SUMMARY MINUTES

CENTER FOR DEVICES AND RADIOLOGICAL HEALTH

MEDICAL DEVICES ADVISORY COMMITTEE

GENERAL AND PLASTIC SURGERY DEVICES PANEL

July 28, 2022

9:00 a.m. EST

Webcast via Microsoft Teams

Attendees:

Chairperson

Hobart W. Harris, M.D., M.P.H.
Professor of Surgery
Division of General Surgery, UCSF — San Francisco, CA

Voting Members

Karla V. Ballman, Ph.D.
Division Chief of Biostatistics & Epidemiology
Cornell Medicine — New York, NY

Mary H. McGrath, M.D., M.P.H.
Professor of Surgery
Division Plastic Surgery, UCSF — San Francisco, CA

Murad Alam, M.D.
Vice Chair, Department of Dermatology
Northwestern University School of Medicine — Chicago, IL

Temporary Non-Voting Members

Karen E. Burke, M.D., Ph.D.
Private Practitioner of Dermatology — New York, NY

Paula Bourelly, M.D.
Dermatologist at Olney Dermatology Associates — Olney, MD

Maral Skelsey, M.D.
Director of Dermatologic Surgery Center of Washington — Chevy Chase, MD

Paul Pisarik, M.D.
Practitioner of Family Medicine — Tulsa, OK

Maria Suarez-Almazor, M.D.
Professor, Department of Health Services Research
UT MD Anderson Cancer Center — Houston, TX

Neil Farber, M.D.
Professor Emeritus of Clinical Medicine
Division of General Internal Medicine — University of California, San Diego

Renata Block, PA-C
Physician's Assistant in Advanced Dermatology & Aesthetic Medicine — Chicago, IL

Laura P. Bush, DMSc., PA-C, DFAAPA

Dermatological Physician's Assistant, Fayette Area Dermatology — Fayetteville, GA

Lisa Gualtieri, Ph.D., ScM.
Professor, Department of Public Health and Community Medicine
Tufts University School of Medicine — Boston, MA

Steven J. Skates, Ph.D.
Associate Professor of Medicine and Biostatistics
Massachusetts General Hospital Clinical School — Boston, MA

Katalin Roth, M.D., J.D.
George Washington University — Washington, D.C.
Professor of Medicine; Geriatrics' Palliative Medicine Specialist

Veronica Rotemberg, M.D., Ph.D.
Memorial Sloan Kettering Cancer Center — New York, NY

Industry Representative

P. LaMont Bryant, Ph.D.
Vice Present of Regulatory Affairs
Ethicon, Inc.; Johnson & Johnson

Consumer Representative

Rachel S. Brummert
Founder, Quinolone Vigilance Foundation

Patient Representative

Deneen Hesser, MSHSA, RN

Food and Drug Administration

Long H. Chen, M.D. — CDRH/OPEQ/OHTIV
Acting Director for the Division of General Surgery Devices
Office of Surgical and Infection Devices

Binita Ashar, M.D. — CDRH/ODE
Director of Office of Surgery and Infection Control Devices
Center for Devices and Radiological Health

Candace Nalls, Designated Federal Officer

CALL TO ORDER

Panel Chairperson Dr. Hobart W. Harris called the meeting to order at 9:00 a.m. He noted the presence of a quorum and stated that present members have received training in FDA device law and regulations. He stated the day's agenda:

1. to discuss the risk and benefits of skin lesion analyzers (SLAs) for external use
2. to recommend whether SLAs should be down-classified from class III to class II
3. to discuss helpful evidence
4. to discuss controls to maintain safety and effectiveness

Chairperson Harris reminded the public and panelists that this is a nonvoting meeting. He then asked members of the Committee and the FDA Staff to introduce themselves.

CONFLICT OF INTEREST STATEMENT
TEMPORARY-NONVOTING MEMBER STATUS STATEMENT
GENERAL ANNOUNCEMENTS

Candace Nalls, Designated Federal Officer, reported that no conflict of interest waivers were issued. She announced that Dr. Bryant would serve as the Industry Representative and that Neil Farber, Paul Pisarik, Katalin Roth, and Maria Suarez-Almazor would serve as a temporary non-voting members. She introduced Audra Harrison as the meeting's press contact.

INTRODUCTORY REMARKS

Dr. Colin Chen delineated the purpose of today's meeting, which is to discuss and provide suggestions regarding three specific questions:

1. Which options for determining the actual, or "ground truth", diagnosis during clinical trials are appropriate – histology, video diagnosis by single or multiple dermatologists, or other means?
2. What are acceptable thresholds for sensitivity and specificity? Should it be different for melanoma versus other skin cancers? Should the threshold be different if the device will be used by dermatologists, versus a primary care physician, versus a layperson?
3. Given the different appearance of skin cancer across the population, what rapid approaches will support accurate devices to market that perform as intended in all the potential patients?

FDA PRESENTATIONS

Dr. Jennifer Bai presented an overview three most common types of skin lesions: basal cell carcinoma, squamous cell carcinoma, and melanoma, the deadliest of the three. She emphasized the statistical impact of early detection on survival rates, and proposes skin lesion analyzes as tools to assist in early diagnosis of skin cancers. She asked the Committee to specifically

consider the threshold for sensitivity and specificity and the clinical impact of false negatives and false positives.

