustic imaging versus a pulse oximeter. They reported no significant differences in SpO2 by skin type.

The third study was a desaturation study performed in the U.S. with 491 subjects classified as dark, medium, or light skin pigmentation. There was a small positive bias in the dark pigmentation group. For more details on each of the lab studies, please see table 12-3at the end of the executive summary.

Now I will discuss the important limitations that should be considered when assessing the published literature and especially real-world data.

First study variables were defined differently by study. There was no standardized cutoff for what SpO2 and SaO2 levels merit classification as occult hypoxemia. This may be defined differently by site based on clinical need. Skin pigmentation data may be captured by different tools, although race slash ethnicity was often used as a proxy.

Secondly, pulse ox accuracy often appears worse in real-world studies than in desaturation studies. Using real world data means that the paired measurements of SaO2 and SpO2 are usually not simultaneous, which may lead to larger disagreements due to normal fluctuations in SaO2 and treatment effects on SaO2.

Also, researchers must often rely on self-reported race ethnicity as skin pigmentation data such as Fitzpatrick scores are typically not collected for hospital patients, and there may be some residual confounding from variables not captured.

Thirdly, the real world population is more heterogeneous than a controlled lab study, typically including very sick patients from an ICU population rather than healthy volunteers. The prevalence of hypoxemia may vary between such groups impacting the positive and negative predictive values and confounding comparisons of the occult hypoxemia rate.

Fourth, heterogeneity of technology used. Pulse ox accuracy will most likely vary depending on which device or product is used. And this data was not provided in every article. It can be assumed that brands, models and the use of reprocessed sensors varied across hospitals and even within the same hospital.

Additionally, technology has advanced over time, and thus pulse ox accuracy may have changed over time. Finally, there is often inherent publication bias in any literature review, which means that results showing statistically significant associations are more likely to be submitted and accepted for publication compared to results that do not show significant associations overall, despite the limitations noted, there appears to be mounting real-world evidence from studies capturing tens of thousands of patients.

The pulse oximeter accuracy may vary by self-reported race and or skin pigmentation. However, there's still a need for prospective studies that utilize standardized measurement of

skin pigmentation. Capture simultaneous measurement of SaO2 and SPO2o paired data and systematically collect data on important confounders in order to have more of robust evidence about the impact of skin pigmentation on real world pulse oximetry.

This has been a high-level overview of the recent relevant literature. Later today, there will be presentations from a number of the authors of real-world evidence studies relevant to the topic. At the end of the meeting, the panel will be asked to discuss the questions shown on the slide.

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 the following. 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 over-the-counter 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, or motion artifacts?

OVERVIEW OF DESATURATION STUDIES IN PULSE OXIMETER 510(K) SUBMISSIONS — DR. JUNG

Dr. Jung: Hello, my name is Mary Jung and I'm an epidemiologist in the Office of Clinical Evidence and Analysis within CDRH.

Today I will be presenting an overview of the desaturation studies used in cleared 510(k) submissions for pulse oximeters. Given concerns about the potential impact of skin pigmentation

on pulse oximeter accuracy, the FDA wanted to better understand how skin pigmentation information is captured and reported in cleared 510(k) submissions.

Thus, the study was conducted to provide initial examination of skin pigmentation information and cleared 510(k)s for prescription use. Pulse oximeters. The study objectives were first to examine pre-market clinical study characteristics and reporting of skin pigmentation classification, and second to assess reporting of factors in device labeling that may potentially impact device accuracy such as skin pigmentation.

On this slide, the process used to ascertain the final sample of examine 510(k) submissions is illustrated in the flow chart out of a pool of 420, 510(k) submissions received between January 1st, 2000 and December 31st, 2020. An approximate 10% sample or 44 submissions with equal numbers of pre and post guidance submissions was identified through an iterative process of random selection and review for eligibility by subject matter experts.

This 10% was selected to provide an initial examination of 510(k) submissions with the post hoc assessment to ensure representation from manufacturers with the largest market share. The 22 pre guidance submissions represented 7% of all pre guidance, 510(k)s in the original pool, and 44% of pre guidance submissions.

While the 22 post guidance submissions represented 20% of post guidance, 510(k)s and 65% of sampled post guidance submissions. Submissions were excluded if clinical data were not required. The device was not a pulse oximeter. The device was for pediatric use only, or the data was not relevant. Pulse oximeters that did not require clinical data included 510(k)s with well defined manufacturer modifications of their own market devices and other words, special 510(k) submissions and submissions, leveraging predicate device data for modifications that did not

affect the oximetry technology or performance. An example of which would be addition of respiratory rate software.

Next are the data elements that were extracted for the 510(k) assessments. For objective one, Examining the clinical desaturation studies. Extracted data elements are listed to the right, including the number of subjects, number of paired oxygen measures per subject, availability of patient line level data, inclusion criteria, reporting of participant characteristics, use of Bland-Altman plots for analysis and skin pigmentation assessment, and categories.

For objective two. device labeling for submissions were reviewed to extract listed factors that were indicated as potentially impacting device accuracy such as skin pigmentation. Descriptive analysis was performed on the data elements to examine their distributions overall and also by pre and post guidance.

Key differences will be highlighted in the subsequent slides. The overall clinical study sample sizes found in the cleared 51(k) submissions are presented in the table. Desaturation studies and the pre and post guidance submissions had a median of 12 study subjects ranging from 6 to 33 and a median of 24 paired SaO2 and SpO2 measurements ranging from 13 to 35 per subject.

The 2013 FDA guidance recommended providing a line listing of data as shown in the figure to the right reporting of patient line level data increased from 54.5% of pre guidance submissions to 95.5% of post guidance per the 2013 guidance recommendations. To characterize the clinical study population common inclusion criteria for all the sample desaturation studies are shown in the left hand figure all pre and post guidance desaturation studies listed being healthy as an inclusion criterion for subject participation.

43.2% of studies specified in their inclusion criteria that participants must be non-smokers. And 50% indicated that participants must not have specific medical conditions for eligibility, such as hypertension, cardiovascular disease, and respiratory disease.

The second figure shows the percentage of sampled studies reporting various participant characteristics. In general, demographics were commonly reported with age, sex, and race, ethnicity being provided more than 72% of studies, while nationality and biometric measurements specifically weight, height, and finger size were reported in less than 28%.

Regarding analyses, a large increase was observed as shown in the figure to the far right and the number of submissions including Bland-Altman plots per the 2013 guidance, specifically 36.4% of pre guidance and a hundred percent of post guidance submissions used Bland-Altman plots.

Clinical studies used inconsistent methods for assessing and categorizing skin pigmentation with many of these classifications, lacking clear and standardized definitions as shown in the table. Categorization of skin pigmentation varied, including singular assessments indicating only subjects with dark skin pigmentation use of two to five skin pigmentation categories ranging from light to dark and application of scales, specifically the von Luschan and Fitzpatrick scales. Although the skin pigmentation categories remain heterogeneous and poorly defined post guidance, skin pigmentation reporting post guidance improved with 95.5% of submissions specifying skin pigmentation categories as compared to 50% of pre guidance.

Shifting our focus to device labeling for objective two. This slide illustrates factors listed in the labeling by the sponsor that may impact device accuracy. The top three listed factors across all examined labels were use of intravascular dyes, excessive light, and the presence of

electromagnetic or electrosurgical sources. The 2013 FDA guidance indicates that labeling should include all applicable safety information, warnings, cautions, and notes.

Subsequently, inclusion of the factors in blue shading increased in the post guidance as compared to the pre guidance submissions when examining inclusion of skin pigmentation.

18.2% of device labeling for all submissions included dark skin pigmentation as a potential factor that may impact device accuracy.

The percentage of labels with dark skin pigmentation increased from 9% among pre guidance labels to 27% of post guidance.

In summary, for objective one, information on pulse oximeter clinical studies that was recommended in FDA guidance were provided in a larger proportion of post guidance submissions, including greater specification of skin pigmentation classifications, availability of patient line level data, and use of Bland-Altman plots.

In addition, a wide variety of skin pigmentation categories was observed, many of which did not include clear or standardized definitions. For objective two, examining device labeling increases in the reporting of factors that may impact accuracy were observed from pre to post guidance, including dark skin pigmentation.

However, the potential impact of dark skin pigmentation on device accuracy was not provided in 73% of post guidance 510(k) submissions. This finding identifies a potential area for consideration regarding required contents for device labeling. The panel will later be asked to deliberate and make recommendations on standardizing, categorizing, and reporting skin pigmentation, as well as potential modifications and recommendations to labeling for prescription use pulse oximeters to better communicate their limitations to clinicians and patients. Thank you everyone for your time and attention.

PREMARKET DESATURATION STUDIES FOR PULSE OXIMETERS — DR. HENDRIX

Dr. Hendrix: Good morning. I'm Kumudhini Hendrix, Anesthesiologist and Medical Officer at CDRH.

My presentation is on desaturation studies required for market pulse oximetry submission. I will discuss when pre-market desaturation studies are required, how they're expected to be conducted, data submissions for FDA review, and additional convenience sample verification for neonatal use. Finally, I will discuss limitations of pre-market desaturation study.

FDA guidance recommends pre-market desaturation testing to be conducted per ISO 80601-2-61. As per the 2013 FDA pulse oximeters, pre-market notifications, submissions, guidance, in vivo pre-market desaturation testing for SpO2 accuracy is recommended under laboratory conditions for all new pulse oximeters as well as for all prior cleared pulse oximeters with significant electro optical sensor modifications and or SpO2 algorithm modifications.

The general purpose of in vivo pre-market desaturation studies is to verify the Sp02 accuracy of pulse oximeter device performance in comparison to the gold standard measurements of blood Sao2 by a co-oximeter over the specified SaO2 range of 70 to 100%. Therefore, typical labeling is for general indication for non-invasive measurement of blood oxygen saturation.

If a manufacturer wishes to seek a specific clinical indication for use of a pulse oximeter, for example, to screen for or diagnose a disease or condition, then additional clinical safety and effectiveness data is requested to be. Per annex EE of the referenced ISO standards, the pre-market desaturation study should include sufficient number of subjects to attain statistical significance necessary to demonstrate a specified SpO2accuracy, a minimum of 200 pooled data pairs.

That's SpO2 and SaO2 distributed evenly over the tested range of 70 to 100% from a minimum of 10 healthy volunteers is recommended. As per FDA guidance. The study should have subjects that range in age, gender, and skin pigmentation. FDA guidance recommends at least two subjects or 15% of subjects, whichever is greater to be dark pigmented.

For ISO standards, carboxy hemoglobin should be less than 3%, met hemoglobin be less than 2%, and total hemoglobin greater than 10 grams per deciliter

For ISO standards, allowable testing conditions for pre-market desaturation Studies include application of warming techniques to improve circulation and pulse amplitude at the pulse oximeter probe site covering of pulse oximeter probes with opaque material to prevent optical interference. And the addition of carbon dioxide to inspired gas mixture to maintain normal curia and to prevent respiratory alkalosis, secondary to hyperventilation caused by hypoxia.

Catheters placed within the artery on all test subjects prior to desaturation A cleared pulse oximeter is placed as reference to detect stable plateaus. The fraction of inspired oxygen is varied in a stepwise manner, so to achieve a series of targeted steady state saturation periods when the reference pulse oximetry stabilizes for 30 seconds or more arterial blood is sampled for comparison of simultaneous data pairs of pre-market device SpO2, as well as SaO2.

Per FDA guidance, all the following data as it pertains to pre-market desaturation testing is requested to be submitted for FDA review pertinent test apparatus used. Inclusion exclusion criteria, number of samples taken per subject, specific conditions of testing such as motion, low pulse amplitude, laboratory conditions, type in frequency of motion testing if applicable criteria and methods for determining stability of reference arterial blood oxygen saturation at the pulse oximeter sensors site

Desaturation profile, including target plateaus and ranges, as well as formula used for determination of root mean square difference that is ARMS. Additionally, individual data pairs as well as pool data pairs are requested to be plotted in a modified Bland-Altman plot to show arterial saturation that is SaO2 on the X axis as it is considered ground truth.

These plots on the right side of the slide have Sao2 on the X axis, and the difference between SpO2 and SaO2 on the Y axis. FDA guidance also requests population mean bias that is mu zero between subject variance. That is sigma mu squared within subject variance that is sigma squared, as well as upper 95% and lower 95% limits of agreement.

FDA asks for discussion and rationale for any points excluded from the analysis. As for FDA guidance, all noticeable outliers are requested to have a discussion of the state of health subject characteristics, test, set up, test, procedure, and any other factors that may have affected these data points. Also, a discussion of how the outliers do not raise safety and performance concerns regarding the accuracy of the device is requested as well.

Devices intended for use with neonates are requested to conduct pre-market desaturation testing with neonatal sensors on adult subjects. Again, any new or significantly modified sensor intended for use in neonates are recommended with such pre-market desaturation testing on adult subjects.

Additionally, convenience arterial blood gas sampling drawn in neonates for clinical decision making in the real-world is requested to verify form, fit and function. The pulse oximeter probe is placed in the same circulatory stream as the sampled artery per ISO standards. FDA requests justification of sample size subjects and analysis technique. Pre-market desaturation studies have limitations that may impact real-world performance.

Subjects are healthy and selected from a pool of limited volunteers. Sample size is typically not large. Limited to 10 to 20 subjects. Skin pigmentation is generally qualified subjectively as light, medium, or dark, and not specific to any particular anatomical site, such as dorsum of hand, forehead, or ventral aspect of forearm.

Additionally, currently pediatric indications are verified only with desaturation studies on smaller adult volunteers, form fit and function remain unverified in non-neonatal pediatric populations. Conditions are ideal and remain quite different from actual clinical use. Hands are warmed up prior to pre-market desaturation studies for optimal performance sensors are shielded to avoid electromagnetic interference.

Data pairs sets SaO2 and SpO2 are obtained simultaneously after a stable plateau of 30 seconds or more. Verified by another cleared reference. Pulse oximeter subjects undergo normobaric hypoxic conditions and therefore experience less adverse effects from respiratory alkalosis. Subjects are desaturated down to tolerable limits only and thus lower deciles, like 70 to 80% may be acquired from a few individuals who can tolerate lower inspired oxygen levels.

Accuracy of pulse oximeter is derived from pool data pairs from all subjects within the entire tested range, typically 70 to 100%. Currently clinically relevant thresholds such as 94%, 88%, or ranges such as 86 to 94% are not analyzed further. Importantly, the study is not powered to determine significant difference between cohorts such as age, sex, or specific pigmentation values.

Given the limitations and the purpose of pre-market desaturation testing, it is important to consider the following questions. Number one, Currently pulse oximeters are not diagnostic medical devices. Instead, they are verified to be estimates of SaO2 within two, a RMS for 95%

of the data pairs. In this function, they are tools and the practice of interpretation and verification of SpO2 is relegated to the clinician or the learned intermediary.

Is the current indication adequate for clinical decision making? Are there clinically relevant ranges or thresholds where a greater degree of certainty is required? And what is the needed accuracy for these critical values and ranges?

Number two, can current pre-market desaturation studies be improved upon to provide clinically relevant pulse oximeter performance for all populations in the clinical setting?

Thank you for your attention.

STATISTICAL CONSIDERATIONS IN THE EVALUATION OF PULSE OXIMETERS — DR. PENNELLO

Dr. Pennello: Good morning. I'm Gene Pennello, a Mathematical Statistician at CDRH. My presentation is on statistical considerations in the evaluation of pulse oximeters, I will first discuss performance metrics. For pulse oximeter studies, common performance metrics include mean bias, the mean of the difference between the pulse oximeter measurement SpO2, and the reference standard measurement SaO2 of arterial oxygen saturation. Precision, the standard deviation of the difference SpO2 minus SaO2 mean absolute deviation and accuracy. Root means square, these metrics are averages across all paired measurements of SaO2, SpO2 from all. If a metric is not constant across all values of SaO2, for example, if SpO2 bias varies with SaO2 value, then a difference in the metric between two groups could be difficult to interpret if the groups differ in SaO2 distribution.

SpO2 bias may be adjusted for, SpO2 distribution using analysis of co variants or ANCOVA. In ANCOVA, SpO2 is regressed on skin color group effect and SaO2. The skin color effects are differences in the SpO2 bias between the skin color groups at the same SaO2 value.

ANCOVA facilitates comparing would can be used to adjust for other co variants besides SaO2, for example, age, sex perfusion, index finger size, et cetera, and adjusts group comparisons for regression to the mean.

I would like to elaborate on regression towards the mean. Regression to the mean is a statistical phenomenon that occurs between two random variables when variation occurs between or within the variables. A classic example of regression to the mean is tall parents tending to have shorter children and short parents tending to have taller children and vice versa.

Another classic example is comparing two methods of measurement. For example, a low SpO2 value of oxygen saturation will tend to be accompanied by a higher SpO2 value of oxygen saturation, even when the two methods are unbiased for each other. Regression to the mean and two methods of measurement occurs because of variation between the measurements made by the methods measurement in precision that is variability in the repeated measurements made by a method or natural biological variation.

If paired measurements are not simultaneous, unless the two measurements are perfectly correlated, they will exhibit regression to the mean to illustrate the data set in which SPO2 was unbiased for any value of EO2 was constructed by promoting the labels of EO2 and SPO2. In the real world data from Ed Meyer at all 2018.

The Quantile plot on the left lies on the line of identity indicating that SPO two is indeed unbiased for all values of S A two. However, the regression fit of SPO2 on s A two, the right panel indicates non-negligible regression. The mean in that low values of S A two, for example, 88% tends to be accompanied by higher values of SPO2expected value of about 90.5%.

And conversely, high values of SpO2, for example, 99% tend to be accompanied by lower values of SPO. Expected value of about 98%. I would now like to discuss box plots of a oh

two at each S P oh two value stratified by skin color. In many publications of real world studies box plots of the A oh two median inter cartel range, and outliers by the SPO two value are shown side by side for each skin color group and the box plots reported by shouting at all 2020.

The left panel for each SPO2 value. The SEO two values tend to be lower for black patients than white patients. In the paper, a cult hypoxemia was defined as SIO two less than 88% when SPO two is between 92 to 96%. The red zone indicates a cold hypoxemia, which occurred more frequently black patients than white patients.

In contrast, in the controlled desaturation study by Barker and Wilson, the right panel for every SPO two value, the difference in the box plot between black patients and white patients, which was much smaller than that observed in the shouting at all real world study and a cold hypoxemia almost never occurred.

The protocol was consistent with ISO 8 0 6 0 1 dash two dash six one, the pulse oximetry standard, which specifies that a third of the, SO two. It will be in each of the decades, 70 to 79, 80 to 89, and 90 to a hundred percent in contrast in the real world. So two distribution is uncontrolled.

Statistically the distribution of SO two, given SpO2. In other words, the box plot depends on the marginal distribution of s A two. The lower the marginal distribution of A two, the lower the distribution of S a two. Given SPO two, all else being equal. The box plot comparisons of two groups, do not consider that the marginal distribution of EO two could be lower sarcastically for one group than another group.

If the groups differ an EO two marginal distribution, then differences between groups in the box plot may be difficult to interpret. In the real world study by Meyer 2018, the SO two distribution for Fitzpatrick five and six subjects is shifted to the left with a longer left tail

compared to the SEO distributions for Fitzpatrick three four and Fitzpatrick one, two patients suggesting that the box plot comparisons between the groups, which is not shown, may be distorted because they have different seo two distributions is box plot comparisons are desired.

Then a suggestion is to standardize the. To the same s A two distribution before drawing the box plot. Two methods for doing this are described here, both based on waiting of the observations.

In many publications of real world studies, authors have focused on the diagnostic accuracy of pulse oximetry to detect hypoxemia. Consider the real world data in Vesu at all 2022. For black infants and white infants in the NICU, hypoxemia is defined as SL two less than 85%. That's the first column in each table.

The pulses defined to be test positive for hypoxemia. If SPO two is less than 90% the first row in each table, and otherwise test negative the second row, the cold hypoxemia rate is defined as the proportion of test negative subjects with hypoxemia. Denominator are shown in light blue, the cold hypoxemia rate was numerically greater in black infants than white infants, 11.8% to 10.2%.

However, pulse oximeter sensitivity to detect hypoxemia was numerically slightly greater for black infants than white infants. 39.7 to 38.2%. And pulse oximeter. Specificity to detect non hypoxemia was also numerically slightly greater for black infants than white infants, 81 to 78%. Note that the prevalence of hypoxemia was greater in black infants than white infants, 15.3 to 12 and half percent.

On the next slide, I show that hypoxemia prevalence influences the AOC cold hypoxemia rate, which could explain the difference in the latter between the skin color groups based on the VAs data. This plot shows that as hypoxemia prevalence increases, so does the hypoxemia rate.

Thus, a difference between groups and a cold hypoxemia rate cannot be interpreted without additional information on the hypoxemia prevalence in each group.

Unfortunately, in many real world studies with a focus on a cold hypoxemia, the first row of data and the contingency tables is not reported precluding the calculation of sensitivity, specificity, and hypoxemia prevalence. In summary, non-randomized comparisons of groups on any metric may be difficult to interpret without adjustment for potential confounders.

Non-randomized comparisons of groups on a cold hypoxemia rate are difficult to interpret because of confounding by hypoxemia prevail. And in general, a pulse oximeter study may be difficult to interpret when the paired measurements of SEO two and SPO two are not simultaneous. Data were excluded from analysis or not reported at all, or limitations exist in study design, conduct analysis, or reporting.

METHODS FOR ASSESSING SKIN PIGMENTATION IN PULSE OXIMETRY STUDIES — DR. PFEFER

Dr. Pfeffer: Thank you. Good morning. I'm Josh Pfeffer, a biomedical research engineer in CDRH. The focus of this presentation is the assessment of skin pigmentation in pulse oximetry studies with subjective and objective methods. Recent clinical studies have provided evidence of racial disparities in pulse oximetry, which have significant implications for public health.

While the root cause of any errors correlated with race and pulse oximetry has not been established, many consider skin pigmentation to be the most significant factor. This presentation focuses on issues that the panel will discuss today, including the acquisition, reporting and standardization of data on skin pigmentation for the evaluation of pulse oximeter, perform.

Pulse oximeters operate based on the principles of tissue optics as light enters tissue. It can be scattered by components such as cells and absorbed by constituents such as hemoglobin, melanin, and water. However, substantial amounts of light will be transmitted through the finger or reflected from other tissues due to the fact that these devices operate at red and near infrared wavelengths such as six 60 and nine 40 nanometers where light attenuation is relatively low. Overall, epidermal melanin is more highly absorbing than even blood at these wavelengths and can reduce detected reflectance at six 60 nanometers by up to 40% per pass through the epidermis. Detected signals are processed to mitigate the effect of melanin absorption, but may not be fully successful.

Anatomic variations in melanin are also significant with sites such as the fingernail and Palmer finger containing relatively low levels compared to sites such as the forehead and the skin proximal to the fingernail. This presentation provides a summary of skin pigmentation assessment methods, which we have reviewed, including subjective and objective techniques.

Starting with some of the less rigorous subjective approaches that are more commonly used in pulse oximetry studies. There is racial ethnic self-identification, skin color descriptors, the well-known Fitzpatrick approaches, as well as more standard color scales. There are more rigorous objective optical methods based on spectroscopy and colorimetry, as well as gold standard validation approaches based on skin biopsy.

To quantify melanin content, perhaps the most common approach for retrospective pulse oximetry studies is to implement self-identification of racial ethnic categories. While relatively easy and inexpensive, there are a number of disadvantages to this method. For example, it is subjective and qualitative, providing little direct pigmentation information.

Also, many people have mixed ethnicities and pigmentation levels can vary significantly within an ethnic group and over time. Researchers have also noted that conflation of race with pigmentation may produce misleading results.

The use of skin color descriptors is another common approach and one which has been implemented in many controlled pulse oximetry studies. This method uses terms such as light, medium, and dark with no further definition of the category. Unfortunately, such an approach is highly subjective and qualitative as it lacks any standardized definitions or scale, and there can be large variations in skin pigmentation within each category. In highly pigmented skin color differences are often overlooked by visual assessment. Additionally, this type of categorization is not highly repeatable or reproducible given reader variability.

One of the most well-known approaches is the Fitzpatrick Skin Prototype Scale. This method was established 35 years ago to assess the propensity of subjects with light skin to incur sunburn based on a questionnaire. This Fourier scale was later extended to six to include subjects with brown and black skin.

Over the years, many have adopted this approach to evaluate skin color. Yet no standardized color charts or definitions for each type have been established. The original methodology tends to show weak correlation with pigmentation and produces results that do not account for variations in melanin content due to tanning.

Physicians typically assign individuals of color to prototypes four through six based on ethnicity, which is unreliable. Although skin pigmentation is a continuum, the color-based FSP method uses relatively co bins and categories are not well standardized. In studies where FSP was compared directly with more objective methods, correlations were highly variable. One type of subjective approach that may be more effective involves the use of skin color scales.

These methods have been more commonly used in regulatory submissions for pulse oximeters than published studies and include the Massey Martin and Valien scales, as well as the Munsell color chart. While theoretically standardized access to well-validated color charts is a common limitation, also, as with other subjective methods, inter observer errors can be a significant problem.

Recently, there has been increasing interest in implementing objective optical approaches to evaluate skin pigmentation for pulse oximetry. These ERs based on visible to neuro infrared reflectance have been studied extensively in the scientific literature. Numerous commercial devices implementing spectroscopy or colorimetry have been marketed in the US and abroad.

These devices can provide data on specific tissue sites, such as the Palmer finger, which contains very little melanoma ERs, implement a range of acquisition and processing approaches and generate different metrics, some of which are not well standardized. One common color imagery output is the individual typology angle, or ITTA, which has been correlated to a range of skin colors.

Overall ERs have shown good inter device agreement indicating that they appear to measure the same biological constituent. The final approach I'd like to cover. Involves acquisition and analysis of a biopsy of the skin. In order to rigorously quantify epidermal melanin content. This often involves optical transmits or high performance liquid chromatography.

Since this is an invasive method, its role is to provide a gold standard reference for melanin content and the validation of a noninvasive assessment approach. The literature indicates that objective devices provide a relatively high level correlation with gold standard measurements of melanin.

Several recent publications on pulse oximetry have included statements related to the rigor of pigmentation assessment approaches. Authors of retrospective clinical studies have noted that subjective methods represent a limiting factor. They are only a surrogate for true skin pigmentation or melanin content measurements, and they have categories that are too blunt or broad.

One recent article even recommended that researchers stop the use of subjective skin tone scales altogether. To summarize the findings of our literature review. Subjective methods for evaluating skin pigmentation are more commonly used, inexpensive, and relatively easy to implement. However, they are less accurate and repeatable than objective methods.

There are a variety of subjective methods with different levels of capability in differentiating skin tone. Some of these approaches may be sufficient for evaluating large differences in pigmentation. Objective methods represent an emerging alternative with greater scientific rigor and accuracy, including strong correlation to gold standard measurements of epidermal melanin.

They may be most useful for site specific measurements, such as on the Palmer finger, but may not be necessary for all types of studies. Objective approaches are typically more expensive to implement and are currently not a standard approach. In dermatology. This afternoon, the panel will be asked the following question.

There are several tools to assess skin pigmentation, including but not limited to color imagery, spectrometry, melanin volume fraction, and skin color scales. Please provide recommendations for studies evaluating pulse oximeters for the following, a standardization of skin pigmentation assessment and B, categorization and reporting of skin pigmentation data. Thank you.

CLARIFYING QUESTIONS FROM PANEL TO FDA

Dr. Nathan: Just want to commend the FDA for the very thorough job that they did in terms of the literature reviews and the various evaluations and providing the panel with a lot of information, a lot to digest there for all of us. Just to remind the panel we are going to. At least two hours or at the most two hours to discuss all of this and try and unpack it and answer the FDA's questions later on.

But I'm going to open it up now to the panel. Four questions to our presenters. We, let's try not to get into too much of a discussion. The questions should just gear to informing our discussion, which will come on, which will come later on this afternoon. And you can do that by raising your hand. And I see Hugh, sorry, I'm referring to you by your first name because that's all I can see. Hugh, you want to go ahead and with your question, please?

Dr. Cassiere: Yeah, no problem. Just real quick, just a question about the de desaturation studies, just to be a little more specific. There's a requirement to have at least 10% of participants have a dark pigmentation. Now those 10% participants they can be put into any category during the desaturation studies. Is that correct? In other words, you could have two pigmented study subjects who are in the category that are receiving normal oxia and you really don't have any data on their desaturation, or is that an incorrect interpretation?

Dr. Eydelman: Hello, my name is Malvina Eydelman and I'm going to ask Dr. Hendrix to address your question please.

Dr. Hendrix: Sure. I'll be happy to answer your question. Thank you so much for your question. So even though I'd say there are 10 subjects that are being desaturated there are many that cannot undergo desaturation to a lower threshold. And so the lower deciles could be a few subjects

when FDA reviews the clinical study we look at through the lens of gender as well as age, as well as pigmentation to see if there's any representation of data pairs there.

It's the same number of data pairs. It just could come from a fewer subjects. So typically it's in that range that we have less individuals, but the same number of data pairs.

Dr. Cassiere: So just to follow up so you can require that they have individuals that are highly pigmented, but they may not fall into the category where they're going to desaturate less than 88%?

Dr. Hendrix: That's correct.

Dr. Cassiere: Okay.

Dr. Hendrix: And there's no specific requirement right now by the FDA to have them do that. What we do is we individually review all the individual blend al and data pairs as well as the table, and we inspect it case by case, and we do ask for more information and justification if that's the case.

Dr. Nathan: Dr. Alam, why don't you go ahead with your question please?

Dr. Alam: Thank you. My question was regarding the use of various devices for rating the pigmentation. I understand we have an opportunity to address FDA specific questions downstream, but I was just trying to understand what FDA was concerned about. When they're talking about melanomas and not using a subjective method like the Fitzpatrick rating Scale.

Are they speaking about what is optimal in the context of a research study or what would be considered something that would be routinely used for patients to assess whether or not a particular pulse oximeter was right for them or how to acquire an appropriate device?

Dr. Eydelman: Dr. Alam, this is Melvin Eydelman again. We presented this comprehensive review so that all panel members are aware of the different methodologies

currently available, some of which are utilized in our clinical studies, which are used for pre-market assessment of pulse s which fall under regulated medical devices. Now, we would like for you to be aware of all of this information when and cons utilize it in deliberation of the questions, which was specifically been opposed for you in the afternoon.

Dr. Nathan: Okay. Thank you, Dr. Loeb.

Dr. Loeb: I was particularly educated by this statistical presentation. I think that, that was excellent. It really points to in my mind, statistical flaws in many of the retrospective studies. And I was just think there were methods presented, statistical methods that will really help tear apart potential influence of other confounding factors such as skin color. I guess that's all I have to say at this moment, but I, yeah I really hope that some of that is taken into consideration in coming up with better metrics than what are used currently.

Dr. Nathan: Thank you, Dr. Connor.

Dr. Connor: Yeah, my question I think it's for Dr. Pennello. I want to make sure I understand part of his presentation. So Jean, you were, describing the way you know, and Nova and Cova is used, and I wondered if it's doing more harm than good here. And like I said, I just want to make sure I understand it, that typically we would use, a another explanatory variable like race or maybe like baseline severity of a disease when we're comparing, say, treatment and a control, because some of the variability in the outcome is explained by baseline severity or maybe, explained by race in some circumstances, things like that.

But in a case like this, by putting in race on the right side of that regression, are we actually ignoring or masking the role that race or skin color has in the outcome. Meaning did we miss we the entirety of public health, a global, We did we miss some of this because we allowed race to be on the right side of that equation and going forward, should we not put it, because then

differences that we see in race will come out that oh, there's measurement error here that we were allowing to, be accounted for somewhere else.

Dr. Pennello: Yeah. Thanks Dr. Connor for the question. I, so I'm presenting the end Cova as just an example of what you could do. The concern I guess is that if you just look at the mean bias across all a two values and the bias varies depending on the A two value that mean bias comparison that isn't adjusted for anything may be misleading or just, maybe confounded by a two, it could be confounded by other factors. I'm suggesting ACO is more possible analysis.

Dr. Conner: Okay. Yeah, and I was just trying to get at whether allowing race to be on the right side was maybe masking some of that variability that, rightfully belongs to the device if the device isn't, know, working right across that. Okay. Trying to understand.

Dr. Nathan: Thank you. Dr. Hennessy.

Dr. Hennessy: Thank you. I have a question for the engineers. It's a theoretical question. I'm wondering whether the potential error that we're seeing in people with dark ski that's dark skin pigmentation is a function of the calibration of the devices. And if it's theoretically possible for a device to both measure skin pigmentation and then use that skin pigmentation reading to calibrate the results of the SpO2.

Dr. Lee: Hi it's James Lee division director here at OHT1. Dr. Hennessy. Thank you. That is one of the ways that a good algorithm, trained algorithm to piece out the dataset and re refinement in what the information. It's basically deconvoluting. It is also an issue that we brought up in our presentation a physics issue, right? The darker pigments darker color schemes will absorb more light and disallow more absorption by the other side by that sensor. So it's a little bit of both. A refinement in the algorithm can certainly improve accuracies.

And if you look at the publications that we have when we clear 510(k)s we do declare different accuracies. But that obviously is both pulled. And as Dr. Hendricks said we do publish Uhland Altman studies as well regarding like the different subpopulations.

Dr. Eydelman: Thank you. And before we leave it, I believe Dr. Pfeffer wanted to add something.

Dr. Pfeffer: Sure. Thank you. Yeah. I think you hit on a couple of interesting points. I'm not quite sure it's really been determined whether any effective of bias due pigmentation can really be calibrated out through use of more patients with darker skin pigmentation or whether there's a need for actual change in the instrumentation or design of these devices.

But, certainly these are things that we're interested in looking into. There, there have been a variety of different mechanisms that have been proposed and potential solutions. And I should mention that there are devices like bi rubio that are known to be impacted by skin pigmentation. And they actually incorporate additional wavelengths in order to estimate the skin pigmentation level and then devolve the optical signals for that the effect of that pigmentation. So there is a precedent for that type of approach of adding wavelengths to improve the measurements and robustness to pigmentation.

Dr. Hennessy: Great. Thanks. This is fascinating.

Dr. Nathan: Thank you. Dr. O'Connor.

Dr. O'Connor: I guess I've got a comment and questions for two different people on the panel. So first and I don't know if even my questions are necessarily appropriate for this point. It's pretty clear to me that this problem will require an additional standard that would be the addition of multiple other LEDs, in fact to deal with a variety of issues.

So let me give you an example. African Americans, one in 10 of them has G six PD deficiency. And of course, in the hospital they're often treated with medications that can produce MET Hemoglobinemia with patients who have an undiagnosed G six PD deficiency. And that could explain a lot of what we're seeing in both the paired and retrospective studies in the clinical context.

And so that's one where the FDA's expectation that you actually not study people with Med Hemoglobinemia and frankly that you not study people with anemia. Both of those I think, have perhaps lent themselves towards some of the problems that we're talking about today. And so my first question actually for James Lee is several of the comments suggested that there are problems with quality.

In probe production and calibration. And that interestingly, none of the discussion thus far from the FDA and its experts has suggested that there could be a problem with probe performance. That is, say they're not really on frequency. They're not really at six 60 and nine 90.

And then my second question will be for Gene Pennello, which is and so if you have that kind of variation, how does that degrade your statistical analysis?

Dr. Lee: Hi thank you, Dr. Connor. Qu a good quality system, a good manufacturing process. Is critical factor in device safety. And as we mentioned in our presentation these are moderate risk devices which we regulate through our five 10 K program. And that being said obviously any device manufacturer is required to have a quality system, manufacturing process, QA, and any corrections needed.

They need to live up to that part of the CFR. We don't do an inspection prior to, clearing a five 10 K, but certainly inspections and looking at a company's quality system in and of, in a very robust way is part of the five 10 K program. Companies do get inspected. Their

documentations are gone over. And it is a requirement when a company perceives something and how they incorporate good QA, good control over the entire device life cycle. That's one of the reasons why a big factor in us being now TPL, where we're looking at a life cycle device from cradle to grave.

And so we're taking a holistic approach and we will certainly take comments today from the panel to heart and evaluate both pre-market and post-market activities. I yield to Dr. Pennello.

Dr. Pennello: Yes, thanks for the question. I guess statistically, if the probe is not on frequency, that might introduce bias, it might introduce measurement and precision in the pulse oximeter. And I guess I'm going to ask the regulators one last question, and that is, it seems to me that the best way to evaluate devices would be in use. That is to say, rather than a small number of volunteers who are selected to be healthy and without problems, we should actually go and perhaps do the equivalent of a phase three trial where we study them in patients at the bedside with, very deliberately timed simultaneous measurement of a saturation and a blood gas.

Dr. Nathan: I was actually going to make the same point. It's just doing healthy volunteers as lack of phase one study, but there's no phase two or phase three study in terms of seeing how these devices perform in patients who really need them, where we make these critical decisions. I think in the interest of time, we're going to cut it off after Dr. Hennessy's question, but Dr. Collop, why don't you go ahead with your question.

Dr. Collop: Okay. Thank you that presentations were awesome. I learned a lot. As I mentioned, I'm a sleep specialist and in sleep medicine the pulse oximeter is like the core signal for diagnosing most sleep disorder, breathing disorders. And I guess my question is because we rely so heavily on it, both now with home sleep testing and even in lab, is there any evaluation or

analysis of the ability of pulse s to adjust rapidly with the trends? That can go up and down quite often during the night. Because I didn't hear anything about that in the evaluation of pulse oximeters. Thank you.

Dr. Lee: Hi Dr. Collop. I could take that. Certainly as we acknowledge, the saturation signal in a sleep study is critical, as you mentioned pulse s generally are very responsive. And the way we use pulse s in home sleep studies or sleep studies in general is that we measure it epoch by epoch. So that desaturation event coupled with a disordered breathing event and measuring that, and as you might know the things that we look at when evaluating sleep study devices either PSG or home sleep studies, is that a trained physician is designed to be the adjudicator and determines the accuracy. So you measure epoch by epoch, event by event, and it should be hand scored by a physician. Any comments today, certainly we will share with all the folks that use pulse oxs not just in anesthesia. Thank you.

Dr. Eydelman: And I believe Dr. Weininger is going to add something. Go ahead, Sandy Weininger.

Dr. Weininger: And just a brief comment, the clinicians can certainly set the sleuthing rate in most medical grade oxs, perhaps not in the finger clips, but in the medical grade. You can set to respond as rapidly or as slowly as you want to smooth out those dips depending on what your clinical application is. Thank you.

Dr. Nathan: All right. I'll put myself in line, so I'll go next. I have a question and probably a comment. I saw right at the beginning there was a mention of reprocessed oximeter, which made me think that oxs are really like any other technology which tend to degrade over time. So is there a half-life on the accuracy of these pulse oxs? Because I'm not sure that they're

replaced as often or checked as often as perhaps they should be. I'll let the FDA address that first, and then I might have one other comment after that.

