FDA Executive Summary
General Issues Panel Meeting
on Skin Lesion Analyzers

Prepared for the Meeting of the
General and Plastic Surgery Devices
Advisory Panel

July 28, 2022

Table of Contents

List of Tables

Abbreviations

ABCD/E/F/U	Asymmetry, border irregularity, color variegation, diameter > 6mm, evolution, funny looking, ugly duckling
AI	Artificial intelligence
AK	Actinic keratosis
BCNS	Basal cell nevus syndrome (Gorlin syndrome)
BCC	Basal cell carcinoma
CI	Confidence interval
CFR	Code of Federal Regulations
FDA	Food and Drug Administration
ML	Machine learning
MM	Malignant melanoma
NMSC	Nonmelanoma skin cancer
OCT	Optical coherence tomography
PMA	Premarket application
RCM	Reflectance confocal spectroscopy
SCC	Squamous cell carcinoma
SLA	Skin lesion analyzer
SR	Systematic review
SSED	Summary of Safety and Effectiveness Data

Section I – Introduction and Background

1.1 Introduction and Purpose of Meeting

Skin cancer, primarily comprised of basal cell carcinoma (BCC), squamous cell carcinoma (SCC), and malignant melanoma (MM), is the most common form of cancer in the United States. [1] [2] Skin cancer affects all Americans regardless of skin color with nearly 5.5 million cases diagnosed each year.

If detected early, skin cancer is treatable and curable. Delays in diagnosis whether due to missed/incorrect diagnosis (false negative assessments), reduced access to healthcare, prohibitive cost, or prolonged wait times for appointments due to dermatologist shortage, can change the prognosis—particularly for melanoma—from curable to deadly.[3] Therefore, the need to identify the risk of malignancy early in a range of skin lesions has increased the public interest in diagnostic skin lesion analyzers. Availability of additional skin lesion assessment tools that support earlier detection and better skin cancer outcomes may benefit public health.

During this one-day Advisory Committee meeting we will discuss computer-aided skin lesion analyzer (SLA) devices intended to assist in the diagnosis of benign and malignant skin lesions. With skin cancer being the most diagnosed cancer in the US, and earlier detection often leading to improved survival rates, there is a clear need to improve access to diagnostic information. Although SLAs may play a role in addressing this need, as with the introduction of any new technologies, there are significant risks that must be considered, including risks of false negatives (missed diagnosis) and false positives (false alarm); potential healthcare advice without a learned intermediary; impacts on health care systems; and questions of health equity. At this panel meeting, the Food and Drug Administration (FDA) is seeking recommendations from physicians, other health care professionals, scientific experts, and patient representatives regarding the benefits, risks, level of performance and evidence needed to demonstrate a reasonable assurance of safety and effectiveness of these adjunctive diagnostic tools.

The Advisory Committee will be asked questions in three areas:

1. Ground truth used to confirm lesion diagnosis in clinical testing of SLA accuracy;
2. Acceptable sensitivity and specificity for different diagnoses and users; and
3. Health equity considerations based on variable incidence of skin lesions in the US population.

As required by section 513(b) of the Federal Food, Drug, and Cosmetic Act (FD&C Act), the FDA is convening the General and Plastic Surgery Devices Advisory Panel (the Panel) for the purposes of discussing and making recommendations regarding the benefits and risks of Skin Lesion Analyzers.

1.2 Structure of Meeting

The panel meeting will be held in a virtual format over the course of one day and includes time for open public comment, questions by the panel, and panel deliberation.
The morning session will focus on describing the current standards of care for treating and diagnosing skin lesions, a description of the various technologies reported, the proposed

performance levels, and regulation of these devices. The afternoon will include industry presentations, public comment, questions for the panel, and panel deliberations.

Section II – Skin Lesions

For the adjunctive diagnosis of skin lesions, this discussion addresses the application of skin lesion analyzers to lesions suspicious for melanoma, basal cell carcinoma, and squamous cell carcinoma.