Dr. Wang spoke on the landscape of SLA devices, distinguishing between devices used by dermatologists, by primary health providers, and by laypeople, and between devices that provide binary classification, multiple classes, and scaled result displays. She described different types of SLAs and methods to diagnose malignancies, including dermatoscopes, multispectral imaging, confocal spectroscopy, OCT, smartphone cameras, electrical impedance analysis, Raman spectra, and algorithms. Given the variety of devices and intended uses, Dr. Wang called for special consideration on their benefits and risks.

Dr. Henry Lee presented on the diagnostic accuracy of SLA devices, defining 'ground truth' diagnostics and the statistical parameters necessary to establish a device as accurate, sensitive, and specific. He provided performance data from literature reviews regarding dermatologists' ability to identify malignancy without optical assistance. He probed the Panel to discuss quantifiable acceptable performance goals for laypeople, primary care providers, and dermatologists. He also proposed the use of alternate diagnostic methods, such as specialist consensus or panel consensus for diagnosing benign-appearing legions.

Dr. Scott Kominsky examined variances in benefit-risk assessments and repercussions of false positives and negatives, suggesting different specificity and sensitivity thresholds may be necessary for different users. He addressed disease prevalence and the variability of accuracy and voiced concerns about usage and availability of SLA devices in low-availability/minority populations.

QUESTIONS FROM THE COMMITTEE

Dr. Suarez-Almanor inquired towards current data on the utilization of SLAs and dermoscopes, to which **Dr. Ashar** affirmed that products with data are not in today's scope and that dermoscopes are highly accepted instrumentation.

Dr. Harris asked **Dr. Kominsky** for his thoughts on the impact of non-professional use of SLAs on the risk-benefit assessment, and on potential psychological risks of false positives.

Deneen Hesser inquired whether patient perspective was considered in clinical trials for the two approved SLAs; the question was deemed in the scope of the next day's discussion.

Dr. Skates inquired whether **Dr. Kominsky** and the FDA would consider making a benefit to risk quantification, to which **Dr. Kominsky** replied affirmatively.

Dr. Bourelly asked whether **Dr. Lee** addressed all comers or only melanocytic lesions in his discussion of specificity/sensitivity; **Dr. Lee** confirmed it applied to all comers.

Dr. Alam inquired towards the purpose of tracking and/or performing a biopsy on potentially benign tumors, to which **Dr. Ashar** stated that this would not be performed in a clinical setting and would be for design purposes.

Dr. Gualtieri wondered if optional/required trainings impact the accuracy of the device. **Dr. Chen** responded that human factors are a big consideration in approving OTC BAR devices.

Dr. McGrath probed for more information regarding the usage of SLAs by the lay public in Australia and New Zealand. **Dr. Wang** was unable to relay reliable data into the matter.

Next, **Dr. Rotemberg** cited a study she found on patient concerns on SLA usage in response to **Ms. Hesser's** prior inquiry. **Dr. Chen** emphasized patient perspectives as fuel for innovations.

Dr. Skelsey asked **Dr. Lee** for clarification on his data comparing performance and specificity in dermatologists to primary care providers. **Dr. Lee** said this will be detailed tomorrow but encompasses the maintenance of high specificity and affects variably provider sensitivity.

Ms. Block inquired into ground truth and the role of each physician in histological diagnosis. As a broad answer, **Dr. Alam** contributed that most samples from dermatologists are checked by dermopathologists, but with primary care physicians, the type of pathologist doing the workups is variable. **Dr. Pisarik** commented that he trusts his biopsies are shipped to people who are appropriately trained. **Dr. Borelli** added that she only trusts board-certified dermatologists with her samples. **Dr. Rotenberg** agreed with Dr. Alam, noting centralized sample processing.

Dr. Skates wondered why randomized trials are infrequently conducted in dermatology; in response, **Dr. Ashar** asserted that the least burdensome approach to answering scientific questions does not itself necessitate randomized trials. **Dr. Skates** voiced his concerns regarding a lack of quantifiable weighing of benefit versus risk in these complex situations.

GUEST SPEAKER PRESENTATIONS

Dr. Glenn Cohen described racial and other social disparities in the application of algorithms to AI-based medical technology. He provided evidence that the current algorithm produces racial bias by excluding race from algorithmic considerations and utilizing cost as a primary factor. To the FDA, **Dr. Cohen** suggested that, at minimum, FDA reviews should be required to show a product performs relatively well as to any skin tone, any race, any age, and any gender. As a secondary option, he suggested that labeling should reflect the limitations of its algorithm, but noted the difficulty of enforcing this. He also called for post-market evaluations across diverse groups and emphasized evaluating performance in real-world use to account for human factors.

Dr. Ade Adamson disclosed his research affiliations and presented data pertaining racial disparities in melanoma care. He explained that AI-powered tools, in general, neglect racial diversity in their datasets, which negatively impacts accuracy in real-world populations. Racial information is also rarely even reported in clinical studies or for real-world usage. To mitigate racial bias, **Dr. Adamson** suggested over-sampling skin lesions and rashes in skin of color, or perhaps a separate algorithm used for darker skin tones.