Dr. Lee: Dr. Lee. Hi. Again thank you, Dr. Nathan. Certainly a labeled shelf live is something that we would review as a part of the 510K process when it comes in. Validated cleaning repurposing instructions would have to show that the device does not deteriorate over time. And so a label life of the device is something that we would use in our review.

Dr. Nathan: Okay. Thank you. And just a comment for Dr. Pfeffer with regards to skin pigmentation and evaluation of that. We typically put the pulse oxs on the fingertip and, or sometimes on the, for if we don't get a good signal on the fingertip. What's the general correlation in terms of pigmentation on the distal aspect of the finger versus other parts of the body? And in terms of evaluating? Pigmentation is the global pigmentation taken into account or specific body areas?

Dr. Pfeffer: I think that's an open question about what's the best way to assess pigmentation. I think there's more of an interest recently in trying to go after the site specific pigmentation level. And I think we're going to start seeing more studies that do that. And I think that's one of the benefits, certainly of the objective approaches is that some of these devices are pretty good at providing that, that information because as you refer to, I think the pigmentation level on the finger is typically much lower than on other locations like the forehead or ear. One would expect that there would be more impact from pigmentation on, on those sites as opposed to the finger.

Dr. Eydelman: Dr. Nathan, we would appreciate greatly if the panel addresses this during their deliberation and lets us know if they have any recommendations on this specific topic as we have pondered it extensively.

Dr. Nathan: Okay, sounds good. We'll do that. I know we are running short of time. We're supposed to break at 10 and there'll be opportunity for other questions. So perhaps we'll just do one or two questions and then we'll break and there'll be opportunity probably in the afternoon to address any further concerns. Dr. Connor, why don't you go ahead.

Dr. Connor: Yeah, Jason Connor. Yeah, a bit of a follow up to Dr. Hennessy. Dr. Pfefer said that bi Rubin devices can measure skin coloration. I know there's also hair laser removal ones that measure it because some of those are actually not made to work for dark skin because they look at skin versus hair and light and dark. And so that in darker skin it can actually do damage. So the device itself says this is been too dark, that I'm made to work on. So devices are certainly capable of that. My question to FDA is, are any of the pulse s currently on the market? Do they attempt to measure in pigmentation and account for that in their algorithm? We saw a lot of general stuff, but not about like specific devices that are approved. So I'm wondering if any actually try to do that these days.

Dr. Eydelman: Hi. It's, thank you, Dr. Eydelman. Not to my knowledge, but I think we'll probably have to follow up if I'm wrong later this afternoon. But generally speaking, their general use, we don't label them for particular subgroup. They're considered general use tools. Yeah, no, and it wasn't that I was implying it would be like one for light, medium or dark skin, but rather would the device itself attempt to say, Oh, this is someone with lighter skin, This is someone with darker skin, and maybe the internal algorithm was, slightly different.

There's obviously computation going on inside the device, but yeah, I what you're suggesting might be the simplest solution. But I was wondering if any actually measured it and accounted for it internally. Not that there would be different ones, but So not to the best of our knowledge. We'll, however, we verify and get back to you after lunch.

Dr. Connor: Okay, perfect.

Dr. Pfeffer: The, if I could one brief comment to, to follow up on you mentioned that there are essentially ERs that are used with laser therapy devices. And actually I think those are the only ones that have gone through the FDA clearance process. Most of the melanomas on the on the market have not, they're not actual medical devices, but the ones that you mentioned I believe have.

Dr. Nathan: Okay. Great. All right, we'll take two more questions because all the questions are so good. And the answers are shedding light on what we're going to have to ponder on. Dr. Hennessy, your question, and the last one will be Dr. Bickler then.

Dr. Hennessy: Great. Thank you. Sean Hennessy, I'm wondering about standardized preclinical testing. So are there any simulation platforms that allow the prediction of the performance under different conditions including different skin pigmentation patterns or is there any standard preclinical testing data that's submitted at all?

Dr. Lee: If I understood your question, Dr. Hennessy, you are asking basically on the preclinical side, on the bench, how do they validate it? Phantoms certainly could be used as a part of product development. But, we don't recommend them. We want the subjects to be studied in the clinical realm. Certainly we use that for verifying the accuracy of the device. But certainly, companies, many different groups could as a first cut do different types of validation, particularly regarding sensitivity to spectra and light.

Dr. Nathan: Dr. Bickler, you'll be the last question of this session.

Dr. Bickler: Okay, thank you Phil Bickler from UCSF. In 2005 we called for a simple switch or provision on oximeters to correct for skin pigment. In subsequent years, we've realized that is a very simplistic view of how to solve the problem. The issue of skin perfusion, the

pulsatile signal is, we have actually found to be crucial in in the in amplifying the bias caused by dark skin. I'll talk about that later, but I hope people don't think that fixing the pigment problem is as simple as selecting a button that indicates dark, medium, or light skin pigmentation.

Dr. Nathan: Okay, thank you Dr. Bickler for that comment. And yeah, there'll be opportunity to address all of this later on. Great session, great questions, very engaging. We're going to take a five minute break now. I have it as 10 after 10, and we'll reconvene at 10:15. We're running a little bit behind, but no worries. There will be opportunity to play catch up a little bit later on. So we'll get together again at 10:15. Thank you everyone.

OPEN PUBLIC HEARING

Dr. Nathan: Good morning again, folks. It is now just after 10 15 and I would like to resume this panel meeting. We'll now proceed with the open public hearing portion of the meeting. Public attendees are given an opportunity to address the panel, excuse me, to present data, information or views relevant to the meeting agenda. Dr. Awojope will read the open public hearing disclosure process statement first.

Dr. Awajope: Once again, this is Dr. Awojope, the designated federal Officer, DFO, for this advisory committee and a sociology and respiratory therapy devices. Our now read the conflict of interest statement regarding the open public care and OPH, both the FDA, the food and the drug demonstration, and the public belief in the transparency process for information gathering and decision making.

To ensure transparency at the open public hearing section of the advisory Committee meeting, FDA believes that it is important to understand the context of individual presentation. For this reason, FDA encourages you the open public hearing speakers at the beginning of your

reading or oral statement to advise the committee of any financial relationship that you may have with any of the company or group that may be affected by the topic at this meeting.

For example, this financial information may include an interest in or group's payment of your travel, lodging, or other expenses in connection with your attendance at the meeting. Likewise, every encourages you at the beginning of your statement to advise the committee. If you do not have any such financial relationship. If you choose not to address this issue of the financial relationship at the beginning of the statement, it'll know preclude you from speaking. Thank you very much. I'll now hand it over. Back to you, Dr. Nathan.

Dr. Nathan: Thank you once again. The FDA has received nine requests. Each speaker will be given five minutes to speak. And for the live speakers, just do keep an eye on the watch, please. We will hear six pre-recorded folks. First, in terms of the open public hearing presentations, and this will be followed by three live presentations. The first speaker is Veronica Hickson from the Urine Bevin University Health Board. Go ahead. Thank you.

DR. VERONICA HICKSON

Dr. Veronica Hickson: I'm Dr. Veronica Hickson, a practicing pediatrician in South Wales and Director of the Electrode Company. I want to talk to you about a clinician's understanding. Many clinicians believe the CIS has the same accuracy level throughout the range of 70 to a hundred. This isn't the case due to magnification of errors.

At lower SPO2 levels, the accuracy is less and it's much better at higher levels in practice. These values are also impacted by the specification of the wavelengths in the L e d in comparison to those used in the clinical trials. High reading bias as seen here can also be the case in high melanin.

Clinicians want safe and effective medical devices that they can rely on. They want clear information on factors that impact SBO two accuracy. They want to know the accuracy for the equipment they're using at high and lower levels. The pulse oximeter is used as a vital sign and they want their equipment to be accurate.

For this, they want to be able to use pulse oximetry as recommended and advised in clinical, but to increase morbidity or mortality. It's often thought that pulse oximetry is a general guide in practice. Clinically, it's used with guidelines and protocols to dictate certain actions, and if they're not taken, then it could contribute to an adverse patient outcome.

Also, the equipment should be complying with regulations and if inaccurate could also contribute to adverse patient outcomes. Pulse oximetry is used in many scenarios from congenital screening in neonates through the NICU, and up to and including covid assessments and triaging patients. Vulnerable patients exist in various patient groups, including neonates, elderly people, veterans.

They all require SPO2 measurements. Many professional groups from paramedics such as ambulance staff and EMTs. The nurses in the ER and on the wards and the medics and doctors and family physicians are using the results of pulse oxs to guide clinical decisions in some situations. Yes, such as operating theater.

The patient is monitored by a variety of different technologies and Pulse oximetry, one of the contributing information points. Users don't want confusing and misleading information. They want sensors that are fit for purpose and enable the pulse oximeter system to continue to work as intended and claimed. They want it recognized that pulse oxs are capable of being and are used as diagnostic tools and inaccuracy can be due to poorly compliant equipment, and not

that the patients are faulty. They don't want to be told it's statistically acceptable to be inaccurate. They want to maintain the accuracy of the system and not have additional bias introduced.

The actual accuracy levels information could be more clearly provided by using discreet ranges. For example, 99 to 90% SPO2 have less than 1.5 90 to 80% less than 2.5 and 80 to 70%, perhaps less than or equal to 3.5 Information is available in manufacturers previous brain down studies and could be used to inform new standard drafting against which they could then prove their current and continued compliance.

The accuracy at extremes of their LED wavelength specification should be assessed and the data included in the files. Good manufacturing practice and quality management can therefore ensure compliance with the accuracy requirements that must be maintained in practice, in production, and not just for the initial regulatory data gathering and accuracy assessment.

Additional biases introduced by a sensor of perhaps more than 3% at any SBO two level should be considered as non-compliance, and therefore not in on the market or in use. I'm happy to take questions.

SAM AJIZIAN

Sam Ajizian: Hi everyone. I'm Sam Ajizian. I'm the Chief Medical Officer for the patient monitoring business at Medtronic.

My medical practice was pediatric critical care for over 20 years, and I've used pulse oximetry in the care of my patients throughout my career, and I'm well familiar with both its strengths and limitations. I want to thank our colleagues at the FDA for inviting us to share our perspectives on how to prioritize and improve health equity and pulse oximetry.

This is the commitment you'll hear in my remarks today. Clinicians like me have relied on Medtronic’s Nucor pulse oximeters for over 40 years to obtain non-invasive estimates of

blood oxygen levels to guide critical interventions. Medtronic stands behind the quality, safety, and performance of our devices, but we acknowledge that more can be done to enhance accuracy and to foster health equity.

Specifically, we stand with the FDA, academics, clinicians, and patient advocates in an effort to further improve standards for pulse oximetry devices and reduce health disparity. When questions emerged about pulse oximetry accuracy, we went back and looked at our own data. We did a pooled analysis from VER previous validation trials with 88 patients and approximately 18,000 patient samples.

We presented these data at the recent evidence-based Perioperative Medicine World Congress in July of this year. Nucor Pulse oximeters perform well overall reporting values within 2% of directly measured oxygen levels. This is consistent with FDA requirements. When you dig more deeply, though, we see accuracy with slightly higher for patients with lighter skin pigmentation. While the overall impact on measured blood oxygen levels was not clinically significant, we see room for improvement including further enhancing accuracy for patients with darker skin pigmentation.

Pulse oximeter accuracy can be affected by melanin, a skin pigment that is present in higher levels in those with darker skin. Melanin concentrations like skin pigment vary widely within racial and ethnic groups. Melanin absorbs certain types of light and can affect the accuracy of pulse Oximeters.

Currently, manufacturers rely on a variety of subjective scales to assess the impact of skin pigmentation on device performance. In order to improve device accuracy for everyone, we must replace these subjective scales with a uniform objective method of assessing skin pigmentation. Now we do see a path forward.

Medtronic is evaluating and sharing data, working with FDA and ISO, and making investments in our own devices and algorithms. At the same time, we believe health equity requires true collaboration across all stakeholders. Specifically, we need to work together to improve information sharing across all key audiences to foster continued innovation in pulse oximetry, and to improve the methods we use to validate pulse oximeters include including standardization of how we assess skin pigmentation and increased diversity in skin pigments and clinical trials.

We also need to work together to educate clinicians and patients around device capabilities, use patterns and limitations, and we need to collaborate to develop reasonable regulatory standards that allow for innovation while at the same time promotes safety efficacy, and of course the health equity of medical device.

As I consider our role in improving health equity generally, I think about my own experience caring for children in the pediatric ICU. Health equity was a priority for myself and all of our team members, and I see our mission at Medtronic to be fully aligned with my clinical experience. We know how our technologies affect patients and their physicians, and as such, the greatest commitment we can make is to consistently design and improve devices that positively influence health outcomes for all patients.

Health equity is absolutely foundational to our mission at Medtronic and our collective successes. And as a company, we're committed to that goal. We thank you for this opportunity to share our data and our perspectives on this issue. Hello and thank you all for the opportunity to speak on this important topic today.

DAVID STRICKEN

David Stricken: My name is David Stricken. I'm a hospital medicine physician at the University of Wisconsin, and I co-wrote a letter that the hospital medicine re-engineering network sent to the FDA on this issue. I'm here today to briefly describe the recommendations in that letter. Number one requires subgroup analyses by objective measures of skin, pigmentation, race, and ethnicity, and gender in the approval process.

Currently, the FDA only requires at least two dark pigmented subjects or 15% of your subject pool, whichever is larger. This means that a device that is accurate in the 85% of subjects with light skin pigmentation could still be approved, even if it is very inaccurate for the 15% of subjects with dark skin pigmentation.

To ensure accuracy for all groups, we believe subgroup analyses should be performed and that subgroups should be large enough to reach statistical and clinical significance for each group. This could be done in conjunction with a recommendation from the pulse oximeter manufacturer Masimo, to include equal distributions of subjects with light and dark skin pigmentation.

Since the relative contributions to these device inaccuracies from objective skin pigmentation versus from the social constructed race are not yet clear as further explaining the letter. At this time, we believe subgroup analyses should be performed both on the basis of skin pigmentation and separately on the basis of race, as well as on gender, given some evidence for gender bias as well.

Number two, prohibit the use of race correction factors. Historically, when different subgroups of people have been found to have different results, it had been common practice to use so-called race correction factors in clinical calculations to supposedly correct for these

differences. For many reasons, however, race correction factors are now considered to be very problematic and are being removed from some calculations such as estimated alum filtration rate.

Instead of using race correction factors to attempt to account for pulse oxs inaccuracies, the devices themselves need to change. Number three, require pre-market data from both healthy and hospitalized patients for device approvals. The recent studies showing racial bias were in hospitalized patients.

Although some studies have also shown racial bias in laboratory conditions, one manufacturer Masimo has recently claimed that their devices do not show racial bias in laboratory conditions. Notably, this claim has not been evaluated by the FDA and it should be. And also Masimo did not perform similar analyses in hospitalized patients despite their FDA submission packet indicating that they do have such data from hospitalized patients.

Nonetheless, it is possible that racial bias is most prevalent in hospitalized patients. Therefore, pre-market testing should be required in hospitalized patients as well, because testing only in healthy subjects might miss some biases. Number four, intensify the overall approval requirements for medical grade or prescription pulse oximeters and consider requirements for over the counter devices.

Currently, the FDA only requires data from 10 subjects for approval of medical grade pulse oximeters. This sample size seems concerningly small for devices that will be used to make life or death decisions for literally millions of patients, and we believe it should be increased accuracy requirements could be intensified as well as Masimo described in their comments.

Over the counter devices, however, undergo no FDA review at all because they're supposedly not intended for medical purposes. It is quite clear, however, that many, if not most

users of over-the-counter pulse oximeters are in fact using them for medical purposes. So we believe some type of FDA review is now needed for these devices in order to protect consumers.

We do, however, believe that the requirements for these devices should be kept separate from those from medical grade devices so that the ladder can have more stringent require. Number five, Have these new requirements apply retroactively to previously approved devices without applying these requirements to previously approved devices.

Manufacturers of racially biased pulse s might choose to continue selling their current devices, which have been grandfathered in rather than try to get new devices approved, which would only serve to perpetuate this racial bias. Number six, investigate other similar devices for the presence of racial bias.

Other optical devices such as temporal thermometers have recently also been shown to demonstrate racial bias likely due to the effects of skin pigmentation on their measurements. We believe the FDA should proactively investigate optical devices for racial biases rather than waiting until independent research has found the bias for them, as has been the case with pulse oximeters.

Lastly, although we recognize that the following two recommendations may not fall within the purview of the FDA specifically, we would like to encourage the FDA to advocate for them given the collaboration that often occurs between agencies. Number seven, create additional safeguards to detect and prevent biases in our other diagnostic tools at the FDA level.

This could include a dedicated committee that reviews all future devices for potential biases prior to their approval. Number eight, invest in independent research. The exact causes of the racial biases seen in these devices are not known. Skin pigmentation almost certainly

contributes, but there may be other contributors as well, such as the health effects of systemic racism on individuals belonging to marginalized groups.

Further research is needed to determine the mechanisms behind these biases and define potential solutions for them. In particular, prospective studies in hospitalized patients are desperately. But would require significant funding. And it will be important for these studies to be independent of industry bias.

So government funding is preferable. The hospital medicine re-engineering network would be happy to help coordinate such studies among its more than 50 participating academic institutions. And with that, I would like to thank the FDA for its time and I'll be happy to answer any questions that you may have.

GRACE BERSON

Grace Berson: Hello everyone. My name is Grace Berson, and I am a science policy fellow with the Federation of American Scientists. I'm delivering this statement on behalf of the Federation of American Scientists and our partners at the University of Maryland Medical System. Thank you so much for giving me a time to speak to the committee today.

As we all know, pulse oximeter has provided rapid non-invasive estimate of the body's ability to deliver oxygen to the tissues and are ubiquitous in medicine. This technology is a fantastic innovation and has eliminated the use of more painful measurements of blood oxygenation like the arterial blood gas.

This technology relies on the absorption of light through the skin and as an optical technology. Anything that changes the amount of light received by the detector impacts this measurement. There's finger fingernail, polish and tattoos, bi anemia, carbon monoxide exposure, anemia, hypothermia, and the concentration of melanin within the skin of the patient.

Collectively, current pulse oximeter models overestimate the oxygen levels in persons of color by an approximately one to three. This is because many models do not adjust for how high quantities of melanin lead to the absorption of a more light within the skin before that light reaches the detector.

Now, this bias may be perceived as small in magnitude, but within a statewide medical system like that of our partners at the University of Maryland Medical System, these become large numbers of people. We estimate that about 1,012 African Americans with apparently normal pulse oximetry readings had hidden, or a cult, a hypoxia when they were discharged from the University of American, Maryland Medical System Emergency Departments.

Over the last year. While this issue was identified as early as the 1990s, the problem of hypoxia in darker skin patients has come to the forefront with the COVID 19 pandemic. Oxygen saturation measurements were an important part. The management of patients Covid 19 and a cult hypoxemia was associated with delays in care and poor outcomes.

In patients with COVID 19 recent retrospective studies have shown that the greater rates of undetected hypoxia are negatively impacted the care of African American patients. With Covid 19 leading to those patients being less likely to receive supplemental oxygen on time and being at greater risk of death, the University of American, Maryland Medical System is doing what?

It's what it can to address this problem. We are preparing our own investigation. The harms associated with a cold hypoxia and the patients treated in our medical system and working to correlate patient outcomes with pulse oximeter models contracted at the time of measurement. University of Maryland medical system is additionally reconsidering how clinicians are trained to consider and apply pulse oximetry readings.

These interventions can provide care in the short term, but they do not solve the underlying problems encoded in the technology. We need improved standards for pulse oximeter devices so that clinicians can rely on the technology to provide accurate results across all populations and in the real world clinical setting.

Thus, we feel that action by the Food and Drug Administration is needed to address this pervasive technological bias. Media and long-term solutions should include requiring the publication of testing data for current pulse oximeters so that clinicians can make best judgment on which technologies are reliable for use in their patient populations.

Reconsidering which models of pulse oximeters can be used as predicates for five 10 K approval pathways and only allowing devices that have been tested on a more diverse range of skin tones. Furthermore, requiring future models be tested in significant proportions of individuals of all skin tone types.

In using objective metrics of skin pigmentation that reflect the broad range of pigmentation, such as skin tone scales, like the Massey scale and the skin tone scale, as well as color metric devices. Additionally, the FDA should consider device manufacturers ensure their devices work in critically ill patient populations where there are not stable SPO two readings embodying the real world care environment or pulse oximetry would be used.

We also think that there should be set a standard for greater accuracy and over the counter in wearable pulse TER devices, given the rapid growth of this market during the COVID 19 pandemic. And finally, in collaborating with funding agencies like the NIH and NSF, the FDA should seek to establish and fund research into low bias and bias-free tools and consider the use of accelerated approval pathways for technologies that have been proven to work effectively on diverse populations.

With these, we believe we can make tremendous progress towards bias-free devices, and I thank you so much for my time.

JEFF MATTHEWS

Jeff Matthews:I'm Jeff Matthews, Director of the Electrode Company. We specialize in pulse oximetry. Back in the 1980s, we invented adaptive noise cancellation in pulse oximetry. We sold this patent to Mamo during the 1990s. We reverse engineered most of the OEM systems out there. And then went on to focus on sensor accuracy.

Not all replacement sensors are as accurate as the manufacturer's claim. There are often mismatches between wavelengths and the calibration data. Lindon can cause high bias in SATs regions. This is because the calibration data for most systems is appropriate for light skin and not for dark skin. Now of these issues will be fixed until it's a better control of breath than June.

In June, the 1990s, we hardly have Sonia accurate sensor. Then CE marking came in. It was as if this was a license sell. Third party sensors, regardless of requirements, they seem to work okay, so customers bought them unaware of any accuracy issues. Orem said, dropped the prices. Quality control on ED selection, suffer sensor accuracy dropped that were even third party sensors from the far east on the US market with without 510(k) clearance.

The number of inaccurate sensors on the market is increasing. There are new OEM sensors that claims be rearward compatible in all systems, but they're not as accurate as claimed in all the systems. There is no OEM that claims that every sensor uniquely calibrated. This is false. Sensor accuracy is not as good as it used to be.

Bad things have happened when manufacturers have been confronted with inaccurate senses. Explanations have been contradictory. Investigations reported to customers do not make sense. Fundamental questions have been dodged.

We need to know what the target wave S sensor are and we need to know what they really are. This tells us her act A sensor is, manufacturers should already have this information.

It makes no sense to try and fix racial bias. Any adjustments will be swamped by manufacturing errors. Regulators have been led to believe that the only way you check in sensor accuracy is by breath. This is false. Sensor accuracy can be checked with a spectrometer a. We should not let commercial interests come before patient safety.

We need to wind the clock back a bit and insist that sensors are at least as accurate as they used to be. Regulatory changes must be retrospective.

DR. STEVEN BARKER

Dr. Barker: Good morning. We at Masimo, been following with great interests, the recent reports on skin pigment related differences in pulse oximeter accuracy, and we commend the FDA for convening this panel on the subject.

As a background, Masimo developed the very first read through motion and low perfusion pulse oximetry using signal extraction technology, or set more than 30 years ago to addressed known pulse oximeter errors. Such as motion and low perfusion set also addresses skin pigmentation and other non-pulsating absorbers in the same way, we applaud the authors of the recently published studies for investigating this important topic of skin pigment.

But we have the following concerns with the methodologies of some of these studies. Number one, most of the studies do not characterize the manufacturer or model of the pulse oximeter. In fact, many of them use several different pulse oximeters in the same data pool. Number two, the A O2 or blood gas analysis COTER data and the SPO2 pulse oximeter values in these studies are not actually simultaneous.

Instead, the readings were often taken more than 10 minutes apart from one another. Number three, dishemoglobins such as carboxyhemoglobin and met-hemoglobin were not measured or accounted for. Number four, artifacts such as low perfusion and patient motion are not reported. And finally, number five, severity of illness and other hemoglobinopathies such as sickle cell anemia are not reported or accounted for.

Our data demonstrate that Masimo set pulse oximeters are equally accurate for individual. Of all skin tones in our volunteer human subject laboratory. An abstract of these data was presented at the Society for Technology and Anesthesiology, known as STA annual meeting in January of this year. And the full manuscript is currently in press and will be, will appear in the Journal of Clinical Monitoring and Computing.

And I expect it to be published online literally any day now. Details of these published results were also submitted with our written comments for this meeting. As I said earlier, we began to address confounders such as low perfusion, motion artifact and skin pigmentation. Decades ago when we first invented Masimo Set Technology, we have also calibrated and validated our oximeters using almost equal numbers of dark-skinned and light-skinned individual volunteers.

We support prospective clinical studies patient studies on this topic, and we are pursuing these now, but bear in mind, controlled desaturation studies with SPO2 values going down to 70% can be ethically done only in a laboratory using healthy volunteers, certainly not in patients In recent years, Masimo has continued to improve our pulse oximeter accuracy and our latest performance.

Improvement coming in 2018 demonstrated an overall uncertainty of plus or minus 1.5% for our RD set sensors compared with the usual industry values reported as either two or even

3%. Uncertainty in our recommendations to the FDA on the skin pigmentation issue, Masimo supports raising the standard on requirements for the percentage of dark skin subjects used in calibration and validation studies.

We also believe it is important that the FDA regulates and applies similar oversight recommendations on all pulse oximeters, including those sold directly to consumers. The full summary of Masimo's recommendations to this FDA panel have been submitted as comments in the docket, and I thank you very much for your attention.

Dr. Nathan: Thanks to our folks or our speakers. That's it for the pre recordings, we now have three live presentations. As a reminder to our presenters, it's five minutes each as a maximum. And let's start off with Dr. Calender please.

DR. EVE CALENDER

Dr. Eve Calender: Good morning. Thank you for the opportunity to speak today. On behalf of the National Center for Health Research, I am Dr. Eve Calender, an OBGYN with a Master's in Public Health, and I'm a senior fellow at a non-profit think tank. Our center conducts, analyzes and scrutinizes research on a range of health issues.

We do not accept any funding from companies that we, that make products that are the subject of our work, so we have no conflicts of interest. Thank you for convening this important meeting to discuss a critical issue of how skin pigmentation impacts the accuracy of pulse oximetry. The pulse oximeter is ubiquitous in medical care in the United States, like blood pressure and respiratory rate.

Pulse Ox is a vital sign used to make critical medical decisions. The accuracy of this device on all patients is a paramount importance, although skin pigmentation was thoughtfully

considered when this technology was first invented, pulse oximeters in use today are dramatically different from the original gold standard device. In early studies, researchers considered many patient specific variables including skin pigmentation that would contribute to pulse oximeter measurements. To accommodate these differences. The original device was standardized for each patient and provided accurate measurements for every individual.

Furthermore, early pulse oxs use eight wavelengths of light, which was considered enough to account for individual patient variations, but today's devices use only two. There is now strong evidence that commonly used pulse oximeters are less reliable for patients with pigmented skin. The inaccuracies related to skin pigmentation vary based on the device used.

While one device may be inaccurate only for pulse measurements under 90%, others may have significant inaccuracies that all ranges. In general, inaccuracies related to skin pigmentation increase as the level of oxygenation decreases. Clinically, this means sicker patients are less likely to get an accurate reading and are therefore less likely to get appropriate care.

The FDA has designed a study to evaluate pulse oximeter performance in hospitalized adult patients of different skin pigmentation levels. This is an important step, but we urge the FDA to do more. Currently, the agency only requires pre-market testing to include at least two dark pigmented individuals or 15% of the subject pool.

This is woefully inadequate. We ask you to urge the FDA to require that manufacturers include individuals with a broad range of skin pigmentations in their studies. Moreover, subjective measurements of skin pigmentation are of little value because there is such wide variation in Hughes and so many contributing factors.

Therefore, only objective tools for assessment of skin pigmentation should be used in studies of how it affects pulse oximetry measurements. The FDA's current guidelines do not

require product labeling to address the impact of skin pigmentation on pulse. A measurements. While labeling is important to alert users to possible inaccuracies, it does not diminish the importance of developing devices that are effective for everyone.

Medical devices must work effectively in all people regardless of skin pigmentation. Products are allowed on the market, should not be accurate for only some races and ethnicities. The FDA should require more scrutiny to minimize bias in medical devices, so they're accurate and reliable for everyone.

Thank you.

Dr. Nathan: Thank you very much for your comments, Dr. Calender. Let's move on to Dr. Abrams, please.

DR. MICHAEL ABRAMS

Dr. Michael Abrams: Good morning. Hi everyone. I'm Dr. Michael Abrams, senior health researcher from Public Citizen's Health Research Group. I have no financial conflicts of interest on this matter today. Public Citizen's Health Research Group has a 50 year history of monitoring activities of the FDA, especially pertaining to the safety and effectiveness of medications in Class II and III medical devices. We believe this advisory committee topic, pulse oximeter accuracy given natural human variations in skin pigmentation levels, is one that exemplifies two ongoing challenges the FDA faces as the principal gatekeeper regarding the safety and effectiveness of medical devices.

The first of these challenges is of course, the issue of racial equity. The second And one that I will emphasize in my remarks are deficiencies of the five 10 K pathway for clearing medical devices concerning racial equity. We want to just briefly remind the committee and the

FDA that lack of diversity in FDA regulated clinical trials has been a longstanding problem that can lead to inferior outcomes for patients of color.

The fact that, for example, as we've heard persons of color have not received proper care related to covid 19 infections because of inaccurate pulse, oximeters is very simply unacceptable. Regarding the five 10 K process, we have long been concerned that evidentiary standards for safety and effectiveness are too lacks for many medical devices, thus giving manufacturers a reckless fast track pathway to marketing.

In the case of pulse oximeters, we know that since the 1980s, hundreds of such devices or their components have been con, have been cleared for marketing. In its briefing materials, the FDA reviewed a sub sample from a total of four hundred and twenty five ten K applications and found that clinical data was not required for approximately 35% of such submissions between the years of 2020.

FDA further. That the measurement of skin pigmentation level was methodologically limited, as we've heard and variable across such device applications since at least 1991, researchers have observed that increasing levels of skin pigmentation correlate with overestimates by pulse ERs of arterial oxygen saturation.

The FDA's briefing packet for this meeting tabulates at least 15 such studies just since in two, in 22, in fact, with thousands of subjects including research that suggests treatments are withheld because of this apparent medical device error. Accordingly, we agree with the FDA that more research is needed on this issue and that better labeling may be useful, but we would go decidedly further than that.

We believe that there are now there is now well supported urgency that the FDA should do the following, at least review all existing pulses for evidence that they accurately measure

blood oxygen saturation levels in persons with increasing levels of pigmentation in their skin, absent evidence, or with evidence of deficiencies.

In that regard, the manufacturers should be required to recall such device. Until they these problems can be corrected. And we know here that we are very concerned that the FDA's briefing packet included a little about the history of pulse activity, recalls in particular, and safety communications as well.

Therefore, we suggest strongly the serious omission is inexcusable from the briefing packet and must be remedied immediately. And finally, the FDA needs to revise the 510(k) pathway standards to ensure that devices will only be marketed after they demonstrate reasonable safety and effectiveness in populations, at least as diverse as the US population and as necessary in populations enriched to make sure relevant minority subgroup issues are not neglected.

We hope the committee will urge the FDA towards these much needed and long overdue regulatory reforms. Thank you very much.

Dr. Nathan: Thank you very much, Dr. Abrams. Our last speaker for this session will be Renee Kohi.

RENEE KOHI

Renee Kohi: Thank you very much. Appreciate the opportunity to speak very briefly. I'm only going to take up a couple of minutes of your time. My name is Renee Kohi. I'm the Vice President for Digital Health, the Consumer Technology Association, a trade association made up of 1500 member companies representing all facets of the tech industry.

I have no financial conflicts of interest with regard to this issue. One of the things that we've been doing at CTA is really focusing on health equity. We were co-founders of the Health

Equity and Access Leadership Coalition of the HEEL Coalition, along with 35 other stakeholders in this space.

And for us here at CTA, there are two general issues we need, we believe, need to be focused on. One is the technological and scientific attempts to rectify the existing problem, which we've been discussing this morning. But the other, which may be more complicated is how do we help to prevent these types of issues from occurring in the future?

So on the technological side of things, I know there are many researchers, scientists, and technologists working on the issue and starting to develop solutions to address the issue. I know for example, the professor Kuson up at Tufts, he's working on solution. Using the same light of available for current pulse oximeters, but also includes technology to measure person's skin tone.

Just today I read an announcement from the company by that has developed a patented alternative to traditional pulse oximetry that accurately measures blood oxygen levels across a full range of very dark skin pigmentation as well as your movement and activity. So there are many people doing great work on this issue, but we may not be fully aware of all that's being done.

So one of the things that we would be interested in at CTA is helping develop a repository of all the work that's being done across the globe, so there's more knowledge and coordination among stakeholders who build these solutions and you use them in clinical situations. I think the larger problem I addressed in my intro and perhaps beyond the remit of this particular panel, it's beyond pulse oximetry.

Because we believe that stakeholders across the ecosystem need to more intentionally think about equity as they develop health technology and medical device solutions. And there are many factors here, right? Inclusive design, by which I mean design that considers the needs and

capabilities of the whole population to help decrease misalignment between the user and the design object, right?

We understand that every design decision has the potential to include or exclude patients, but the concept of inclusive design emphasizes making informed decisions by better understanding user diversity. We think we need better recruitment of diverse populations into the testing of medical devices and health technology.

I think the FDA is doing good work there. We need research teams, quite frankly, engineer scientists and others that are more diverse because I think the diverse stakeholders bring diverse perspectives to the table. And I think the other thing is we need to be developing more quality and safety peer networks for technologists, researchers, clinicians, and patients can exchange information about the potential bias that may impact patient safety.

The question for the agency is how to better incorporate all these factors in its decision making, in ways in which stakeholders can help. And just to finish up, other governments around the world are recognizing this. For example, the UK go. Earlier this year initiated an independent review on equity in all medical devices, seeking to identify and address bias.

The reviews, looking at the device's entire life cycle from analyzing the evaluation process all the way to marketing and implementation. That's going to lead to a set of recommendations. Now that may be beyond the remit of this particular panel, but I think the agency may want to look at those kinds of reviews to ensure that we don't have the situation with other categories of medical devices.

I thank you for your time.

CLARIFYING QUESTIONS FROM PANEL TO OPEN PUBLIC HEARING SPEAKERS

r. Nathan: Thank you very much and thanks for the speakers for respecting the time limits, you're all well within it. We now have 50 minutes to ask questions to the speakers, and as a reminder, you should be or can be asking questions, the speakers can't ask each other any questions. So let's open it up to the panelists.

And see if there are any questions that anyone might have of any of our current slate of speakers. I do not see any hands as yet. I'll give it another few seconds, see if anyone has any questions. Going once, going twice. All right. Oh, one question. All right. It just takes one question to get the show going. Dr. Loeb, go ahead. Two questions, there we go. There we go, I started a trend. Robert Loeb.

Dr. Loeb: I do have questions about the wavelength inaccuracy, specifically how might the FDA be able to check wavelength accuracy of already- released products, or -- you know, I believe wavelength accuracy is an inherent problem with pulse oximeter technology. We've heard testimony that it's an increasing problem. It I believe could be something looked at by biomedical engineers in hospitals and medical centers, much more difficult to do in doctors' offices and by private individuals.

But my question is how could FDA, in a sense, release this if it's a problem? Is there someone on the FDA who can try and address it? I guess it's not a question for our speakers. Is the FDA able to address that question in this current session?

Dr. Eydelman: So this is Dr. Eydelman. I believe this session is dedicated to answering questions from the open public hearing. Why don't we save it later on. There might be opportunity to inquire to the FDA or recommend to the FDA later on. Doctor !!! why don't you go ahead with your question.

I apologize, I don't recall the name of the speaker, but hopefully he's present and will identify himself. One of the speakers stated their concerns that even though this is a moderate risk device the 510(K) process only requires clinical information from ten subjects in a study, at least that was my understanding of what the speaker was saying. And the question for the speaker would be, if he believes this to be inadequate, does he have some suggestion regarding how the regulatory regime may be changed and what sort of additional information or how many subjects may be required or would he have concerns about this in fact going to a 510(K) process and want the process to be completely altered. Thank you.

Dr. Nathan: And they might not be with us. So unless someone wants to -- unless any one of our speakers want to try to address it, I think unfortunately, that one will go unanswered.

Dr. Eydelman: Thank you.

Dr. Nathan: Okay. Sure thing. All right. Well, that puts us back on track, not that we're too far off track. I now pronounce the first public open hearing session to be officially closed and we will proceed with today's agenda. We'll now proceed with invited speakers' presentations, during this invited speaker session we will be hearing from various prospectus from patients, industry and researchers. In addition, the FDA has also invited a wide spectrum of professional societies representing clinicians who are most directly impacted by the issue of potential bias in pulse oximeter readings due to skin pigmentation. And all of the accepted invitation and completed the necessary paperwork have been given time to present during the session.

We are pleased to be hearing from eight of these professional societies today, and we will begin with Rekha Hagen's presentation, who will address the adult patient's perspective.

INVITED SPEAKERS' PRESENTATIONS

REKHA HAGEN — ADULT PATIENTS PERSPECTIVE ABOUT PULSE OXIMETRY

Rekha Hagen: Hello, everyone. I'm Rekha Hagen and I'm happy to be speaking with you about this important issue. I joined the patient family advisory council of the hospital of the University of Pennsylvania last year on the recommendation of a neighbor and a friend who was a long time member and who said they were a great group to work with.

They are a great group. With compelling individual stories who have had a long association with the hospital. I'm one of the few people on the council without a medical background, nor have I had a serious disease or been a caregiver to someone with a serious disease. I bring the perspective of a layperson and community member who desperately avoids doctors and hospitals. Apologies to all you doctors out there.

As you can see, I'm Asian Indian with a darker than average skin tone. My husband is white and our three kids are often judged to be Hispanic or middle eastern. In other words, we are many shades of brown and white. Furthermore, we live in west Philadelphia, a diverse and dynamic community. So, the first time I had ever heard of a pulse oximeter was about two years ago. A friend from India asked if we had had -- if we had an extra one that we could lone to her daughter who was in college in Philadelphia. I had never heard of such a device, and frankly thought it sounded too good to be true.

But, about a year- and- a- half ago I was standing in line at CVS and saw pulse oximeters arranged on a shelf at the end of the aisle. I bought one. I brought it home. All five of us tried it. It was almost like a toy and simple to use. It went into our medicine cabinet where it has been ever since.

A few months after I had bought it, I heard that people were being reimbursed by insurance for purchasing one. So I felt more confident about the device, thinking it had been given the seal of approval. The first I heard of the oximeter not being accurate was from the council about two months ago. I was shocked to hear that this device could lead to false conclusions. As a layperson, I would compare it to using a thermometer or weight scale whereby we believe the number to be absolute and not swayed by other factors.

In this time of COVID, it's very important to have an accurate reading because people are acting or not acting on this information. For example, if your thermometer says you have a temp of 105, I would treat it differently from a temperature of 101. I think of the pulse oximeter reading in the same way and, frankly, if the reading was acceptable I would not go to the hospital or seek help. Of course, this could be dangerous.