2.1 Epidemiology

Skin cancers can be broadly categorized into non-melanoma skin cancer (NMSC) and melanoma. Basal cell carcinoma accounts for 75% of NMSC and squamous cell carcinoma accounts for 20% of NMSC cases with more than one million cases diagnosed in the US each year. [4] The risk factors for skin cancer include light Fitzpatrick skin type, marked by light skin pigmentation and high susceptibility to sunburn; excessive exposure to ultraviolet light; family history; and, for some skin cancers, immune suppression, particularly after organ transplantation. For more information on Fitzpatrick skin types, see Appendix A.

Melanoma is a public health challenge and incidence rates are rising. [5] The cost of melanoma for the healthcare sector is estimated to be $3 billion a year, with an indirect individual cost of 20 years of potential life lost and the intangible cost of individual patient pain and suffering. [6] While melanoma is the third most common skin cancer, it spreads rapidly and results in the greatest number of skin cancer deaths. The incidence of melanoma in non-Hispanic white individuals is nearly 20-fold greater than that in non-Hispanic Black or in Asian/Pacific Islander individuals.[3] However, non-Hispanic black people were reported to be more likely (16% vs 5%) to have melanoma with distant metastasis after being diagnosed with melanoma, compared to non-Hispanic whites.[7]

2.2 Skin cancers: Natural History, Diagnosis, and Treatment

This section will discuss common skin lesions that cause most skin cancer-related morbidity and mortality in the United States population and will focus primarily on basal cell carcinoma, squamous cell carcinoma, and melanoma.

- Basal Cell Carcinoma

Basal cell carcinoma (BCC), the most common type of skin cancer, is a small, skin-colored lesion that arises from keratinocytes at the basal layer of the epidermis. Most BCCs occur spontaneously with no precursor lesion and appear most commonly on the face due to chronic sun exposure.[8, 9] However, there is a high rate of developing BCC within an uncommon congenital lesion called a sebaceous nevus. BCC is also part of several inherited syndromes, including basal cell nevus syndrome (BCNS, or Gorlin syndrome); xeroderma pigmentosum (XP); Bazex syndrome; and albinism. In these disorders, multiple BCC may develop every year, starting early in life. Benign lesions that can mimic BCC include nevi, sebaceous hyperplasia, and benign tumors.

Although BCC typically grows slowly in place, aggressive subtypes may metastasize with poor prognosis. Diagnosis is confirmed by biopsy, which is performed to differentiate BCC from

similar-appearing lesions such as benign nevi, sebaceous gland hyperplasia, sebaceous hyperplasia, small inflammatory lesions, and amelanotic melanoma. In some cases, superficial BCC may be treated with topical drugs. BCC is treated by excision with 95-99% cure rates.[10] All patients with BCC require regular monitoring for new or recurring BCC.

- Squamous Cell Carcinoma

Cutaneous squamous cell carcinoma (SCC), which originates from more differentiated epidermal keratinocytes, also commonly occurs in sun-exposed areas but can develop anywhere, including the lip, anogenital skin, and within scars. Compared to BCC, SCC can have higher rates of metastasis (2-6%), especially for cancers that develop in the "H zone," the area demarcated by the ears and central face.[10]

SCC commonly appears as scaly, thin erythematous lesions, which grow with variable speed. SCC may begin as actinic keratoses (AK), which are small, white, scaly foci of roughness that arise on chronically sun-damaged areas. AK are generally treated with liquid nitrogen or a topical drug without confirmatory biopsy. Larger AK lesions require biopsy to differentiate SCC from inflammatory disorders, common warts, or inflamed benign lesions. Very thin lesions (SCC in situ / Bowen's disease) may be treated non-surgically. Invasive SCC is surgically excised, with 95-99% cure rates.[11] All patients with SCC also require regular monitoring for new or recurring SCC.