QUESTIONS FOR THE COMMITTEE

Dr. Harris first noted that **Dr. Cohen** is unavailable to receive questions.

Dr. Alam wondered about **Dr. Adamson's** thoughts on how to handle the potential approval of technologies that do not have diverse, representative datasets.
Noting the negative impact on patients who may need the technology from preventing approval entirely on this basis, **Dr. Adamson** decided the middle-ground solution is labeled disclaimers of study limitations.

Dr. Rotemberg questioned the weight of data from reader studies versus from in-person evaluations. **Dr. Adamson** replied that most dermatological studies have been reader studies, but demographic metadata could help account for that.

Ms. Block asked if **Dr. Adamson** was concerned that FDA-approved devices on the market are not approved for acral palmar and plantar surfaces. **Dr. Adamson** replied that companies need to put forth good faith efforts to get diverse populations in their datasets. He confirmed there should be more emphasis on acral lesions and an effort in that area would also benefit other races.

Dr. Farber requested thoughts on ways to increase specificity in darker-skinner persons and address the incorrect popular belief that darker-skinner persons do not need to worry about skin cancers from sun exposure.
In response, **Dr. Adamson** posited that he is uncertain if greater specificity would help given an overall lack of obtainable data for darker skin tones.

Dr. Bryant requested clarification on the use of the term "not feasible" in obtaining representation from all skin types, which **Dr. Adamson** rephrased as "challenging."

Dr. Bourelly voiced concern over a potential overlook of medium skin tones, and pointed out that preventative screening during all appointments, regardless of a patient's initial concern, may spread awareness and increase the number of people of color willing to participate in data collection.

In a final question, **Dr. Bush** asked if studies should mirror the U.S. population in terms of skin color in order and adapt these to changes over time.
Dr. Adamson commented broadly that some race-specific parameters are necessary in dermatology and included that an app could preserve dermatological resources if able to classify potentially harmful skin lesions correctly.

OPEN PUBLIC HEARING

Dr. Harris announced the beginning of the Open Public Hearing, and **Ms. Nalls** read the Open Public Hearing Disclosure Process Statement and advised all speakers to begin by disclosing their financial affiliations.

A private-practice dermatologist, **William Steffes**, voiced his positive experience with Nevisense. He asserted that he does not use the technology them for strongly suspicious melanomas, and he is hypervigilant about the detriments of false negatives. He only uses Nevisense as a supplement to other diagnostic tools and considerations, but he assured that he has found it very useful in his clinical practice in diagnosing intermediate and subtle melanomas.
Dr. Steffes called upon the FDA to rigorously test and review all skin lesion analyzers

and to maintain their Class III categorization. In the right hands, he said, the technology can be helpful, but inaccurate devices could be devastating.

Simon Grant, the CEO of SciBase, which developed Nevisense, spoke next. He urged the FDA to keep SLAs highly regulated and maintain their Class III status. SLA studies are complex and high-risk, which is one of the reasons Nevisense is the only U.S.-approved device, he argued. He continued that patient risk has not changed and called upon the FDA to cease discussions of down-regulations, positing that standardized approval guidelines will risk patient safety. For melanoma, he said, it is too early.

QUESTIONS FROM PANEL

Dr. Alam asked **Mr. Grant** if he had a suggestion for where the benchmark for sensitivity should be if he rejected the proposed 90% benchmark. **Mr. Grant** responded that it was not his position to say exactly, but he would like it to be higher.

Ms. Block posed concerns about use of Nevisense by primary care physicians and laypeople, despite its indication that it is not for clinical use on obvious melanoma. **Dr. Grant** stated that obvious melanomas should go straight to biopsy, and he assured that for now, Nevisense only sells to dermatologists, and that use by other groups may be possible down the line, but this would require very complex studies.

Dr. Skates asked for Nevisense's ratio of false positives. **Dr. Grant** confirmed a value of 7.3 in their clinical studies and offered access to the evidentiary documents.

For a final query, **Ms. Hesser** asked **Dr. Grant** if Nevisense offered training specific to physician's assistants and/or nurse practitioners. **Dr. Grant** responded that Nevisense had both a general training program and programs that tailor to a provider's specific usage of the device.

Dr. Harris then pronounced the Open Public Hearing closed at 12:00 p.m. and broke for lunch.

PANEL DELIBERATIONS

Dr. Harris resumed the Panel meeting at 1:00 p.m. and began the Panel Deliberations.

Dr. Skates elaborated on statistical models for quantitative benefit-risk ratios that should be considered in FDA regulatory approaches.

Dr. Pisarik asked if there is evidence of decreased wait times and/or decreased mortality in skin cancer cases resulting from SLA implementation. **Dr. Asher** and **Dr. Rotenberg** agreed that there are not data in that area, and the Panel can and should discuss how to obtain that data. **Dr. Skates** pointed out to this regard the difficulty of performing long-term, involved studies that meet a "gold standard."