My question would be, since we have many skin tones in our immediate family, who would we use this device on? As for current solutions for the FDA, perhaps you could have a skin tone color chart on the box whereby you are advised not to use the product if you are darker than a certain skin tone. Or sell the oximeter behind the pharmacy counter so that the pharmacist can explain usage to the patient.

And lastly, honestly, until I had heard about this issue from the Patient Family Advisory Council, I didn't know about the difference in pulse oximeter outcomes. As a frequent watcher and reader of the news, I'm sure that if I don't know, large sections of my community don't either. The FDA has time to fix this communication. They should start now.

Thank you very much.

JESSICA COCOLIN — PEDIATRIC PATIENTS PERSPECTIVE ABOUT PULSE OXIMETRY

Jessica Cocolin: Good morning, good afternoon, members of the panel and distinguished guests. I'd like to start with a brief introduction of myself. My name is Jessica Cocolin, I'm a certified registered nurse anesthetist. I have a background in neonatal critical care nursing as well. Most importantly, I'm a mother to a child that had a significant congenital heart defect. Consequently, our day- to- day lives relied heavily on pulse oximetry. In speaking with you today I raise concern about pulse oximetry both as a practicing clinician, as well as a mother whose child was critically ill. I declare that I have no interests or involvements with any of the pulse oximetry products mentioned hereafter.

Since the invention of the pulse oximetry device in 1974, there has been a growing body of evidence that it's technology may not accurately capture low hemoglobin saturation levels in individuals with darker skin pigmentation.

To date, pulse oximeters in the United States are required by the Food & Drug Administration to be tested and certified to be no less than plus or minus 3% root mean square error at arterial saturation values.

Most pulse oximetry technology involves a minimum of two wavelengths of light emitting diodes that are differentially absorbed by oxygenated hemoglobin and deoxygenated hemoglobin; however, a vast majority of calibration and confirmation tests have been conducted on volunteer subjects of light skin pigmentation at baseline, known issues exist, such that pulse oximetry will have decreased accuracy with body temperature, altitude, barometric pressure, () motion artifact, pulse variations in tissue thickness, the age of the patient and patient comorbidities. In the 2005 study by Bikler, et al., greater difference in bias existed between light

and dark skin subjects in compare Nelcore, Novametrics and known products. In 2022, () et al both conducted systematic reviews of the literature to find that dark skin pigmentation influences the accuracy and performance of pulse oximeters.

Overestimations by approximately one percent are found in people with high levels of skin pigmentation, meaning those of darker skin tone. The findings from these studies become particularly concerning when these overestimations place individuals at risk for occult hypoxemia or when the blood is less than 88% when the pulse oximetry persists with a normal reading of 92% or more.

Pulse oximeter guides care on our most delicate populations. Decisions regarding the care of critically ill neonates balances the benefits of supplemental oxygen with the risk of retinopathy, of prematurity bronchopulmonary dysplasia and even death. Pulse oximeter can be used to detect infants with congenital heart disease.

While the COVID- 19 pandemic made pulse oximetry a household product by the guidelines in the World Health Organization, home and hospital pulse oximetry are not created equal. The care of a child with a critical congenital heart disease necessitated the continuous use of a medical grade pulse oximetry device at home. The reliance on the accuracy of this monitor literally meant life or death, life at home, or life in a hospital.

The number on a pulse oximeter display is not merely a number. It is implications for the next medical intervention a child receives, the timing of the next surgical procedure a child must endure, or even the loss a parent and family must suffer.

Consequently, I urge you to reassess the standards and parameters to which a current pulse oximetry devices are tested and certified.

Thank you again to this panel for allowing me to express my concerns regarding pulse oximetry and how their accuracy may affect individuals of darker skin pigmentation. Thank you for your time.

BOB KOPOTIC, *ADVAMED* — INDUSTRY PERSPECTIVE ON PULSE OXIMETRY

Bob Kopotic: My name is Bob Kopotic and I'm presenting on behalf of AdvaMed. I will emphasize the relationship between industry, clinical users and associated device standards, both domestic and international, which guide manufacturers of pulse oximeters. Three relevant attributes and priorities of AdvaMed are listed. We asked for this voluntary presentation owing to my clinical, industry and standards development experience noted in the bullet points.

As my current employer, Edwards Life Sciences does not manufacture or market a pulse oximeter, I have no conflict of interest.

The domestic and international regulations were detailed by others, missing was the clinical perspective for understanding and use of a pulse oximeter, which I shall address. In the United States, this medical device requires a prescription for use on a patient. The pulse oximeter is a supplemental nondiagnostic device for clinical assessment of patients. It provides an indirect and inherently imperfect estimate of arterial oxygenation and pulse rate.

So if it's imprecise, why use a pulse oximeter? It's convenient, requires little user sophistication, and provides a continuous noninvasive display of estimated arterial oxygenation and pulse rate. To improve understanding and use of a pulse oximeter, clinical use guidelines are available from various professional societies. These guidelines are published in peer- reviewed journals, commonly available online, and are frequently updated. You already heard some facets of user education and technology shortcomings that will soon be published to further improve user understanding and prompt manufacturer improvements.

What are the assumptions that create the two required displayed values of the pulse oximeter? Pulse rate and arterial oxygenation. Pulse rate is assumed in counts of detected light peeks which are believed related to expansion and relaxation of arteries from contractions of the heart or heart rate. The display of SpO2 is calculated from the amount of light lost in different wavelengths of light emitted into a pulsing volume of tissue commonly across the fingertip. Only a small fraction of the light emitted into tissues is detected by the censor, and yet, it is that tiny amount of light which the SpO2 calculates.

These are the common causes of pulse oximeter imprecision mentioned in the medical literature and clinical use guidelines. Others today have detailed these issues. International standards establish expectations of medical device manufacturers; indeed, the title of each standard commonly includes the words "essential performance and safety." The first pulse oximeter standard was released by ASTM in 1992. Dr. Goldman, Sandy Winer and I were involved then, but with a much smaller committee. With the transition to ISO, the committee of all volunteers includes international expert clinicians, key researchers, and device regulators, plus most major manufacturers.

These standards require regular review and updating to meet additional clinical needs and advancements in technology.

Five of the major items already under development for the next version of the ISO standard are listed here. Subgroups of our committee are busy addressing those major items with the goal of the next version of the ISO standard being released next year.

Thank you for your attention. Hopefully you have a better appreciation of this technology, which has benefited millions. And as with other aspects of healthcare, we will improve as we understand ways to make things better.

PAUL BATCHELDER – RESEARCHER PERSPECTIVE ON THE CONDUCT OF PULSE OXIMETERS DESATURATION STUDIES

Paul Batchelder: Hello, I'm Paul Batchelder, Chief Clinical Officer of Clinimark, today I will share pulse oximeter accuracy assessment from hypoxia trials in multiple systems. From controlled clinical testing, which evaluates the basic design and engineering of pulse oximetry.

I'll first review patient data that stimulated increased investigation into the pulse oximetry in dark skin. Then describe conventional pulse oximeter statistics, and finally present data from controlled clinical lab testing meant to evaluate the basic design and engineering of pulse oximetry over the range of 70 to 100%.

On the right is the plot of interest from the letter to New England Journal of Medicine. In the data presented all pulse oximetry readings were from 89 to 96%, while in many instances a significant number of coreference readings were below 88%. The red rectangle is the occult hypoxia reading. On the left is the clinical trials. You will note that the red outside hypoxia rectangle has few points in that region. Our goal is to understand why there is such a difference between clinical lab and patient data.

To provide an apples to apples comparison I will utilize ARMS and regression plots. ARMS accuracy root mean squared combines bias with standard deviation into one number. Neither bias nor standard deviation can be very large in order to meet an ARMS of 3.

Here's is example of an informative regression plot. You can see the 45 degree line of identity along which all accurate data clusters. It is easy to identify a significant consistent bias in one person here.

This plot quickly identifies a very poorly performing device. This is a plot from a tight sensor placement study. All subjects were first study with correctly placed sensors with data falling below the 45 degree line. Then the sensor was tightened to all subjects. You can see all subjects with tight placement read inaccurately high, with black patients reading the highest.

Now we'll look at data from sensors tested over 4644 paired data sets using the same protocol with non-- motion, carefully placed sensors in normal perfusion conditions. This data is from a finger clip sensor with black and white data sets collected at the same time. The black data is on the left with white on the right. The difference in ARMS from black to white is ARMS difference of 0.31 and a bias difference of 0.51.

This slide shows a forehead reflected sensor in the top graphs and eight different finger clip sensors in the bottom graphs. On top the forehead sensor showed a black to white difference of 0.5 and a bias difference of 0.32. The bottom plots of eight different finger clip sensors shows a black to white difference in ARMS of 0.8 and a bias of 0.73.

Here we have inferred the data from the New England Journal of Medicine letter and then plotted those data in a regression format. In the left plot you'll noted that the dotted rectangle outlines numerous occult hypoxia readings from the patient- inferred data and has no similarity to that of the control clinical laboratory plot on the right.

It is known that the Palmer service of the hand in many black persons is much lighter than the rest of the body. In these two plots we compared the ARMS and bias for black persons with a finger clip sensor to the same black group with a forehead sensor. You'll note that there's little difference between the finger and forehead. With an ARMS difference of 0.12 and a bias difference of 0.18.

Now, under carefully controlled clinical laboratory conditions, we found that when comparing to light skin over the range of 70 to 100%, the bias in dark skin is less than 1% SpO2. This indicates that the source of bias as reported in patients may be something other than the basic design and engineering of the pulse oximeter.

Thank you.

PROFESSIONAL SOCIETIES PERSPECTIVES ON PULSE OXIMETRY

DR. AMAL JUBRAN — AMERICAN ACADEMY OF SLEEP MEDICINE

Dr. Amal Jubran: Pulse oximeters determine oxygen saturation by shining a light through the skin at two wavelength, 660 in the right region and 940 in the infrared region and measuring the difference light absorbance at these two wavelengths to estimate the arterial oxygen saturation. The ratio of absorb ban sees at these two wavelengths -- direct measurements of arterial oxygen saturation obtained in a group of healthy people.

The resulting calibration curve is then used to generate the pulse oximeter's estimate of arterial saturation. Shortly after the introduction of pulse oximetry, we observed that readings were less reliable in black patients. Here you see the arterial oxygen saturation plotted against the pulse oximeter saturation SpO2 in white and black patients. Inaccurate readings were two- and-a- half more common in black patients than in white patients.

We also observed that a pulse oximeter target of 92% insured a safe arterial oxygenation which we defined as PA 02 safety and -- but in only 50% of our black patients. We surmise that the inaccuracy in black patients is related to skin pigmentation, which may interfere with the absorption of light at the wavelengths used in pulse oximeters. Studies have shown that black

nail polish causes an increase in the difference in light absorbance between the red and infrared wavelengths leading pulse oximeter to register falsely local saturations.

In addition, the calibration curve incorporated in the electronics of the instrument may be suboptimal for black patients. It is possible that the calibration data in pulse oximeters are derived largely from white volunteers. So it's reasonable to expect that a calibration curve based on data obtained in black volunteers would show better performance in black patients than the present devices.

The lesser accuracy of pulse oximetry in dark skin patients was brought to the fore during the COVID- 19 pandemic. Studies conducted by showing in the U.S. and crooks in the UK documented that pulse oximetry was less accurate in black patients than in white patients as had been shown in the past.

In the Cooks study the differences between paired measurements of pulse oximetry and the true invasive arterial saturation in black patients was plus 5.4%, indicating that pulse oximetry tends to provide falsely high oxygen saturations.

The reported 95% confidence limits, however, range from minus 25.9% to 36.8%, indicating that pulse oximetry also provides falsely low oxygen saturation in black be patients. As we highlighted in a recent letter to the European Respiratory Journal.

Managing patients with unreliable measurements of oxygenation is hazardous, partly because dangerously low saturations are missed, but also because patients are erroneously judged to have lower oxygen saturation than is really the case, which results in unnecessary intubation.

Clinical decisions based on faulty pulse oximeter readings most likely contributed to the seven fold greater number of deaths from COVID- 19 in ethnic minority patients than in white patients. In the 32 years since we first reported the greater inaccuracy with pulse oximetry in

black patients, the physical structure of pulse oximeters have undergone tremendous change, but the inaccuracy in black patients has not lessened. The mean difference between pulse oximetry and arterial oxygen saturations was 3.3% in black patients and 5.4% in the black patients of Crooks Et Al. Although the twofold population differed in several respects. The data collected in 2020 to 2021 indicate that there has been no improvement in the accuracy of pulse oximeters in black patients since 1990.

In the interval, we are not aware of any manufacturer introducing adjusted algorithms into the software of pulse oximeters to resolve the inferior performance in dark skin patients.

DR. ERIC GARTMAN — AMERICAN COLLEGE OF CHEST PHYSICIANS

Dr. Eric Gartman: Hi, I'm Dr. Eric Gartman. I'm a pulmonary and critical care physician, and I'm speaking today on behalf of the American College of Chest Physicians on the clinical implications of flawed pulse ox imagery based on skin tone. I have no financial disclosures related to this talk, so I think we'll jump right in.

To some of the clinical research that's been released in the last couple years many of you may be aware of some of this research, but most of it was related to what's called hidden hypoxemia. So the pulse ox is reading a normal level on the screen, but in indeed when they check it in the arterial blood gas they may have a low level and sometimes a profoundly low level.

And what's been found is that there is a disparate nature between the black and white patients surrogate for skin tone, which is not a perfect one, but this is self-defined race. So anywhere between 15 and 30% in a very reliable way in multiple studies over the last two years showing that much more black patients are displaying a normal oxygen level on the pulse ox, but indeed actually have hypoxemia present.

And diving a little deeper into this curve here from this study if you look towards the lower levels of normal in the pulse ox, so in this case, 91% will take a. Over 50% of the black patients here were hypoxemic and didn't know it which potentially is having significant clinical implications, both for treatment and qualifications, for different levels of care, et cetera.

This is a study from this year in Covid, patients looking for eligibility for Covid-specific treatment, which a lot of times is based on pulse oximetry. Noting that black and Hispanic patients were roughly a quarter to a 30% less likely to be offered Covid-specific treatment. Presumably because they were reading higher normal oxygen levels than they actually had again showing again the difference between hidden hypoxemia rates between whites and non-whites.

This study took things a little bit further and looking at harder outcomes showing that those with hidden hypoxemia, despite having similar levels of organ dysfunction and sickness ahead of time in that presentation, had excess mortality simply because of their hidden hypoxemia and worsen organ dysfunction scores and those with hidden hypoxemia.

So significant clinical outcomes also seen just in that group with hidden hypoxemia. What is very striking to me also in the table from this study is showing the rates again at borderline levels of normal oxygen levels, how poorly pulse oximetry performed, and it was pretty much bad for every group. It was just worse with those with potentially with darker skin tones.

So a quarter to even higher than that had normal oxygen levels on, or what we would consider okay, oxygen levels in the hospital. But actually were hypoxemic and sometimes profoundly. So that's the inpatient side of things, but what's going on in the outpatient setting could be even more profound and affect a lot more people.

And pulse oximetry is tied into lots of things and guidelines for different criteria. And one, particularly one that we see a lot is the qualification for supplemental oxygen qualification for supportive equipment or disability parking placards. Do we know how severe our patient's disease is and whether they're progressing or not?

And can we trust that pulse oximeter reading that we're seeing? And one very significant whole other field of this is sleep apnea testing, another sleep study testing where the pulse oximetry criteria are very much tied into the severity and the definition of the, of whether disease is present or not, and how bad it is.

So this is potentially a big issue for sleep medicine as well. So why is this potentially a big issue? Is that when you look at the laboratory versus the real world calibration and performance of these pulse oximeters that they perform pretty well when it's done in a normal population that is artificially made hypoxemic.

And the splay is not there in the data. Interestingly, the dark and light skin tone difference is preserved, but not dramatically but this is from the prior studies showing and shown again in multiple other studies how poorly performing the pulse ox imagery is in those with chronic diseases, those with potential severe illness, with acute illness, with circulatory issues, and potentially a range of skin tones.

So what I would advocate and what I think we should advocate. Is the calibration cohort for pulse oximetry needs to be done in the population that it's going to be used in, not in an artificial normal population. So those diverse skin tones, diverse lung disease levels of illness. So that's who we really are concerned about oxygen levels mostly in.

And that's not the population potentially that all these devices are being tested in, and we wouldn’t tolerate that for a medicine. So I'm not sure we should tolerate that for a device either.

So where do we go from here? So first I think we need to recognize that there's a problem and the potentially a big problem, and that we need to test these devices and ask these device companies to show us who they're testing these devices in and making sure it's in the people that we're using the devices in.

And that population is diverse, not only in an objective skin tone way, but also in a variety of disease process ways and inpatients, outpatients, et cetera. And in that way, we can start to trust the levels that we're seeing and hopefully more appropriately trust the accuracy of this device in taking care of our patients.

DR. ANN RIZZO — AMERICAN COLLEGE OF SURGEONS

Dr. Ann Rizzo: I'm from the Guthrie Clinics. The question at hand is whether pulse oximetry is affected by skin pigmentation and color. For treatment in the ICU as well as with anesthesia the pulse oximeter is something very important to our treatment of patients. And the American College of Surgeons not only lives by the motto but performs by the motto to heal all with skill and trust.

It is understood that there are several factors that can affect a pulse oximeter reading and the most accurate way to obtain oxygen levels in blood is to actually do a blood gas sample. Things like ambient light, the size of a person's finger, the presence of nail polish can all affect pulse oximetry in patients, and these are well- known factors in the ICU.

Because of this, more sensitive probes, like a probe on the ear, can help ascertain this risk and in other issues in the ICU, again, just obtaining a blood sample to confirm oximetry is important.

Going forward, if we can improve the accuracy of pulse oximetry, which at present it's only fairly accurate between 70 and 100% saturation, this would help this issue, especially since

the studies have shown that the inaccuracies increase with the decrease in saturation; thus, the sicker the patient, this may be affected.

I would be happy to take any questions, but the American College of Surgeons is standing behind treating all patients equally and accounting for any disparities in equipment with blood testing and blood gases in the ICU. Thank you.

JESSE EHRENFELD — AMERICAN MEDICAL ASSOCIATION

Jesse Ehrenfeld: I'm Jesse Ehrenfeld, President Elect of the American Medical Association, a board certified anesthesiologist, and a routine end user of pulse oximetry.

The AMA has long been aware of concerns regarding the accuracy of pulse oximetry in pigmented skin. Variations in the accuracy of pulse oximetry readings by patient race in critically ill patients has been reported as early as 1990, an overestimate of oxygen levels of patients with darker skin pigmentation. Increased use and reliance on pulse oximetry broadly during the COVID- 19 pandemic has highlighted the race- based inaccuracies of these devices.

The journal of the American Medical Association has reported an increased incidence of hidden hypoxemia in racially and ethnic minorities, specifically black, Hispanic, and Asian patients. Failure to appropriately identify a patient with very low oxygen levels can result in patients not receiving supplemental oxygen and potentially life- saving medical treatment leading to an increase in major organ dysfunction at 24 hours and increases in hospital mortality. The British medical journal reported black patients having higher odds than white patients of having occult hypoxemia noted in arterial blood gas but not detected by pulse oximetry.

There is significant additional literature demonstrating significant issues of racial bias and inaccuracies of pulse oximetry measurement in those with darker skin pigmentation.

Additionally, the FDA has noted its awareness that skin pigmentation can affect the accuracy of pulse oximetry readings.

Now, the AMA would like to applaud the FDA for recognition of this serious issue, that disproportionately -- amongst these populations. We appreciate the FDA's convening of this meeting today and very much look forward to working together to correct this issue. It is critical that the FDA ensures the accuracy and reliability of pulse oximetry readings in patients with diverse degrees of skin pigmentation.

The FDA should require quantitative data on device performance across a range of skin pigmentations in clinical studies, particularly devices with color- sensing technologies, to identify and mitigate the impact on historically minorized communities. Studies illuminating the varying medical treatment responses based on the color of skin warrant the design of regulatory pulse oximetry accuracy standards. Pulse oximeters that have been shown to reproduce racial bias should require a warning label that warns the end user of risks of inaccurate readings. The FDA should also require a additional collection of real- world evidence and increased post-market surveillance of these devices to ensure appropriate performance and help mitigate potential bias.

Finally, the AMA also recommends that all healthcare providers be made aware of the limitations of pulse oximetry technology and be trained to account for the systematic measurement error when developing diagnosis and treatment plans.

Earlier today when I applied the pulse oximeter to a patient in the operating room, I paused. I asked myself, how could I know that this FDA- approved device will give me an accurate reading during the critical moments ahead in surgery. We need to take appropriate steps

to remove the growing uncertainty in these devices and ensure the health and safety of the public. Thank you..

DR. STEPHEN GAY — AMERICAN THORACIC SOCIETY

Dr. Steven Gay: I'm Dr. Steven Gay, clinical professor of medicine, division of pulmonary and critical care, Michigan medicine, and interim associate dean of medical student education, University of Michigan Medical School. The issue under discussion is complex and begins with a disparity in the accuracy of pulse oximeters that affects a unique group in our society. This complexity begins with the concern that a population may be experiencing greater morbidity and mortality related to the accuracy of these devices.

That complexity is amplified by the concern that the segment of the population has been historically underrepresented in medicine and has experienced a complicated history with both access to care and fair and equitable treatment. This evaluation and investigation are not about assessing blame. We are clinicians, scientists, researchers and care providers. It is our ethical duty to investigate these issues. We do not want there to be doubt in our interventions and our approach to care.

When there is doubt about an intervention or evaluation in one portion of our society, the likelihood is that doubt spreads to others who may not be directly related simply because something isn't working in the way we expect and hope it will work.

We've had significant amounts of historical precedent in healthcare where individuals who are underrepresented in our society have been affected in such a way where doubt has occurred in their minds that they are being evaluated, treated and cared for with equity. These are two separate and distinct phenomena. There is doubt that occurs when you feel an accepted practice does not include your best interests at the forefront of its actions, but there is also doubt

from a scientific and investigational standpoint that a commonly used device is giving us correct information to best care for a population of our patients.

We should not and cannot perpetuate this doubt in any way. We cannot simply dismiss the concerns and the unique findings of this data. As individuals at all levels who participate in the delivery of healthcare in all facets of our society, we must strive to ensure that we are always working to deliver the highest standard of care, while putting our patients, their complexity and their concerns first. We must work to ensure society's trust in our practice of medicine, especially in times such as the COVID- 19 pandemic.

Considerations going forward that are directly related to this issue include, one, the ATS notes that pulse oximeter can offer valuable clinical information to guide clinical care in the outpatient and inpatient settings. Two, the medical and scientific community is aware that differences in skin pigmentation affect and impact pulse oximetry results and that such affects may adversely impact clinical decision- making. Three, the COVID pandemic and increased reliance on pulse oximetry monitoring to initiate and adjust treatment for patients with COVID expose the problems clinicians see when treating patients with darker skin pigmentation when they must rely on pulse oximeter readings for decisions on care. Four, unless the skin pigmentation pulse oximeter issues are resolved, interim guidance on the appropriate use of pulse oximeter monitoring for patients with darker skin pigmentation should be developed. And five, the pulmonary community in partnership with medical device industry and federal agencies must collect appropriate data to understand how skin pigmentation impacts pulse oximetry monitoring and must develop approved methods to ensure accurate interpretation of pulse oximetry saturation levels for all patients, including patients with darker skin pigmentation.

I am grateful to the committee for your diligent consideration on this complicated matter and I thank you for the opportunity to address you.

JULIAN GOLDMAN — ANESTHESIA PATIENT SAFFETY FOUNDATION

Julian Goldman: My name is Julian Goldman. I'm a physician at the Massachusetts General Hospital, and presenting to the panel today on behalf of the Anesthesia Patient Safety Foundation. I have no conflicts of interest to disclose.

In June of 2021, the Anesthesia Patient Safety Foundation published an online statement on pulse oximetry and skin tone, to address recent findings regarding pulse oximetry inaccuracy in darkly pigmented patients.

Pulse oximetry is an essential aspect of modern anesthetic practice. In 1986, the American Society of Anesthesiologists published standards for basic anesthetic monitoring, as the standards were adopted they led to five- fold monitoring. The current ASA monitoring standard requires the use of a quantitative method of assessing oxygenation such as pulse oximetry, naturally pulse oximetry is what is used due to its widespread availability, low cost, and relatively good performance.

When we use pulse oximeters, we consider three aspects of information. One is the trending of the saturation values, for example, we look at the patient's response to changes in inspired oxygen concentration, changes in ventilator settings or other therapeutic interventions. The threshold valued coupled with an alarm is vital to alert caregivers to respond to the patient's bedside. That response may be by anesthesiologist in the patient's OR, or to the bedside on a hospital ward by a rapid response team.

We have to ask the question, when we have a low saturation setting of 85% that's used for the alarm, what is the patient's actual saturation? How much lower is it than 85% in the

presence of pigmentation bias? And should the set saturation for the alarm be adjusted to compensate for that error? Without further information, it's unclear what to do.

Pulse oximetry is used, of course, for spot- checks as well. A single value could be used for screening, diagnosis, longitudinal care or repeat visits to a clinic and study inclusion and exclusion criterion. A very important use of spot check pulse oximetry is to assess criteria for home oxygen therapy, which is an 89% saturation value.

So if a patient is evaluated and if their actual blood oxygen saturation is 80% but their displayed oxygen saturation is 90%, that patient would not be eligible for home oxygen therapy, despite needing it. The repeatability of oxygen saturation measurements is important, and not very clear. Pulse oximetry statistics use ARMS, which is an inter-subject population statistic. But in a single patient we need to consider extra patient repeatability. We move probes from digit to digit, hand to hand as a patient is repositioned or to protect skin integrity, and we are making an assumption that in a single patient ARMS is zero. Is this assumption valid? Can we depend on repeated? This should be investigated further.

The foundation recommends the continued use of pulse oximetry as an essential patient safety device and that that SpO2 value is but one piece of information in an overall clinical picture. Patients would not be well served if clinicians ignore pulse oximetry data or default to arterial blood gasses in all darkly pigmented patients.

The APSF further recommends improved education on the proper use of pulse oximetry, including the importance of following manufacturer recommendations for proper use, proper site selection, probe selection and placement, interpretation of data, and a deeper understanding of conditions that could impair performance, such as low perfusion.

Furthermore, additional research is needed, better prospective well- controlled studies to obtain real- world evidence, better population definitions, patients with broad representation of skin tones are included and that studies are sufficiently powered to support statistical analysis.

We need objective methods to measure skin pigmentation and record that as part of the clinical studies. As far as the types of data analysis ARMS is useful but may be inadequate because it may not represent the patient population on the device will be used. Perhaps ROC curves and Clark Error Grid analyses could provide insight into clinical performance. In conclusion the APSF supports the renewed attention to the accuracy of pulse oximeter, which has revolutionized medical care and augmented patient safety, and we should use this opportunity to more closely examine how we could improve pulse oximetry performance and its use and work together with clinicians, manufacturers, and regulators to achieve this.

Thank you very much.

GARETT BURNETT — SOCIETY OF TECHNOLOGY IN ANESTHESIA

Dr. Garrett Burnett: Hello, my name is Garrett Burnett, anesthesiologist, I've prepared a short talk on pulse oximetry and skin pigmentation for you on behalf of the society for technology and anesthesia.

First, I'd like to discuss some basics of pulse oximetry. Pulse oximeters provide a noninvasive estimate of the arterial oxygen saturation or the SAO 2. And continue to be an essential monitor for anesthesiologists caring for patients in the per operative period. They function through red and infrared light through perfect fused tissue, a patient's tissue or ear. Pulse oximeters are designed to focus on the pulsatile component and oxygen saturation is provided by comparing the observance of light of oxygenated hemoglobin to deoxygenated hemoglobin in.

Anytime you're using pulse oximeters it's important to recognize the limitations of that monitor, so factors such as motion, ambient light and temperature can affect pulse oximeter function. Additionally, cardiac dysfunction irregular heartbeat or low perfusion states can impact the function. Although pulse oximeters should focus on the pulse tilt of the finger and not the finger itself nail polish has been recognized to interfere with the function. This is particularly true with nail colors such as blue, green and black. Similarly skin pigmentation and -- have been shown to limit function which has become increasingly popular topic of discussion.

Although this is kind of been a more recent discussion there's been concerns regarding the impact of skin pigmentation as early as 1993. Further, volunteer studies in 2005 and 2007 demonstrated pulse oximeter inaccuracies with patients with dark skin pigmentation, mostly in the range of pulse oximeters less than 80%, so kind of in a hypoxic state.

Now, the COVID- 19 pandemic and the importance of pulse oximeters has driven considerable attention towards the impact of skin pigmentation on pulse oximeter accuracy, particularly in the more commonly encountered oxygen saturations above 90%. The first notable study focused on ICU patients where they found a higher rate of occult hypoxemia in black patients compared to white. The term occult hypoxemia means value greater than 92% with actual oxygen saturation of less than 88%.

In one study comparing black and Hispanic patients with white patients in the OR, black patients had 44% higher odds and Hispanic patients had a 31% higher odds of occult hypoxemia compared to white patients. Regarding outcomes of this discrepancy, there was one study that demonstrated higher rates of hidden hypoxemia in black and Hispanic patients. Hidden hypoxemia, though similar to occult hypoxemia has a slightly different definition. It means that the pulse oximeter value is greater than 88% while the actual oxygen saturation is less than 88%.

Further, in this study they demonstrated that in patients with hidden hypoxemia there was a higher incidence of in- hospital mortality as well as organ dysfunction. Finally, in COVID- 19 patients one study demonstrated that the overestimation of blood oxygen levels were shown to delay or withhold COVID- 19 treatment in black patients and Hispanic patients.

In conclusion, the influence of dark skin pigmentation on pulse oximeters may impact clinical care and outcomes disproportionately, yet pulse oximeters will continue to be a vital tool for the practice of anesthesiology. This limitation should be recognized by clinicians and future technologies to adjust for this impact may be necessary.

Evidence has shown that pulse oximeters have varying degrees of bias across all manufacturers, and while not every manufacturer has the same level of discrepancy, it's imperative that device manufacturers be held accountable for the performance of their product across patients of all ethnicities. Finally, considerations for increasing diversity of FDA validation and premarket testing should be considered to further address these disparities.

Here are my references and thank you for having me.

ELIZABETH BRIDGES — AMERICAN ASSOCIATION OF CRITICAL CARE NURSES

Dr. Liz Bridges: I'm Dr. Liz Bridges, past president of the American Association of Critical Care Nurses. Thank you for the invitation to provide testimony.

AACN is the largest specialty care nursing organization in the world, providing education and advocacy for acute and critical care nurses in the U.S. Why is this issue important to AACN? Because acute and critical care nurses are the professionals providing direct care to these individuals. We are the first line in the detection of clinical deterioration or improvement.

Our advocacy on this issue reflects our over ten year journey related to equity, diversity and inclusion. In a recent EDI assessment of over 4,000 of our members, they told us they expect AACN to address these issues.

Additionally, AACN is a member of the critical care society's collaborative, which is a partnership of the four major professional societies whose members care for America's critically ill and injured patients. AACN was the signatory to the CCSC letters in 2021 and 2022 urging the FDA to direct developers and manufacturers of FDA- regulated pulse oximeters to test all devices to ensure accurate and reliable readings for patients with diverse degrees of skin pigmentation.

We continue to advocate for this action. As we address this complex issue, there are numerous areas for consideration. In the absence of immediate technical solution, the conversation on what to do right now in clinical practice is complex. A solution that has been recommended is a race- based adjustment, which based on the literature may mitigate or exacerbate the error. In considering this action we need to ask who would be benefited and who would be burdened. The other solution is to increase the SpO2 to avoid occult hypoxemia. Two recent studies demonstrated this study was -- increased in black individuals which also has adverse effects.

In clinical studies, there was a greater incidence of occult hypoxemia compared to laboratory studies of healthy subjects. We continue to advocate for post- market evaluations. It is imperative that we explore the effects of the complexities of critical illness and social determinants of health, include comorbidities on the observed bias.

In 2007, Finer published a laboratory study that found in addition to race effect in five of the six oximeter probe combinations females had a greater bias in saturation estimates. In

preparing for this testimony I reviewed pulse oximetry studies. In the past two years less than 50% of studies reported any gender data and only two studies included gender as a model variable. None of the stud teas used gender disparities as a main effect.

Finally, while acute and critical care are considered as places of care, a more appropriate conceptualization is that acute and critical care follows the patient. As evidenced during the pandemic, individuals were monitoring themselves at home and the decision to advise a patient to come to the hospital was based on pulse oximetry. To reiterate what we stated in our 2021 letter to the FDA, we remain concerned about the inaccuracies and we continue to urge the FDA to use its influence with device manufacturers to make this issue as they are making critical healthcare decisions.

Pulse oximetry is a ubiquitous technology but the discussion today is the tip of the iceberg in the bias and inaccuracies, in algorithms and clinical trials in the care of acutely and critically ill patients. This is an opportunity to set precedent to inform study designs, to ensure the avoidance of racial and ethnic bias through the use of diagnostic assessments and algorithms appropriate for all patients. The FDA can identify discrepancy across the board. This is an opportunity for the FDA to partner with professional societies, journal editors and organizations responsible for the development of guidelines and guideline standards to ensure we are not perpetuating bias through our dissemination efforts.

Beyond the regulatory solution we are considering today the FDA has an opportunity to set precedent through its powerful role as a convener, to bring together the scientific and professional communities. AACN as a leader in critical care stands ready to support this partnership. Thank you for the opportunity to present.

CLARIFYING QUESTIONS FROM PANEL TO INVITED SPEAKERS

Dr. Nathan: Thank you very much. And thanks to all of our speakers for sending presentations. And I'd like to open it up now to anyone on the panel who might have any clarifying questions for our speakers. We have about 20 minutes for questions. Yes, Dr. O'Brien, go ahead, Dr. O'Brien.

Dr. O'Brien: Joel O'Brien, patient representative. Thank you very much for all the presenters this morning. I have two questions, if I might ask. One for Ms. Hagen and Paul Batchelder. First of all, thank you for speaking up and being willing to speak to us today about your patient experience. I don't know if I heard correctly, I was just curious as to why your friend was asking for her daughter, what was their concern that they wanted the oximeter for?

Secondly is that you had indicated that you didn't know that there was any issues, you were new with the oximeter, your whole family had taken it to see, it was sort of a toy and then it was only after you went to your patient advisory group that you found out that there was some issues there, and I was curious, was there any instructions that came with it, was there any labeling that came with the device when you bought it from CVS.

Ms. Hagen: So to your second question, the box is very bland and I will admit that I didn't really look at the box. I tore it open. And then everybody proceeded to put it on their finger. So the only reason I bought it is because I had heard of it. And that's really, I guess, the only thing I can tell you. It wasn't a deep -- there wasn't a lot of deep thought that went into it when I purchased it. And I'm sorry, your first question again was?

Dr. O'Brien: You had said your friend from India had asked you about it for their daughter. I was curious as to what they were concerned about that they were looking for the pulse oximeter.

Ms. Hagen: Well, I think maybe it's more widely used there. It's a device that they knew about that I had never, ever heard of. And I think that she had packed up a box of medicines and whatnot for her daughter to have and that's one thing that she did not pack up. And so she asked if I had one and could I provide it. And the answer was no, I had never heard of it. And I could not provide one at that time.

Dr. O'Brien: Well, again, thank you very much.

Ms. Hagen: Sure.

Dr. Nathan: And Paul Batchelder from Clinimark.

Dr. O'Brien: And thank you very much again for your presentation as well. I was very -- thank you actually for discussing accuracy of root management squared because that's not something I knew and it was very interesting to learn about that. I appreciate that.

And talking about the difference between what you've heard several times now about the clinical lab versus patient data, et cetera, what I was interested in was -- I mean, it's clear that there's an issue with pigmentation but I was looking at causation. And you said something in your last slide, I believe it was, saying that maybe something else other than design or engineering was happening here. And I just wanted to get clarified from you what exactly you're talking about. I don't know if they can pull that slide up that you were referring to when you had made that statement. I think it was contrasting.

With that, as I read through the pre-material data that was given to me, and I was surprised to find out that in the finger it's only 5% of melanin there and yet we have this disparity

that seems to be equal other places in the body and I was very curious about that in terms of is it -- clearly the observation with skin pigmentation has an issue, but is that really the cause of the variance we're getting.

Mr. Batchelder: That's a good question. The -- your question about what I meant when I said maybe it is something other than the design and engineering, controlled clinical laboratory testing is designed to test the best case performance of the device, and then from that, after that, we might -- we're entertaining other types of tests as well. But the big question is, when these devices are performing this well on deeply pigmented subjects, you know, and we had 1,000, 14 paired data points from deeply pigmented subjects in that data pool, what is it that creates this vast disparity in readings in the deeply pigmented subjects. And I believe Dr. Bickler will mention some things about that, but there may be combinations of low pigmentation -- I mean deep pigmentation with low perfusion or possibly if you refer to one of my earlier slides, sensor placement can also cause that.

So these are the conditions that we start with, we know the basic performance of the device in the clinical laboratory testing. Now, what's next and how do we provide controlled laboratory reproducible testing that represents the egregious issues that we're talking about today.

Dr. O'Brien: Okay. All right. Thank you.

Dr. Nathan: Dr. O'Connor.

Dr. O'Connor: Great. So my question is nominally for Jessica Cocolin. And before I ask it I want to make it very clear that I'm asking this question as much to get it on the table rather than any expectation that she'll be able to answer it. And the question is, you know, neonates have a lot of hemoglobin F which has a lower P50 than adult hemoglobin. So my question is, are we confident in the performance of pulse oximetry in neonatal patients, and once again, the issue

of skin tone matters here because African Americans and people of color in general are at far higher risk of having children prematurely. And so as we go down the pathway of trying to understand this, we may require a different expectation or standard for preemies that are neonates or for the neonatal population altogether than what we do for the adults. And I don't see that she's here to answer it.

Dr. Cocolin: Hi. Sorry. Just hopping on the video now. So while I did practice in neonates, I haven't in a long time. It's been about ten years. And I -- I don't think I can say for sure what the current research is and I don't -- I don't know how you can accurately measure fetal hemoglobin each ratio in each infant. So I'm not -- I'll be honest, I'm not sure exactly how accurate a pulse oximetry would be in that patient population. It is a great concern, though, especially in those neonates when you're trying to treat, again, those issues like retinopathy and prematurity and --

Dr. O'Connor: Yeah. I didn't expect you to have the answer. I just wanted to get the question on the table. Thank you so much.

Dr. Nathan: Thank you. Dr. Collop.

Dr. Collop: My question is for Mr. Batchelder again, and it gets to the issue about the probe, I guess. You were insinuating I guess with the one slide that the tightness of the probe can make a substantial differences. I know at least one home testing devices has a probe that's one size fits all and I can imagine that may have some implications on its accuracy. I also wondered like current oximeter probes, there are lots and lots of choices out there.