If treatment is delayed, however, SCC can metastasize particularly on the head and neck. Once metastasis occurs, the five-year cure rate for SCC is 34%.[10] Patients who are immunosuppressed are at high risk for metastasis and SCC-related mortality. Immunosuppressed patients, such as organ transplant recipients, have been reported to have 65-250 times higher risk of developing SCC.[12] In addition to immunosuppression, SCC with deeper invasion, perineural or lymphovascular invasion, and poor differentiation have higher risks of metastasis.[13, 14]

- Malignant Melanoma (MM)

Melanoma has the greatest impact on public health because it metastasizes quickly and carries a high mortality rate. Melanoma can develop in nevi, particularly dysplastic nevi. However, 70% of melanomas develop on normal skin.[15] Risk factors for developing melanoma include Fitzpatrick skin phototypes I-III (light pigmentation); high susceptibility to or history of blistering sunburns; a high nevus count; the presence of clinically atypical nevi; and family history. Despite its association with UV exposure and light skin pigmentation, melanoma can develop in any individual and on any part of the body, including anogenital skin, palms, soles, and under the nail. People of color have higher percentages of acral melanoma (melanoma of the palm or sole) and subungual melanoma (melanoma under the nail plate) than non-Hispanic white individuals; these subtypes have poor prognosis. [9] Understanding the different clinical presentations of melanoma among all skin phenotypes and ethnicities represented within the US is fundamental to ensuring timely diagnosis and effective treatment for the entire US population.

The combined average five-year survival rate of all stages of melanoma in the US is 93%. As shown in Table 1, according to the National Cancer Institute Surveillance, Epidemiology, and End Results (SEER) program, in people diagnosed with melanoma between 2012 and 2018, the

five-year relative survival for localized disease was 99%; this dramatically decreases with later stages of diagnosis, as the 5-year survival decreased to 71% for regional disease and 32% for distant metastatic disease.[3]

Table 1. Melanoma 5-Year Relative Survival Based on Stage

Stage	5-year relative survival rate
Localized	99%
Regional	71%
Distant	32%

These numbers are based on people diagnosed with melanoma between 2012 and 2018
From the National Cancer Institute – Surveillance, Epidemiology, and End Results (SEER) Program [3]

Section III – Skin Lesion Analyzer Background

3.1 History of Dermatologic Devices

Various devices have been used throughout history to assess skin pathology. Skin lesion examination lenses have been used to examine and diagnose skin lesions since the 1660's and were improved with the use of immersion oil and illumination in the 1880's. In 1989 the hand-held dermatoscope featuring an achromatic lens with a 10-fold magnification was developed, to assess skin lesions and help diagnose pigmented skin lesions more quickly and easily.

Dermatoscopes are FDA regulated as class I devices under the regulations for examination lights, with today's dermatoscopes incorporating polarization allowing for visualization of deeper skin structures and non-polarized light providing information about the superficial skin. Most modern dermatoscopes allow the user to toggle between the two modes, providing complementary information.

Some other examination scopes with image analysis capability provide quantitative information about skin lesions, such as melanin, hemoglobin, and collagen content. These devices typically emit light and collect light signals from the target tissue and then analyze the energy to provide information about the lesion without yielding direct diagnostic information and are regulated under Class II devices. Additional details are in Appendix C.

3.2 Devices Approved for Adjunctive Use in Diagnosis of Melanoma

FDA has approved two computer-aided devices that are indicated for adjunctive diagnosis of lesions deemed suspicious for melanoma by a dermatologist, for use in making the medical decision to biopsy. MelaFind, approved in 2012 under P090012, is a class III device that uses multispectral imaging to analyze skin lesions. [16] Nevisense, a class III device approved in 2017 under P150046, measures impedance levels to analyze skin lesions. [17]

MelaFind is a multi-spectral, non-invasive and automated imaging system that captures the image of a pigmented skin lesion, calculates a risk score, and classifies it based upon degree of 3-dimensional morphological disorganization: MelaFind Positive (high degree of morphological disorganization) or MelaFind Negative (low degree of morphological disorganization).