Dr. Alam weighed in with some considerations:

a. whether the device is being used by a dermatologist or non-dermatologist, and higher thresholds of accuracy required for non-professional use

b. patient cost of a false positive is lowered by the fact that melanoma biopsies are minimally invasive
c. benign and malignant tumors are not clearly delineated, as some benign lesions can progress to melanomas
d. there are benefits to removing lesions that are not technically true positives

In consideration of all that, **Dr. Alam** strongly proposed that approval processes ensure very high sensitivity and/or a very high risk-benefit ratio.

Dr. Alam and **Dr. Skates** further discussed hypothetical scenarios of risk-benefit analysis in melanoma cases for use in statistical models. **Dr. Skelsey** then requested clarification on the specific types of devices and intended uses that were in the discussion's scope, to which **Dr. Asher** responded that the Committee members should specify this amongst themselves. **Dr. Farber** reiterated the importance of specifying the type of cancer, as different cancers have different risk-benefit profiles.

Dr. Rotemberg reiterated that specificity/sensitivity in the intended-use setting is more important than absolute specificity of static analyses, and suggested that future studies utilize a percentage-based scale of intended accuracy as a measured improvement over the current standard of care. She added that the design and phrasing of a SLA's display screen could impact doctor and patient perception of risk, as with binary versus multi-class output devices.
Dr. Chen responded with mention of a few previous studies that suggest qualitatively that devices enhance patient outcome, but there is no data regarding effectiveness of specific features.

Dr. Burke proceeded by pointing out MelaFind's 6 different categories and how categorization can impact false positive rates and suggesting a fine-tuning of the data. She reiterated that biopsies are minimal, and it is better to be over-suspicious and perform a biopsy than to miss a melanoma. She also noted that patient anxiety over potential procedures is inevitable with any condition and should be handled by doctor-patient relationships.

Dr. Bourelly voiced concerns over patients putting pressure on dermatologists to perform biopsies just because a device told them to. **Dr. Ballman** contributed that the word "biopsy" should be left out of the user interface for laypeople, and countered with "see a specialist," which **Dr. Farber** seconded.

Dr. Asher requested that the Panel converge on specific, measurable criteria the FDA can specify to manufacturers seeking device approval.

Dr. Rotenberg asserted that prospective studies are the key to narrowing down criteria. She suggested defining prospective information and the intended use setting. She emphasized the importance of prospective studies, which she asserted are non-burdensome, can be small, and would help understand burden, cost, and potential harms.

Dr. Alam, too, believes that caution must be exercised in what information devices make available to laypeople to avoid anxiety and, in particular, a false sense of security.
Dr. Farber briefly added that patient anxiety is a complex issue in itself.

Dr. Suarez questioned the FDA on the classifications of the approval process, inquiring whether there is a two-layer process, where first, accuracy is assessed, and second, real-world performance is assessed. **Dr. Asher** weighed in that clinical trials in the pre-market space give

an idea about the device's capabilities, and that in post-market studies, real-world evidence is obtained for understanding and future improvements. She requested specific ideas on how the FDA might put patient anxieties into context within the benefit-risk framework to assess real-world evaluations. Continuing, **Dr. Skelsey** mentioned the false sense of security that SLAs may give laypeople, and suggested that companies give an idea of other common tumors the device can and cannot assess. **Dr. Asher** replied that the FDA will keep this important point in mind.

Dr. Rotemberg pointed out that, due to levels of comprehension, labelling for laypeople and PCPs may need to be broader than those labels on products intended for dermatologic use, which should include labeling for specific indications that includes data used for training and validation. She emphasized the importance of considering how devices communicate, and inquired how algorithms can account for decreased accuracy over time due to data drift.

Dr. Bush expressed concerns over patients 'doctor-shopping' as a result of their SLA devices to find a surgeon that will excise their lesion. **Dr. Bust** responded to a request from **Dr. Asher** for Panel suggestions on mitigation and communication of risks, suggesting that SLAs could provide printouts of data sheets or provide information along with test results.

Dr. Skates commented that he would like to see the Panel discuss thresholds and risk-benefit ratios of SLAs more quantitatively in this discussion.

Ms. Block added that it is not the appropriate time to have discussions about future devices, given current patient risks and liability issues for practices. She also drew attention to the high cost of SLAs: providers may be currently unwilling to invest in these devices until regulatory processes are stricter and the process is more streamlined and/or affordable. She calls for the FDA to focus on existing devices before assessing whether SLAs should be used by non-dermatologists.

Ms. Hesser followed up on **Dr. Rotemberg's** comment on capturing prospective patient information in pre-market developments, and asserted that FDA has patient preference information tools already developed. None are specific to skin cancer, so she pointed out this is an opportunity to develop a tool to inform psychological impact, communication effectiveness, and risk tolerance.

Dr. Roth touched on economic disparities and lack of access to dermatologic services, especially in patients without insurance or in rural areas, and urged the FDA to consider this closely.

Dr. Ballman reiterated the importance of prospective trials, noting that layperson use will differ significantly from trained users in clinical studies. She proposes that prospective, pre-market studies utilize a population sample that is similar to real-world device users.