Do you have any comments about like how probes themselves should be analyzed, if that's important? And one other question about like skin thickness, could that actually have a difference as well.

Mr. Batchelder: Certainly. I'll start at the end. Skin thickness definitely can. And I want to caution against using simulators with sensors because it's the sensor skin interface that can only be tested in humans. And the performance of the entire system has to be looked at in humans in this area. Is there another area of your questioning that you have?

Dr. Collop: Yeah. Just, I guess, the probes themselves, like how well they fit or how they fit, you know, how should that be examined?

Mr. Batchelder: Right. Yeah. So that's a real big question. And the information that I provided in that one slide was taken from a -- Dr. Manheimer's dissertation which is publicly available, where if the sensor is wrapped too tightly, you know, in a disposable sensor, that can and what has been shown in his studies to increase the reading, increase the bias in all subjects, and in black subjects as you saw in that study much higher.

So sensor placement has also been long known that where it is placed on the finger can make a big difference in the SpO2 reading in a single subject. Dr. Goldman mentioned, you know, intersubject variability and that is something that needs to be evaluated. And that is one of the things that can be part of this issue. When we move from carefully controlled laboratory studies where the sensor placement is assured before the hypoxia trial starts, then what happens in the clinic, when we're in a hurry the sensor may move. And then another aspect of sensors that you brought up is they can be significantly different. Some sensors have white background on the sensor, some are black, and those make a big difference in how the sensor will perform.

And then also, if the sensor does not inject the light directly into the tissue but allows some of it to go along the tissue and hit the detector without going through the tissue can make a significant inaccuracy. So all of these things are combinations of what might be happening from

our carefully controlled testing of deeply pigmented and lightly pigmented subjects in the laboratory to the hospital environment.

Dr. Collop: Thank you.

Dr. Nathan: Also, as you're talking to us hearing other factors, made me think of things I hadn't thought about before, for example, morbid obesity, I don't know how much subcutaneous fat there is, but how does that affect the saturations. And I think, I didn't know this previously, but when you don't have a good signal, the knee jerk is to squeeze the probe harder or wrap it tighter until you get a signal which appears that falsely elevates the SpO2. Would that be correct?

Mr. Batchelder: That is correct. That happens often. And that can cause a significant issue with the actual accuracy performance of the device.

Dr. Nathan: Thank you. Dr. Loeb.

Dr. Loeb: I guess I'd like to pick on Dr. Ehrenfeld. One of his statements is he now goes into the operating room and wonders is this pulse oximeter going to give me accurate readings in this patient. And I would propose that is actually a question that all clinicians should ask themselves about every noninvasive monitor. Pulse oximetry, noninvasive blood pressure, temperature readings are -- should be known to be inaccurate or have limits of accuracy, I guess I should say.

And to me one of the problems is -- so, yeah. I would ask Dr. Ehrenfeld has his questioning the accuracy changed because of recent studies or is this something that will I wonder whether he asked himself all along. And I'd be interested in actually a response from any of the technology- savvy clinicians, many of which are on this call.

Dr. Nathan: I don't see anyone responding to that. I guess it's quite provocative for all of us to think about. I don't know if there's --

Dr. Gay: I'm happy to take that, Steve Gay, pulmonary care, University of Michigan. I think we have always assumed that none of the devices that we use from a noninvasive standpoint are completely and totally accurate. I think the issue at hand here is there seems to be a greater discrepancy in specific groups of patients as opposed to an overall variability that we would expect in every patient. And I think that is one of the things that gives us a little bit more concern in our approach of interpreting the data.

We have now had to begin to wonder at what point do we look at the normal amount of variability and it's okay, at what point is -- do I start to consider is this the amount of variability that's expected or is this more than the amount of variability that's expected. I think because it's not a standard across the board in all patients, it makes it more difficult to interpret and use from a clinical standpoint from that approach.

Dr. Nathan: That's interesting, though, that, you know, as clinicians we like these hard guidelines. So if you think about the prescription for supplemental oxygen, if you're 88% or less and there's no kind of continuum of that and so people kind of hang their hats on that, and I think that perpetuates the fact that the devices are more accurate than is realized.

Dr. Rizzo: So this is Dr. Rizzo from the American College of Surgeons, I'm a surgeon intensivist and I would like to echo what was just said. We as clinicians are naturally dubious of any answer and continue to check it. I tell people you want someone who's paranoid and checks it again and again. And frankly, I don't prescribe home oxygen or not without getting a blood gas to corroborate the evidence.

And it's with all of our devices, it's not just with pulse oximetry. We have to understand that there is a natural tendency for them to not be 100% accurate. And anything that industry can do to make these devices more accurate is an obviously welcome development for us as clinicians. We strive for accuracy but we also check and double- check. And I'm sorry, I can't share my video because it's giving me the "you can't share video" message. So thank you.

Dr. Nathan: Mr. Branson.

Mr. Branson: So I may be getting a little bit ahead of our discussion this afternoon. Health disparities are real in every part of healthcare, but as an editor of a journal, if any of these studies that have been presented outside of the ones by Bickler and Jobran and Toban back in the '80s were submitted to my journal as an evaluation of pulse oximetry, I would have rejected all of them because the time of the recording of the SpO2 value and the recording of the blood gas value was anywhere from five to 30 minutes or closest one available in electronic medical record.

I'm not saying that this is not real, but it does make -- it makes some sense that there are these large disparities that, again, we see from some of the other presenters that there's no difference in studies related to race. And it clearly includes skin pigment perfusion, temperature, and there are so many other things. And as a respiratory therapist, I find the root mean square to mean absolutely nothing to people at the bedside. You know, you can all test it this afternoon if you go into the hospital and ask the therapist what the root mean square is of the pulse oximeter device they're using. They don't know. They're taught that if the heart rate on the oximeter reads the same as the heart rate on the EKG, then it's probably accurate.

So I think there needs to be -- I'd be anxious to hear from any of the speakers that is there a better way to discuss what accuracy is so that people are well- informed? And I was listening to

Dr. Bridges, who I've known for several years, and I think about first principles and that's that if there is a gender bias, what's the -- you know, with skin pigment we think it's the melanin. If it's a gender bias the first thing that comes to my mind is the probes don't fit as well on women who tend to be smaller than men and there is this effect that Paul describes where the light shines around the finger and goes to the sensor.

So I'm sorry if that's too much rambling, I've been sitting here thinking about this that some of the data we see from these studies really begs for, as other people have said, prospective data collection with measurement of skin pigmentation.

Dr. Nathan: Thank you. Dr. Loeb, go ahead.

Dr. Loeb: I believe in the last presentation we had some of the authors of the papers of the retrospective studies, but I was wondering about this outcome of occult hypoxemia, which I wonder how that was chosen. Seems like a lot of the studies have occult hypoxemia, which is a very compelling outcome to a clinician. But as was shown earlier by the statistician, positive predictive value is really the much more important outcome. It speaks to the same affect but does it in a more statistically correct way than just looking at two boxes in the four- box of outcomes.

So this is a question for any of the people who have published using occult hypoxemia how they picked that and why it seems to be, in my mind, perpetuated in many of these studies.

Dr. Nathan: Any of our speakers want to tackle that?

Mr. Goldman: This is Julian Goldman. Regarding Dr. Loeb's question, I wanted to respond to the prior question as well. I think that there is a consideration that we've missed in the discussion and that is that as clinicians we look -- we have a high index of suspicion that perhaps we cannot trust data from any monitor, any device, including a lab test. But I think the unique aspect of today's discussion is that our historical cues as to the performance of a pulse oximeter,

as Dr. Branson said, for example, matching the heart rate display with the ECG, the pulse rate rather with the ECG, or looking at the quality of the pletisgraphic wave form or the graphic we've all come to learn over time may not help and does not seem to help provide insight into this aspect of performance or the failure in performance in some subjects.

I think that's really what we're wrestling with uniquely and that has been pointed out, this is unique and it's partially unique because our other cues no longer apply. And if we had other information that could flag or identify when a patient's measurements might be more inaccurate because of the skin tone or the pigmentation, that in and of itself could help assess or help us interpret the data. Thank you.

Dr. Nathan: Sure thing. Dr. Yarmus, you had your hand up and then you lowered it. Okay. Well, if there are no more questions from our panelists to any of the speakers, then we will close this session out. We'll take a 40- minute lunch break, or just under 40 minutes, and we'll reconvene at 1:00 p.m. Panel members, please do not discuss the meeting topic during the break amongst yourselves or with anyone else attending virtually. And as I mentioned, we'll resume the meeting at 1:00 p.m. Thank you very much

INVITED SPEAKERS' PRESENTATION — REAL-WORLD EVIDENCE AND PULSE OXIMETRY

Dr. Nathan: Good afternoon, everyone. It is now 1:00 p.m. and I would like to resume this panel meeting. It is now time for the second scheduled open public hearing as posted in the Federal Register Notice. However, all requests received from the public to present today panel were covered during the morning open public hearing session. So we gain an hour even without the official time change.

We will therefore move forward with today's agenda. So there's open public hearing that was scheduled between 1:00 and 2:00. Nothing to fill there. We had a lot of comments in the morning, a lot of information to process for the panel. And so we're going to move on to the invited speakers presentation.

This next session is covering real-world evidence related to pulse oximetry. During this invited speakers session we will be hearing from several researchers who have conducted studies to assess real- world performance of pulse oximeters and will present their findings to the panel. Starting with Dr. Sjoding.

DR. MICHAEL SJODING

Dr. Sjoding: Hello. My name is Michael Sjoding I'm an associate professor of Internal Medicine at the University of Michigan and I'll be speaking about our research in racial bias in pulse oximeters. We started to get interested in this topic during the COVID-19 surge at the University of Michigan. Myself and some of my colleagues noticed anecdotally a couple of times that we were caring for a patient who seemed to have a normal saturation based on a pulse oximeter value on the monitor, but during routine care when we performed an arterial blood gas analysis we saw the oxygen saturation was low and we didn't understand why this discrepancy existed.

When we dug a little further, we learned about prior studies demonstrating pulse oximeters had different accuracy in white and black patients. We got really concerned that's what we were seeing at Michigan when caring for patients with COVID- 19.

And so for that reason we set out to perform a study to evaluate pulse oximeter accuracy in clinical practice. And so we analyzed routinely collected electronic health record data, including patients who self-reported as white or black and were hospitalized in 2020. We

compared pulse oximeter recordings to arterial blood gas measurements that were performed within ten minutes. When using this criteria we arrived at a population of 1339 white adult patients and 276 black adult patients, including over 10,000 pulse oximeter and arterial blood gas pairs.

When we plotted the data, we saw that there was a clear difference in pulse oximeter performance in black patients than white patients. That the same pulse oximeter reading value, say 92%, the arterial oxygen saturation levels in black patients were lower than white patients and this was true across the pulse oximeter range. And this suggested that pulse oximeters seem to be overestimating the arterial oxygen saturation in black patients compared to white patients.

And to really make this clinically relevant, we asked how often was pulse oximeters missing hypoxemia when they were reading at 92% to 96%? And this occurred really infrequently in white patients, 3.6% of the time. But in black patients, 11.7% of the time when the pulse oximeter was reading normal in a range of 92 to 96%, the actual arterial oxygen saturation was less than 88%. And when we also performed this analysis adjusting for differences in age, sex, and cardiovascular SOFA score we arrived at the same conclusion, that there was a significantly higher rate of occult or hidden hypoxemia where the pulse oximeter was reading normally but the actual arterial oxygen saturation was less than 88%.

We also asked another related question, which was, how well was the pulse oximeter at discriminating low oxygen, defined as an arterial oxygen saturation less than 88%? We found that in black patients the area under the receiver operating characteristic curve for low oxygen was 0.84, whereas in white patients the area under the receiver operating characteristic curve was 0.89, again demonstrating that the pulse oximeter was less able to identify low oxygen levels in black patients compared to white patients.

And finally, we replicated these findings in a multicenter cohort from 2014 and 2015 and we published these findings in our research letter in the New England Journal of Medicine in 2020. And since that publication, we've replicated these main findings in several other patient cohorts. We found again that pulse oximeters were missing hypoxemia 21.5% of the time in black patients and only 10.2% of the time in white critically ill adults getting evaluated for ECMO therapy. We also found that in black children pulse oximeters were missing significant hypoxemia 21.1% of the time in black and 15.6% of the time in white children.

And finally, in a recent analysis we found again that pulse oximeters were missing hypoxemia 19.6% of the time in black and 15.6% of the time in white Veterans. And this analysis was performed across all U.S. Veteran Affairs Hospitals.

And in doing these analyses, one thing that we've come to recognize is that potentially small differences in pulse oximeter accuracy can translate into really significant population disparities in hidden hypoxia and we're illustrating that in this figure in a publication that recently came out in the American Journal of Respiratory and Critical Care Medicine. So what this figure is illustrating is the possible values for arterial oxygen saturation when a pulse oximeter is reading 92%. And so first, let's take a pulse oximeter that's very accurate, with a precision of 2% and zero bias, and that pulse oximeter would have an accuracy root mean square of 2%. And for that pulse oximeter, this occurrence of hidden hypoxia would be quite rare.

When the pulse oximeter was reading at 92%, the rate of hidden hypoxia with the pulse oximeter missing low oxygen is 2%. But if that pulse oximeter was slightly less accurate, if the precision increased to 2.5% and the bias increased to 1%, the rate of hidden hypoxia becomes quite a bit higher. Now when the pulse oximeter is reading 92%, that could represent a true arterial oxygen saturation less than 80%, 12% of the time. So this again just demonstrates how

small differences in pulse oximeters can translate into significant population disparities in hidden hypoxia. So that's why it's important to recognize that these small differences at a population level could potentially have a big impact.

So these analyses that we performed were all performed using retrospective electronic health record data and this type of data has many opportunities and strengths but it also has some limitations. It provides an order of magnitude more patient data to analyze, so you can do analyses that we've never been able to do before. And it also provides real- world understanding performance and potential clinical impact of inaccuracies.

However, pulse oximeter and arterial blood gas recordings are not performed simultaneously in this type of study, using this type of data, and that will introduce some random noise because from the time that the pulse oximeter is recorded to the time the arterial blood gas is collected, there can be some random drift in the oxygen saturation which will introduce some random noise.

But importantly, in our New England Journal of Medicine study the difference in time between the pulse oximeter and arterial blood gas on average, was really identical in both black and white patients. So the same type and amount of noise would be introduced, and so that just cannot explain why we would see these differences between black and white patients.

And then finally, these types of analysis are based on self-reported race, which is an imperfect measure for skin pigment level. And, of course, we all tend to think that differences in skin pigment level are what's explaining these differences; however, we do think it's really important to recognize that that those most impacted are racial minority groups.

So I'd like to close by just offering some actions. I really think the FDA should consider based on these analysis. First, it’s paramount that manufacturers report pulse oximeter

performance across racial groups. And then we must also power pulse oximeter studies, to be able to detect small but clinically important performance differences across these groups.

And we have to ensure that pulse oximeter testing is aligning with our use in clinical practice. So in clinical practice, when I am caring for a patient what I want to know is if that pulse oximeter is reading normally, how often am I actually missing real low oxygen saturation and that's why we came up with this concept of the occult or hidden hypoxia rate. And finally, we have to start using routinely collected real- world post-market surveillance data more and more, including both retrospective and prospective data to evaluate how pulse oximeters are performing in clinical practices because it's critical these pulse oximeters perform the same for everyone because they're such a important device for providing good patient care.

DR. IAN WONG

Dr. Ian Wong: Hi, My name is Ian Wong, and I'm pulmonary and critical care doctor, in the Pulmonary and Critical Care Medicine, and the Division of Translational Biomedical Informatics, at Duke University. Thank you again for the invitation to speak on pulse oximetry and hidden hypoxemia.

Here are my disclosures and funding. I've been invited to speak to you today about some of my recent work published in JAMA open about hidden hypoxemia and pulse oximetry. Thank you also to my colleague, Michael Sjoding who helps inspire our team in this line of work.

To briefly review to ensure we are aligned on terms, I'd like to discuss both the pulse oximeter and the arterial blood gas. On the left, the pulse oximeter estimates oxygen saturation by shining a light through tissue. Arterial blood gasses on the right measure true oxygen saturation with blood taken directly from your artery. Pulse oximetry values are represented by SpO2 and ABG by SaO2. For this study we studied hidden hypoxemia where patients had a

pulse ox value of 88% or higher but an ABG value strictly less than 88%. As you know 88% is the threshold for when we consider patients hypoxemic clinically. For hidden hypoxemia we sought to answer two questions. The first, what's the clinical impact of hidden hypoxemia and the second, can we extend this work beyond black and white patients.

Our results across 87, 971 patients in search of these answers let us define clinical implications in terms of mortality, where the presence of hidden hypoxemia was associated with 70% higher mortality in all comers. Beyond mortalities we found effects on organ function, as evidenced by 31% slower lactate clearance, which is effected by oxygenation, and 14% higher organ dysfunction as measured by SOFA scores. We sought to extend this work beyond black and white patients and not only found that hidden hypoxemia was 38% more frequent in black patients, but also 23% higher in Hispanic patients.

And another unexpected finding associated the patient's raise to 40 to 66% fewer chances of having their ABG, the gold standard test, measured in Asian, black and Hispanic patients. We found the gold standard test was less frequent in patients of color. These are three big takeaways. The presence of hidden hypoxemia is associated with 70% higher mortality, that hidden hypoxemia is 23 to 38% more common in Hispanic and black patients; and that Asian, black and Hispanic patients had 40 to 66% fewer ABG tests.

This makes sense given the physics behind pulse oximetry. Since it's a simplified version of spectroscopy as you see here. We see spectra of blood, mortar and hemoglobin here the most common skin pigment molecules, the pulse ox compares the red and green lines you see here in this chart and you see that melanin overlaps with these lines here. The only way around this is to remove the melanin again, the ABG, which removes blood out of the body and measures it. So although the pulse ox is more convenient, the ABG has inherent accuracy benefits.

We've known that these complications have been present for years in both healthy and ICU patients. So what can be done about this? Are we hopeless? We believe that there are both potentials for solutions in hardware and in software. And we're happy to discuss this at a later point.

But for the purpose of this talk, let's ask, what could the FDA do? The FDA had kindly released this pulse oximeter safety communication in February of 2021, but there were no conclusions and no recommendations off that. Just caution in considering the pulse oximetry results.

So while these improvements are being developed, how can we help the FDA build the framework by which we can judge these improvements? Based on the results of this work, I believe we can consider reevaluating what the standards for accuracy are. Per the communication, this generally means that the accuracy requirements of 66% of SpO2 values within 2 to 3% of blood gas values, 95% of SpO2 values were within 4 to 6% of blood gas values respectively.

So if we look at this bell chart, looking at standard deviations and with the pink showing 95% ??? interval, given the plus minus 6% anywhere from a true ABG stat of 94% range from 88% - 100% by the pulse oximeter. If we're reevaluating accuracy standards we can consider tightening accuracy requirements. A true value of plus minus 6% is permitted in SaO2 94% anywhere between 88 and 100% which means that is my entire clinical range of treating a patient and keeping their oxygen saturations.

Additionally, to reevaluate accuracy standards. The FDA guidance suggests that these clinical studies were not statistically powered to detect differences in accuracy between demographic groups. Can we considered testing enough patients to do for example, just as we

consider testing patients in drug trials, could we hold device trials to similar standards or perhaps standards that focus on results in specific subpopulations. Furthermore, what context do we need to evaluate and test these patients.

We note that pulse oximeter accuracy is highest at saturations of 90 to 100%, intermediate at 80 to 90% and lowest below 80%. Pulse oximeter accuracy is affected by value here. So can we challenge them to test across sufficient SpO2 ranges throughout clinical care. Specifically, we need to test ranges to put the lower range accuracies or in this case inaccuracies to the test, the lower ranges, especially around 88%, are the most important in changing clinical management. For example, the patient is at 90%, perhaps current standard of care may not consider turning up oxygen therapy, whereas if they were 86% they might.

Furthermore, another consideration includes that the FDA guidance recommends that there are at least a range of participants skin pigmentations within the accuracy testing, including at least two darkly pigmented participants or 15% of the participant pool, whichever is larger. But let's ask, what does darkly pigmented mean? How do we quantify that? And finally, how do we do so in numbers that allow us to draw significant conclusions from that data?

We do need to test across skin chromophore for ranges. As we noted, skin chromophores like melanin have reasonable ways they can bias our measurements. On one side Fitzpatrick himself noticed ethnicity and race have no scientific basis but the FDA publication on reducing bias in AI/ML- based research itself notes that there are many contexts to race and ethnicity.

We need to test for accuracy within a range of skin chromophore concentrations to know when our results might be valid. Could we consider looking at both performance of pulse oximeters and downstream care of as possible results in metrics. But how do we do this? Is it the Fitzpatrick skin scale which was developed originally across four skin types and skip types four,

five and six were added for black patients. But how do we do this? Is it the skin scale? It was developed with ranges one through four primarily for white patients and types five and six were generated just for black patients decades later.

Is it the Vonlution scale with 36 different options or is it the ?? skin tone scale which was developed to help improve research in AI and ML algorithms. Or conversely, do we realize some of these may actually be subjective gradations and that would objective measures be perhaps beneficial for helping us to remove the subjectivity and helping us deliver more objective equal care to all patients.

So in summary, I think that the FDA's biggest effort here is to consider reevaluating accuracy standards. Could we begin to consider pulse oximeter disparities and the downstream implications as a quality improvement metric. The fact of life issues, for example, of () were resolved once we considered them as part of our QI system. Are these different ways by which we would consider judging both the technology and the systems of care themselves?

I'd like to thank my colleagues as shown here. Thank you for the opportunity to present with you. I remain ready to answer any questions you may have. Thank you.

DR. ASHRAF FAWZY

Dr. Ashraf Fawzy: Hello. I'm from Johns Hopkins University. I'll be speaking about the consequences of racial bias in pulse oximetry on clinical decision-making, highlighting our recently published research in patients infected with COVID- 19. I want to thank the organizers for inviting me to speak on this important topic. I have no financial interest or professional relationship to disclose.

Informed by Dr. Michael Sjoding's study published in December of 2020 showing higher prevalence of occult hypoxemia among black patients I became aware of the issue in my clinical

practice. This graph displayed the oxygen saturation of a 50- year- old black man with COVID-19 infection who I cared for in the intensive care unit. On the Y axis you see the oxygen saturation and on the X axis the time and hours. The readings from the pulse oximeter are represented by the blue line, while the gray line represents the actual oxygen saturation as measured by an arterial blood gas. As a consequent of consistent overestimation of the true oxygen saturation by the pulse oximeter and an error of up to 20%, the patient experiences ten hours of occult hypoxemia defined as oxygen saturation below 88% despite normal pulse oximeter readings highlighted in red before a supplemental oxygen is increased to a hundred percent.

This 85-year-old black man with COVID-19 spends over five hours with occult hypoxemia as a consequence of pulse oximeter error of up to 17% before this is recognized and the amount of supplemental oxygen is increased to 80% improving the patient's oxygen saturation from 76% to 96%.

Informed by our clinical experiences my colleagues and I conducted a study using curated clinical data from patients infected with COVID- 19 within the five hospitals of the Johns Hopkins health system. We first aimed to determine whether and to what extent pulse oximeter accuracy differed by race and ethnicity. Second, we aimed to evaluate whether pulse oximeter inaccuracy by race and ethnicity impacted recognition of eligibility for oxygen threshold-specific therapy.

By studying a group of patients who all had the same disease, in this case COVID- 19, we were able to evaluate clinical decision- making. To accomplish the first aim of the study we identified 1216 patients with over 32,000 paired oxygen saturation measurements within a 10 minute window. We show that while on average oxygen saturation is typically underestimated by pulse

oximeters for white patients, pulse oximeters routinely overestimate the true oxygen saturation in Black, Hispanic and Asian patients, which can lead to underdiagnosis and undertreatment.

We also note that extreme overestimation is more common among patients from racial and ethnic minority groups with almost a quarter of pulse oximeter readings in black patients overestimating the true saturation by 4% or more.

Using a statistical model to adjust for demographics, comorbidities and other factors potentially associated with pulse oximeter inaccuracy, such as temperature, blood pressure, hemoglobin and bilirubin level pulse oximeters continue to signifigantly overestimate true oxygen saturation among Black, Hispanic and Asian patients compared to white patients by an average of 1 – 2%.

As I'll demonstrate in the subsequent slides the seemingly small average difference resulted in substantial clinical consequences.

To accomplish the second objective of our study which was to determine the clinical impact of pulse oximeter inaccuracy, we looked to the larger sample of patients with COVID- 19 who were monitored by pulse oximetry but did not necessarily have an arterial blood gas measurement.

Using the statistical model we created we predicted patients' true oxygen saturation. Since the joint guidelines of the CDC and the infectious disease society of America recommends starting COVID- 19 targeted medications, such as remdesivir and dexamethadone at an oxygen saturation threshold of 94% or below we identified 1903 patients breathing room air with predicted oxygen saturation fell to 94% or below while their pulse oximeter reading was above 94%.

In essence, these 1,903 patients qualified for COVID- 19 therapy based on guideline recommendation. However, this was not picked up by the pulse oximeter. These patients were further divided into two groups, those with pulse oximeter readings, never fell to 94% or below

and were thus never recognized as needing COVID- 19 therapy and a larger group who eventually had their need for treatment recognized.

Of the 451 patients who never had their need for treatment recognized, meaning that their pulse oximeter readings never fell to 94% or below, despite their predicted oxygen saturation falling to that level, half were black, a quarter were Hispanic, but only 17% were white. The remaining 1452 patients eventually had a pulse oximeter reading of 94% or below. Of this group, black and Hispanic patients were significantly more likely to have a delay in recognizing their need for COVID- 19 treatment. Black patients were 29% less likely to receive timely therapy, and Hispanic patients 23% less likely because of systematic overestimation of oxygen saturation by pulse oximeters.

In conclusion, our study builds upon mounting evidence showing that pulse oximetry overestimates oxygen saturation in patients from racial and ethnic minority groups, by demonstrating the dispite small average differences there were substantial delay in recognizing need for COVID- 19 therapy for black and Hispanic patients as a consequence of differential pulse oximeter inaccuracy. This may have contributed to racial disparities and COVID- 19 outcomes reported in the literature.

Limitations are similar to other retrospective analyses using clinical data with the added limitation, the treatment delay estimates depend on when pulse oximeter was measured or recorded. The differential overestimation of oxygen saturation by pulse oximeters is part of a consistent pattern of systematic underdiagnosis of disease and withholding of therapy from patients of racial and ethnic minority groups.

For example, determining kidney and lung function relies on equations that incorporate race leading to underdiagnosis and either delay or withholding of life- saving treatments or

interventions such as transplants. However, unlike these examples, pulse oximeters appear to have a fundamental flaw in the acquisition rather than the interpretation of data.

The evidence that pulse oximeters overestimate oxygen saturation among black, Hispanic and Asian patients in clinical settings has implications beyond COVID 19 and acute hypoxic respiratory failure. Diagnosis of chronically low oxygen saturation among individuals with COPD or interstitial lung disease in the outpatient setting relies on pulse oximetry. And insurance coverage for long- term oxygen therapy is based on strict oxygen saturation cut- offs which can lead to inappropriately withholding this therapy from certain populations.

Finally, it is imperative to ensure that pulse oximeters perform equitably before further expanding their use. For example, the diagnosis and classification of acute respiratory distress syndrome, or ARDS, a condition that can affect patients with severe COVID- 19 and has a high mortality rate is currently based on partial pressure of oxygen which must be measured by an arterial blood gas.

Some researchers have recently suggested loosening this definition to allow estimating partial pressure of oxygen from pulse oximetry. Based on the results of our study and others, it is likely that the scenario depicted here where oxygen saturation is overestimated by 3% resulting in misclassification of ARDS as mild instead of moderate would occur more commonly belong Black and Hispanic patients leading to withholding life- saving treatments.

I would like to acknowledge and thank my coauthors who made this research possible. We urge the FDA to strongly consider the real- world evidence presented today using data from thousands of patients in the clinical setting that highlights the challenge of relying on differentially inaccurate pulse oximeters. These findings raise significant concerns that the device class has

substantial shortcomings that lead to disparities in the way patients from racial and ethnic minority groups are treated.

Enhanced regulatory requirements are urgently needed to ensure equitable patient care. Thank you for your attention.

DR. ERIC GOTTLIEB

Dr. Gottlieb: I'm Dr. Eric Gottlieb, a hospitalist at Mount Auburn Hospital in Cambridge, Massachusetts, the MIT Laboratory for Computational Physiology and Instructor of Medicine at Harvard Medical School and I'm going to share today our research on racial disparities and pulse oximeter performance and oxygen supplementation rates. I have no relevant disclosures.

As we know there is a growing body of research on pulse oximeter performance disparities by race, and in addition to studies over the past few decades that have showed that these disparities exist, a number of more recent studies have shown that these differences by race are associated with adverse outcomes, including mortality.

However, what has not been explored until recently are the factors that mediate these adverse outcomes; in other words, if you know you have pulse oximeter performance disparities and you eventually get these adverse clinical outcomes what are the factors in between that get you from point A to point B?

This question led to a recent study which was published in JAMA Internal Medicine. The premise of this study was that if there are disparities in pulse oximeter performance and some patients have pulse oximeters that read artificially high, does that lead those patients to receive lower delivery rates of supplemental oxygen in the hospital?

So again, the objective of this study was to determine if there are differences in supplemental oxygen delivery rates across races when controlling for blood hemoglobin oxygen saturation and other confounders. This study was based on data from the Medical Information Mart For Intensive Care, or MIMIC-IV, which is a collaboration between the Beth Israel Deaconess Medical Center in Boston and MIT. For this study we identified a cohort of patients from MIMIC-IV who were in the ICU for at least twelve hours prior to intubation, noninvasive positive pressure ventilation, or BiPAP, high flow nasal cannula or tracheostomy or went for five days or more, and five days being the maximum amount of data in this study without receiving any of these treatments.

For this study we used time weighted averages for the variables of interest, and we did this using an area under the curve, and these variables included oxygen delivery, hemoglobin oxygen saturation or SaO2 and SpO2. For our ableists we used linear regression and mediation analysis to assess associates between SaO2 and SpO2 and oxygen delivery rates according to race.

This figure shows the selection criteria for our final cohort of 3069 patients and the reasons that patients would have been excluded from this analysis and this includes race other than Asian, Black, Hispanic or White. Patients missing key data points that were used for analysis. And then patients with an index period, which is the relevant included data period of less than 12 hours.

This is our table one which shows the clinical characteristics of the patients in this study and important things to note here are, first, the racial breakdown of this study up top, which includes 83 Asian patients, 207 Black patients, 112 Hispanic and 2,667 White patients, and the other thing that's very important down here is the SaO2 averages, the hemoglobin oxygen

saturation and the SpO2, and we can see, again, differences according to race. And what's notable is that if you compare, for example, Black patients and White patients, black patients have lower average blood oxygen saturations, SaO2 compared with White patients, and Black patients also had higher SpO2 conversely than White patients.

This is table two, which provides some additional information on the clinical course of patients in this study including when blood gasses are drawn and how many for various patients. For those patients who required advanced respiratory support what that consisted of, and also the care units in the hospital where patients were admitted.

This is a visual summary of our main findings. And walking through this figure in panel A, we see on the X axis the average hemoglobin oxygen saturation and on the Y axis is the average SpO2. And if you look at this figure, the light blue is Black patients and the dark gray is White patients. And what we can see here is that black patients were given hemoglobin oxygen saturation have a higher SpO2 than White patients.

In panel B, we have box plots showing the average oxygen delivery rates and we can see that there are many more White patients receiving high rates of supplemental oxygen compared to Asian, Black and Hispanic patients.

In figure C, we have hemoglobin oxygen saturation versus average oxygen delivery rates, and it should be noted here that for this visualization we excluded some outliers so the cut off for oxygen delivery is zero to ten liters whereas in panel B we showed up to 20 liters, and we can see that for – when looking at average hemoglobin oxygen saturation versus oxygen delivery, Black patients who had higher SpO2 than White patients have lower oxygen delivery rates.

And then in panel D, we again are looking at oxygen delivery rates on the Y axis, but on the X axis is SpO2 pulse oximeter instead of SaO2, or hemoglobin oxygen saturation from the

blood gas. And here we don't see a significant difference across races. And what this suggests is that patients are being treated according to SpO2 readings, so it's not that a doctor is saying, well, this patient is White, so I'm going to give them two liters of oxygen and they're Black, so I'm going to give them one liter of oxygen; but it's what these providers are not seeing that leads to these disparities across race.

This table is showing the linear regression results supporting our above findings and we have five models that we share. Model one to model five. Model one has the outcome of SpO2, corresponding to the prior panel A. And this is showing that Asian, Black and Hispanic patients, again, have higher SpO2 than White patients, controlling for hemoglobin oxygen saturation and other confounders.

The next three models, model two to model four, all have the outcome of oxygen delivery rates. And we show that these three minority races all have lower oxygen delivery rates than white patients. In model three we have a similar regression except we include the additional term of the discrepancy between the SpO2 and the SaO2, or the gap between the pulse oximeter and the blood gas, and we see that when we include this the differences according to race are no longer significant. So this explains that this difference between race manifests itself in terms of this gap in performance and there's not some other inherent quality about race that affects the ultimate results.

Model four, then, corresponds to panel D, which is the SpO2 on the X axis and oxygen delivery on the Y axis. And again, we show that when we're titrating oxygen according to SpO2, which is generally what's done, differences according to race are not significant. And model five is what's called a falsification test where we use a totally different variable, instead of oxygen delivery rates, in this case Heparin rates that show our findings are not chance findings, and with

Heparin rates we would not expect there to be a difference between races there, and in fact there is not a difference between races and that's what we saw here.

We then conducted a mediation analysis, which is a type of simulation to show that the discrepancy between SpO2 and SaO2 mediates the difference according to race we're seeing. In this here ACME is the Average Causal Mediation Effect and the significant P value less than 0.05 shows that this gap does mediate the differences in oxygen delivery that we see according to race.

So these are our conclusions. First, we confirmed the previously reported pulse oximeter disparities according to race and ethnicity. We show that minority patients received less supplemental oxygen than White patients for a given hemoglobin oxygen saturation, but not for a given pulse oximeter reading. And then we showed that the gap between the pulse oximeter reading, the SpO2, and the hemoglobin oxygen saturation mediates the observed racial disparities.

Limitations include that this is from a single center. It only includes data on nasal cannula oxygen delivery and not other forms of support. There is no minute to minute temporal matching of blood gas and pulse oximeter and that race is self-reported and that's imperfect.

With that I can take any questions. Thank you very much.

DR. PHIL BICKLER

Dr. Bickler: I'm Dr. Phil Bickler representing The EquiOx Study Group, at the University of California at San Francisco. The EquiOx Study is a prospective clinical study of pulse oximeter errors in hospitalized patients. Today I will provide the background for the study and the trial design. My coinvestigators are Dr. Michael Lipnick, Dr. Carolyn Hendrickson, and Dr. John Feiner, we all hail from the UCSF Hypoxia Research Laboratory at UCSF And our

partners include The Center for Excellence in Regulatory Science and Innovation at UCSF and Stanford, UCF Hypoxemia Lab, Health Equity and Surgery in Anesthesia and the U.S. Food & Drug Administration.

Why are we doing the EquiOx study, EquiOx is in part a response to five retrospective clinical studies showing apparent racial bias in pulse oximetry . In each pulse oximeters were reading in the 90s in black patients when their oxygen level in the blood was actually in the 80s, missing a diagnosis of hypoxia, or occult hypoxemia. Data from the first of these studies at the University of Michigan is seen here. I must say that in some quarters such results were explained away by limitations of retrospective trials including imprecise pairing of saturation measured by the pulse oximeter in blood, hemoximeter set to fractional and not functional saturation, self-reported race without skin pigmentation data, and the presence of interfering piments, anemia, low perfusion and motion. However, this bias albeit smaller was described in the 1990s and detailed by our work from UCSF in 2005 and 2007.

Pulse oximeters are a marvelous invention but they require calibration involving healthy subjects, arterial catheterization and controlled hypoxia. Conditions only achieved in a handful of labs in the world. In lab testing pulse oximeters can be calibrated quite accurately and perform well across all skin pigment types. As in the analysis of 4,000 paired samples from Clinimark, and in collaboration with Paul Batchelder, which we presented in a paper by Okinoa and her UCSF colleagues earlier this year.

So in these data we can identify a gap between the real- world and the lab. And it has been our mission at UCSF to identify what is amplifying this gap. Notice that I said "amplify" not create this gap has been seen before.

Next I want to tell you about results from a new UCSF study that goes a long way toward explaining this gap between testing in the lab and in the hospital. We examined in detail the performance of () and five nine five and radical seven. Oximeters which are in widespread use in North America, Europe and the Asia Pacific region. We put together 9,800 data pairs that match hemoximeter values and corresponding pulse oximeter readings in the Nelcor and Masimo and we subjected the data to a multivariable model including skin class measured by Fitzpatrick perfusion index, hypoxia, sex and age and we segregated these data according to the perfusion, and perfusion index was chosen by a sensitivity analysis and its relationship to bias and furthermore corresponds to our clinical experience with these devices.

The study that I'm going to describe has just been posted on med archive so you can refer to it for the details. It's the first author is () and the study has been critical to our design of the EquiOx study. The results of the multivariable analysis were striking. The bias in both the Nelcor and Maximo were highly associated with skin pigment, perfusion and hypoxia. Let me explain these complex plots.

On the horizontal axis in each we have the arterial oxygen saturation expanding the entire region we studied from 68% up to 100%. The vertical axis in each are the bias in the oximeter, that's the difference in the reading between the pulse oximeter and the true value in the blood sample. The data are segregated by skin category, so light, medium and dark, and also importantly by the perfusion index, the red points indicate perfusion index greater than 2%, the cyan color is 1 to 2% and the X marks low perfusion where the perfusion index is less than 1% and you can see by simple visual inspection how in the darkly pigmented group there is a much wider spread of data and the particular outliers on the high side are represented by data points with low perfusion index.

The relationship between perfusion level and bias was also striking. Here we plot perfusion level on the horizontal axis and the bias on the vertical. And in the darkly pigmented subjects you can see the distribution of bias is clearly in the positive range, where it is more neutral on the lighter subjects and intermediate in those with intermediately pigmented skin. Indeed low perfusion explains the misdiagnosis of hypoxemia in our data. We are convinced that low perfusion is the key factor that is amplifying bias in those with dark skin and low perfusion.

Here hemoxcimeter readings and saturation values are compared showing that in dark subjects the rate of misdiagnosis of hypoxemia in healthy subjects in the lab environment are up to 30%. It was 30% in the Masimo and 8% in the Nelcor.