Nevisense measures electrical impedance in skin lesions and provides an output called the electrical impedance spectroscopy (EIS) score. Electrical impedance is a measure of a material's overall resistance to the flow of alternating electric currents of various frequencies, which differs in normal versus abnormal tissue.

3.3 Skin Lesion Analyzers in Development

We will refer to the non-FDA-cleared or approved computer-aided medical devices (including apps) that are currently in development for adjunctive diagnosis of skin lesions including lesions suspicious for skin cancers as Skin Lesion Analyzers (SLA).

There are presently no legally marketed, FDA cleared or approved SLA devices indicated for use by non-dermatology healthcare providers or the lay public.

SLA device technologies reported in literature range from imaging modalities that visualize cutaneous structures with different sources of contrast (e.g. autofluorescence, fluorescence with exogenous contrast agents, photoacoustic tomography, high frequency ultrasound imaging, thermal imaging, etc.), blood flow, or tissue perfusion (e.g. multi- or hyper-spectral imaging, laser doppler, speckle contrast imaging, photolethysmography) to non-imaging technologies such as Raman spectroscopy, and skin electrical impedance detection (similar to Nevisense). For more description of these technologies, please see Appendix C

The imaging modality may be a smart-phone camera or a high-tech imaging device and involve artificial-intelligence and machine learning (AI/ML) based algorithms that analyze images, detect abnormalities, classify lesion types, or assess lesion risks of malignancy. The algorithm may either be built-into the imaging device, or a stand-alone software device also known as software as a medical device (SaMD). The proposed intended user for SLA devices being developed may be a lay person, non-dermatologist healthcare provider, or dermatologist.

3.4 Algorithm Development and Testing: Impact on Generalizability to Full US Population

AI/ML-based diagnostic devices are developed in three sequential stages. The first step is training of the software algorithm: the device is provided a training data set of labeled data (for example, images of melanoma that are paired to the diagnosis "melanoma" and images of BCC that are paired to the diagnosis "BCC"), and the software teaches itself to identify visual patterns that correlate with the tagged diagnosis (machine learning). Next, the algorithm is run with a second image set (validation or tuning data set) of images. Finally, the accuracy of the algorithm in classifying images is measured with a third set of images (test data set). In this testing phase, the goal is to establish, for each test image, how often the output was correct or incorrect.

Robust data sets are crucial for training AI/ML algorithms; however, training datasets can be narrow in focus, particularly if the data does not represent a diverse set of patients. The publicly-available skin lesion image databases and proprietary image sets currently used for SLA training and testing may have limited distribution of skin types and skin lesion types. [18] This could affect the generalizability – the accuracy with which results can be transferred to people other than those originally studied – of device performance to the full US population.

Similarly, testing of device performance may be needed across the range of individuals in the US population because skin lesions can look different (to the unaided eye and to the AI/ML algorithm). Specifically, testing in all expected patients may be needed to ensure that the device performs with sufficient accuracy in all intended patients. However, recruitment of patients in lower-prevalence of skin cancer groups for performance testing may lead to longer device development timelines and delays of the potential benefit of those SLAs coming to market sooner. These aspects will be discussed with the panel in Question 3 with the intention of balancing device availability and ensuring health equity.

3.5 Clinical Workflow

Lesions suspicious for skin cancer may be detected by a patient or by a non-dermatology healthcare provider during routine physical examination. This section will describe the pathways

from initial identification of a lesion as suspicious to final diagnosis and provide a brief description of sensitivity and specificity at each step.