Dr. McGrath questioned the FDA: are manufacturers currently asking for FDA review of future devices marketed to patients; if so, what are they asking? Secondly, he inquired if primary care physicians have been included in the FDA's discussions and suggested that their input could give insight into community preferences and knowledge. **Dr. Asher**, while unable to provide specifics of interactions, hinted that there is interest in approving devices for patient use and that the day's discussion is timely to their current conversations with manufacturers.

Dr. Alam briefly elaborated on previous comments. To **Dr. Skate's** request for hard numbers, he showed data that PCPs, dermatologists, and oncologists would, across the board, prefer 95% or greater specificity in SLA devices. He noted a general expectation that a free-standing device should be more specific/sensitive than one used to supplement a dermatologist's assessment. **Dr. Alam** added that he thought the current ratio of false positives, around 100 to 700, is acceptable. He also emphasized, with regard to false senses of security, that lay devices should have a disclaimer that patients often present cancerous lesions in areas other than the one they suspect. Ideally, he said, an SLA could include scan a large portion of the body and/or many lesions.

Addressing patient anxiety, **Dr. Alam** mentioned anxiety from SLA results may cause potential backlogs in dermatologist's office; he also clarified that the window of concern for cancerous lesions in typically longer than a year and they are rarely of deadly urgency.

Dr. Burke expressed her concerns about the wide range of device usages, and reiterated that this technology is expensive, and, especially in the context of economic disparity, SLAs are unlikely to be used by PCPs and/or laypeople at this time due to high cost.

Dr. Rotemberg argued that the statistical analyses performed to date are nearly irrelevant because they do not capture real-world usage data. She called for more prospective trials that actively incorporate the patient experience into baseline-oriented quantifications of efficacy. To this, **Dr. Skates** responded that the purpose of quantifying weighted risks of false results as compared to benefits of true positives/negatives is to provide PCPs with a standard to try to meet when using SLAs that is similar to dermatologists' accuracy.

Dr. Rotemberg pointed out that the studies have a pre-selected population and exclude bodily lesions that may have been missed by dermatologists, and they do not provide false negative rates on patient outcomes, to which **Dr. Skates** asserted that comparing dermatologists' performances in prospective studies provides false negative rates of the device and gives a general baseline to assessing safety. **Dr. Rotemberg** agreed that weighted ratios should be used in conjunction with reports of increased standards of care.

Dr. Burke weighed in on the potential usefulness of false positives of melanocytic lesions; she stated that clinically, if the cells are highly abnormal but not malignant, most dermatologists excise it anyway. **Dr. Skates** clarified that these would typically be classified as true positives, but the technical definition can be relative.

Dr. Suarez wondered about the false positive rate in melanoma diagnoses for the current standard of care. **Dr. Rotemberg** answered: from selected samples to rule out melanoma, approximately 2 to 30 benign lesions are biopsied for every one malignant melanoma. She notes that the data is more difficult to analyze for other cancer types but dermatologists are even more sensitive and specific in identifying non-melanoma skin cancer; she estimated a ratio of 1 in 2 to 8 in 10. She reiterated that, in terms of cost-benefit, dermatologists are far more willing to biopsy a legion suspicious for melanoma than any other skin cancer due to its high risk.
Dr. Lee addressed figures on sensitivity/specificity and cited that providers with dermoscopic experience have a 92% sensitivity and 95% specificity for melanoma, giving a 5% false positive rate.

In a final remark, **Dr. Skelsey** expressed that companies should discuss whether their devices can assess lesion evolution, as this would increase utility and safety of the SLA device. She also called for more robust data on intermediate morphologies rather than exclusively malignant/benign legions.

FDA QUESTIONS

Question 1

Rudy Andriani introduced Question 1, covering ground truth. He defined ground truth as "the gold standard that will be used to determine the diagnosis of a lesion"; he defined accuracy as "the measured sensitivity and specificity of a device compared to ground truth." Question 1's sub-questions were:

a. Should histological diagnosis be required for obtaining ground truth diagnoses and all lesions in SLA clinical trials?
b. Are there scenarios for which alternate means or a combination (e.g., histopathology for suspected malignant lesions and consensus opinion of experts for suspected benign lesions) of ground truth that would be acceptable?

Dr. Rotemberg began in agreement with histopathologic diagnosis of melanocytic lesions when there is not expert consensus. She also commented that to establish good performance in SLAs, many benign lesion samples must be available for training and validation. **Dr. Alam** concurred.

Dr. Farber voiced that, while important for standardizing SLAs, it is unethical to watch a patient over a period of time when melanoma may be present. For basal and squamous cell carcinomas, he worried about unnecessary biopsies simply for the sake of a study, and called for a middle-ground biopsy approach in clinical studies.

Dr. Ballman confirmed that all lesion suspicious for melanoma should be sent for biopsy and histopathology during pivotal clinical trials. **Dr. Bourelly** also favored histopathology even in benign-looking lesions, and added that incorporating benign lesions into the algorithm can make the SLA smarter and more accurate. **Dr. Skates** concurred with histopathology as a gold standard due to variations between studies/study designs. **Dr. Rotemberg** and **Dr. Bush** agreed that histopathology is appropriate for all suspicious lesions, but **Dr. Rotemberg** urged the Panel to consider alternatives to histopathologic diagnoses of questionably malignant legions in clinical trials, in order to reduce cost and burden while increasing sample size for benign legions to improve algorithmic models.