With that as a foundation, the EquiOx study is designed to address the most pressing questions related to pulse oximeter accuracy in a real- world setting, our prospective study includes synchronous paired samples so we have a tight knowledge of what the oximeter is reading at the time the blood sample is taken, functional saturation is measured, what the pulse oximeter sees, skin pigment is quantified by the Fitzpatrick chart, the Vonlution chart and by the Monk scale as well as quantified by the Conika, Minolta -- perfusion is quantified, optical signals are measured. We have an inclusive network of collaborators, stakeholders and statistician and our target population is one that is representative of a full range of skin pigmentation. Dr. Carolyn Hendrickson is the clinical lead at San Francisco General Hospital.

The study designs includes the primary aim of measuring bias in pulse oximeter performance in hypoxemic patients across varying I didn't think skin pigmentation. Our secondary aims include determining what skin pigmentation metrics best correlate with pulse oximeter bias, determining if pulse oximeters perform in clinical use at regulatory standard level.

And third, we will test if low perfusion state as a possible mediator that might explain differences in oximeter performance across categories of different skin pigmentation.

Our study timeline is rapid. We began this study in August. We obtained IRB approval, we have National Clinical Trial Registration. We had 55 patients enrolled as of now. Our interim data analysis with the aid of our FDA partners will be in November 2022. We're considering possible site expansion depending on enrollment numbers and data distribution and we anticipate that the study will be completed in July of 2023.

All our data is being posted on the Open Oximetry Project website, and we want to thank our collaborators, CERSI at UCSF Stanford and UCSF hypoxia lab as well as the U.S. Food & Drug Administration, the Gordon and Betty Moore foundation, Patrick J. McGovern Foundation for essential help with the project.

Finally, I want to acknowledge my enormously talented research group. Without which we would not be able to tackle this task. They have been incredibly helpful as we initiate the project and will curate forward. So special thanks to all of them. Thank you very much.

DR. CHRISTOPHER ALMOND

Dr. Almond: Hi. It's a tremendous pleasure to be here, my name is Chris Almond, I'm from Stanford University and I work in the Division of Pediatric Cardiology at Children's Hospital. I'm going to take a few minutes and just talk about the launch of a prospective clinical study to evaluate the accuracy of pulse oximeters in children. I should mention that none of the study investigators has any relevant financial disclosures.

Just to introduce some of those investigators, this includes, Desiree Conrad, who is a 3rd year cardiac fellow with us, two cardiac anesthesiologist, Chandra Ramamoorthy, and Manchula Navaratnum, Rebecca Kameny, is in the cardia ICU, and also leads our DEI initiative. Greg

Adamson in the cath lab, Rohan Teneja is our study coordinator. This began with a conversation with Michelle Tarver and we consider her to be part of our team as well.

So I'll briefly review the motivation behind the pediatric study to describe its design and finish up with reasons why this project may be a little more challenging to conduct in children relative to adults.

So I think as has been mentioned before, this is also true in children that pulse oximeter is used widely to determine whether children are adequately oxygenated but recent studies suggest that pulse oximeter systematically overstates the true oxygen saturation in children with darker skin pigment. Now, this error puts children with darker skin pigment at considerable risk for health disparities by failing to detect important levels of hypoxemia that may drive critical treatment decisions like COVID medication usage, hospital admission, ICU transfer, intubation and even consideration for ECMO or life support therapies. And prior studies have understandable limitations related to the retrospective design including using race and ethnicity as a proxy for skin pigmentation. O2 saturations that appear to be in the electronic health record but might not be neither simultaneous or performed at study state and it isn't very good at analyzing technical factors like motion artifact and perfusion that could also affect the pulse oximetry values.

As has been discussed there are a number of studies in adults that have rarely looked at this especially in the COVID era and probably the best pediatric study was conducted by Erica Andrews at the University of Michigan who looked at it in a thousand children and found similar findings, the discrepancy in the accuracy in patients with darker skin pigment.

The primary rational for the pediatric study is to overcome these limitations with a carefully designed prospective study and determine if hospital grade pulse oximeters perform outside FDA specifications in children with darker skin pigment.

So the design in the pediatric study actually is parallels the adult study with a primary hypothesis that the mean bias or error in FDA- cleared pulse oximeters increases with increased skin pigmentation as measured by colorimetry and standard pigmentation scales, these most notably the Fitzpatrick and the Vonlution. We'll look at other pediatric specific factors like age and the presence of cyanotic heart disease and important subset of children in how it influences the degree of error.

So this is a prospective single arm clinical trial in children in a real- world setting which is Lucille Packard Children's Hospital. And the total sample size is 154 subjects.

Children will be eligible for inclusion in the study if they are up to and including the age of 21, which is the definition of pediatrics at the FDA's center for devices and radiologic health. They must have already have an arterial vascular access line as part of their routine clinical care and they must not be anemic, have hemoglobinemia or a skin condition that could complicate probe attachment or interpretation and they cannot be non-pulsatile, this is relevant for our heart failure patients on left on a ventricular device.

Patients will be recruited generally in the clinic a day prior to their scheduled cath procedure. They'll provide written informed consent and for older kids there will be an opportunity to provide assent for a participation. Skin pigmentation will be measured three different ways at the potential site of pulse oximeter, this includes colorimeter, and two pigmentation scales, and this will be performed at up to five to eight different locations where the probe may be affixed. We will then conduct five paired measurements of SpO2 and SaO2

obtained simultaneously under steady state conditions and technical factors such as perfusion and temperature will be captured at that time.

The total sample size for the study is 154 subjects, you can see the assumptions listed here, perhaps most importantly the sample size depends on the interclass correlation co efficient of the cluster measurements within a patient. And since this is something, the sample size can vary considerably, the plan is to actually start this study, verify that the ICC is roughly around 0.15, and if needed, adjust the sample size either upward or downward based on the preliminary results. So to this extent the study has somewhat of an adaptive component to it.

We've been lucky with our great team in the collaboration with FDA and UCSF to have been able to get this study almost ready to be up and running. We received the notice of a grant award from the FDA in August. We had Stanford IRB approval in September. We have a study coordinator that has now been hired, we had two site visits with UCSF both to the hypoxemia lab and to San Francisco General where the study is enrolling, and we are nearly ready to enroll.

Lastly I wanted to touch briefly on a few reasons that this project may be harder to conduct in children so we need to be prepared. The first is that unlike adults, pediatric diseases in general are relatively rare, especially including children who require arterial lines. And this could complicate enrollment. Now we think we have more than enough patients at Stanford to enroll the study, even if we need to augment the sample size, but it's something that we want to be aware of.

Secondly, there's quite a bit of color variability in children, most of which is related to vasa mode motor reactivity that's normal. So children and especially infants and newborns can be very red. Perhaps someone has seen a, independent who's crying or is stooling, and they can turn a very beat red, which is not in the volution or Fitzpatrick scales.

They can also be muddled, so a kind of a variable skin tone. And they can also be pale, depending upon their state. And this could complicate things somewhat. It's also very age-appropriate for infants and toddlers not to follow instructions.

So for example, in adult study, you might be able to say place hold your hand steady or steadily while we make careful measurements to make ensure reproducibility at steady state, and of course this can't really be done in children

And then cyanotic heart disease which represents an important sub-set of children and probably introduces another set of variability that can increase the degree of bias. Just to highlight this group, many of the clinicians are familiar, but these are children born with cyanotic heart disease and typically have saturation in the 70's and 80s rather than 90s where the most set probes have been well validated. If this is typically caused by structural shunt inside the heart that redirects blue blood out to the body rather than sending it to it lungs where it gets oxygenated.

On a historical note the first blue baby surgery to redirect blood back to the lungs was initially performed at Johns Hopkins University and is credited to Dr. Alfred Blalock and Helen Taussig and later credited to Dr. Vivian Thomas who was a research lab assistant who actually perfected the technique. Side note, the story has been very well retold in the 2004 movie "Something the Lord made" which I would recommend to people for review. So at Stanford, we conducted two studies in hypoxemic children looking at the accuracy of pulse oximeters and found that FDA pulse Oximeters overestimate the arteriole saturation in children with cyanotic heart disease very much in the way that's purported to occur in children with darker skin pigment. So there may be couple of layers that we see here, which needs to be trouble shot in this population.

So in conclusion, we suspect the problem of Pulse Oximetry errors may be as prevalent in children as it is in adults and this study will help to confirm it using perspective study data collection. We are working closely with FDA and the UCSF teams which has really been wonderful in helping us get up and running as quickly as possible to ensure our methodologies are harmonized so we can pull our data across age groups and for analysis and make them available to the public for secondary analysis.

And we do acknowledge several factors may make the study harder to perform in children related to cooperativity vasomotor reactivity and cyanotic heart disease This may require an adjustment to the sample size population and perhaps some troubleshooting of methodology along the way. But we have a great team and are certainly prepared for that.

So I want to thank the U.S. Food and Drug Administration, especially Nilsa and Michelle, for their wonderful support and the rest of the team. And then the UCSF group that has been exceptionally helpful in getting us up and running and we really credit this study being where it is really because of their support through the UCSF Stanford.. So thank you very much.

DR. MICHAEL LIPNICK

Dr. Lipnick: Thank you for the opportunity to speak today. I'm Michael Lipnick and I'm an anesthesiologist at UCSF, I am also co-director of our World Health Organization Collaborating Center, for Emergency Critical and Operative Care, as well as our Center for Health Equity in Surgery and Anesthesia.

And I'm also an investigator in the UCSF Hypoxia Lab, which I will be sharing a bit more about. Most of the focus of my talk today is going to be about a new project which we hope will accelerate and support improvement in Pulse Oximetry performance, especially as it relates

to performance bias due to dark skin pigment. And at the end of the talk, I'll invite collaboration and share information about how to do so.

I have no conflict to disclose. First, just a brief background. Our lab was founded by John Severinghouse and has been doing healthy human study subject of pulse oximetry performance for 50 years. And we studied thousands of subjects and thousands of devices. Most of the focus historically has been on device development and performance verification testing. Though over the past 10 years and especially the past few years, we've been working with number of organizations operating at the intersection of pulse oximetry and health equity.

Through Pulse Oximetry, there are many opportunities to improve patient safety and reduce health and healthcare disparities not only in the U.S. but worldwide. Global scale access remains an issue and is worth noting. Not just access to any device but safe and reliable device. We work with collaborators like Lifebox, the WFSA as well as leading aid agencies and global health authorities to advocate, explore ways to improve access to pulse oximeters as an essential patient safety tool.

More recently the impact of oximeter and more specifically poor performance in populations with dark skin pigment has become a focus for us. Issue of performance and impact on health outcomes have likely been covered adequately by other speakers so I wanted to share couple of statistics highlighting the for potential of oximetry to improve safety in this issue of access. About 30 million surgeries are performed worldwide without pulse oximeters in 70 to 100,000, operating rooms and intensive care unit beds worldwide, to not have pulse oximetry and for a large proportion of the world’s population access to pulse oximetry is not there in primary care settings.

With all this in mind approach the problem concerning all of these objectives, trying to find a balanced approach that maximizes safety for all communities. But with special attention to populations at most risk. Now, during the COVID-19 pandemic, our team was inundated with inquiries to the lab which made us double down in efforts and look for new strategies to accelerate work in this area and hear are a few examples.

The skin pigment effect seems small. Does this make a clinical difference and should I be concerned? What methods are you using for skin pigment quantification in lab and what is the best, what is the most reliable and pros and cons of each technique. Do you have independent data that you could share, raw data on device performance? Have you tested certain devices, fingertip devices, hand-held devices can you share the results? From another team, we figured out how to solve the issue of performance of bias in dark skin but we need help verifying it, can you help. And then finally, this example during the biggest surge in south Asia, NGOs looking to donate a large number of units and asking if we could test several units rapidly for donation.

One of the devices in the center, you can see setting in setting in the mid-80s when the real sat was 70%. This poor performance was very concerning, but more concerning was the uncertainty around how this would be, this information would be utilized, how it would be shared with the communities receiving the donations and whether or not it would be impactful and changing the donors mind.

So collectively all these experiences highlighted to us the need to explore better ways to share data we have. Figure out how we can open up our lab infrastructure and make it more accessible for other teams. And also to cross-pollinate with other fields and disciplines that were approaching this problem in different ways. So this was the Genesis of the Open Oximetry project which we are launching thanks to multiple supporters including Catalytic Investment

from the Gordon and Betty Moore Foundation, and the Patrick J. McGovern Foundation. The project has six components, and we're actively building collaborations in each of the core domains. I'll walk through these going clock wise from the top left. First is real world clinical perspective studies. EquiOx is the first of these which you have already heard about today ,this is supported by the FDA Cersi and is a real-world clinical study in San Francisco General Hospital in adult intensive care unit patients, trying to understand the magnitude of potential bias as well as potential root causes of this bias.

In the lab we're exploring new protocols to better account for factors like skin pigment, low profusion and severe anemia and other factors that we know, things we hypothesize that if better accounted for in device development and regulatory guidance could potentially improve device safety as well as equity. We're also working on new techniques in vitro techniques, phantoms and device testing, products that could help us identify in less resource intensive ways, poor performing devices and potentially some of the root causes for those.

Now for these three core domains I've shared, all the protocols we are also sharing online in co-developing with other collaborators. We hope to be able to use the data collected and the tests to work with regulatory bodies and to test proposed frameworks and better understand the impacts. The final two components of the project, and a large part of why I am talking today, is our plan to perform, or rather to create a mechanism for collaboration and an open data repository both attempts to harmonize efforts and create transparency around pulse oximetry safety and try to accelerate progress on improving safety. The next two slides before I wrap-up I'll go into this in a bit more detail.

So this network we're trying to build for collaboration really is a mechanism for those who are already working in this space or perhaps starting to work in this space, of improving

pulse oximetry safety or improve access to Pulse Oximetry with an emphasis on the issue at hand which has to do with performance bias and skin pigment. This will be a forum that meets regularly. We recognize that there are many efforts working in this space. Unfortunately, often in parallel and we're actively trying to figure out how we can leverage the strength and resources of other ongoing efforts to prevent unnecessary duplication moving forward.

The forum is going to include international, multi-stakeholder engagement. This includes diverse disciplines, geographies, industries, viewpoints, and populations. Again, trying to avoid duplicating but rather be additive and complementary with existing groups. We're already actively working closely with the leads for the ISO and IEC working group on oximetry performance.

Some preliminary objectives for this forum include developing a research agenda for both regulatory and technology standpoints. We hope to not only identify but also test them and view data in real-time with the group.

As well as to account for international perspective and attention to communities at risk in the U.S. and as well as globally. The group hopefully will account for total product lifecycle, design, distribution, utilization. This would include not just the performance, but also procurement guidance as well as education and communication and implementation guidance. We recognize there are number of entities working in that space and we are again actively working trying to find ways to work together across these different groups.

And then finally, promoting clear and consistent communication, not only in the education piece and advocacy pieces, but also as it relates to data collection, definitions, best practices for data. And doing all this in a way that doesn't unnecessarily erode confidence in this essential patient safety tool.

This forum will have its first of regular meetings in December of 2022. Through this QR Code or through the open oximetry website, you can sign up or forward this to other groups or individuals who might be interested and we welcome participation from anyone. So please check that out and share.

The final piece of the project I'll be sharing about, again, in this effort to leverage multiple collaboration and stakeholders is data sharing. We're building a couple of different mechanisms to try and promote data sharing. One is an online visual summary that would be more consumer-oriented. This would show up device data for large number of devices that will be testing that's already begin in the lab and will continue for next two years.

The second is a raw database. This will be raw performance data. Again, we'll be collecting these data in our real world clinical trial infrastructure, and also in the laboratory infrastructure. And putting all this into one dataset where individuals who are doing research and development can access the raw data and query for study questions that they want to ask. We also invite other researchers to contribute to both of these datasets and please reach out and we can share more information about how to do so.

Note the current site is a demo site, and although we'll be adding and updating to it continuously, full new version will be launched in 2023. With that I'll conclude. I's like to thank multiple supporters as well as the numerous collaborators at UCSF and multiple organizations around the world who are contributing to this work. Thank you very much.

CLARIFYING QUESTIONS FROM PANEL TO INVITED REAL-WORLD EVIDENCE SPEAKERS

Dr. Nathan: Thank you Dr. Lipnick and speakers and outstanding talk to everyone. I like to open up to the panel to see if anyone has clarification questions they would like to address to the speakers. We have about 15 minutes for this.

I know once we get going, people will start asking questions. I'll ask the first couple. I was taking notes or questions as I was going. Question to Dr. Wong. And I guess Dr. Sjoding can answer it. For this entity or hidden hypoxia with increased mortality rate was 70% higher mortality rate. Who was the reference group? Was it other patients who were not hypoxic or was it patients who were hypoxic but it was evident? I think that becomes important, because I don't know if the hidden hypoxia is just a surrogate for disease severity or if the inference is, if it was treated, then it would have made a difference in the outcome?

Dr. Wong: Thank you so much Dr. Nathan for your question. With regards to those, the comparison of hidden hypoxemia, the reference was – so all patients had a pulse oximetry value that was greater than 88% as you well understand, but the reference group were the intervention group or the case group with a hidden hypoxemia, so those with an arterial set point below 88. The reference was those with both a arteriole sat and a pulse ox sat that were both above 80 as the comparator.

Dr. Nathan: Because I guess that gets to my point that these are sicker patients and you just didn't know about it. It would have been interesting to compare the mortality to other patients who were hypoxic where it was known to make the next assessment or the next inference that, if they had been treated for their hypoxia they would have had a different

outcome. Maybe all you can say at this point, is they had a worse prognosis that was unknown, because their OT set was falsely elevated.

Dr. Wong: Correct. So I believe that perhaps Dr. Gottlieb or others who were perhaps working on causal inference techniques as we and many understand, this is an inherent limitation to retrospective analysis. But the best we can do I believe is through cause [indiscernible] techniques as I'm sure he's willing to talk about.

Dr. Nathan: That doesn't diminish the importance and the impact of the study. But thank you for clarifying that. Okay, Dr. Yarmus, you're next.

Dr. Yarmus: I think this is probably directed at Dr. Almond, but this is open for anyone, as an intensivist, I think most of this has been focused, on critically ill patients given probably the access to arteriole blood gas monitoring. Clearly there's historical data out there like the study that Dr. Almond quoted from Harris for looks like outpatient oximetry monitoring in less severe illness patients.

I'm just curious, probably for everybody, you know, are there working solutions out there where, for example, in that study you have outpatients being monitored, so has there been workarounds or any types of parameters that clinically has been utilized to help sort of adjudicate some of these issues? And that also transcends beyond other diseases like sleep medicine for example, has been recognized previously and are there any other efforts for that patient population, interstitial lung disease, 6-minute walk testing, all those geared towards outpatient monitoring program?

Dr. Almond: Thank you very much for that question. I think most of the attention so far has been focused on prescribing the scope of the problem. Like many things, we're always behind a little bit in the adult world. I would say that in the study that Brunwin Harris did, three

different oximeters were tested, The Delco Massimo, and one called Massimo Blue, which is appropriate for blue babies. And that one did have relatively good performance, or better performance I should say relative to the two.

So in some ways, there's some sense there might be some oximeters that perform better. You're right about the outpatient oximetry study, which I didn't mention but was looking at the whole monitor and I think it's been pointed out elsewhere, this is a large business in pediatrics with worried families related to SIDS and other things where these over-the-counter machines are used quite frequently. I think as the point was made before, they're not being used kind of for casual reasons, but really for health monitoring.

And the results were relatively similar in that there was still this bias in the bluer kids. And what we're hoping to do in the context of the prospective trial is to sample both those kids that are coming in for elective procedures that are "well" but they have other diseases or heart disease or other reasons that they are getting [indiscernible] as well as patients in the cardiac ICU, and those undergoing cardiac operations. So, hopefully, we'll have [Indiscernible]. And I think some other people and some distinguished speakers have more experience in terms of monitor development. I would defer to them.

Dr. Nathan: Thank you, Dr. Almond. You were a little bit soft there but I think everybody heard when you had to say. Just in case, if you can get closer to the mic, please for next time. Dr. Loeb, please go ahead.

Dr. Loeb: I have a question for Dr. Wong: One of your slides showed the absorption spectra of water and melanin and I wondered if that was accurate and where that spectra came from? Because I noticed that they're quite similar, melanin and water. If that's true, I'm just trying to get my head around the difference between laboratory findings and the clinical studies

and how it can be that melanin makes such a big difference, especially, if it has a similar absorption spectrum to water.

Dr. Wong: Thank you so much for your question. In reference, I believe that was from a Baxter, that was a therapeutic laser reference, one that would actually examine the optical quality of tissues, and that's where that came from. In terms of the reference of melanin to water, I think the question -- so maybe I might ask or rephrase your question. There are effects of melanin and water. And to some extent, we have compensated for water and the presence of water in the design of public symmetry. But one of the differences that we may not always control for is, many people of different – all of the spectra can overlap. The concentration, and as you may differ, and that as the concentration of melanin differs, right? And is I believe inverse to the proportion to the, or the concentration, the Lambert might dictate how much affect it might have on the resulting spectral comparison.

Dr. Loeb: Okay. yeah, certainly concentration makes a difference. I have one follow-up question if that's okay for Dr. Bickler. Because I thought that was very well described, how, again, the lab studies and the clinical studies could have very different results. Profusion and therefore the signal-to-noise ratio, obviously, is very important. And I noticed that's one of the variables that you're looking at. I think that's great. I wonder about the ratio of DC absorption at the two different – at 660 and 90, I'm sure I got the number wrong there. But the red and infrared, DC absorption, if the ratio of those has been looked at as an explanatory factor? And if that could be, you know, put in the model? And Dr. Lipnick, I think the hypoxemia lab probably knows the physics and knows this all works better. But I wonder if that's been looked at as an explanatory factor.

Dr. Bickler: Thank you for that question. Yes, the ratio of the pulsatile and non-pulsatile light signals in the red and infrared forms the basis for pulse oximetry. And as I mentions, all pulse oximeters need to be calibrated in humans in order to translate the optical signals that is seen by the electronics to saturation reading. And you can imagine that id all calibrations procedures are done in subjects with low skin melanin, you produce one marker that would produce a pulse oximeter that would be accurate in individuals with lightly pigmented skin. And what has become apparent, it's been insufficient to account for the presence of melanin. Now, you can do another calibration for subjects with darkly pigmented skin and you would get a different calibration curve. So that is possible.

And almost 20 years ago, we advocated for something like that. But something that's really important to point out from our recent work is that it is more complicated than just skin pigment. Because our work on low profusion shows there's a third factor, which is incredibly important. And that is the profusion index. So if you don't account for that, the simple correction for skin pigment alone would be insufficient. So that represents both, I think a fundamental discovery and also an engineering challenge to incorporate multivariable in the way that pulse oximeters determine the saturation values.

Dr. Nathan: Dr. Collop?

Dr. Collop: Yeah, I want to keep going on that subject if you don't mind. Can you explain, I guess, how you calculate the profusion index? If you said, and I'm sorry if I missed it and I apologize. But I was trying to understand the relationship between dark skin and the profusion. Are you suggesting that people with dark skin have low profusion or is that a complete other variable evaluating the oximeters?

Dr. Bickler: Thank you for the question, that really gets at the heart of what is going on here. So, First of all with respect to understanding profusion index, you can think of it as pulsatility of the blood. Because the pulse oximeter requires seeing pulsatile blood in order to estimate saturation and it's based on ratio of ratio of pulsatile and non-pulsatile blood, at the red an infrared wavelengths. If there's no pulsatility, it can't determine that ratio and it will produce a widely inaccurate estimation of saturation. Now, skin pigment – melanin absorbs much more red light than infrared. So the effect on that ratio of ratio is to shift the ratio to lower values and that means higher saturation. So you can understand the absorption how the absorption of red light would shift the calibration curve to the right produce abnormally and inaccurately high saturation value. So that's the essence of it. So think of perfusion index as just the amount of pulsatility of blood in your finger. Now think about sick hospital patients, they don't have normal perfusion. They are presenting will all sorts of illnesses that make them have less pulsatility. They could have low blood pressure, they could be in shock or dehydrated or critically ill. There are numerous factors which produce decreases in profusion in hospitalized patients, especially those in the ICU. So it's easy to imagine that in a population like that pulse oximeters would face much greater challenges to reading accurately.

In fact, you can imagine how that would bias their readings toward artificially high saturations. It's informative to think about some of the other retrospective clinical studies. We heard from Dr. Burnett at the University of Michigan. He studied patients in the operating room. It's our clinical experience that anesthetized patient's actually have higher profusion because of vasodilating effects of the anesthetics. What he found was the errors were intermediate. They weren't as grave as identified by Dr. Sjoding, Dr. Fawzy and Dr. Wong. And that fits with the idea that profusion is a critical factor in performance oximeters.

Dr. Nathan: Thank you, Joe O'Brien.

Dr. O'Brien: Yes, Thank you very much, Joe O'Brien, Patient Representative National Scoliosis Foundation: Another question for Dr. Bickler following up. As I said earlier, ever since there was only 5% melanin in the finger, I was very confused as to what the causative factor was with skin pigmentation. And as I listen to your presentation as we discussed, I was thinking "okay, profusion, that makes sense, that may resolve it." One of the studies I read, I was just wondering if you can address it was Hisan, and his group in the Netherlands and they looked at critically ill people in ICU and looked at profusion index versus the gap between what we're talking about. And they found there was no correlation between it, that they could not use it for any indication whatsoever. So I was wondering if you can address that?

Dr. Bickler: I'm sorry, I'm not familiar with that study. But I would say in general, that low profusion produces errors in pulse oximeter. Anybody who uses them clinically knows that. And clinicians often have to change the site where we're performing the oximetry measurement. We often use the ears or nose or other places on the body if the oximeter is not producing a good signal from the finger. So this is a common part of the clinical experience with pulse oximeters in general.

Dr. O'Brien: Okay, thank you.

Dr. Nathan: That was going to lead into my question actually, in the context of the EquiOx study, are you looking at just one site, or are you comparing sites, like a finger probe versus an ear probe, and does that help with the profusion aspect of this and I don't know an individual patients, the profusion might be a little bit different?

Dr. Bickler: That's right. The EquiOx study is designed to capture whatever is the prevailing profusion index. And some of the patients that we've already enrolled have ear

oximeter because the oximeter has not been working well in the fingers. In the course of this study we're going to capture different sites, and different levels of profusion and we will incorporate that into our analysis.

Now that complicates matters statistically. Because our study design is, you know, confined in time and space. And the FDA wants an answer from us by next summer. And I think that some of these detailed questions like the one you posed perhaps will not be answered in that frame of time. We'll have some ideas, but probably no firm conclusions.

Dr. Nathan: Wouldn't it be easiest if you just try to get 3 readings on the same patients once you have them consented? That might enable to answer additional questions rather than one patient's ear, and another patient's finger.

Dr. Bickler: We're actually doing that. We're primarily focusing on standard clinical care. So we're not walking into the patient's room saying, oh, can you please switch the oximeter? Because what we want is a real world clinical study. We want to understand what's happening in real clinical environments without that kind of interference from the study.

Dr. Nathan: Sorry to keep you on the hot seat, but being that it's real world between the two centers primarily, is it a going to be a standardized pulse ox that's going to be used, or it's whatever on the patient?

Dr. Bickler: It's whatever is on the patient. And this is one of our goals in expanding clinical sites where these kinds of studies are done. We're sharing our protocol on Open Oximetry, and we're inviting anyone in the world to adopt these protocols and incorporate their specific patient populations and clinical environments into the information that will really benefit a more global understanding than just two centers, like you've said.

Dr. Nathan: I think that's valuable. Dr. O'Connor.

Dr. O'Connor: Yeah, I want to, again, a question for Dr. Bickler. There's literature that suggests for example, there may be difference in accuracy between the finger and the ear and people with heart failure. And so you can imagine the heart failure is more prevalent in some populations than others. So I'm going to encourage you to prospectively get much data as you can, not just about the patient's demographics, but also their comorbidities. Because particularly with the heart failure, it's a small study. But it's a real signal, or seems to be a real signal. And it would be worth exploring on a larger-scale. Just like the profusion index, if we can say, if your patient has a heart failure, put it on the ear. That would be very useful guidance to providers like me.

Dr. Bickler: Thank you Dr. O'Connor. That's an excellent point. As you can imagine, we are recording a wealth of demographic and medical information about all the patients in our study. I think maybe Dr. Almond might want to comment about that he has a population with a high degree of cardiac problems and he alluded to some of the challenges, maybe he can provide some additional comments.

Dr. Nathan: Sure, Dr. Almond, if you want to weigh in.

Dr. Almond: Yeah, I think we do find in pediatrics that we need to oftentimes move the probe around. A lot of the kids that we have that are blue, some of them are quite healthy running around the clinic and at school. And then others are critically ill in the ICU with low cardiac output or other profusion problems. So that is such a common thing to do in pediatrics to basically rotate the probe around until you find the signal. And I can't say we've seen that especially true in heart failure patients, pediatric heart failure patients, but that certainly is a common practice. Thank you.

Dr. Nathan: Dr. Klein.

Dr. Klein: Yeah, I have more of a comment than a question. Really for Dr. Almond, as we talked a lot today, we talked a lot about adults if little bit about neonate, So I wanted to thank you for the pediatric perspective as a fellow pediatric person and thinking about those kids with congenital sciatic heart disease and a patient population that probably has skin thickness somewhere between the neonate and the adult, and all these confounding pieces. So I just wanted to say I look forward to the results of your pediatric work.

Dr. Nathan: Thank you Dr. Klein. I just have one additional comm, I'll give other folks time to think if they have any last questions they want to ask. As you're talking about the pulse ox flow, it's not an aspect of pulse oximeter fully understood or appreciated, but there are populations, what has been eluded to already are heart failure patients with vasodilator and we have patients with pulmonary hypertension who are on very potent vasodilators and we base therapy base on the pulse oximetry reading sometimes. And in other population of interest would be COVID patients, with left ventricular assist devices, who don't have any pulsatile flow. And I am guessing, I hadn't thought about this you can't get a pulse oximetry on those patients in?

Dr. Bickler: That's right. You're generally out of luck. Sometimes you'll get a little bit of a signal on LVAD patients, but I would most of the time I think they're highly unreliable.

Dr. Nathan: And even patients who are on VA ECMO, depending on how much, pulsatile activity there is, I imagine that effects the signal as well, the sickest of the sick patients.

Dr. Bickler: That's right.

Dr. Nathan: All right, Dr. O'Connor.

Dr. O'Connor: With the LVAD patients if the blood pressure cuff works the pulse ox is probably not bad if you can't get a blood pressure on them, then you should take the pulse ox off the patients. That's how we take care of these. And I want to ask Dr. Almond a question since I

was inspired by the other question about pediatrics. What are you going to do with hemoglobin F in your very young patients? Is there a way to deal you are going to deal with that in your data collection?

Dr. Almond: Yeah, that is a great question. My general sense is that most of the studies that have been out, that are out there have not found at the bias that seems worse in the population. A lot of the kids we see do have significant congenital heart disease so their saturations are abnormal, especially in that newborn period. And we're just launching the study and it may be that we want to consider collecting information about hemoglobin F. I think at the same time, most of the cardiac anesthesiologist here have that they just haven't really empirical observed the issue related to that. So that's something we definitely should take under consideration.

Dr. Nathan: Dr. Kopotic. you can't ask a question as a speaker, but you if want to pass a comment in response to one of the other questions please go ahead.

Dr. Kopotic: Yes, it was related to the neonatal issues where preterm may have some level of fetal hemoglobin. That was well investigated in the mid-80s and shown not to have an effect. However, in Dr. Almond's condition where there's cyanotic heart defect. Polycythemia can be an issue, and with polycythemia there is a reduced flow to peripheral sites. So, some of your challenges in placing a sensor and getting an adequate value or at least a signal sufficient to give you a derived saturation may be associated with that poor peripheral profusion much like a sickle-cell child not being able to get hemoglobin to flow well because it is so thick in the polycythemia state.

And just to follow to the ECMO situation, having been an ECMO coordinator at a few hospitals, it is very much the case in VA ECMO that you will not get an adequate signal, and

again, no saturation can be derived by a pulse oximeter. But that's also the case on any patient that's on cardiopulmonary bypass in the operating theater.

Dr. Nathan: Thank you, Dr. Hennessy, we'll make yours the last question before we close out the session.

Dr. Hennessy: Thank you. This is question for Dr. Sjoding. I wonder if any higher rates of occult hypoxemia in Black patients compares to white patients is due to a higher presence of hypoxemia in Black patients versus White patients or if it's all due to the pulse oximeter functioning worse in that population?

Dr. Sjoding: Thank you for that important question. Which was also discussed I heard earlier among the statistical analysis. So the occult hypoxemia rate, is also known as false emission rate can be influenced by the prevalence of hypoxemia, and so if as a population, a group of patients had a lower or higher prevalence of hypoxemia that would affect the results.

That's one of the reasons why we did the additional analysis by looking at receiver operating characteristic curves showing there was still a signal. I would want to make one other point about this, which is we heard from others that in hospitalized patients, the rate of pulse oximeters, SpO2 being low is extremely rare. So once a patient comes to the hospital, there are – the oxygen saturation is immediately titrated to get the pulse oximeter to a normal range. And we heard from Dr. Gottlieb that everyone, titrates oxygen to pulse oximeter So what it means then is that if there is residual occult hypoxemia, it's probably related to the fact that we're not actually titrating oxygen to the SaO2, which is low, we're titrating oxygen to the SpO2. I think in general, that's the more likely mechanism here rather than of difference in baseline prevalence of hypoxemia.

Dr. Hennessy: Can I ask one follow-up question?

Dr. Nathan: Sure.

Dr. Hennessy: You noted there was a lower use of arteriole blood gas in Black patients and White patients. Do you have a hypothesis as to why that was?

Dr. Sjoding: Yes, that was one of my colleagues. I can't remember who that was.

Dr. Hennessy: I'm sorry.

Dr. Nathan: That was Dr. Wong.

Dr. Hennessy: Okay, Dr. Wong?

Dr. Wong: Yes, thank you. So we were unable to -- we were unable to distinguish why this happened. So rather in short, in terms of adjusting for the risk of – for example, like SOFA score dysfunction for other patient characteristics, we were still unable to show difference in the underlying population. So in short, we don't have a clinical reason as to why this happened that would otherwise explain it besides some of the other factors we have here.

Dr. Hennessy: Thank you.

Dr. Nathan: Okay. Thank you. Dr. Gottlieb, you may answer any of the questions posed by the panel but please don't add any new questions. You're muted. Was your hand up inadvertently?

Dr. Gottlieb: I'm sorry, thank you. Yeah, I just wanted to follow-up on few of the comments that were just made in response to Mr. Hennessy and Dr. Sjoding comments. We also showed in our table 1, and this is somewhat generalized data, because it's such a just a single median with Wilcoxon's test. We did show overall lower SaO2s and higher – so Black patients had lower SaO2 than White patients, but had higher SpO2.

So overall, I don't think the main effect is the Black patients were more hypoxemic. And one other thing I wanted to mention in response to what Dr. Sjoding had said was in terms of

oxygen titration. There's were a couple of papers and some work underway and I don't know if this is actually in the clinic yet, is devices that would auto titrate supplemental oxygen rates according to SpO2. So I think it's really important to stress that with these new sophisticated technologies, if these disparities become very much baked in and could really be amplified with the devices like this.

Dr. Nathan: Okay. Thank you very much. And that brings us to end of this session. It is now 2:44 on my watch. So we'll take a 5 and a half break and let's reconvene at 2:50 p.m. Eastern Time. Thank you.

PANEL DELIBERATIONS/DISCUSSION

Dr. Nathan: Welcome back, everyone. This next session is going to be devoted to panel deliberation. Before we get to that, there was a question this morning that was addressed that hasn't been addressed that I like to ask Dr. Lee from the FDA to answer.

Dr. Lee: Thank you Dr. Nathan, yes, I believe question the question came from Dr. Loeb regarding whether the devices are on market cleared that simultaneously measure skin pigmentation. While devices and LED can adjust the intensity, it does not do that actively. Meaning basically, a finished device that FDA reviewed, the intensity level of the LED can vary. That is something that a medical device developer would choose and make design specifications. But they don't actively adjust due to any perceived differences in pigmentation or other adjustments. I hope that makes sense. Thank you.

Dr. O'Connor: Yeah, that was mine. Thank you, Dr. Lee.

Dr. Nathan: Thank you, Dr. O'Connor.

Dr. Nathan: Thank you, we will now we'll begin the panel deliberations. This portion of the meeting is open to public observers, public attendees may not participate except at the specific request of the panel chair, again, that would be me. And additionally we request all persons asked to speak identify themselves each time. This will help the transcriptionist identify the speakers.

Let's begin. Do any panel members have a question or comment to discuss? Panel, please turn on your video if you don't have them on already and unmute yourself when you are asked to speak, raise your hand, I think everyone has the hang of that with Zoom, and I'll call on you in turn. So we have up to 2 hours for the deliberations. Let's see how it goes and I'm looking forward to a lively discussion from our panelists. Dr. Alam. I see you have your hand up first.

Dr. Alam: My name is Murad Alam, and I have couple of comments to make. I'm a dermatologist, so many unlike many of you, I have limited expertise with using these devices in highly acute settings. But there's several things I might be able to contribute.

First was regarding pigmentation. I want to give background, I'm a dermatologist and I'm also a laser person, I'm the president-elect of the America Society of Laser Medicine and Surgery. I do clinical research in that area. We in fact, have devices here, and I went and looked at one during the breaks. An IPL device which does have a mechanism that's used for skin treatment of red and brown dyspigmentation and it has a mechanism for doing what I think Dr. O'Connor was referring to where it picks up what the person's skin type is and adjusts the filters and the influence of the IPL so it doesn't burn someone for instance, with a darker pigment. Those sort of devices do exist, and they're relatively crude. But I also want to raise the issue of the position of pigmentation. FDA was giving us description of devices that are very precise in measuring pigmentation and being somewhat skeptical about the Fitzpatrick rating scale which is

of course, subjective and somewhat imprecise. But I did want to raise the issue that we might be trying to be measure precisely, at least hypothetically measure precisely something that's inherently not precise.

So I think there's a danger of having a very sensitive instrument for measuring something that varies dramatically in the same individual given the same type of day, given how it's applied. For instance, I'll give you an example. Some Asian people are very photo protected because they're out of the sun a lot. But they might go to the beach one day. And as a consequence, their pigment can vary quite dramatically.

So unless you were having a device that was ascertaining this every time very precisely, I don't – I think you'd have to be cautious in trying to slice something very finely, when, really it's big categories and that's probably as accurate as you can be. So that's one observation.

In terms of the devices, there's were a couple of types considered. One would be something that would have an algorithm. Dr. O'Connor referred to this. Where it would be working the same way, but it would use information about your skin type to somehow adjust the number that was the output number to compensate for that. So I think if you were doing something like that, that might sound a little bit more feasible? I think if you were actually having a device that had different mechanisms for every one with every different skin type based on what little I know how these devices work, that's a very complicated optical imaging device.