- Laypersons

To aid laypersons in self-examination, the ABCD rule was developed in the 1980s as a guide for features that suggest a pigmented lesion could be suspicious for melanoma. [19] The mnemonic was originally ABCD (asymmetry, border irregularity, color variegation, diameter >6 mm), and has since garnered the additional letters E (evolution, or change in the lesion), F (funny looking), and U ("ugly duckling" – one lesion that looks different than the rest). [20] However, the ABCD rule has primarily been used by medical providers, and its significance in the general community is not clear. Studies show that the ABCD rule approaches 80% sensitivity depending on the number of features identified.[19]

- Non-Dermatologist Health Care Providers

When patients identify a lesion of concern, they may seek evaluation by a primary care provider (which we define as any physician, nurse practitioners (NP), or physician assistants (PA) in any non-dermatology specialty) or by a dermatologist. If the lesion is clearly benign, the patient is reassured, and no further management is needed. If a dermatology consultation identifies a lesion as suspicious, it is biopsied to provide a definitive pathologic diagnosis and to determine treatment. If the lesion appears benign but may have the potential to become malignant in the future, continued surveillance is recommended with regular monitoring for any visual changes or development of suspicious features. The cardinal feature of development of possible malignancy is change in size, shape, or color of the skin lesion, which prompts clinical evaluation and biopsy.

- Dermatologists

Access to dermatology assessment can be the rate-limiting step in diagnosis of skin cancer due to physician shortage. Dermatologist availability is limited in nearly all states in the country and access to skin specialists generally involves long wait times. [21] According to Awadalla *et al*, only 30%-40% of dermatological concerns are reviewed by a dermatologist expert [22]. In addition, for those who do undergo a dermatologist review, the average wait for an appointment is 38 days.[21] Even in metropolitan areas, the mean wait time for a dermatology appointment is 15.6 days (+/-12). [21].

3.6 Diagnostic Accuracy

- Diagnostic Accuracy Overview

The diagnostic accuracy of clinicians varies significantly between specialties and within a specialty and may depend on a variety of additional factors, such as the provider's training, experience, and practice. Local geographic factors, such as the incidence of skin cancers in a particular region, may also affect a clinician's accuracy when diagnosing skin lesions.

The Cochrane Skin Cancer Diagnostic Test Accuracy Group published a systematic review (SR) and meta-analysis of visual diagnosis in 2018. [23] The SR identified articles

that assessed the accuracy of visual examination of lesions suspicious for melanoma compared to a reference standard of histological confirmation or of clinical follow-up. The SR included 49 publications that provided data on 51 study cohorts, with a combined 34,351 skin lesions, including 2,499 melanomas. Meta-analysis of reports that measured in-person evaluations exhibited sensitivity as high as 92.4% with specificity of 79.7% or better. However, there tended to be an inverse correlation between level of sensitivity and level of specificity.

A parallel Cochrane meta-analysis assessed the accuracy of visual inspection, with and without dermoscopy, for diagnosing melanoma. Meta-analysis of the included 86 publications demonstrated dermoscopy to be more accurate than unaided visual examination, increasing sensitivity and specificity from 75% and 76% to 92% and 95%, respectively.[24] Dermoscopic equipment is typically used by specially trained dermatologists and this standard component of skin lesion analysis increases the dermatologist's sensitivity and specificity of diagnosing lesions.

Table 2. Visual Examination Without and With Dermoscopy for Diagnosing Melanoma

	Sensitivity (95% CI)	**Specificity** (95% CI)
Visual examination	**75%** (66-85%)	**76%** (57-87%)
Visual examination with dermoscopy	**92%** (87-95%)	**95%** (90-98%)

Source: Cochrane Skin Cancer Diagnostic Test Accuracy Group [27]

The Cochrane group also assessed the reported accuracy of teledermatology in diagnosing skin cancer.[25] This systematic review included 22 studies reporting diagnostic accuracy of teledermatology assessment compared to final histology. The studies covered 4,057 lesions and 879 malignant cases. Correct categorization of lesions as malignant based on photographic images was high, with summary sensitivity of 94.9% and specificity 84.3%. Sensitivities and specificities were more variable for melanoma (sensitivity range: 59% to 100%; specificity range: 30% to 100%). The importance of this teledermatology data is to demonstrate the relative accuracy of trained dermatologists' evaluation of lesions based upon photographic or dermoscopic images alone.