Dr. McGrath advocated for a hybrid model, as well. He emphasized that there should be a clear delineation of who decides whether a lesion is benign or malignant, and that defined follow-up procedures should be part of clinical trials for those who are not biopsied.

Dr. Skelsey supported histopathological analysis of all clinical samples, even benign ones.

At the request of **Dr. Harris**, **Dr. Farber** expounded upon his ethical and practical concerns for biopsying benign tumors: this process involves subjecting patients to potentially unnecessary procedures for the sake of algorithm development. He suggested this may deter study participation, and called for, at minimum, a more rigorous informed consent process for patients who are receiving biopsies on suspected benign lesions.

Dr. Alam agreed with **Dr. Farber's** perspective but re-emphasized that the risk of a biopsy is low. The inconvenience of informed consent, he argued, is not a valid reason to alter a study's design. He also asserted that if benign-presenting legions are not biopsied, some malignancies may be missed, even with professional consensus. He therefore advocated for the biopsy of all legions, with consideration to the minimal pain and inconvenience of this protocol to study participants.

Dr. Farber followed up by reiterating that manufacturers need to be very careful with ethics and informed consent in these studies; patients must understand the tissue samples are for study purposes and not diagnostics.

Dr. Roth pointed out that writing a consent form explaining why some people are not biopsied would be more complex and deterring. She supported biopsy for all participants to improve data.

Dr. Ballman put forth the idea of excluding certain lesions upfront in participation requirements, and she probed for ideas on patient burden: would burden be increased by having to return for monitoring in cases where lesions were not biopsied?

Ms. Block astutely proposed that, with patient consent, a data bank could be created from years of prior samples from many doctors to improve algorithmic accuracy.

Dr. Skates shared results from the MelaFind study: out of 83 lesions dermatologists were certain were not melanoma, two were melanoma. He explained this type of human error is his main reason for supporting biopsy of all lesions in clinical studies.

Dr. Bourelly addressed **Ms. Block**'s comment, noting that SLA analysis needs to be done ahead of time, rather than on stored samples, to mimic real-world usage. **Dr. Bourelly** also added that not analyzing all tissue from all skin types, even when benign, omits important missing information regarding lesions on darker skin types.

Dr. Asher noted the consensus of the Panel that histopathology should be used under almost all circumstances, with some excepting benign lesions. She urged participants to provide thoughts specifically pertaining to manufacturers putting out a device/app explicitly not intended for diagnosis.

Dr. Rotemberg added some thoughts: trials for practitioners should tend towards biopsy for all samples, but trials for laypeople, including those for apps, should include some flexibility in clearly benign cases to prevent overwhelming sub-specialty healthcare resources. She also refuted a prior comment: biopsies are not minimally invasive, especially when sampling uncommon anatomical sites like hands/feet or sampling patients prone to scarring, etc.
Dr. Burke fully agreed with **Dr. Rotemberg**'s ideas, expanding that even the most minimally suspicious lesions warrant biopsy, but absolutely clearly benign cases should not be biopsied.

Dr. Suarez-Almazor furthered the idea of a hybrid approach, affirming that, in studies with large numbers of benign lesions, it is appropriate to neglect biopsying certainly benign samples.

Dr. Alam weighed in on **Dr. Ashar**'s query and suggested that non-medically-marketed SLAs will evade FDA regulations but label their devices to mislead users into thinking it proven to be medically reliable.

Dr. Ballman wrapped up comments for Question 1 by saying that established baselines are necessary, but histopathology needn't be the gold standard in all cases.

Dr. Harris summarized the discussion for the FDA representatives:

The Panel, in response to part A, generally agreed that histological diagnosis should be required for obtaining ground diagnosis in all lesions in SLA clinical trials, with important caveats and reasons to adopt a hybrid approach. Other considerations included:

- Setting of device use (dermatologist, layperson, primary physician)
- Trial design
- Ethics of biopsying benign lesions
- Benefits of biopsying benign lesions (algorithmic improvements)
- SLAs' potential to bias dermatologists' opinions

After making sure the Panel had nothing to add, **Dr. Harris** reiterated the above towards the second point, and he highlighted that appropriate use of consensus opinion depends on indication, setting, intended user, and weighting of the device.

The Panel mostly agreed there should be heavy reliance on histopathology as the source of ground truth; however, consensus opinion may be valuable in certain cases.

Dr. Skates added that some Panel members mentioned clear exceptions to the use of histopathology. No other Panel members added to Dr. Harris's summary. The summary was deemed sufficient by Dr. Chen.

Question 2

An unnamed speaker read the slides containing question two's components.

Part A:

i. Should the performance threshold of SLA devices intended for standalone use be a pre-defined sensitivity and specificity across all SLAs, or should performance be compared to another metric, such as the performance of the study dermatologists without the use of the SLA?
ii. If preset threshold are preferable, are the proposed thresholds for sensitivity and specificity appropriate? If not, what sensitivity and specificity thresholds do you propose?
iii. Should the performance thresholds differ if the device is intended for use by dermatologists or by non-dermatology healthcare providers? If so, what performance thresholds do you recommend for each?

iv. Should the performance thresholds differ based on the target diagnosis (melanoma, BCC, SCC)?