I did some optical imaging research too, and to have that in a relatively small box. Some of the stuff we have costs several hundred thousands of dollars. So I'm not sure if they would be very robust devices if they had 15 different laser frequencies embedded in something somewhere and deciding which to use based on fine distinction and pigmentation. I think the algorithmic

mode makes more sense. And I don't think you should slice pigment too finely. I think you'll get into trouble.

The other comment I had and I'll try to be brief was pertaining to the many devices that FDA considers non-medical that are available at CVS or Walgreens. And I think much of the discussion, the research that's been done today and presented today, rather, has focused on devices for use in the hospital for patients who are very ill in some cases, who are on ECMO or LVAZ, et cetera. But we shouldn't forget the very large percentage of the population comes in contact with those inexpensive "non-medical" pulse oximeters that they buy-in department stores and grocery stores, and pharmacies. And they rely on those to even decide whether or not they should present to the physician.

So I think we have to really think about those as well. There's a risk of ignoring them, and there's also a risk of saying, well, those are so bad, let's just stop making them, which I fear that might be throwing the baby out with the bathwater. Because they are provide some information to people, and it might be a matter of figuring out some way to make them not as dangerous. Instead of giving a number, it should give a range, 91 to 96 or something like that.

But some way to also to think about those devices, because while they're not maybe what you all who are very smart are dealing with, a lot of us lower tech people, that is what we deal with. I can speak from personal experience. I have skin of color and I had COVID and that's what I relied on a lot, that is what we have access to. And we need to have more safeguards there, because you are all who are intensivist are highly sophisticated and able to compensate for the errors in your devices in the hospital just based on your own intelligence and experience.

But the average user isn't. And what you don't want to do is either make those devices go away or improve them so much, the common man's devices, that they get out of the reach of the

average person, right? If we made them so much better and each of them cost $1,000, there would be an access issue where people with not be able to get a pulse-ox at CVS anymore.

Dr. Nathan: I think you make a lot of a good points, I want to give other folks a chance to weigh in but, just to pick up on what you said, I think, obviously, an objective assessment of skin pigmentation is key. Your point is very well-taken in terms of complicating it and making it too complex and having different devices for different folks. And I think it's going to work. And what you alluded to, and I'm not sure whether the technology is at some king of internal device that measures or gives some measure and corrects for the pigmentation that it senses.

I think it's also helpful if you have a patient who you put it on the finger and you're not getting a good read, and then you put it on the ear, and you get a good read there. And the pigmentation might not be the same, but if you have some kind of internal correction through available technology, I think that's probably going to be the best way to go.

I don't know if anyone else wants to weigh in on that particular aspect of pulse-oxs and where the future should go. We'll talk about the consumer device as well. That's a whole other can of worms. I think they're out there and you can't stop them being out there. I don't think I have ever bought one over-the-counter, they should imagine they come with a big warning. But like anyone else, you tear open the box and you throw away the warning and you take the pulse-ox out of the box and you're taking the number for what it is. I think it's natural for most people. I don't think there's actually a warning of any sort on the device itself. Anybody else want to weigh in about the pigmentation issue, and how best to go about evaluating that. Dr. O'Connor, I see you next, I don't know if you wanted to stay on topic or wanted to say something else? I think maybe we should try to – Let's stay on topic right now and have everyone say about this.

Who was that? Ms. Edwards, did you want to comment on the pigment issue?

Ms. Edwards: I wanted to comment on the over-the-counter oximeter.

Dr. Nathan: Okay.

Ms. Edwards: So my mom recently passed and she had a heart disease. And she had an oximeter at home. And this was in May. And the morning that I called the ambulance to come and take her to the hospital, I remember that, I don't remember what the oximeter read, but that was not what calls me to call them. It was the blood pressure reading and it was her obviously labored breathing. And we did have the oximeter and we did check her. And I guess my question would be, or my concern would be that, yes, I do believe that over-the-counter oximeters should be re-evaluated too and maybe come with some kind of warning. Or someone suggested earlier, maybe it should be behind the counter in the pharmacy, and explain the use and what the consumer should consider as far as the readings.

Dr. Nathan: Yeah, first of all, I'm sorry about your loss. It's always very difficult losing a parent. So my condolences to you on that. You bring up another very important aspect that low oxygen is not any cause or shortness of breath. And I think for the consumer, they shouldn't get freaked out by thinking I'm short of breath and my saturation reading is 92%, therefore I'm okay. There could be many other things going on that require a trip to the hospital. So I'm not sure if consumers are lulled into a false sense of security by that. Thank you for that comment. All right, Dr. O'Connor. Over to you.

Dr. O'Connor: I'm going to summarize a lot of what I think for the group here. First of all I think that the ARMS, of 3% is too permissive. We should instruct the FDA to think about using an ARMS of 2% or 1.5%. Second thing is, it sounds silly, but many of the studies done to validate these devices all the data is north of 92% saturation, which is frankly a domain in which the differences make no difference with respect to clinical decision-making. So I'm going to say

that they need to have a statistically significant dataset between the saturation of 75 and 92%, because that's the domain where pulse oximetry will in fact inform changes in decision-making.

The third thing I want to say is, as I've already said, they should ask people to do Phase 3 trial rather than try to subject volunteers to ridiculous hazards, we should, in fact instead conduct Phase 3 trials in ICUs and operating rooms where we're really looking for data from patients who are hypoxemic and we have very precisely matched specimens from blood gas that's contemporaneous with a pulse oximetry reading.

I don't think you're going to solve the problem with 2 LEDs. I think the only way you can make forward progress it both with respect to dealing with pigmentation and frankly getting a better signal is to in fact add LED. Co-oximeters don't have two LEDs they have multiple LEDs

And I think some of our problem is met Hemoglobin, some of it is carboxy hemoglobin. And frankly these devices need to perform reliably in the setting of anemia which is also another, as yet, unrecognized confounder of this data. And I think they need to think very seriously about a standard that is derived from data and patients, with multiple LEDs, and the other thing is I think they need to do better signal processing on the signal they get, particularly, for things like venous pulsation in patients with pulmonary hypertension which is increasingly prevalent problem in society.

Finally speaking as an anesthesiologist and intensivist, designs that mitigate the influence the ambient light on the signal, both in adult and pediatrics. And designs that mitigate RF interference from pulse oximetry, which is a huge problem in the operating room, it would be hugely helpful. And finally, displaying SQI, Senior Quality Index and profusion prominently alongside the pulse-Ox and perhaps some sort of guidance for end-users in the clinical domain would be very useful to front-end providers like myself.

So I think that is a pretty good summary of all the things I wanted to say. One last thing, with respect to ARMS, the other thing that manufacturers should provide us with is a positive predictive value at basically every point between 92% and 77%. So if the pulse-ox reads 77%, was the positive predictive value that this patient is going to have pulse ox readings saturation that's within an arterial blood gas saturation 2 or 3% of this?

Dr. Nathan: I think you make a lot of valuable points and summarize a lot perhaps many of us are thinking in terms of really needing to tighten it up in terms of the ARMS and I don't know what the right number is, 1% or 2%, but clearly when you look at distribution curves, even if it's 2% of patients who are below that 88%, 2% of tens of thousands of readings is a lot of patients. So that really needs to be tightened up. I don't know – to your point, rather than getting a single number, I don't know if the Gaussian distribution can be shown. This is where we think you are at 94%, but there's a 10% chance you could be 88%. Just to give people an idea of a spectrum that it's not an absolute number.

And your other point, I think that you made and I reinforced in the earlier session is these healthy control studies are like Phase 1 studies of medications, and then you go to Phase 2 and then you need Phase 3. So I really think the FDA needs to, before they endorse any of these devices, needs to have real world well controlled studies in sick populations of diverse population or various disease severity, various illnesses to validate, that, yes, to your point, they perform well over 92. That's all the data lies. But what about less than 90 and going into the 70s? How do they perform there? Because that's where the interventions starts. It's interesting to me that the interventions that start when they desaturate, all you have to go through very rigid Phase 3 clinical trials.

You look at remdesivir and whatever disease entity you want to pick, putting patients on ante tropes. So event though, it's not just the device, very important decisions about medications that come into much greater scrutiny are going to be implemented or not based on the reading of these devices. So your point is well-taken.

Dr. O'Connor: I also want to point out when patients are profoundly hypoxemic, if somebody is 87 you can wait for a blood gas, but if they're 55, you're going to act. So the idea that it's not a diagnostic tool is, in fact, not really correct in the clinical context. The reality is that if the sat's 65%, I have to do something.

Dr. Nathan: And to the same point, it might be 85, but you're going to watch them because you don't want to be 75 in an hour's time as well. So the trend and accuracy of the trend becomes very important. Dr. Wilson, you were next up.

Dr. Wilson: Thank you. First, let me say I agree with all the points that were just made by Dr. O'Connor. But there are couple of additional points. So I'm not going to amplify on that. Except one regarding the ARMS just to emphasis critical that not only do we narrow it, and the FDA has to determine is it 2% or 1.5%? But it's important that that continues to be within that range in the important portion of the oxihemoglobin [indiscernible] So when we are in sat between 70 and 80 just as well as between 80, 90 and 100, so that was just that emphasis.

The other point is related to how we evaluate skin pigment and what the percentage or number of patients required for studies. So currently, it's 15% or two subjects. And that number needs to increase. FDA needs to determine how much. But I would say we need to at least double so that you can get data that is going to cross more bends in terms of skin pigment.

Now, a related point to that, in addition to increasing the percent and number of dark skin patients in the subject pool is how do you evaluate that? And I think Josh Fefert did a nice job

early in reviewing, how we look at skin pigmentation. And clearly, a quantitative methodology that takes the viewer out of it is optimal. But I do think that we need to have some sort of you have Fitzpatrick or Massie scale or VanLucian scale also as follows. In the laboratory data and the data that goes to the FDA, quantitative data with some sort of corresponding table is important. But in the clinical space, if we're going to give some guidance to individuals, they're not going to have, at least for some time, quantitative measurements on patient skin, and they are not going to have a spectrometer necessarily. So it would be helpful for us to have that. So I would just like to explore a discussion about which of those three that have been mentioned that the Massie scale, Fitzpatrick, and Von Luschan. The Von Luschan is nice, it has a lot of gradation, but as you look at it and as you go from 1 to 36 as I recall, sometimes you go to increase in tone, and then have a drop in tone. So it's a mix and it's a little bit problematic. And with Fitzpatrick, the problem is there's not enough gradation and that it gets over emphasizing the light skin. So from that perspective, and I don't know if Josh is still on or if he wants to show the image, I can bring it up for the group of the three that we mentioned. The Massie does a better job. It's not perfect, but it goes from line to darker to darker with 10 bins so you can better differentiate where those patients are and perform histograms and ensure that those patients that may be in certain ethnicity also have a spectrum of color, skin color subjects that is appropriate.

Dr. Nathan: Thank you for those points. I just want to build on what you just said. And that 15% number, I think maybe is historical and derived from the proportion of the population that is minority. I'm not sure where it comes from. But is that the right number especially when the 15% is so potentially different in terms of where they're at and what their pulse-ox will read in relation to everyone else who is more lighter toned. And for those 15%, it's 100% when the pulse-ox is put on them. So I would venture to say there should be equal representation so they

have equal chance, any of the different skin tone has an equal chance of having as accurate measurement as someone of a lighter tone than the 15% because that's what the proportion is.

Dr. O'Connor: Could we have some discussion on comparing those charts that Josh mentioned? Fitzpatrick scale was determined initially as a way of determining how people would become sunburn in the first three, and in some versions, even the fourth version of the scale goes from 1 to 6. But 1- 4 is light skin and 5 and 6 is dark skin. It doesn't seem to really cover the spectrum.

Dr. Nathan: I just want to give a caution about the scales which can be subjective. And the more categories you have, the more opportunity there is for discordant reads and less chance for concordance. And there's the CAPA statistic, and I think if you go from 5 to 30, there's going to be less concordance. And I don't know from Dr. Alam now, our dermatologist if anyone has done a look at CAPA concordance assessing skin coloration because I don't know if it's standardized that you look in this area versus that area. Some people might look different, and looks a little bit different there versus here. I think really where it needs to go is some kind of more objective, some of these technologies that were mentioned rather than a visual scale. Anyone else care to comment on that?

Dr. Alam: I'm happy to comment on that. I think the concordance, those studies have not been scientifically done, in part, because the Fitzpatrick scale as someone said earlier, I think it might have been Dr. Wilson, is not really meant to be based on skin color. It's supposed to ask people how they burn or tan and under what circumstances. It's more about the reactivity of the skin than the color of the skin. Two people could have skin of identical color and actually have different Fitzpatrick types.

However, in common usage now it's sometimes also used as a scale not pertaining to sun burns, but just pertaining to what skin colored people have. I think one other issue of complexity you might wish it consider is that color and ethnicity are actually different. And so you could have two people whose skin is exact same color. One is maybe dark-skinned Filipino Indonesian and one is light-skinned African-American patient. The color number might be the exact same, but the structure of their skin, its ability to transmit light, it's likelihood of reacting to sun or getting an injury might be quite distinct.

So I wonder if, really, when we're working on this, we might need to expand our concept beyond just color, right? Because, really, what we're talking about, the reason we care about color is somehow whatever is going, whatever LED is transmitting through the thumb or the ear, or the forehead is not getting the same signal because the color varies. But I suspect the thickness of the skin, the density in which, I don't know, the keratinocyte are packed together, A variety of other factors could probably also affect transmission. We have to be careful there. Even if we came up with the world's most precise calorimetry, we still wouldn't take into ethnicity, anatomic site and other elements.

Dr. Nathan: Thank you for that. Dr. Wilson.

Dr. Wilson: Yeah, I don't want to monopolize, but I just wanted to respond to your really good point about some problems with concordance. And if you have too many choices, people are going to make more mistakes. The one point that maybe I didn't articulate well is Massies, is that of the 6, 4 -- I'm sorry, Fitzpatrick, of the 6, 4 of them are light-skinned and 2 of them are dark. So you're just skewing the data one way. So if there was something maybe was an 8 point scale, but it was just more graded and uniformed the way that the Massie scale is fairly uniform, that has some attractiveness.

But I do agree that a quantitative methodology is going to be essential for FDA, I hope the FDA would require that for future submissions. Do you not also agree that initially there should be some correlation between the quantitative scale and some nicely gradated version of a color categorization?

Dr. Nathan: I think that would be a good study to do just to see how the different scales perform. Dr. Cassiere.

Dr. Cassiere: Yes, I'm going to be a contrarian, and say self-identifying what your color and race is has generating pretty good data. And we have a meeting with FDA around it. What I'm worried about, and this speaks to the fact that, different skin tones from different races and ethnicity might also matter. We don't know that yet. So jumping right to asking the FDA to use a scale, whether it's a visual scale or the device that measures your melatonin level and assuming that's going to translate to some of this disparity data were seeing on pulse oximetry is a leap.

And I think we're doing a pretty good job having people self-identify themselves what they are in terms of skin color and ethnicity. And I think we should stick with that until the studies shows that can correlate and you can put it into a numerical value or look as a Massie scale a Fitzgerald scale. I am going to be a contrarian and say I don't think we have the data yet to translate what the data's been driven, that there's disparity in pulse oximetry in certain populations and now we're taking the leap that we can pick those populations out now using a numeric skill and even a visual scale.

Dr. Nathan: Let me be a double contrarian and say, do we need a scale? As long as the manufacturers can show that it performs adequately and as well in people of different colorations, then it's up to the manufacturers how they get there rather than using this scale and that scale.

Dr. Cassiere: Agreed. That was my point. We're doing a pretty good job showing such discrepancy between dark pigmented, self-identified race and ethnicity, why are we leaving that until we have studies that actually go head to head comparing I designate myself as a certain race? Here is a numeric value or scale and that kina correlates and the data still generates.

I'm agreeing with you on that. We should not abandon the self-identification. There's a lot of other things in the skin besides the melanin. And my dermatology colleague, excuse me for forgetting your name, pointed out succinctly.

Dr. Nathan: As long as the device performs in everyone threw a wide range, I think everyone is going to be satisfied with it.

Dr. Cassiere: The other thing I want to say, again, I just want to make this other point. All the money is around abnormal, low PAO2s. We should have an expectation that the device companies look at those levels to show that in the lower ranges of pulse oximetry, it's reliable, it's precise, and it's accurate, and we should definitely lower the expectation of a 1.5 to 2%. Again, we need to have some type of number, call it a reliability index of how reliable is that number that I see based upon all that information?

Dr. Nathan: Okay. Thank you for that. I'm not going to go in order because I want to make sure I give everyone an equal say. I see Dr. Connor has his hands up. Do you want to go next, please?

Dr. Connor: Sure. I think just sharing some ideas about what Dr. O'Connor mentioned the idea that not all differences are the same. If you're a 95 and you're off by 2, it's not clinically relevant. But if you're truly an 87 and you're off by 5, then that's where you're certainly now treating patients differently. I was just going to point out, there's some literature on this methods, too, that Ken Sherlock and myself and Brad we altogether at Cleveland clinic did INR point of

care devices. When INR devices were first being produced, point of care we looked at this versus going in [indiscernible] clinics and saw different values result in different clinical decision-making depending upon where you are on that scale. So we basically took a [indiscernible] plot and created regions that led to discrepant clinical decisions, not just discordant numerical values, publish that and also ran some trials comparing different point of care devices using that methodology versus just looking at concordance.

So I just wanted to put on record maybe this is a shameless plug. There are ideas that just don't look at ARMF and look at really when do clinical decisions differ and you can use that sort of methodology by race and see when things are breaking down or by race or skin tone and that sort of thing.

Dr. Nathan: Thank you. Let's go back to Dr. O'Connor.

Dr. O'Connor: Yeah, I just wanted to make one additional comment. It's about the coloration issue. 15% is, in fact, the wrong metric. What we need instead say they need to have a statistically significant number of patients. One that allows them to assert a 95% confidence interval that's no different than Caucasian patients. So it doesn't matter -- other people who are wiser than me will decide what we're going to do about coloration, but for myself, we need to have statistical criteria, not percentage of population. And what that may mean is you end up with 200 Caucasian, 200 African-American, 200 Asians, 200 Hispanic. 200 grade skin 1, 200 grade skin three, 200 grade skin 24.

Dr. Nathan: Agree. Thank you. Dr. Loeb.

Dr. Loeb: I'm going to be a little contrarian about the coloration, because in my mind, all of the studies do show gradation between darkly pigmented, intermediate pigmented,

and lightly pigmented people. Unless you are building a model that's going to count on a per person basis why you need to put the intermediate people in there.

In my mind, there are two problems with the way that the [indiscernible] studies are done, one is they predominantly have been done in lightly colored people. And I would propose that equal numbers of darkly pigmented and lightly pigmented people be used. By darkly pigmented, I'm saying the people who are on that highest or the next to highest in terms of pigmenting.

And secondly, they're done under best possible conditions of good profusion. I think that's vitally important that the studies be done in poor profusion conditions and for instance, that could be done by cooling the hand or possibly having a cuff above the pulse oximeter up to something that will increase venous congestion that will decrease the pulsatile amount. So I think that's important because each pulse oximeter should demonstrate given an equal distribution. It's important what Dr. O' Connor said, an equally bad distribution in different racial groups.

Dr. Nathan: Yes, I agree and I think, that point is getting to the FDA. It needs to, these devices need to be studied and validated across a wide spectrum of coloration and a wide spectrum of disease severity. It's interesting, if you compare, we have made some comparisons to typical drug studies where you have a phase III study and even in drug studies with a phase III, they're not really real world because they have so much inclusion and exclusion criteria and it's a highly selected population. What I would venture to say is in the equivalent of a phase three study with the pulse oximetry devices, you need them to be the phase three to be real-world. You can't have a population with no peripheral disease or circulatory problem or vasodilators. The device might not perform as well, but at least let’s to have some transparency around it and some

and warnings around it, hey, if this patient's blood pressure is less than, 80, whatever the case may be, then your reading might now be accurate.

So I think the equivalent phase III study needs to be real world, not a highly selected population and oh by the way, after approval, you need to do your real world study. All right, let's see we're at Dr. Wilson, who you had hand up.

Dr. Wilson: Thank you, I wanted the opportunity to agree with the counter contrarian earlier, about it may not just be the pigment. I totally agree and maybe not even just be the skin. We know within every race, there's a wide gradation of skin pigments but in terms of the static absorbers which can be even more important than skin pigmentation according to our research, bone density, thickness of the digits and all of these factors, that's the reason why it's so important to have the full spectrum across however you can characterize various races and ethnicities so I just wanted to agree there.

Also, one aspect that became very prevalent with COVID is we found in critically ill patients, a number of publications that showed very high hemoglobin numbers, higher than expected. And carboxyhemoglobin, which is related to breakdown of hemoglobin and we'll need to make sure, if we do phase III trials, there has to be a precise quantification of the hemoglobin above some number would need to be evaluated so you could have a fractional, not just a functional saturation.

Dr. Nathan: And I want to bring up another aspect. Before I do, Dr. Kirsch, I have not heard from you, and I see that you have your hand up.

Dr. Kirsch: Oh, thanks, I agree with all of the great comments. I want to emphasize the disparities of weight, BMI across different populations and I would strongly encourage us to

consider having the FDA take it into consideration when they look at different types of patients, not only their skin tone but their BMI which adds significantly to the problems of measurement.

Dr. Nathan: Yes, I agree. That would be the broad swath of patients that we would suggest, no BMI cut offs as patients are hopefully included in these studies. Dr. Collop.

Dr. Collop: Yes, just sitting here and a lot of the focus is on critically ill patients and healthy patients but another area that – relatively healthy patients, that have significant oxygen desaturations are in the sleep lab. We have patients every night that desat into 80s and 70s and that might be another area that would be fruitful to look at because as far as the phase III trials, there's a lot of variability in the sleep lab with regards to saturation and BMI and you know, there's obviously a diverse patient population as well.

Dr. Nathan: Thank you for that comment because that does remind me as well over the patient population we see in our clinic, patients with advance Lyme disease, at least one- third are on oxygen and we base it on a 6 minute walk test for the most part, and it's a big decision. Do you put patients on a therapy which is a significant imposition in terms of quality of life having to lag around oxygen or not, or are you withholding supplemental oxygen to patients who are desaturating and may develop an organ dysfunction, or pulmonary hypertension because they haven't been put on oxygen in a timely fashion so when we think about it and mention broad populations, it also has to spend disease severity from normal healthy, where it has always resided to patients who are on the fringe of needing and need supplemental oxygen to the critically ill patients so this is another aspect of the broad designs to validate these devices.

Dr. Collop: I would add in a CPAP machine too. We base a lot of it on when to give someone a CPAP machine.

Dr. Nathan: Dr. Branson, Mr. Branson, I'm sorry. I like honoring your doctor –

Mr. Branson: That's okay, I always tell the guys at work, just give me the doctor's salary and you can call me whatever you want.

I really don't like the idea of using self-identification of race and I have talked about it enough that I use the same examples. I am always looking for, who are two people who people may know who they are, but think about Colin Powell who self-identifies as Black and Tyreek Hill who plays wide receiver for the Miami Dolphins who is a young man from south Georgia who is very Black. And those two individuals in self- identification have significantly different skin colors and I really do – I'm worried that, I'm not saying we don't collect the self-identification data. I'm just saying something about the skin color is, I think, as important or more important. And I'm with you. one of my other tasks recently has been working on new orders for home oxygen therapy for CMS and I would hate for us to finish this and find out that oximeters aren't accurate enough to prescribe oxygen therapy for patients, or as you said, Dr. Nathan, are depriving some patients of oxygen who should be getting it and we have to revert to having to have a blood guess to determine who needs, or who is going to be allowed or permitted by CMS to get home oxygen therapy.

So the thing that is for sure, we need to have better designed oximeters that take these factors into effect.

Dr. Nathan: Thank you, Dr. Connor.

Dr. Connor: Yes, I was going to reiterate. The ways these are tested could be much more rigorously done. I agree that the ideal situation like Dr. O' Connor said is to test these, in ICU, or the OR, I think it's hard to do and I can imagine [indiscernible] not granting permission for that because there's little advantage for the patient in having this done just for research

purposes but there wouldn't be potentially any medical value to them given that the timing of the blood reading may not help them especially in the OR.

There's plenty of examples. The gentleman from Massimo who described and put down the research we saw earlier but I think he validated the point. I have no doubt that Massimo is really really good at getting very accurate measurements and better than anyone else in the world – meaning in the lab. So that, when we're actually using these in emergency situations like a hospital or OR, all of these situations where the readings are not nearly as good.

But I think the FDA can now mandate the ways that these are tested. I think, analogous situations – I have a friend who is a pilot, and trains other pilots. So in the simulator, he doesn't take them up and simulate windless cloudless days, he simulates really hard situations and if they can't handle that, he can ground them and they cannot fly commercially.

Similarly on the opioid side in CDER when a new long lasting opioid comes out, the company basically has to try to break it meaning they have to try to break it down so the long lasting opioid can't be used all at once which is what used to happen so people could essentially make their own home heroin and get a much more rapid high than it would allow. So there are instances, even within the FDA, where a company is forced to challenge their own product and I think that needs to be done here with tightness, more light, less light, different color skin. The example of using the cuff to limit oxygen flow. I think all of these situations, we understand they won't be perfect but at least this way the clinicians for particular patients can realize, oh, for my types of patients, I should choose this device over another device.

Dr. Nathan: Thank you for those comments. Let's go to the consumer representative, Ms. Edwards.

Ms. Edwards: Okay, I just wanted to talk about the “darkly pigmented”, that the FDA has asked for in the -- what is the 15%, right? I guess I want to know how FDA defines darkly pigmented because in the African- American community, the pigmentation is as diverse as the people are. So I think that would be interesting to get a definition of what they, define as darkly pigmented.

Dr. Nathan: I think that's a very valid point and one of the things that comes out of it, because I think going back, it just says exactly that. Without being more specific than exactly that. And I think that in terms of new guidance in addition to increasing the 15% and I think a couple of us have thrown out, should be 50% or thereabouts. There needs to be a specific guidance, not just dark skin but maybe that's where the scales come in, which ever scale you choose has got to be a broad representation across that scale.

I think the companies can use the scale to say, yes, we did study X number of individuals across this spectrum of this scale. But I don't think the scale itself can be incorporated into any kind of clinical practice. It just helps them to validate their device. So I think that's very helpful information or suggestion, thank you very much. Sorry, go ahead.

Ms. Edwards: So why not have a test of just people with color? To figure out where the issue lies? We already know that it's in the people of color, right? Do you get what I'm asking? [multiple speakers]

Dr. Nathan: Well, maybe this is a separate study in people of color in yes, I think you can have it as a separate study or wrap it in one study where it's equally divided or you know, or nuanced that way. But I can see Dr. Eydelman has her hand up, maybe she wants to address this issue from the FDA standpoint. You're muted.

Dr. Eydelman: Yes, thank you, Ms. Edwards. I wanted to clarify you're correct. As is Dr. Nathan that up until now, we did not have a definition of what that meant and that is why we believe this is one of the key questions that we would like the panel to deliberate and that's why Josh presented that, had the presentation where he discussed all of the different ways to subjectively and objectively assess skin pigmentation and we're very interested in hearing panels.

The foundation as far as preclinical studies for premarket assessment, in other words, and not just after they get on the market but what, if any, of these skills should be utilized in the assessment of how good the pulse ox is before we make the decision whether it gets to the market or not. Thank you.

Dr. Nathan: Thank you for that and I guess that will be one of the questions we'll be addressing but just to plant the seed as we try and answer this question. I'm not sure if we as a panel can come up and say, well, this scale is better than that scale and we recommend that the FDA uses this scale. And maybe the important message is, use a scale. Whether it is a 6 graded scale or a 30 graded scale, as long as there's good representation across the spectrum, maybe therein lies the answer rather than being too prescriptive.

That's my viewpoint and maybe that's because, I don't know enough about these individual scales. I have never used them or seen them in action or know what the concordance is and how difficult or easy they are to use. But any way, and I know, we have someone who has their hand up. His hand must be getting tired. Let’s get it over to him.

Dr. Yarmus: I put my hand down, thanks. I'm going to shift a little bit, hopefully that is okay, which is to talk about the outpatient setting again. We know medicine in general is moving very much towards outpatient practice. And ideally keeping patients out of the hospital and more and more of the hospitals are being taken up by ICU beds for the critically ill. So I just want to

be mindful of that as we kind of continue to navigate through the pandemic where many of us have utilized these out-patient pulse oximeters as indicators for outpatient management and potentially inpatient admissions.

Many of these devices are not as regulated, right? They're the ones that you buy at CVS. And we have looked at it through the years. I know our CF group for example has looked at different devices to assess patients in the outpatient setting. So I think it would be remiss and I don't think we're doing it but just to sort of put a side those devices and the importance of them. If you look at watches and everyone who has an Apple watch can monitor their pulse oximetry, and how we manage those.

So I think part of this discussion should be focused on the regulatory aspects of how those devices are assessed because that may actually, be an equal or greater problem as this progresses.

Dr. Nathan: I agree. And I think as that becomes part of the question in terms of differentiating between the consumer and medical grade devices. This lays a foundation for us to try to answer that one.

You're right, this is probably used much more than when we use in the hospital and people make their own decisions about triaging and what they're going to do best and what their readings from those devices are.

Dr. Yarmus: Sorry, Thanks. I think it's also, they're not -- I'm not sure it's just the patients, I think providers are increasingly utilizing that information and may or may not understand the regulatory aspects of the differences of a pharmaceutically purchased device in a drugstore versus one in the hospital.

Dr. Nathan: Absolutely right. I know what we do for the lung plant transplantation population, everyone goes home with a pulse ox and if they are short of breath, the first thing you ask them what is your pulse ox rating and you're absolutely right. Dr. Hudson.

Dr. Hudson: Thank you, As I was listening to this discussion, I was thinking all of the same thoughts. I think the iceberg is the outpatient devices and I will second and this is why I raise my hand, I don't think most physicians know that these are fully unregulated devices and untested and all of that is now feeding into apps and into medical decision making.

They are not used recreationally. They are used for medical decisions, and I think we would indeed be remised to not address that issue with this panel. I will also say that as clinicians, we use pulse oximetry very differently than patients do. We use it as a device to monitor trends. We use this information in conjunction with a lot of other clinical information. I think the devices that are sent home with patients and that are purchased, should not be a number. It should be something that takes into consideration the profusion averaging over time to give the indication whether there's concern or not for connecting back with their care team. And I think that would be a great advance for home use of these devices.

Dr. Nathan: Thank you, Dr. Hudson. Dr. O' Connor?

Dr. O'Connor: I agree with what everybody said about the outpatient devices. Two things astonish me with COVID. One is how many patients were profoundly hypoxemic, were not [indiscernible], right? And secondly, I think we should think about offering the FDA the opportunity to come up with a third class device which is the device intended for use at home by patients in collaboration with their care providers for decision making.

That is to say, we may not need all of the features for an inpatient device but we may well need to provide them with guidance because you know, one of the other lessons we learned in

the COVID 19 pandemic is when the hospitals get filled with people who are mechanically ventilates, we send people home on oxygen by nasal canula using these devices to guide their care and anything we can do to improve that performance is a benefit to the public.

Dr. Nathan: Yes, I think, we can certainly suggest that. I'm not sure the feasibility of a third category of device but for one, the medical grade ones versus the consumer ones need to be clearly marked. I have bought a couple of pulse oximeters, to be honest, and maybe I should be embarrassed by this, but I didn't know there's more medical grade ones versus consumer ones. I just thought there were cheaper ones and more expensive ones, and you get what you pay for, and I thought the cheap ones just didn't perform as well.

I think it needs to be clearly delineated that this is consumer grade versus medical grade is a very good first step. Dr. O'Brien, Before we go there. Dr. Eydelman, Do you want to respond to one of the comments or answer a question?

Dr. Eydelman: I just wanted to clarify that prescription used devices can indeed be used at home with prescription.

Dr. Nathan: So that's the difference. If we want to make sure that the patients are going to get medical grade, you need to write a prescription.

Dr. Eydelman: You are allowed to do that so there's no reason to create a third class.

Dr. Nathan: Okay, thank you. Mr. O'Brien.

Mr. O'Brien: Yes, thank you, I just wanted to weigh in with the discussion about the use of it and clearly from a patient perspective, I think we have to include the home devices that are being done. The proliferation is just amazing among the patient population. I have my watch, my spirometer. I am a 72 year old male with anemia, with a pacemaker, atrial fibrillation, sleep apnea, and I am also on pian meds, opioids for a significant amount of time I have a lot of issues

with it and not to mention, that thank God, I'm not female or Black because I'm concerned about what do I do with this? And part of the discussion we're going to have to have is education around that, and what it means and I just realized, Eileen Hudson, I think you're the one that is one of the authors for what I saw in the Annals of Family Medicine which seemed to suggest these non-medical devices could be used to rule out hypoxemia, so I think there's education in a lot of levels at the physician level and patient level in terms of what this means.

There was a discussion earlier about perhaps putting that on the pharmacist. I think that would be a tragic mistake because the pharmacists are overburdened and I don't think we can rely on the patient education from the pharmacist perspective on each individual purchase.

First of all, they're on the shelf the Pharmacists are going to know, and I think it's an unbearable burden and I hope to God, we don't come up with another label that is 19 pages long that you need a magnifying class, and most people just throw it out, and it doesn't address what is happening in today's world.

Today's world is about audio- visual. It's about a one minute video that can explain the critical issue, this is how to use it, how not to use it. This is what it means. This is your skin color, et cetera. It should be something easily identified for the education today what it means, not what we have used 100 years ago, et cetera. I think that has to be a part of the discussion for the FDA going forward.

Dr. Nathan: Thank you for that. Dr. Cassiere.

Dr. Cassiere: I am going to be a contrarian again. I'm going to say what we're trying to talk about is regulating devices for non-intended uses. And in other words, these consumer grade oximeters are not medical grade oximeters and there should be no expectation they are. And it's

very difficult to go to the consumer grade oximeter companies and say, you have to make oximeters close to medical grade because the public is using them in unintended ways.

I don't think that makes regulatory sense, and I think you should not side-step your provider, whether it's a nurse practitioner, a physician or PA, just because you have access to something that can simulate a medical device, doesn't mean you should use it outside of medical supervision. So again, I'm going to be a contrarian that I think, to speak to what the FDA said, you can order and prescribe a medical grade oximeter for patients at home. So we are talking about making another category, that actually doesn't make much clinical sense to me at this point.

Dr. Nathan: Your point is well taken. I think the consumer devices will do their own thing. What is important is the education and the limitations and perhaps, better labeling. I don't know what the labels look like now. But some kind of warning on the box in terms of caveats, about the reads, and they may have them now.

I'm not sure how bold they are, to the point of, you don't want to read the 19 page little insert and on the bottom of page 8, there's a little warning that says it may not be accurate and to call your doctor. A little bit more visibility which goes to the education of the consumer; about the pulse oximeter is exceedingly difficult.

How can you educate them? It has to be written in bold with the little box and even then, people are going to throw away the box and not remember a month later. This is a difficult issue. Dr. Hudson?

Dr. Hudson: Yes, I would say I know a lot of individuals that have pulse oximeters at home and I don't know anyone who uses it for recreational purposes. They all use it for medical decision making so I think it is not in the public's best interest to have these devices that we all

do understand are using for medical decision making and turn a blind eye and say it's for recreational use. I don't think it's in the public's best interest.

Dr. Nathan: I have a question for the FDA then. For these wearable devices. How much oversight does the FDA over these devices in terms of requirements? Is it the same? Can you put in the requirement that are needed or because they consumer wearables, it falls under a different category? I'm going to ask Dr. Lee.

Dr. Lee: Thank you for the question. So you know, as we discussed in the beginning actually, we do regulate some of these products under different product codes, and as you can imagine, particularly if there's importation of the devices. There are requirements and registration and listing so they are falling on their – particularly even if they are, you know, sports and aviation, we have defined that. But as you can imagine, there's other devices that you are discussing that could fall under general wellness where a company is not claiming medical use but that it's only for health and wellness and that would fall under our wellness guidance.

So even if though they would have to oblige by the wellness guidance and fallow the parameters in which non-medical devices are regulated for health and wellness or they have to register and fall under sports and aviation. It Is a little more complicated and we certainly appreciate the discussion. Because there are a lot of different label activities out there but Dr. Nathan, does that answer your question?

Dr. Nathan: Yes, it does. It seems like this is quite complicated and you know, your reach only extends so far if it's under wellness, then it seems like you might not be able to mandate exact requirements as my understanding. I'm sure it's very complicated and it's probably workarounds that these companies can go to, to get their device on the market, I suspect.

Dr. Eydelman: If I can just add one second. This goes back to the comment made earlier. We can ask for a particular labeling on OTC pulse ox, if that's the recommendation in the panel and there's a way for us to require that. We just did it for OTC hearing aids. I would say we would be interested to hear if that's what the panel is recommending.

Dr. Nathan: It sounds like that, let me speak what I think for the panel. It sounds like there should be verbiage on these devices that are not as accurate, that they are not meant for medical purposes. Medical – for medical usage. I don't know what defines recreation, looking at what your oxygen level is. I'm not sure what the right verbiage would be but if you have a low saturation level or you are short of breath or you have any other symptoms of concern, you need to follow up with your PCP. This goes back to the point I said earlier, not all causes of shortness of breath are accompanied by low pulse ox. I think there needs to be more warning around them, that these are recreational and should not be used for health purposes.

Dr. Eydelman: I just one last comment before we leave this, we're mostly interested to hear what the panel recommends and then we will take all of your recommendations into account and try to delineate the best path forward.

Dr. Nathan: Sounds good. We have a bunch of people lined up to weigh in. So let's start with Dr. Hennessy.

Dr. Hennessy: If I go to a pharmacy and buy a digital thermometer, I assume it's accurate. If I go to the pharmacy and buy a pulse ox, that gives me a digital reading, I will assume it's accurate despite what you put in the box or on the machine. I think it sends a mixed message to allow companies to sell something, yet not have confidence in their results of what it puts out, particularly because we know it’s being used for health reasons.

So when these things were first sold, it's before the COVID- 19 pandemic and they didn't have the use they do now. I think the environment has changed and as a result, the regulatory environment should change.

Dr. Nathan: In an ideal world, we would want these devices to be held to the same standard as the medical grade ones and we might recommend to the FDA they tighten that up. So is it feasible to mandate that of these consumer devices in terms of rising to become medical grade because people might use them for medical purposes, is that pragmatic or feasible?

Dr. Hennessy: So somebody else proposed, and I think it makes sense is, if it doesn't have the degree of precision to warrant giving a number that goes from 75 to 100, to band it and say, it's within this range, or within this range, or within this range, without giving a digital number or give a range.

Dr. Nathan: Okay. Ms. Edwards, why don't you go ahead?

Ms. Edwards: I just would say from the consumer perspective and back on my mom. So my mom was under a doctor's care and I am almost certain that her doctor was the one that suggested she get the oximeter and blood pressure machine to measure it daily. As I said, up until the very day that I called the ambulance, she had not had any signs that her oxygen levels were low, not by the oximeter. So I guess, I go back to, there needs to be something on there that alerts the consumer to the deficiencies in the oximeter, or make it the same grade s the medical, but then that too has deficiencies, I guess in correcting it, they both need to be corrected.

Dr. Nathan: I agree. To what level do you mandate that the wearables be corrected. We can discuss it further. And try to help the FDA find direction in that regard. Dr. Kline.

Dr. Klein: Just tacking on a little more to the discussion about the consumer products. Initially when the suggestion was made about they should be behind the pharmacists

and there should be a conversation. I think that's a good suggestion although I recognize it's challenging, and it's already onerous and we have a workforce shortage for pharmacists and that's additive work but I'm concerned about, we are talking about pretty complicated messaging.

If we have consumers who are low literacy, unable to read, how will we message all of this so they can be informed when you see these products. So that was one comment. This is something that might be a little in the weeds but unrelated to that comment. This may be more specific to pediatrics, I'm not sure. But for pulse oximeters, with a disposal probe so now I'm getting away from the actual device but talking more about the probe. But in this cost contained environment post COVID and health systems, trying to use things longer and longer so, it can get complicated by, is the probe wearing out? It may not be the actual oximeter machine. Maybe a new fresh probe was put on the patient or for patients that are at home, and also trying to extend the use of these, typically they are wrapped probes for pediatric patients. If they're trying to extend the use so they don't have to buy a new probe. I just wanted to add those two cents.

Dr. Nathan: Thank you, Dr. Alam.

Dr. Alam: Thank you, I agree with a lot of what has been said already. I think I want to call a spade a spade if you will. Bear with me, which is that I'm not sure what the history is but I think it's clear based on the panel's views and certainly mine that the OTC devices are misclassified. These are not health and wellness items. This is not a toy. This just isn't. The only people use pulse ox is because they, their doctors or loved ones are concerned about their health and want to be alerted as to when their health is deteriorated to a point where they need to seek immediate medical attention. So this is a misclassification problem. I don't know how it happened but this is not shampoo.

So FDA will need to find some way to regulate these and if the current scheme is insufficient to dramatically improve the products or regulate them with a level of vigor that would be sufficient to actually make some difference, then FDA may have to reclassify them. There is a caveat which I think is very important. FDA should not complicate the reclassification so much that the cost of the pulse ox goes from ten dollars to a thousand dollars, because that will create an access issue. We'll have pulse oximeters, but if 99% of people who could benefit from even a not very good pulse ox machine, will not have access to any.

So there has to be some nuance to hitting a range or clarifying numbers, 85 to 90. If you're lower number is 85 you should give someone a call. Maybe make them a little better to work slightly better overall and be slightly more reliable, slightly better in skin of color patients but not make them completely disappear. But I think the main thing is, reclassification is essential. I just don't see how these are health and wellness devices.

Dr. Nathan: I think you hit on a very important point which is to find the balance between being too idealistic and being pragmatic at the same time. If we impose the same hurdles in terms of these consumer devices that are medical grade, we're going to put them out of reach for people who need it the most. And then people don't have access to these devices unless they get a prescription which may or may not be paid for. So the FDA should continue what they're doing, in terms of having a different threshold for the consumer devices, but much like we recommend tightening the requirements for medical grade, certainly it seems like the same holds true for consumer grade devices that the requirements for approval need to be a little bit more stringent which makes them a little more expensive but it's the question of trying to thread the needle and find the balance which is going to be quite difficult. Mr. O'Brien.

Mr. O'Brien: Joe O' Brian, patient representative. I will echo what was just said ten times and then I think it's absolutely essential. And just in terms of looking at it, I don't think there's any finger pointing here. What has changed is the environment. All of the sudden, a problem that existed for thirty years, the risk benefit ratio changed. We now know there's a life threatening risk that is associated with misrepresentation of a number with the accuracy of a number. That changes the dynamic. That changes the use perspective that we had, so previous panels, previous decisions about wearables and non-wearables is looking at one thing.

Now we know perhaps, this is a much different perspective that we have to look at and protect the patient in that particular environment going forward and a lot of that is -- and I absolutely agree. We cannot over price to the market and access to healthcare is essential for the patient population. Otherwise, we defeat the purpose for that. But if you look through the internet and the web sites for both the IT companies that are using wearables and that are developing wearables and you look at the major companies that are in the field already, that share the space, they do wonderful jobs on videos for marketing and advertising and et cetera and I think certainly if they use their brain and their money with the right influence, they can come up with very specific appropriate for the population, educational material that can be given quite quickly without giving all of that documentation that we typically do and have it so we can have a much better device and a much better patient going forward.

Dr. Nathan: That is nice. I want to take an important point which is the undertone of what everyone is saying, which is that even though the initial, the mandate of this advisory group is to address the disparity in terms of pulse-ox reading of folks who have different skin tones.

The number one thing that resonated with me is one of the studies presented said that for people White people there was a 10% inaccuracy rate and for Black people, there was 15%. We

shouldn't strive to get Black folks to 10% but strive to get everyone to 1%. So even though the initial mandate is to look at the disparities based on pigmentation, what you're hearing from the panelist who have taken it up a notch for everyone and make everyone equal and better in terms of the pulse oximeters that are out there. Dr. Hudson?

Dr. Hudson: I would like to add to the discussion that the companies that make the medical grade pulse oximeters are the same who make the non-medical grade oximeters, the recreational ones. I don't think it's a big leap for them to up their game. It's more of a marketing issue than anything else. I don't think we would be asking too much of the companies. I think one thing that we would potentially consider is taking the number off of those devices that are used at home to not give individuals a false sense of security which is exactly the same thing we used in a hospital setting. And just give some kind of a warning that is taking into consideration these other factors and then you eliminate the issue with literacy, and understanding and complicated labeling for the FDA. If it's something as simple as red, yellow and green.

Dr. Nathan: Thank you. Dr. Yarmus.

Dr. Yarmus: So maybe I will play the contrarian, for a little bit. And just refocus a little bit, why this may not be as importantly for us as physicians and the FDA because it's already out there. So the question may be focused to the FDA, if we're here to opine on these inpatient parameters and research design, do we have the power to do that for the outpatient realm? Should we look at these study designs for outpatient patients? I know there's some studies that have not shown much difference but that's in a select population. Should we look at it more broadly for all oximeters, I guess?

Dr. Nathan: I think that's where the conversation has gone. Once you start looking, that we're going to render some advice on all oximeters that are out there? Dr. Yarmus, you had one more point?

Dr. Yarmus: Yes, I guess where I was trying to head is, there's certainly discussion that most of the onus is on the industry. But in fairness, I feel it swung a little bit for the outpatient where maybe it's not -- they should be equal. Inpatient or outpatient, I think that assessment should be treated the same and we should hold these standards equally.

Dr. Nathan: Absolutely, I think it's hard to not address the whole thing but to pass on that would complicate that. Mr. O' Brian?

Mr. O'Brien: My apologies. I forgot to take my hand down.

Dr. Nathan: Okay, then Dr. Loeb?

Dr. Loeb: I hope I'm not being simplistic but to me, the problem both with the racial disparities and consumer grade versus non consumer grade is the signal to noise ratio. The problems with skin color, the problems with anemia, the problems with low profusion, all of that gives a lower signal. Some companies call the signal to noise ratio the profusion index and other companies have another name for it. As that number goes down, the accuracy goes down. To me, the consumer grades are supposed to be used in outpatients who are otherwise healthy and so they should just be regulated to be accurate with good profusion and good signal and not give a value at all when there's low signal.

As opposed to hospital grade which would give a value when it's low signal and try to deal with those types of patients. And to me, the oximeters should have a display of signal quality because as Dr. Goldman said, the waveform might be right but the signal quality is low and that's clue that you need that is not accurate. Again, my guess is that the amount of static

absorption, that DC component at the two different wavelengths could be used as a correction factor as the signal gets lower.

Dr. Nathan: The questions and comments have been very engaging to the point that I have neglected checking on the chat, but I see someone who would be very appropriate to weigh in. I want to ask Dr. Pfefer from the FDA, who's not on the panel, but perhaps can weigh in, He gave us a talk initially on skin pigmentation and based on the discussion we have been having, he might have some important insights to add.

Dr. Pfefer: Well, yes, there's a lot of issues that have come through over the past half hour or so. I'm not sure where to start or what you would like me to address first. There's a fair bit of discussion on the objective versus subjective approaches for evaluating skin pigmentation and what is really necessary.

Certainly one of the problems we see with approaches like Fitzpatrick, is there seems to be a lot of variation in the pigmentation level or melanin content within the highest levels. Like the types five and six. Whereas some of the objective approaches might be able to give you a finer gradation based on the actual melanin content, and maybe event correlate it direction with the concentration of the melanin in the epidermis especially, if it's well standardized with respect to some of these biopsy approaches.

So I think there is a potential to improve the ability to ensure we're doing a comparison against the people who have the greatest skin pigmentation and not missing those folks.

There's was also a discussion on whether we need to keep this concept of the racial or ethnic categorization because we don't really know that skin pigmentation is the main factor and I would just point out that the vast majority of the literature that we see in optics shows a lot of the approaches that are impacted in terms of optical measurements are impacted by melanin.

There's a lot of evidence on that. For other ethically or racially correlated biological factors, I'm not aware of many studies that show there's a correlation between those factors and optical signals so I don't think we should toss out the idea of evaluating skin pigmentation. Possibly you could evaluate both ethnicities as well as skin pigmentation but I think it could be quite useful.

And of course, different scenarios –one of the points I was trying to make in my presentation is different that different scenarios may require different approaches. There may be certain studies where you find gradation of skin pigmentation and others, it may be less important so I think we need to be somewhat flexible about that. And part of – [unintelligible].

Dr. Nathan: Thank you Dr. Pfefer Thank you very much. Was there a particular slide that Dr. Nathan, I believe, that that's how we asked Dr. Pfefer. Did anyone on the panel want to present a particular slide, that would be helpful in our deliberations?

Dr. Pfefer: So this on an the following one showed some of the subjective scale. So certainly, the Fitzpatrick scale, one of the problems of using it, I consider Fitzpatrick to be two different scales. One is about the sunburn scale and it's based on interviewing the subjects and asking them question about their sunburn, suntan habits and what their skin does. This was used to evaluate light skin people. So some studies have been used that approach which is the original Fitzpatrick approach, and more recently there's a trend to kind of convert this approach into colors and these are not well standardized colors, certainly. I think often people may take a certain range of colors and print them out on a piece of paper and compare them to patients that I'm not really sure how it's done but certainly, the colors can't be assumed to be very consistent.

If you go to the next slide. In some of these, the approaches and color levels may be a little more standardized, for example, this one, Massey-Martin I don't think there's an official Massey Martin scale printed in a consistent color level that can be distributed. I believe, people

are just taking a paper that has been published and printing it out and comparing it to people's skin color. So it's certainly not ideal in terms of standardization and I believe one of the panelist talk about the VonLucion scale and pointed out that it's monotonic, You're not increasing the pigmentation level as the numbers go up. It seems to be combining the effect of both melanin and erythema and that's actually something that we've in our evaluation of objective methods, we have looked at erythema, as a confounding factor and we try to -- there's a lot of prior studies that have looked at the degree to which they can evaluate changes in melanin, or differences in melanin content without being affected by the redness of the skin, the blood content essentially. So there's flaws and certainly there could be a way to come up with a more optimal subjective approach. And maybe something like Munsell is a little more well standardized where you can actually get a Munsell color chart that's more well standardized but I don't know.

Dr. Nathan: This just reinforces what I said in terms of the subjectivity around this. It looks fine on the chart, but you can just look at 4, 5, and 6 and imagine that in practical application and people asking to differentiating between 4, 5 and 6 when you are actually faced with someone and where to you look or event 7, 8, 9 in differentiating there.

So I think what you alluded to in terms of those objective assessments I think are going to be much better at the end of the day in terms of trying to standardize it. Going back to what I said, earlier, should we really care about being prescriptive, as long as the companies can show the devices work for everyone across a broad range from 1 to 10 or 1 to 4 scale and across the different range of pulse oximeter readings, it doesn't matter how they got there as far as I'm concerned.

Dr. Connor: But the problem is, you need to make sure it works in the people with the highest pigmentation level and if you have a large bend for that top range, you might miss the

people towards the top at those highest concentration levels. The highest epidermal melanin concentration levels and that is the group I think we need to make sure are included in these studies.

Dr. Nathan: All right, thank you very much, thank you for weighing in there. Dr. Connor.

Dr. Connor: Yes, thank you, Jason Connor. I wanted to politely but strongly disagree with the idea of recommending the home use devices or commercial grade. The idea of a range, red, yellow, green situation. As a statistician, giving people less information is never really that valuable but also, it doesn't fix the problem. If there are true 88 in white but say 93 in darker skin, the color or range doesn't fix that problem. It's still wrong with those people. More so, if there's a situation where we would show yellow when we hit 88, it may still be yellow but someone else is 78. And if they are at the bottom of that range there may be a very different decision. So I think the idea of A, giving less information is bad and I also said this to FDA, there times when I have been involved in diagnostic devices and FDA wants to read out for the clinician as high or low and that's generally a terrible idea. Almost universally a terrible idea. But the idea it still doesn't fix the problem of bias here. That's the key. It's still, an error even if we give ranges.

Dr. Nathan: Thank you, Dr. Connor, Let’s go to Dr. O’Connor.

Dr. O’Connor: I'm just going to push back, as well but from a slightly different direction. It's 2022, LEDs have never been less expensive. Electronics have never been less expensive and manufacturing has never been less expensive. Frankly, we can design something we think of as a garbage device with modern technology and it's better than the best pulsometer 15 years ago and will be better than any clinician.

So I think honestly, this is an instance where the reality is this. The patients who are asking to use these things at home are quite sick. They'll have poor profusion and a risk for real hypoxemia. The price on all of this stuff has plummeted, I can tell you what the home pulse ox cost five or ten years ago, it was out of reach. The reality is, in the modern world, we're managing so many people at home that the price on these devices will be within reach within a year or two and I think if you set the standard now, the device manufacturers will be able to hit it. Not today or tomorrow, but in a few years, they can hit it and they'll hit it very well and I think our public deserves that.

Dr. Nathan: Thank you Dr. Alam.

Dr. Alam: My comment is about something else, but I would like to agree with the previous speaker, Dr. Connor. I think that's very wise and probably quite achievable and to your point. Dr. Nathan, we shouldn't prescribe how people should make it better but create standards and let them work towards them.

I wanted to speak to the issue of the Fitzpatrick scales which is something I know a little bit about. There's concern that it's hard to classify people but it's not. Most dermatologist I think almost always classify people correctly on that scale. Which is only 6 items, basically. If you're a red head, you're one, if you're blond, you're two, other Caucasian is 3, and everyone else is 4, 5, or 6 and it’s straightforward.

But I agree with your point, which is that it’s probably not necessary to again, be prescriptive or use that, it’s a relatively blunt tool to really assess and apply the technology and make sure it's working for everybody. I wouldn't use it for that purpose, but I agree with the previous speaker, we should probably look at both skin color and ethnicity. In terms of showing that whatever mechanism companies, manufacturers choose to use to increase the efficiency and accuracy of

their devices, that at the end of the day, they have made a change that is sufficient, such that now pulse oximeters work well for people of all ethnicities.

And the reason for that is because now this is out there, that they don't. So we have to reassure the public that they do. To that extent, Fitzpatrick skin types should be used just to make sure that when these devices are tested, we can say, well, sure enough, they work for all of these Fitzpatrick types and then say sure enough, they work for all ethnicities. So, again, it wouldn't be a method of using the skin type to create the technology but once the technology is fully operational, it just has to be shown that it does work for all skin types and ethnicities, thank you.

Dr. Nathan: Let me ask the panel this, because we'll come up with recommendations for the FDA, at the end of the day, based on the question, and I'm not sure if this is one of the embedded questions, or not but let's say, FDA takes some or all of our advice and there are new standards for medical grade, which hopefully they will be and consumer grade as well. That takes a while to operationalize and become available. What do we tell the FDA to do in the mean time? These devices are out there.

Dr. Alam: If I can interject, I don't think they should take any devices out of commission that are currently available, because that would get us to the issue where there wouldn't be access to any device in the interim. So maybe the approach mentioned sounds reasonable. To create some specific standards that are reasonable, and achievable given current technology, and then come up with a date, like the car manufacturers do with regard to gas mileage. Let's say, by 2030, your device should meet this performance criteria.

Dr. Nathan: Okay, sounds good. Dr. O'Connor?

Dr. O'Connor: Yes, I will answer my first question. This is just general and maybe someone can answer it another time, but I wonder how close the at home devices were, I just

went on Amazon and cheapest one was ten bucks and on average they seem to be about 25 dollars. I wonder how different the technology was from what you have on in the hospital. It's so cheap these days as Dr. O' Connor said. I was just curious, is the technology fundamentally different. I assume it's different than the Apple watch because that has so many purposes but in a dedicated at home device, I'm wondering if anyone can speak to how different it was. But retrospectively, prospective, what should we do? I don't think we should ignore the retrospective. I 100% agree with Dr. Alam, we don't want to take the device out of rural hospitals that don't have the ability to buy brand new devices really quickly, but I think recommending a situation where makers have one year or whatever it is to test under these more real life situations and then that's either published or FDA publishes a list of people who ignored that guidance, seems relevant.

Dr. Nathan: We all talk about health disparities. Usually it's not very clear what the root cause of those disparities are, or it's a whole bunch of things contributing to the disparities. Here, we have a pretty – we agree that this is a problem. And we understand what the problem is and the bias due to skin tone. To ignore these devices which I assume get used for years and years without replacement, seems to be ignoring the problem that will continue to contribute to health disparity so I think something needs to be done even retrospectively.

I think just to layer on another level of complexity, the devices has been approved based on studies in healthy patients, and now we are going to ask them to take it up a step and really regard it very similar to drug trials where we want study in sick patients, the patients that it will apply to.

It takes a while and it will cost money for these companies. And even though the technology might be a whole lot cheaper, it becomes similar to a drug study in that you need

centers and approval and IOB. It costs a lot of money and eventually they will have to recoup the money so it will edge and layer onto the expense of the pulse oximeters you have to get over these additional hurdles in the future.

Dr. Connor: And I think it's appropriate to do the lab studies but just do a more real world setting where you're not warming the finger up. Hospitals are cold but I guarantee you, Massimo is warming people's fingers before they do the official test. I'm willing to take that bet. So I think, just the FDA prescribing a real world list, putting the cuff on to limit blood flow. These sorts of things seem appropriate and actually wouldn't increase cost much beyond what they're doing today. Because you know, they're bringing in, maybe a few more people with darker skin but it's still in a control setting where you have to pay to bring people in but doing a few more iterations under real life circumstances.

Dr. Nathan: I am thinking if you want real life you are going to have to go to the hospital, you can simulate real life but it's not real life and maybe I'm being too idealistic, I don't know. Dr. Yarmus.

Dr. Yarmus: Thanks, I guess a couple of points questions for the inpatient data, we have just heard all these presentation and great studies that have illustrated – I think everybody agrees that there's clearly a problem. So in the interim while it's studied prospectively, what is the downside and or process of putting a black box warning as an example for education. And the same in the outpatient study we've all got that data, but there's – I would argue in the outpatient setting the messaging that seems to be more important is to the physician.

We have a whole group of experts here and most of us are not familiar enough with how these things were regulated. So messaging on both ends and that's probably a direct question to FDA.

Dr. Nathan: Dr. Eydelman, Do you want to address that? What is the equivalent of a black box? What does it entail?

Dr. Eydelman: So, rather than discussing the mechanism by which it's executed, it's most beneficial for us to hear what exactly is the intent and messaging and then once we hear that. We'll figure out the best way to communicate it.

Dr. Nathan: Okay, thank you.

Dr. Yarmus: Can I follow up on that? So I think part of the issue – I'm not going to speak for the group but there's a clear indication for prospective data that will take time. And we know from retrospective that patients are being harmed and that's the urgency. In the time frame where we clearly have to figure it out, is that an appropriate mechanism to use for patient protection?

Dr. Eydelman: So, Dr. Yarmus, I believe what you're recommending is some labeling or warnings in the labeling and specific education? So if you wanted to give us more insight on what type of education, to whom it should be directed, I think that is very beneficial.

Dr. Yarmus: So for outpatient from my perspective off these conversations, clearly physician education which presumedly would be an easier access point. So if I get a message from FDA that says, look, these things are not as regulated and they should be used under caution and under what they're intended for, I'm going to take note of that. On the inpatient setting, we don't have that time because we already have data that suggests that there is patient harm. So in thinking of timelines and regulatory aspects, I would think a much stricter and more straightforward message to clinicians and patients, that says these devices don't have the accuracy that we thought they did under these circumstances and they should be used under extreme caution. In my world, I will just give an example as I'm an interventional pulmonologist

by trade and airway stinting is an important patient care aspect. When stints were first introduced, they were primarily uncovered stints in the airway that caused significant problems. That was recognized, that led to an FDA black box warning which almost overnight led to a significant change in practice. I would use the same analogy here.

Dr. Eydelman: Thank you very much. So the more specific the panel could be about the key messages that you would like us to communicate as they better will be in position to hopefully make it happen.

Dr. Nathan: Okay. Maybe we can formulate something during the question period and come up with some guidance on the statement. I do wonder as well and I should know this but I don't. Between all of the great societies that are presented to us today, I don't know if anyone has a position statement on the use of pulse oximetry, but it might in addition as a bookend to the FDA coming out with something. Maybe there should be a position statement from all of the societies together and it may have a little bit of weight as well in the medical community. Okay. Dr. Yarmus, you had your turn. Dr. Wilson?

Dr. Wilson: Yes, just in response to the question you asked a moment ago of what the FDA could do in terms of consumer products. In addition to developing the new standards which has been mentioned by a number of our colleagues. The FDA could send out a notice warning to physicians and patients that if the patients are being monitored at home for a medical condition, it should be with a medical grade pulse oximeter. That would at least elevate further the concern to those who just think they're getting something off the shelf that should work, like a thermometer.

Dr. Nathan: I agree and I think that could be a recommendation but the thing that I don't know is, medical grade, are you always going to get one if you write a prescription?

Because people might be under the impression that this is a big device and I paid a lot, and it may not be medical grade. So aside from getting it by prescription, is there a clear designation on the box or otherwise this is medical grade?

Dr. Wilson: Perhaps, Dr. Eydelman might be some advice on that? Maybe some new requirements for labeling.

Dr. Nathan: Yes, Dr. Eydelman.

Dr. Eydelman: So prescription use devices, once the physician writes a prescription for a prescription use device, only the -- sorry, my mic doesn't work well. Dr. Lee, why don't you answer while I fix my mic.

Dr. Lee: Thank you, So certainly, I think, the requirements on accuracy, that is something we would look at in the final labeling for a cleared medical grade device. I think we discussed this on the top. You know, we do publish, we have the manufacture publish even the Bland-Altman information so as a clinician, you can see information and some companies do disclose a very granularly information about performance across different [indiscernible]. Did this answer your question?

Dr. Nathan: Okay, Mr. O'Brien.

Mr. O'Brien: We have a problem where we have a diversity problem due to a medical inaccuracy in a device that is a real contentious issue because it's life threatening and it's a disparity that shouldn't exist that we're trying to eliminate in all aspects of our life that is there. I'm afraid that having a simple solution that just says, listen, if you're using this for medical reasons, you have to go back and get a medical prescription or a nonprescription and pay X number of dollars for it and we're going to create access problems to people who need it and

create a political diversity that is really going to be problematic. I there has to be a lot of thought given to what seems to be a simple action before it takes place.

Dr. Nathan: Thank you, I'm keeping an eye on how long we have gone. We have had gone about an hour and 45 minutes of discussion. We can spend a long time discussing it. It was allotted for two hours so we'll go for fifteen minutes more. I do want to make sure that everyone on the panel feels like their voice is being heard and I think everyone has participated. I just want to put it out there for folks who maybe haven't said as much. I just don't want to leave anyone out of this discussion. Dr. Kirsch, I see you have your hand up.

Dr. Kirsch: One thing I don't think has been touched on is for those of us who have a non-medical grade pulse oximeter which I have one. It can result in bad information that results in additional testing that is costly and may actually cause harm. So the case study is, I happen to enjoy skiing at high altitude and I took my portable non-medical grade pulse oximeter with me at high altitude skiing, and found, despite any shortness of breath or any problems skiing pretty aggressively, my pulse oximeter reading was the mid-70s while I was skiing and resulted in – I just continued to ski but I went back to my internist afterwards and it resulted in a variety of tests that were negative. A bad result, resulted in tests that were unnecessary, very costly and theoretically dangerous. might have caused me a complication from the additional test itself. Just another perspective on why I think tightening up the quality of the device is important.

Dr. Nathan: A very good and valid point. The cascade effect you find when you have an erroneous reading and then before you know it, you have a major work up. I was going through a list of participants and I want to give Dr. Lynch an opportunity to weigh in.

Dr. Lynch: Hi, thanks, I agree with all the comments, I think one thing so far, in thinking about how it can be approved, I think the dermatologist said, you don't need to have an

algorithm correction and not an instrumentation correction so just understanding how that would play out and someone else commented on structural heart disease and the pulse oximeter and the heart disease, potentially, missing the diagnosis, which I think that's a theoretical problem and not an actual problem because you're comparing differences in a pre-post and post doc stat in someone with a certain skin pigmentation so you're comparing two values, on that same patient, So that is probably more of a theoretical issue than an actual issue. The bigger problem is, when we set these guidelines, below a sat of 92, or below a sat of 87, you need to get this treatment, that's the absolute number and it needs to be verified in the different skin pigmentations.

Dr. Nathan: Thank you, and Dr. Katz, anything to weigh in?

Dr. Katz: There's a couple of things that struck me. The first is that we have a lot of children or preemies who are hypoxemic, or not hypoxemic, but when they're ready to leave the hospital, they're often at the [indiscernible] the physiologic nature of their hemoglobin levels, so not only are the potentially hypoxemic, they're also anemic. So that is a population of kids who are never going to be able to be studied. Because they don't have arterial lines, we're never going to be able to get data on that and then we have a large population of older kids who are asthmatics and have various conditions also who are not going to be amenable to this study. So the other thing I wanted to point out is, we have experimented in our sleep lab, and looking at data which has taken five or ten minutes apart is very foreign to us in pediatrics because they're moving around so much that data just has to be very inaccurate so it's going to be really important to get this FDA sponsored data out of California to see if you can get some clarity on this issue.

To me, the point it not that the over counter products will work less well and some have been taken off the market because parents were using them in pediatrics as apnea monitors and

they cause more anxiety and have no beneficial effect on outpatient and infants so I think, the bigger problem from my standpoint is the fact that even the medical grade, machines are demonstrating errors in 15 to 20% of patients and even in Caucasian patients, they're causing significant harm potentially. So I do think that, a warning is indicated because if people are relying on these devices, and are comforted by the fact these things are accurate, we become so accustomed to using them that we have forgotten there are strengths and weaknesses of this.

Dr. Nathan: Thank you very much, I'm not seeing any other hands. I think there will be opportunities to --

Dr. Cassiere: I'm sorry, I have had my hand up for literally twenty minutes.

Dr. Nathan: I apologize. It's hidden in the corner there.

Dr. Cassiere: I don't mean to interject but I just want to make a comment and a reset. I think the comment and reset is, just looking at the total data we're looking at. We're looking at and trying to tell companies FDA recommendations on what to do moving forward and what to do retrospectively. I think and I hope people can agree with this, what we're really looking for is desaturation studies under low profusion states. That can be simulated in healthy people and that should be the expectation on the bar that these companies need to meet in all skin tones. And there should be an even distribution of all of these normal volunteers under desaturating states and hypo profusion states with equal representations in every single one of those groups and in particular, desaturation states where it will make a difference in terms of therapy and that is, we can agree, maybe like 94% or 92% but I think that is really what the FDA is looking for.

What do we want the companies to do prospectively, that will prove their device and what are the expectations going to be retrospectively for the devices that are out there? And I think looking at all of the data presented, it's not just about skin tone. It's about skin tone and

hypo profusion. I think that's been hammered home a bunch of times with the speakers and some of my colleagues on the line. I just wanted to reset. That's what we need to look at moving forward.

Dr. Nathan: I agree. And I think right from the get go, we had outlined a broader group of patients for sure under different conditions, now whether you can accomplish that in normal healthy volunteers versus in a hospital setting I'm not sure we need to be too prescriptive I'm worried about making patients normals hypoxic, and hypoperfusion them, although you can do it locally, but I'm not sure how much that replicates in hospital conditions but it is a start.

As I say, I don't think we can be too prescriptive in the exact studies you need to do and maybe I don't know enough about what can be done outside of clinical research facilities where normals come in all of the time for the phase I studies. maybe this can be accomplished or in a hospital setting, which ever works best. As long as there's good data to support beyond the 90, 92% and above, where most of the data sits and to represent all racial tones I think is the appropriate messaging that will come out of all of this, and a tighter range in terms of arms so I think we started with that at the beginning of this very productive discussion, and we can bookend with the same message with a lot of great discussion in-between. If everyone is okay with it, I don't see any other hands up. Oh, we have two hands, okay, Mr. Branson?

Mr. Branson: I agree with what was just said. In deference to the physicians in the hospital, the nurse and the respiratory therapist are the two people who put the oximeter on and look to it and the messaging about this inaccuracy needs to go directly to them and to know what they should do about it. That's an important point of whatever FDA does. I agree with making the ARMs be better. I did want to comment with, for one of the speakers talking about skiing, if you're healthy enough to ski at fairly high elevations in the mountains, I spent a lot of time in the

back of an aircraft and an altitude chamber at eighteen or twenty thousand with an O2 sat of 8s% and had no shortness of breath whatsoever. So it's not just a matter of hypoxia, it's also a matter of cardiac output and the ability to hyperventilate. Is may have been possible that his SPO2 was that. But he's healthy enough it didn't affect him at all and the device may have been accurate.

Dr. Nathan: And there's the entity of high altitude pulmonary edema and it could have been that as well. All right. Dr. Wilson, our industry representative in terms of what we are saying and what the FDA will say to industry. Maybe you can speak to that.

Dr. Wilson: Thank you, So the first point is of course, not all pulse oximeters are created equal, some are better than others but we can hold our manufacturers to a higher standard. That should be the goal. The FDA believes is the new percent or number, or as Dr. O'Connor said, let's make sure that the P value is powered properly so that we can ensure that we have the right mix, that has to happen.

The FDA has to make a decision on how to tighten the ARMS because as someone said, even some of the commercial over the counter pulsometers are better than what we had a couple of decades ago because of the miniaturization and so forth. But the point I wanted to make though, but thank you Steven for inviting that comment. But the point I did want to make was, a little bit related. It has to do with the FDA does allow for extremities to be warmed and I think the FDA should tighten that requirement.

In general, they should not allow any warming of extremities if it's -- it should be basic conditions about room temperature and so forth but then, the endpoint should be, what is the profusion index. So if the goal is to have higher than normal profusion, then the FDA could allow some warming in order to get the profusion number but in general, it shouldn't be allowed unless you're trying to augment that upwards and if you're trying to decrease profusion, there's

other factors that one would need to do in a laboratory setting to decrease that. It might be colder, ambient temperatures, and things like that in order to bring the profusion index down. But those would be better endpoints and I think just a little bit more scrutiny about if we want to be able to compare apples to apples, then the experimental conditions need to be clear.

If we do use profusion index, let's say as a measure, then that can cross over to the clinical domain and in terms of the measure of how well the extremity is profused.

Dr. Nathan: Thank you for those comments. We'll end with Ms. Edwards and adjourn this session. Go ahead Ms. Edwards.

Ms. Edwards: I just have a question. So are we moving away from the ethnic implications to more of a pigmentation? I guess the reason I'm asking that is because I'm thinking about COVID-19 and how we got here to this point. You get what I'm saying, that's my question and my concern. Pigmentation versus ethnicity. I'm wondering what the implications are there.

Dr. Nathan: I think the two are closely linked for the most part. I think ultimately the goal is to make sure that anyone who has a pulse ox put on them has an accurate representation, whether it's different ethnicities or skin coloration, or pigmentation, so I am not sure if we need to divorce the two. I think if we get at one, we'll get at the other. But ultimately, if you look at the objective measure which is measuring the amount of melanin in the skin, that's skin pigmentation thing which is linked in many ways to ethnicity. Hopefully I addressed that okay.

Ms. Edwards: Yes, that was my concern because the African- American community, the disparities in care there is how we got here. I just wanted to make sure it wasn't lost.

Dr. Nathan: Absolutely not at all. But I think a point that has come across is, a rising tide carries all boats and we don't want to get to the African- American community to where the White community is here. We want to get everybody better than that and equally.

Ms. Edwards: I get you.

Dr. Nathan: With that, let's adjourn the panel discussions it's been very interested. A lot of great comments and I have three minutes to five. Let's make it 5 after 5 to get back together and we'll address the FDA questions at that time. Thank you.

FDA QUESTIONS TO THE PANEL

Dr Nathan: Welcome back everyone. We are on the final leg of a busy day, but a very interesting day. At this time, we're going to focus our attention on the FDA questions. Panel members have received electronic copies of the FDA questions and they'll it post on the FDA website.

I would ask that each panel member identify him or herself each time he or she speaks to facilitate transcription once again. All right. Please present the first question.

QUESTION ONE

Dr. Nathan: I'll read the first 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 and we'll break this out, but specifically please address, this will be the first part.

Does the currently available clinical evidence demonstrate disparate performance in patients with darker skin pigmentation? Maybe we can break it down even further and that'll be the first question for us and.

I'll speak for the group because I think you, you heard you didn't hear any dissent, but I think the current available clinical evidence does demonstrate disparate performance in patients with darker skin pigmentation.

I think we saw a lot of studies and everything pointed to the same. I know that they're, some of them were retrospective and they, some of them, the ABG wasn't done exactly at the same time. But how much can change in 10 minutes? In some cases it might, but for the most part, it's a relatively stable thing. So in my mind, and I don't want to speak for the group, if anyone else from the group has a dissenting opinion about that, I think the evidence we saw was quite clear.

Anyone want to add to that or reinforce that?

Dr. Loeb: Absolutely. Robert. Robert Loeb. I would only add that the early literature. I did have questions about co-founding confounding variables and covariates and things like that. The, but I was convinced in the presentation as the presenters went and they generally went in order of when studies were done. I am now convinced, I wasn't initially, but the latter studies that were doing more sophisticated analyses. Yes I agree that it, there is a problem. It has clinical implications.

Dr. Nathan: And I think they, I think they brought us along very nicely in terms of the presentation, but there's biologic plausibility, why that should be the case. It's not just, it happened and it was a chance finding. There's good reason for this. Dr. Hennessy, do you want to weigh in on that?

Dr. Hennessy: Yeah, I agree with everything that you said. In addition to being a higher risk for hypoxemia Black patients – the performance of the test is measured by the area under the curve is worse in Black patients than White patients.

Dr. Nathan: Mr. Branson.

Mr. Branson: I agree. There's definitely a single signal there. All I would say is that I think it is exacerbated in the covid, in COVID- 19 because so many of those patients were not only hypoxemic, but they also had viral sepsis and poor perfusion.

Dr. Nathan: I think certainly COVID brought this issue out to much more to the open. Dr. Connor.

Dr. Connor: Yeah, just to strongly agree, but also to reiterate the point I heard, during some of the open session, I think some of the makers of these devices point out some of the imperfections in the real world clinical studies. But I think the fact is, while that adds noise, we repeatedly saw that the performance was worse in darker skin patients.

I just wanted to point out that we shouldn't let some of the imperfections in those studies, which are really just real world experience, negate from what is a clear signal with skin pig mutation leading to poor performance.

Dr. Nathan: All right. Thank you. So it seems actually, Dr. Color, let me address the second part of the question, because there may be, I should have done that initially because they, they're built into each other and then I'll come back to you.

If so, do you believe such disparate performance may lead to increased risk?

And I think My bias is that the evidence was there, that it does lead to increased risk. Dr. O'Connor, why don't you go ahead and weigh in.

Dr. Collop: Yeah I guess I have more of a question than a comment. It looks like most of the data that we heard today, which I agree, is overwhelmingly that the darker skin pigmentation makes a difference, was on prescription use oximeters, and I don't know how much data we got on the over the counter pulse oximeters, and if I missed it, I'm sorry, but it seemed like all of the, real world data was on the prescription use. So I don't know how best to comment on the over the counter part.

Dr. Nathan: Yeah, I think, yeah, you're right because all of this was in hospital data, so it had to have been prescription use. I think the assumption is if that's how prescription use oximetry performed, then consumer devices which aren't held to the same standard have to perform just as bad, if not worse.

Dr. Collop: No, we would assume so, but , I don't know that we know that for sure. , but I agree.

Dr. Nathan: Okay. Does anyone descend from the fact that these devices and their performance in folks of darker skin lead to increased risk and worse outcomes?

Okay, so I think we do have consensus there.

Please include prescription use, actually, this is getting to what Dr. O'Connor mentioned, and over the counter pulse oximeters when used for medical purposes in your deliberations.

We had a very robust discussion about the over the counter pulse oximeters and so actually, Dr. Color, I think you helped me answer that. The third part of that question, I think the assumption is that over the counter devices perform in all likelihood, even worse, without seeing studies specifically attesting to that because they're held to a very different standard.

Does anyone else want to weigh on over the counter devices?

Okay, let's move on to Part B. Do you believe the reported disparate performance or increased risk may be explained by factors other than darker skin pigmentation, such as perfusion, index and motion artifacts? Dr. Cassiere, do you want to try and tackle that first? You muted if you're talking. I don't know if you wanted to weigh in or if your hand –

Dr. Cassiere: Can you hear me now? Can you hear me now?

Dr. Nathan: Now we can hear you.

Dr. Cassiere: My computer literally booted for an update on a new computer.

Dr. Nathan: Oh gosh, sorry. Okay.

Dr. Cassiere: But what I wanted to say is that at least the way I read the data and the way it's played out, it looks like it's not just the difference in skin pigmentation, but combined with a low profusion index. In other words, patients who are critically ill or sick in the hospital, by definition, there's profusion issues. And this is amplified by the fact that there's also racial skin pigmentation issues that exacerbate it. So I think that those two hand in hand are the source of the problem.

Dr. Nathan: Yeah, I do remember that. I think that one slide with the Bland-Altman plots showing the combination of pigmentation and perfusion really caused a very wide dispersion there. So I think the one feeds into the other. I would agree with that. Dr. Yarmus.

Dr. Yarmus: I agree and I would just add that I think the signal of that 13% versus 10% right, that 10% difference in White patient still shows there's an issue beyond race and that should be based on what we heard include these types of parameters.

Dr. Nathan: All right. Can you–oh, some more comments. Dr. Connor?

Dr. Connor: Yeah, strongly agree with the perfusion index comment. I think one other thing that would be interesting to make sure is studied though is, finger width. We saw in some

of the papers that we were given to review that there was less of an effect with race or skin pigmentation in infants or kids, which has a smaller path length for the laser to travel. And because obesity in, at least in the U.S. is higher in some, non-Whites. I think that could confound that. I certainly believe that skin pigmentation is the root of the problem, but we also wouldn't want to ignore, for example, if this didn't work as well in, light skin people or White people who were heavier as well.