Part B:

i. Should the performance threshold of SLA devices intended for standalone use be a pre-defined sensitivity and specificity across all SLAs, or should performance be compared to another metric, such as the performance of the study dermatologists without the use of the SLA?
ii. If preset threshold are preferable, are the proposed thresholds for sensitivity and specificity appropriate? If not, what sensitivity and specificity thresholds do you propose?
iii. Should the performance thresholds differ if the device is intended for use by lay users versus dermatologists or by non-dermatology healthcare providers? If so, what performance thresholds do you recommend for each?
iv. Should the performance thresholds differ based on the target diagnosis? If so, what sensitivity and specificity thresholds do you propose?

Dr. Rotemberg began by emphasizing the importance of prospective studies in the intended use setting rather than placing thresholds on retrospective data.

Dr. Alam suggested 95% sensitivity and 80-90% specificity as parameters for lay devices, and that practitioners could perhaps use devices with lower parameters. He elaborated that he would like to see the thresholds vary based upon lesion type, permitting a 5% lower threshold for BCC and SCC.

Dr. Skates pushed for high sensitivity to ensure safety and effectiveness, and to use the benefit-risk ratio from a dermatologic setting and apply it. When asked if he did or did not favor a scenario where diagnostic specificity of a prover could be doubled by using an SLA, but the device fell short of the performance level of a dermatologist, **Dr. Skates** replied that it does not necessarily need to be exactly at that level of performance, but it should be close, and the bar should be clearly defined.

Dr. Ballman reiterated the previous comments and suggested the risk-benefit ratio to give specificity, but that all studies should show improvement in accuracy with using the device over current standard of care.

Dr. Alam voiced that FDA has a handle on adjunctive device usage, and his primary concerns revolve around devices that will be mass-marketed towards a lay audience; people will forget about the nuances of the device's capabilities once it can claim FDA approval. He reiterated that he would like to see sensitivity similar to that of a dermatologist's, or ideally more. He also stated that he would like to see extremely low specificity.

Dr. Farber agreed with these comments and noted the potential for adverse psychological effects when devices are used by laypeople. **Dr. Bourelly** noted that for laypeople and non-dermatologist providers, the standard should be at least that of a dermatologist's.

Dr. Alam called upon FDA to come up with disclaimers that are suitably clear to the public to ensure average users know the device's intended use and what it beyond its approved

capabilities. **Ms. Hesser** put forth that all FDA-approved SLA devices should meet the same gold standard. **Dr. Rotemberg** emphasized that the questions posed cannot be answered without prospective trials and real world assessment. **Dr. Roth** agreed with most of what others' said, especially that the standard for PCPs should be to bring the PCP up to the performance of a dermatologist; she also reiterated a need for extensive testing in darker-skinned populations.

Dr. Harris proposed a question to the Panel: does the Panel agree that if devices should be required to perform at the level of a dermatologist; and if so, why, then, would we not want it to give diagnoses or differentials to laypeople? **Dr. Alam** answered that the idea is to get patients into a dermatologist's office and mitigate patient anxiety. **Dr. Harris** asked for **Dr. Farber**'s professional opinion on whether there is a substantial difference in anxiety levels if a device gives diagnostic data versus not. **Dr. Farber** replied that the extent of the difference depends largely on phrasing, but, in general, it would be prudent to exclude diagnostic data from a layperson's devices' user interface and only provide the recommendation to see a physician.

Dr. Asher asked the Panel two questions: first, should sub-studies be done in each case with prospective trials, characterizing dermatologist performance and proposing something equivalent or improved over that; and second, could risk be communicated to patients more effectively such that sensitivity thresholds might not have to be inherently higher to compensate for communication issues?

To this point, **Dr. Suarez-Almazor** added that, in some patients, less information might cause more anxiety, and for other, too little information might not be enough to prompt them to see a doctor. She also pointed out, in regards to diversity, that more information on a population's reaction is needed before deciding whether higher thresholds of sensitivity are necessary for certain populations. Next, **Dr. Skelsey** expressed her wishes that the FDA require follow-up studies on device usage by users with different backgrounds and levels of education.

Dr. Rotemberg pressed that the way users respond to SLA apps is entirely unstudied and the Panel's assumption that lay users would necessarily experience anxiety is overly-ambitious. She called for new studies to answer these questions in order to be able to evaluate risks. Responding directly to **Dr. Asher**, **Dr. Rotemberg** asserted that the concern is less about communication and more about the low prevalence of skin cancer; high standards are necessary because prevalence is low.

Dr. Bush proposed the use of broad categories to mitigate anxiety and false senses of security. **Dr. Skates** provided his vision of SLA evaluations being based on whether or not the device can improve what's currently being done. **Dr. Farber** echoed **Dr. Rotemberg**'s comments and also desired to mitigate patient anxiety through appropriate phrasing. **Dr. Alam** further agreed with **Dr. Rotemberg** and **Dr. Skates** and called for tests that are better than dermatologist performance in order to benefit high-risk patients. **Dr. Ballman**, **Dr. Rotemberg**, and **Dr. Burke** all made final comments in support of requiring devices to exceed current dermatological care standards.

Wrapping up, **Dr. Harris** issued a final call for ideas about hard numbers to use for sensitivity and specificity thresholds. **Dr. Rotemberg** suggested a 10% minimum improvement in performance. **Dr. Skates** and **Dr. Alam** concurred. **Dr. Harris** also noted for the FDA that the Panel concurs that SLA devices need to exceed what is currently available in the specific practice setting and that there should be separate safety criteria. He makes an important note that

nobody is comfortable with providing any pre-ordained or across-the-board sensitivity/specificity thresholds.

QUESTION 3

Question 3 was presented:

Should FDA allow skin lesion analyzers to be marketed based on study data from a limited U.S. demographic – for example, in higher incidence populations – with subsequent data collection in lower incidence populations to explain the indications for use? Or, should the FDA require the training of AI/machine learning (ML)-based skin lesion analyzer technologies in all populations regardless of specific cancer incidence?

Additionally, to ensure generalizability across the entire US population, should FDA require all AI/ML-based skin lesion analyzers indicated for use beyond cancerous lesions to be trained and tested in a representative U.S. population?

Dr. Bourelly announced that all skin types should be included so as to set an example of inclusivity that can impact other healthcare efforts. **Dr. Alam, Dr. Ballman**, and **Dr. Rotemberg** concurred. **Dr. Rotemberg** stressed that devices must clearly communicate warnings about skin tone diversity in their labeling if it is not feasible to obtain a truly representative group of skin types in skin cancer patients for clinical studies, as would be ideal.

Dr. Farber, **Dr. Roth**, **Dr. Skelsey**, **Dr. Burke** and **Dr. Skates** all stressed the importance of maximizing inclusivity. **Dr. Roth** stated that it is possible to design algorithms from voluntary user input to mitigate potential tone bias. **Dr. Skelsey** questioned if devices should be excluded if they cannot assess areas of high risk in patients of color. **Dr. Burke** also voiced concern about equitable access to anatomical sites, and she suggested that one or two generations of ethnicity data be collected for clinical data.

Dr. Skates requested clarification, which **Dr. Asher** provided: FDA asked the Panel to give recommendations on how to achieve the ideal, which is studying the population representative of the diverse U.S. population.

Dr. Alam affirmed the use of disclaimers and acknowledged that due to enrollment limitations, it may not be possible to have data for patients with the darkest skin types that is as good as data for other patients. Answering **Dr. Asher**, **Dr. Alam** responded that sooner access is ideal to allow for treatment in high-risk patients, and that, if companies are permitted to conduct clinical studies without sufficient numbers of patients of darker skin toner, there should be accountability measures/consequences to ensure diverse enrollment happens over time.

Dr. Skates asked the FDA what "sticks" are in place to ensure studies continue in darker skin types. In response, **Dr. Asher** solicited further advice from the Panel on how to regulate post-approval studies. **Ms. Hesser** suggested frontloading patient engagement to offset reliance on post-approval studies.

Dr. Rotemberg put forth the consideration that for other diseases, like psoriasis, it may not be appropriate to approve a study that has not already accumulated a diverse/representative population. However, melanoma disproportionately affects lighter skin tones, she urged the FDA to consider the difficulty of recruiting the darkest Fitzpatrick skin tones when designing their study requirements. She also called upon the FDA to facilitate data transparency around

sampling and design decisions and to aid in communicating this information to patients and end users.

Dr. Harris ensured the Panel had no other comments before summarizing for FDA:

The Panel unanimously endorsed the inclusion of patients of all skin types. If a study cannot encompass all skin types prior to market, the Panel made allowances for the study to proceed, with the added concern that companies may lose motivation to continue collecting the less accessible data from darker-skinned populations. Some members called for FDA involvement in requiring post-market studies or other measures to ensure robust data collection.

Dr. Bryant and **Ms. Hesser** thanked the FDA and the Panel and stated they had no additional comments.

FDA SUMMATION

Dr. Chen asked the Panel to clarify the extent of sensitivity/specificity/accuracy that FDA needs to pay attention to. **Dr. Harris** presented that the consensus was that devices should meet essentially the same standards as devices that would be used by non-laypersons. **Dr. Alam** amended that some thought the standards for layperson use should be even more rigorous.

FINAL REMARKS
ADJOURNMENT

Panel members voiced additional thoughts and various points of contention on the stated consensus presented in response to Dr. Chen's question. **Dr. Harris**, **Dr. Chen**, and **Dr. Ashar** confirmed they had no additional comments prior to meeting's end.

Chairperson Harris thanked the Panel, the FDA, guest presenters, and Open Public Hearing speakers and adjourned the meeting at 4:20 p.m.

I certify that I attended this meeting on July 28, 2022 and that these minutes accurately reflect what transpired.

Candace J. Nalls -S

Digitally signed by Candace J. Nalls -S
Date: 2022.10.12 15:57:36 -04'00'

Candace Nalls

Designated Federal Officer

I approve the minutes of this meeting
as recorded in this summary.

Hobart Harris, M.D., M.P.H.
Chairperson