So I think that's something else that was unclear to me or not yet quantified. That's worth considering.

Dr. Nathan: I think it's worthwhile reinforcing, and it did come up in terms of just studying these devices in a wider population, including different body habits, especially because of the obesity issue in the U.S. in particular?

Can you go down a little bit on the question? I just want to make sure we're not missing anything. There isn't a part. C, that's all there is. Okay. I can see. Dr. Collop once again. Why don't you weigh in?

Dr. Collop: Yeah. The only other thing I think we heard also about maybe the probe itself may play a role in some disparate performance, which gets to the prior person's comments, about maybe bigger fingers, littler fingers, what have you. But I think one of the presenters suggested that, if it's too tight or too loose or there's light coming in, those also seem to, may cause some air.

Dr. Nathan: All right. Let me try and summarize this. Dr. Eydelman from the FDA with regard to question one the panel believes clearly there is a disparate performance in patients with darker skin pigmentation, and this does increase these patients risk for whatever disease

they might be having. Even though we didn't see data on over the counter usage, we know that these are used by patients on the outside to gauge whatever illness they might have.

So we believe that there is this effect that we are seeing from the numerous studies on inpatients probably has ramifications for the outpatient, over the counter devices as well. And so that's a summary with regards to a, we do believe that other factors and clearly perfusion plays a role, not just pigmentation, but perfusion.

And I think something that the panel picked up on not only perfusion pigmentation, but. perhaps demographic factors in terms of the width of the finger, the breadth of the finger, obesity being an issue, and so there are factors just beyond pigmentation, but certainly pigmentation is the main reason why there is this disparity.

But as the FDA looks at this, I guess we would encourage the FDA to consider these other issues as well that can be addressed perhaps in some fashion at the same time.

Dr. Eydelman: Thank you very much, Dr. Mason.

Dr. Nathan: All right, Dr. Eydelman, is this adequate or is there anything further you wish us to address with regards to question one?

Dr. Eydelman: No, that is definitely adequate. Thank you very much,.

Dr. Wilson: Dr. Nathan. May I add?

Dr. Nathan: Is that Dr. Wilson, please? Yeah, please go ahead.

Dr. Wilson: Just from, so two points. First that we're all in agreement that. The preponderance of data shows that there is a signal however, and that's despite having some methodological concerns. However, all of the clinical data had all of the various manufacturers combined, they were not separated out. So it would be appropriate to say that there could be

differences between different manufacturers that you don't really know. On average there's a signal –

Dr. Nathan: Yeah, I don't know that we need on that –

Dr. Wilson: I just want to finish the point, which is that since we think that perfusion and motion is important, that some manufacturers have worked on those problems and they're less problematic. So it is important to not call out any, not to say that any company has a solution just yet, but that there may be there, there may be more work to be done by some manufacturers than others. That's essentially the point.

Dr. Nathan: And as I said, I don't know that we need to drill down to that. I think everyone will be held to the same standard eventually. Some companies might be closer to that standard or at that standard already. And, but your point is taken that different devices were probably used in these clinical studies doesn't take away from the signal that was ubiquitous across these studies. Okay. Let's move on to question two then.

QUESTION TWO

Question two. There are several tools to assess skin pigmentation, including, but not limited to colorimetry, spectrophotometer, melanosome volume fraction, and skin color scales like the Fitzpatrick and van Luschan color scales. Please provide recommendations for studies evaluating pulse oximeters for the following: standardization of skin pigmentation assessment.

Dr. Nathan: Dr. Cassiere, why don't you go ahead.

Dr. Cassiere: Yeah, I'm going to be the contrarian again and say that I think looking at the totality of data and what's been reported, it's mostly self-reporting. We could, we should have an expectation of self-reporting and if you'd pick a color scale to go along with that, to report both, to maybe have some kind of standardization because we just do not have enough studies that

point to skin color scales to find – pointing out the problem that we found. We found the problem through self-identification, not through skin tone studies.

Dr. Nathan: Dr. Loeb, why don't you go ahead.

Dr. Loeb: I agree. And I think that the patient advocate Ms. Edwards brought up a very good point that this began or really is seen as a racial/ethnic issue and that it would be wrong to just go with some standardized assessment of color.

I agree that two standards should be used, both some sort of standardized assessment of color and I would add color at the site of the measurement. And also self-described ethnicity, and that the analysis of results would be on both subgroups.

Dr. Nathan: Okay. I'd venture to say that, I'm not sure how prescriptive the FDA should be, as long as the assessment spans a spectrum of coloration and skin pigmentation. I'm not sure primarily because I'm not sure what the best scale is, but whatever scale is used, it should go from 0 to 100 lightest to darkest. And I'm not sure if the FDA wants to be prescriptive to the manufacturers of "this is the scale to use".

I think the objective measurements that were mentioned have a lot of attraction in them because they are objective and even though the color scales perform well I do think that there's probably some advantage to using these more objective scoring systems.

Perhaps in terms of standardization of skin pigmentation assessment. Perhaps it's more important to say that the full range of skin pigmentation be evaluated. Would that be fair and encompass everyone's opinions or do you feel that we as a committee need to come down on one thing versus another? Dr. Hennessy, why don't you go at it?

Dr. Hennessy: Thanks. So I'd say two things. I'd say that race is a social construct and that it, in my mind, it doesn't make sense to evaluate the comparative performance of a medical

device based on a social construct. I think it makes more sense to look at skin pigmentation. I think that spectrophotometers are an objective way to measure skin pigmentation. It could be applied to the site where the pulse ox is going to be applied, and it's not a particularly expensive piece of equipment. So I don't think that's a conclusion, but that's the direction that I would lean in.

Dr. Nathan: Dr. Yarmus.

Dr. Yarmus: I am going to just mirror what Dr. Hennessy just said. I think it was perfect. I think we have to hedge a little bit just because of that, the panel expertise, right? We have one, one dermatologist. If I go to Home Depot and I want to match a color of a paint, I can have them do that, right? So I think from a technological standpoint, there are ways to do this that I think are beyond my expertise. But I do think there should be both a subjective and detailed assessment on both ends.

Dr. Nathan: Ms. Edwards.

Ms. Edwards: Hi. Yeah, so I agree with Dr. Loeb, and while I agree that race is a social construct, it’s one that's used to discriminate and it's caused a lot of disparity in healthcare for African Americans. So that cannot be I think, made to be a small issue when it's really a great one.

Dr. Nathan: I think it clearly is, and I think what I would say is if we go with the pigmentation, which is more objective, I think that encompasses the race aspect and it takes care of African Americans. But to the point that was made earlier, folks who identify as African American in terms of skin pigmentation can be very light or can be very dark. But if we objectify it with one of these different measures, the spectrophotometer being one example that I'm

unfamiliar with, it's not my era of expertise. I think it takes race a little bit out of the equation, but addresses race at the same time, if that makes sense.

Ms. Edwards: I guess my fear is that historically, the disparity in healthcare is like we – it never stops. And so when you try, when you lump African Americans in with everybody else, we always end up on the short end. And so I just, I guess I have a problem there, but I that's just my perspective. I because this started 30 years ago, right? And here we are today, We're just addressing it. Yeah. That's just, that's my opinion.

Dr. Nathan: These studies can be done to address both. I think certainly the part of any study is to identify ethnicity or self-identified race, but at the same time to have the objective measure at the same time. And I think what you'll see is if you address the spectrum of pigmentation, you're going to have great African representation as well as other ethnicities within the group to get these devices validated.

Ms. Edwards: And I guess FDA would just have to make sure they give some pretty strict guidelines as far as tests – the testing rather than, the two darker.

Dr. Nathan: I think that became clear. I think that there needs to be equal representation because certainly that's where the problem lies, is with folks who have greater pigmentation.

Ms. Edwards: And to hold manufacturers accountable. Because I think that's where too, it gets a little murky too, because I think in some of the testing, it came out that some of the manufacturers didn't even follow that guideline.

Dr. Nathan: Right, Dr. Collop.

Dr. Collop: Yeah, just want to agree. I think right now most of the literature would suggest that it is race related based on how the studies have been done. So I think we need more studies to see if it is truly race related or if it's pigment. Physiologically, we think it's probably

pigment, but we don't know and I think we need to confirm. I think both should be identified, both race and some measurement of pigment to confirm that's the case.

Dr. Nathan: I agree. I think a bigger danger is if you say 50% African American and you get 50% very light skinned African American, then it doesn't address the issue. So I agree with both being addressed. Dr. Connor.

Dr. Connor: Yeah, you just made the point I was going to make that is playing the skeptic about device manufacturers and I think most have patient best interest in mind, but I would worry that, if you need X African Americans, you may recruit because this, if it's not in the clinic, and if it is in the lab, they can recruit the people who, they know their system may be less biased in, so therefore they're checking the box, but they're not really solving the problem.

So I think making sure we have enough people, on, for instance, the Massey Martin Scale that are dark skinned. Maybe we also, given this is the United States FDA would also want to make sure that those, darker skin types are represented by enough African Americans, not just dark Indians or dark people of other races. That way we make sure that African Americans are included in that to address Ms. Edwards very important concerns. But we're also, like you said, making sure we have enough, dark skin people so that manufacturers can't game the evidence they need to bring.

Dr. Nathan: Okay. Let's end off with Mr. Branson and I'll try and summarize.

Mr. Branson: Again, I appreciate Ms. Edwards viewpoint and she's the person of color on the group, but I also know currently we are finally doing away with race-based equations for predicting pulmonary function and we are doing away with race based equations for renal disease. And in the past, race based differences were exploited by White supremacist to suggest there's actually a difference between us, other than the color of our skin. So I think we have to be

careful there. Again, and I still, objective majors are for science. But I do appreciate exactly what she says, race definitely has to be included.

Dr. Nathan: Okay, thank you. Dr. Branson, I'll try and summarize this with regards to question two. The panel generally believes that yes, there should be standardization of skin pigmentation assessment, which should include an objective measure ideally, but can also include one of these more subjective Fitzpatrick scale.

I think it should be easy enough to include both of them just at something else that the companies can do just to see what the correlation is. I think that'll be – move the field forward a little bit. But definitely race identified race should be included as part of the demographics captured, and there should be equal representation across the spectrum of skin pigmentation.

In terms of categorization, reporting of skin pigmentation, I think I've alluded to that already. Categorization ideally based on objective measures, but also can include these visual analog scales to complement that. And the reporting of skin pigmentation can be categorized by all the mechanisms including the subjective color scales as well as easy enough to do subgroup analyses based on race. If there is the desire to do this, but to the point that Dr. Branson made the data will exist, but you made an important point in terms of trying to bring the races together in the context of all these things and just differentiate people based on pigmentation.

Dr. Eydelman, I feel like I danced around that a little bit, but did the group, and did I summarize this adequately?

Dr. Eydelman: Yes. Thank you.

Dr. Nathan: Thank you.

Dr. Collop: Can I add you did not in your summary, say where the color imagery would be performed, and I do think it's important that it be done at the site of the measurement.

Dr. Nathan: Yes, I agree. I think that is very important. What I didn't also add is, and I'm not sure how far the FDA should go with this, because to the point I had made earlier, you put it on the finger, don't get a good signal. Put it on the earlobe, put it on the forehead. I would encourage the inclusion of multiple sites to fully evaluate the performance characteristics of the devices.

All right. Let's move on to question number three.

QUESTION THREE

FDA currently recommends assessment of the effectiveness of pulse oximeters using ARMS, which is the root mean square of pool data pairs and adherence to the currently recognized ISO 80601-2-61:2017 standard for this variable. Currently, pulse oximeters are expected to have accuracy within one standard deviation 66% of the time, and within two standard deviations 95% of the time.

Please address the following, Please discuss how accurate pulse oximeters should be for clinical use. In your discussion, please address whether the accuracy varies based on the clinical setting or the levels of oxygen saturation.

Dr. Nathan: Well, I think, I'll start off the group. I think what we heard is clearly the ARMS need to be tightened. How much it should be tightened? I don't know. I don't want to throw out a number. I don’t know if statisticians have a better idea of how tight we should recommend to the FDA for the arms to be tightened. We, can get to that, but I'll let Dr. O’Connor weigh in at this point.

Dr. O’Connor: First, I'm sorry, I lost my internet. That's why I was absent for the first question. I think that the clinical data suggest strongly that in ARMS of, between one and a half

and 2% in the range of say 70% to 92% or 94% would be significantly better clinically than in arms of 3%, which is the current FDA standard.

I think the second thing the panel's likely to agree upon is that we need to have data about positive predictive value for each saturation that a pulse ox might report, and a range of 95% confidence interval. And we, [audio dropping] the criteria for developing a pulse oximeter should require a sufficiently large population of patients with different pigments. However the FDA decides, or the manufacturers decide to determine that, that the 95% confidence intervals for the Caucasian and the African American, or the very darkest skin to people will be very similar, both in terms of their accuracy and their precision.

Dr. Cassiere: Yeah I'm in agree with Dr. O'Connor. I just want to make a – I think we really need to tighten up in the O2 SAT ranges especially, below 92%. I think the 2% is definitely reasonable and also want to highlight we have to, pulse oximeters in the past were used as monitors. Now we're using them for therapeutic interventions, that's a very different mindset. I am going to start or stop a therapy on you based upon your peripheral O2 saturation. That should be held to a very different standard, and most therapeutic interventions are going to happen in the lower oximeter ranges. So really need to tighten that up. In terms of, that 2% sounds fine with me. The lower the better. I don't know the technology enough to say how low can you go? But we could be California in turn, all electric cars in 2033. We could tell manufacturers by X date, we expect your oximeter to be within 0.5 to 1%.

But I'm saying really low cause we're making therapeutic interventions, withdrawing therapy and adding therapy based upon this. It's no longer used as a monitor.

Dr. Nathan: So let me try and summarize, unless anyone else wants to weigh in on A, which kind of we've addressed a little bit of B already in terms of folks with different skin pigmentations. Dr. Lowe I see you have your hand up and then Dr. Katz.

Dr. Lowe: I know I talk a lot. Sorry about that. I think there is a problem with ARMS and that was presented, and it's not the way clinicians think. It was brought up and I agree with Dr. O'Connor, that the problems that we see in the clinical studies are that people have a normal saturation, but they're hypoxemic. That's a big problem. It should be – it's doable and it should absolutely happen that high saturation, if you have a pulse oximeter with a high saturation, there's virtually no chance that you're going to have a true low SaO2 function, a functional low saturation.

So the high levels it can be, and it should be very very accurate in a subgroup analysis just at those levels. As the levels go down, technologically, unfortunately, I'm not sure that it's achievable, that it, that level of accuracy is maintained. So I would be okay with the level of accuracy being somewhat lower at the lower saturations, especially if the curves are going to have to be done to account for disparities in pigmentation.

I would say that, I agree that there should be something more clinical in terms of clinical thresholds and making sure that people don't get misclassified at 92% and 85% or something like that. I think we can choose almost whatever number, but something up in the nineties and something down in the eighties.

Dr. Nathan: Thank you, Dr. Katz.

Dr. Katz: Yeah, I think the discussion mostly deals with 0.2, the levels of SaO2, which I agree with. A-1 is the clinical setting, and I think what they're referring to is low profusion states. So if a company is going to make a claim that the oximeter is accurate in low

profusion states, they should be required to study it under those conditions, which may be very different than providing the FDA with data during wakefulness of wake under normal conditions.

Dr. Nathan: Thank you. Dr. Lynch.

Dr. Lynch: Just want to comment on what Dr. Loeb just said about him being okay with worse accuracy at lower saturations. I feel like that – this entire day we've been focusing on hypoxia as like the pathological state and people that would normally be at 99%, but there's a growing population of congenital heart population where their resting saturation is 80% or 75%.

And we also need to know when they have pulmonary illness, so we would need an accurate pulse for these patients in that, in those percentiles as well. So I would just wanted to point that out that it's not that everyone needs a more accurate pulse oximeter at the higher percentages.

Dr. Nathan: Dr. Olson?

Dr. Olson: Sorry, I initially raised my hand to counter the point that Dr. Low made about it's okay to have a wider ARMS or whatever the statistical methodology is at the lower saturations. That's where it really matters the most. And currently I believe all manufacturers can easily have a 2% ARMS across the entire spectrum, 70 to 100%. And if, in the future, you made it 1.5%, something that they would try and achieve in the next couple years, that's very doable as well.

Dr. Nathan: Thank you for that insight. Dr. Connor.

Dr. Connor: Yeah, I think one other metric that I would recommend, because the issue is now we know there's this racial or skin tone discrepancy that I would recommend, say, a minimum number at 88 or at 92, I would let the clinicians recommend what the right number is, but then show what is the mean bias, for instance, between Massey Martin 2s and Massey Martin

8s or something like that. Cause I think the problem is right now you can hit the ARMS limit because you don't need many dark-skinned people. So you know, the 34% of people you're allowed to have outside the range, or 34% of measurements can just all be in a type of person that leads to racial discrepancies.

So now, if we have doctors who care about racial differences, and hopefully 100% do. This one, enable them to look at the label and say, Oh, this one has the smallest bias between light skin and dark skin patients. So I think I totally agree with what others have said, but I think we need a clear, here at 88% saturation according to an arterial measure. This is the bias between light and dark skin patients, and then we can comparison shop accordingly.

Dr. Nathan: Thank you, Dr. Cassiere.

Dr. Cassiere: I just wanted to step off on that. Dr. Connor, thanks for pointing that. I think that the error rate should be an individual skin tone racial groups as well, not just the whole group. And that's calling the company out and saying that you can't just mix the data up and you made the threshold. I want to see the threshold in all the different color tones. I think that should be the expectation.

Dr. Connor: Right, and maybe do two fives and eights. We don't need all 10 on the Massey Martin scale, for instance, but right, at least two fives and eights is a minimum. Absolutely agree.

Dr. Nathan: I see Dr. Eydelman has her hand up. Dr. Eydelman. Are you muted?

Dr. Eydelman: Can you hear me?

Dr. Nathan: Yes.

Dr. Eydelman: Okay. I was just trying to respond to something Dr. Connor said. I just wanted to make sure it's clear that we are trying to get a recommendation from the panel for the

minimum acceptable parameters that would be – that would be adequate for a new pulses oximeter to go on the U.S. market.

So in other words, it is not for the purposes of comparison shopping, but what is minimally good enough. Thank you.

Dr. Nathan: Thank you.

Dr. Connor: No, and I agree. If I can address that, but I think showing what the difference is in light and dark-skinned people. So maybe it's not comparison shopping, but maybe it's to say, This is good enough and I'll buy it, or this is not good enough, and I won't use it in my hospital. So I, I didn't necessarily mean comparison shop, but it's, I think, the key metric that this whole panel's about. There's a difference in light and dark-skinned people, put that number on the label.

Dr. Nathan: Okay, thanks. Dr. Hennessy.

Dr. Hennessy: So we've been talking mostly about ARMS, but we haven't talked about its uses diagnostic tests to identify hypoxemia. So should we recommend that sensitivity, specificity, positive and negative predictive values for identifying SpO2o less than 88% be a requirement and both overall and in subgroups based on skin pigmentation.

Dr. Nathan: I'm not sure we should go there. 88 is the number that's charted as being this magic threshold, but it's really a spectrum as well. In terms of oxygen saturation, let me see if I can try. There's a lot of information to process and this is going to be a difficult one to summarize, but let me see if I can and try and I'll give it my best shot.

So actually, yes, it should be based on the clinical setting. We, healthy volunteers is one thing, but in the hospital setting or at least simulated hospital setting, I think becomes very important. We heard about the importance and the influence of hypoperfusion, and certainly

you've heard a lot about the level of oxygen saturation. So there needs to be validation across the SpO2 spectrum from 70% up to a 100%.

I heard I think from Dr. Connor reporting on bias between folks of different skin pigmentation. But what the FDA wants is the minimal requirement to get approved. And what I would say is if we do, if the FDA does mandate that the ARMS gets tightened to, we heard that 2 is achievable, 1.5 is more achievable. As a clinician, I have a tough time contextualizing what this means for my patients in terms of how reliable the oxygen saturation becomes.

All I know is that it does become even more reliable and I don't know how prescriptive we need to be beyond that. But what I would say is that threshold needs to be applied not only 90 and above, but 80 to 90. And I don't know if it's feasible to have the same kind of tight range between 70 and 80. So what I don't know – we know that pulse ox gets more inaccurate as you go lower. What I don't know is, does that make the ARMS of 1.5 less achievable as you get lower? And I think someone with technical expertise can answer to that.

I think that what we should strive for is not the bias between different pigmentation. I think that all pigmentation types should be represented equally. And what I would also guard against is having a label with all these different things, bias ARMS, that becomes uninterpretable to the average clinician.

So once again, this is a difficult one. I think the FDA has a lot of information and input from us, and hopefully we've addressed at least 3-A adequately. I think 3-B, we've clearly addressed that there should be equal representation across different skin pigmentation. So that's an easy one. Let's, let me see what 3-C says and see if we've addressed that.

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. Discuss alternative methods to assess accuracy of a pulse oximeter.

I heard mention of reporting the performance characteristic, sensitivity, specificity, positive predictive value, et cetera. That's something that resonates more with clinicians because that's what we're used to reading in clinical trials. But I would think that if you have a tight ARMS that can easily be converted into those kind of performance characteristics. But I'm not sure that there's rationale for looking at. The sensitivity specificity was to predict specific thresholds and cut points like 88%.

Dr. O'Connor, do you want to weigh in?

Dr. O'Connor: You said it perfectly. I think what we should do is for 88%, 89%, 99% whatever the oximeter is reporting, the company should have a positive predictive value for that value. How likely is it that it's that number? How likely is it that it's one or two higher, or one or two lower and I think they should do that across the entire domain.

Dr. Nathan: Thank you. Maybe we can stick with their recommendations of an ARMS of 1.502, but then it's up to the companies to translate that to something that's more clinically meaningful as just articulated. Does that seem reasonable to everyone?

All right, Dr. Eydelman, with regards to question 3, have we addressed this adequately?

Dr. Eydelman: Yes. Thank you for your thoughtful comments.

Dr. Nathan: Thank you. All right, let's move on to question four.

QUESTION FOUR

I believe question four is our last 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 labeling modifications to address inaccuracies that may be associated with skin pigmentation.

I think I'll weigh in first, and then I'll let the group weigh in after me. I think that what we would hope to strive for is that there is accuracy across skin pigmentation, but certainly if there isn't any inaccuracy, this needs to be, there needs to be transparency around it. How much inaccuracy is permissible? I'm not sure that there's the right answer, and maybe, to say that the pulse ox needs to perform exactly the same for – with a saturation of 75% across the spectrum of pigmentation might be unrealistic. My hope is that the ARMS incorporates the accuracy across the spectrum.

But I think whatever the inaccuracy is, there needs to be some transparency so that at least people who are using the pulse ox are aware of this as an issue. I'll let the panel weigh in with maybe a different perspective on this or how better to frame this. Dr. O'Connor.

Dr. O'Connor: So once again, I, the criteria should be that the 95% confidence intervals for all of the different identified skin types skin pigmentation should be overlapping. They don't need to be identical. They need to be statistically indistinguishable, and you need to have enough people, enough subjects to have really tight confidence intervals. I don't want to see 10 and 10 with confidence intervals that span 25%, which is crazy. I'd like to see confidence intervals in the measurements that are in the range of the ARMS. Is the aspiration here or a little bit smaller so that they are indistinguishable.

Dr. Nathan: Thank you. Dr. Yarmus.

Dr. Yarmus: Another question to FDA. What, and perhaps I should know this, but as of now, what does the label address for inaccuracies for pulse oximetry?

Dr. Nathan: Dr. Eydelman, can you address that?

Dr. Eydelman: So currently the labeling summarizes the performance, but I'm going to ask Dr. Lee to comment further.

Dr. Connor: And I guess just for clarity, I think it would be helpful, right? Hopefully he's intuitive, but understanding that I think will help frame the recommendations we can give associate –

Dr. Eydelman: Yes. There's also, 510(k) summary performance, which is a public document, which shows the data upon which as a five 10 K decision was made for all of the prescription use devices. Dr. Lee, is there anything you wanted to add? Okay. I believe he's having an issue with his camera.

Dr. Lee: Can I just answer that question?

Dr. Nathan: Oh, go ahead James. Go ahead, Dr. Lee.

Dr. Lee: Okay. Yes, so there are two things. One, we have the regulatory disclosures where on the 510(k) websites we would disclose in the summary for the clearance, and also certainly the labeling. And that's something that we would also highly encourage and have looked at making sure disclosures about accuracy particularly in some cases companies disclose accuracy and what decile. But certainly I think these comments are welcome and it will fuel all decision making on the pre-market and post-market side. Thank you.

Dr. Nathan: Thank you Dr. Connor.

Dr. Connor: Yeah, I just wanted to clarify or add to something that Dr. O'Connor said about overlapping confidence intervals. So I think we need to be careful because things can be pretty far apart and still have confidence intervals overlap. So I think the meaning of what you're suggesting is a hundred percent spot on. I think executing it that way is probably too liberal. For

example, if we measured a whole bunch of people who had arterial oxygen 90 and we saw the mean was 88.1 in light skinned and 91.9 in darker skin, both of those have, would have an ARMS less than two and the confidence intervals would overlap, but they would almost be four percentage points apart.

So I don't think we're saying that's okay, but I think given other things we've said today, it would actually meet the criteria for okay. I just wanted to point out that just using overlapping confidence intervals is probably too liberal. That we would want to say, that the bias, that average difference between things that are truly the same, the arterial pressure is the same. The mean difference we see between light and dark skin is sufficiently low, like 2% or like some number that, I would defer to you that is clinically important.

Dr. Nathan: Okay. That's a good way of framing it. Thank you. So coming to 4-A: labeling modifications to address inaccuracies that may be associated with skin pigmentation.

I think what Dr. Connor just articulated resonated with me in terms of the transparency that should be around it. It's maybe too high a bar to us that everyone would be exactly the same, but that the bias and the bias difference between Caucasian or Whites and darker pigmented folks be being the labeling.

How much permissible bias it should be really minimal. You mentioned I think two or three and – but I think the main thing is as the ARMS is tightened, make sure that there isn't too much of a discrepancy between folks of different color. All right. Recommendations for the content of labeling for lay users who may use pulse oximeters at home? Anyone?

I think we did discuss this. Dr. Hennessy, do you want to weigh in on this?

Dr. Hennessy: Sure. I think that for the products that receive little regulatory oversight for which we don't have much of an idea about accuracy that the label and the device itself should,

accurately portray what's known about the accuracy, even if nothing is known about the accuracy.

Dr. Nathan: Okay, thank you. Mr. O'Brien.

Mr. O'Brien: Yes. I don't know what that was about. Yes. Joe O'Brien, outpatient rep. I clearly would not discuss ARMS or any other of that technical information within the lay user content. I think what's important here is to just, to what first, the message is given for these devices that they're not for medical use, That's clear. But knowing that they are being used in a variety of different ways I think we have to address the fact that it would – what works it seems, in the patient community is plus or minus. So that, and it is given in various areas I saw on websites, et cetera, that they're being told or communicated already that there's a plus minus 4% variance.

So if you have a 90, it may be 86 or maybe, and some of them use a 6% variance. I think in that process, that's probably a good way to communicate it, to say that there is a – to be aware that there is overall a plus or minus four or 5% variance. What that means, it could be 5% more, 5% less and that this variance increases with people of color. So that you should be aware of the variance and to ensure that you communicate with your physician to follow up if there's any question, something to that effect.

And I think, as I said earlier in this day and age, I think we have to go beyond just. The typical labeling, I think it is important to have, whether it's on the FDA website or on the company website, but a clear audio visual that explains to them how to use it, what it means, what that variance is, shows the color, shades, et cetera, in a very succinct way that they're used to doing right now. And that's the way they get their information. I think it would behoove us to, to use that same approach to make sure that the patient population is properly advised.

Dr. Nathan: Okay. Thank you, Dr. Cassiere.

Dr. Cassiere: Yeah I'm going to echo that. Mr. Bryant, you know my thoughts exactly. I think it should be clearly labeled not for medical use, and I would just put it in lay terms. You could have a little package insert there that says not for medical use. Basically you could have normal oximeter readings and still have a low oxygen level, seek medical advice, something along that line. It has to be, Mr. O'Brien opened up that sheet who could read that? It's just lines and lines of minutia.

There should be a requirement where there's a small package insert that is there that says not for medical use and some verbiage on there as to why. Real plain English. And I think that should go a long way. And if consumer devices want to reach the threshold to catch up to medical in hospital use, then let them market that and let them get FDA approval for that, and they could win market share if consumers want to use that product.

Dr. Nathan: I think you articulated it very nicely. It just needs to be clear in plain language, big letters that this isn't a medical device and the reasons why not. So thank you for that, Dr. Connor.

Dr. Collop: Yeah, just a comment. I do think it's a little tricky. I looked at a couple package inserts and they don't really even, and just from like family members, like who've gotten these, they're like what's normal? People don't really know, is 85. Okay. They don't really know that, they don't know what's normal. So when you say a plus or minus 5% that's a big spread. I agree a hundred percent I should say that. But I think this whole labeling for lay users is really tricky because as everybody's iterated, it is being used for medical, but they don't say in the labeling not what's normal, what's abnormal so I think it is, it's a little challenging for sure.

Dr. Nathan: Thank you. Lemme go to Dr. Eydelman and see that she has a hand up.

Dr. Eydelman: Thank you. I just was hoping that the panel in their deliberation in part B also tells communicates their thoughts on content of labeling for prescription use for home. In other words, prescription devices, not just OTC. When they are given for home use, is there something, should there be patient labeling and what should be in that labeling? Thank you.

Dr. Nathan: Yeah, that's a tough question that I don't know that we addressed at all. But what I would say is that, if they get a prescription device to take home, it really becomes the interpretation of the physician who's prescribing it for them. That becomes very important. I don't know that the label should say anything other than, Please refer to your physician's instruction on how best to use this and when to inform him of change, or what he or she would regard as significant change. The physician is writing it for a reason that they're concerned, don't know that a physician would just give it to the patient and say, Here go.

There's got to be accompanying prescription instructions – take this medication three times a day. There has to be some kind of instruction on how often to monitor and maybe that's what the prescription needs to say. Maybe it's not just a provide a pulse ox and maybe this comes to the pharmacist that, okay, I'm giving the pulse doc, but what are the instructions for use and when do you call or when do you go to the ED? I don't know if anyone else went. Oh, a lot of good hands here. Mr. O'Brien I'll come back to you, Dr. Klein. Go ahead.

Dr. Klein: Sure. I actually was going to comment on the OTC portion. I totally agree with the simple messaging. I am a little worried since we don't know exactly how well these perform are really giving a, within 5% up or down. But then also the people don't know what these mean.

But also when we think about some populations that don't, are not normally saturated at baseline. So one person's normal is not another person's normal. So just getting into that consideration as well. So thinking about the cyanotic heart disease, children specifically.

Dr. Nathan: I think yes, where doctors' instruction become important. Not just call if your sets 80%, because they might be walking around with sets of 80%. Dr. Yarmus.

Dr. Yarmus: Thanks. Two, two thoughts on one of each group. I guess the first is for the over the counter. Aspect, I think, again, I'll advocate for I think there needs to be some communication directly to physicians about this, because I think a lot of that is where this lies, right? Better understanding from physician groups that these are proposed to not be as accurate and should not be used under routine care for patient monitoring, which then gets to the second for the prescribed devices. Are we working under the assumption that because those devices are FDA cleared, they are of the same standard as those used in the hospital?

Dr. Nathan: That’s my assumption.

Dr. Yarmus: Okay. Then if that's the assumption we're going with, then I totally agree with what you said, Dr. Nathan, that this should fall under the purview of the prescribing physician.

Dr. Nathan: Thank you, Dr. Alam.

Dr. Alam: I would agree with what you said Dr. Nathan about – for the prescription products. That's really the purview of the physician. For the OTC, I agree with this general sentiment, but I can't say I completely agree with what's been said before. Regarding, it should just say not for medical use, because first of all, that's a peculiar thing since we've already said that these are being used for medical use, since no one's really using these as toy or as in form of entertainment. To the extent that they're not to be used by physicians in lieu of prescription

products, that's certainly true, but that's not a message for the general public. The general public are not physicians and they don't really even know what medical use means.

So since these devices are going to continue to be used by the lay public as a screening device, essentially as a home screening device, I think we should give more detailed instructions on what they're supposed to do with them. I understand we don't know exactly what their accuracy is but I would suggest that FDA might want to consider a labeling chain so they can find out. So they have then the purview to find out what the accuracy is and potentially to create some benchmarks, which even if they're able to be implemented in the very short term are reasonable, do not super increase the cost of the product but are some benchmarks that might be implemented further downstream.

But in the meantime would allow the FDA the ability to investigate these products and give consumers some guidance as to how to use them and have some power to, at the very least, compel the manufacturers to convey that simplified and expanded labeling to patients who will in fact be using these to try to save their own lives.

Dr. Nathan: Mr. Edwards, do you want to weigh in?

Ms. Edwards: Yeah, I was just thinking as a consumer, simple language – I was thinking that to address it, you could say that, put in there that the levels of melanin in the skin may cause discrepancies or inaccuracies in the measurements, it may be too simple, but I was –

Dr. Nathan: I think that there's a lot that can go in a label and maybe that's something that is a quicker fix, at least for now. And certainly for the consumer devices that there's less regulatory requirements around. There should be some kind of warning around that, I would think, to the point of the consumers habits and not – I still agree that the labeling needs to be very simple and you could nuance it, not for medical use, not for medical decision making.

It's a warning, if your stats go down, don't make a decision, but you might want to act on it and call your doctor and it needs validation, but you're not going to make a big decision around it. So I'm not sure if there's a middle road around this, but I think the warning clearly needs to be there both about it shouldn't – you can't rely on this, it might guide you a little bit not only because it's inaccurate in everyone, but especially in patients who have increased melanoma content. So getting back to B. Dr. Loeb, bail me out.

Dr. Loeb: In thinking about both of these labeling questions I think about pharmaceuticals and first of all, if I get vitamins or some other non, a pill that you take that's not regulated by the FDA, it says on it, not regulated by the FDA, and I think it's appropriate for things that devices that FDA does not check on that it says the accuracy of this device has not been verified by the FD, basically, in, in addition to, and it's not for medical use because it's not been, it's accuracy has not been evaluated.

I also think that when a prescription – I go to the pharmacy, I get a thing, and they gimme a whole big sheet of things that are mostly things that your doctor was supposed to tell you about, side effects and how to take it, et cetera. And I think that all, and I think they do, all pulse oximeters come with labeling that says it's affected by motion, it's affected by dys-hemoglobinemia, all this sort of stuff.

And I think the layperson should know these are the things that affect its accuracy, including pigmentation. And lastly, I was really surprised that the FDA presented in their very, in the very first session, what labeling is on pulse oximeters, and I did not know, it looks like it's up to the company, what they put in about what affects their pulse oximeters. But the same things affect all pulse oximeters pretty much.

So I don't know why the FDA doesn't have a list of all these things. Dys-hemoglobinemia and certain nail polishes and motion and pigmentation. All these should be – I would think the FDA would have this list and the companies, it be incumbent on the companies to say, No, ours is not affected by that, otherwise, you got to put it in there. It's affected by that.

Dr. Nathan: I'm trying to think of a simple way to help the FDA in terms of the labeling. Now with the consumer devices, do they go through the FDA at any point? It's just a different threshold or standard that they held against, because maybe a simple way is for the label to say, not approved by the FDA for medical use or something like that.

And I think keeping the warning simple that people can see and it's very evident and it's not buried in the background. I think is effective should be a simple sentence, something like that, or something else. But to that point of that 19 page thing and all these caveats and then – but I think that it needs to be out there that people can see a message like that.

Anyone want to endorse that recommendation. Very simple, not approved by the FDA for medical use.

Dr. O'Connor: Sounds great. [Hands go up all around]

Dr. Nathan: All right, Dr. Eydelman, with regards to question 4, did we address this adequately?

Dr. Eydelman: Yes. Thank you.

Dr. Nathan: Thank very much. I'd now like to ask our three representatives, Beverly Edwards, our consumer representative Dr. William Wilson, our industry representative, Joseph O'Brien, our patient representative, if they have any additional comments. Ms. Edwards, you first.

Ms. Edwards: Okay. I thank you for inviting me to be on this panel and I do think this is just something that's of – we need to discuss and in the discrepancies and the disparities in oximeters and just to the FDA. It really happens in, it's not just oximeters, it's across the board in other areas of devices. Again I thank you and I think I've made enough comments today.

Dr. Nathan: Thank you for that. But I think you bring up a point that came up in the morning discussions by one of the speakers, is that. It's not just pulse oximetry, it might be other devices as well. And perhaps one of the recommendations is that the FDA should be proactive and look at these other devices where skin coloration may play a role in inaccuracy. So thank you for that. Dr. Wilson, do you have any comments?

Dr. Wilson: First I'd just like to thank the panel for including me and during all of my many comments as well. But, really my parting words would be that the public is best served when we ensure the quality of the devices by raising the bar. That doesn't unfairly bias any particular company. Everybody can, if they want, they want to do the work, is going to now provide devices that are going to more accurately measure pulse oximetry for our patients. So that's the goal and really support the discussion that we've had here.

Dr. Nathan: Thank you very much, Mr. O'Brien, any comments?

Mr. O'Brien: Yes as all the others, I would thank both the panel members and the FDA today in addressing the issue. Over the last several years the risks associated with these devices has been redefined and it's clear from the evidence, et cetera, that if we don't redefine the accuracy, then we'll continue to potentially harm all populations and in particularly certain populations that they are in the black community and the darker skin.

And we've got to start eliminating the disparity. And I think that this type of dialogue and these type of days are what's absolutely essential to begin. And I compliment the FDA and I, once again, after another panel say, I'm certainly glad I'm not them.

Dr. Nathan: Thank you very much. I'll now turn the discussion over to Dr. Eydelman from the FDA for summary remarks before we adjourn the meeting.

FDA SUMMATION

Dr. Eydelman: Thank you. I want to thank all of our invited speakers and a special thank you to all of the open public hearing speakers. We truly wanted to hear opinions from everybody, and I'm really excited that so many people have signed up and have taken the time to come and share your valuable input with us.

I also want to thank all of you, the panel members for very thoughtful and thorough deliberations during this very long day. We really appreciate your commitment to the public health. And last but certainly not least, I want to thank a very extensive FDA team that made today possible. I now pronounce this panel meeting complete.

ADJOURNMENT

Dr. Nathan: Thank you. Thank you. I'd like to thank the panel, the FDA, the invited speakers, and all of the open public hearing speakers for their contribution to today's panel meeting. This meeting of the Anesthesiology and Respiratory Devices Panel is now adjourned.

Have a pleasant evening everyone. Thank you all very much.