Table 3. Teledermatology Accuracy for All Skin Cancer and for Melanoma

Method	**Sensitivity (95% CI)**	**Specificity (95% CI)**
Teledermatology using photographs, for any skin cancer: malignant vs. benign	**94.9%** (90.1-97.4%)	**84.3%** (48.5-96.8%)

Teledermatology with dermoscopic images, for any skin cancer: malignant vs. benign	**100%**	**Range: 25% to 92%**
Teledermatology using photographs: melanoma	**Range: 59% to 100%**	**Range: 30% to 100%**
Teledermatology with dermoscopic images: melanoma	**85.4%** (68.3-94.1%)	**91.6%** (81.1-96.5%)

Source: Cochrane Skin Cancer Diagnostic Test Accuracy Group [26, 27]

- Diagnostic Accuracy of Dermatologists versus Primary Care Providers

Non-dermatologist health care providers such as primary care physicians (PCPs), nurse practitioners, and physician assistants often evaluate and treat the majority of skin conditions in practice. [28, 29] The diagnostic sensitivity of general practitioners assessing melanoma has been reported to be between 29-98%, compared to 67.2-100% for dermatologists (Table 4) [30-34] Research has shown that the referral process by general practitioners is sometimes inconsistent, leading to a relatively large number of unnecessary referrals.

A literature search was performed to investigate the sensitivity and specificity of clinicians when diagnosing skin malignancies, particularly melanoma, squamous cell carcinoma, and basal cell carcinoma. The following table (Table 4) describes the sensitivity and specificity ranges for dermatologists and primary care providers. [30-34]

Table 4. Diagnostic Accuracy of Dermatologists versus Primary Care Providers

Lesion Type	Specialty	Sensitivity	Specificity
Melanoma	Dermatologist	67.2-100%	54-95.6%
	Primary Care Provider	29-98%	49-98%
Squamous Cell Carcinoma	Dermatologist	65.8%	95.6%
	Primary Care Provider	42-69%	86-93%
Basal Cell Carcinoma	Dermatologist	74-97%	87-98.9%
	Primary Care Provider	79-89%	76-83%
Binary Outcome (e.g. malignant versus benign; biopsy versus observation)	Dermatologist	65.8-94.8%	59.8-95.6%
	Primary Care Provider	87.8-95.7%	57-90.6%

As expected, there is a wide range of sensitivities and specificities noted in the literature for melanoma diagnosis without clear delineation of the skin phenotypes tested. Dermatologists were found to have a sensitivity of 67.2-100% and a specificity of 54-95.6% for melanoma. In contrast to dermatologists, there were few publications regarding the diagnostic accuracy of

primary care providers. Sensitivity for general practitioners ranges from 29-98%, and specificity ranges from 49-98%.[31, 35-37]

Overall, it appears that PCPs can perform as well as dermatologists, particularly in making a binary determination of whether a lesion may be malignant, but that diagnostic performance is less variable among dermatologists than in PCPs.

- Diagnostic Accuracy of Dermatopathologists

Histopathology has traditionally been the diagnostic ground truth for skin lesions. Evaluation of the accuracy of histopathology is essential in order to provide additional context to the gold standard for ground truth. This will allow the panel to determine if there are acceptable alternatives to histopathology for ground truth in specific situations. Reported alternatives to histopathology for other lesions that are not suspicious for melanoma have included a clinical diagnosis made by a specialist such as a dermatologist, a consensus diagnosis by a panel of dermatologists, or a confirmed benign diagnosis as evidenced by long term follow-up over a period of months to years.

For pigmented lesions concerning for melanoma, histological diagnosis was the ground truth for the pivotal studies by MelaFind and Nevisense, which used a consensus diagnosis from a panel of core study dermatopathologists, though three months of clinical follow-up was also allowed for the MelaFind study for non-suspicious lesions. After approval, Braun et al. assessed the interobserver variability of individual dermatopathologists assessing clinically difficult melanocytic lesions in the MelaFind study in two-category classification (benign vs. malignant) and found a high – but not perfect – correlation, with a kappa of 0.80. [38] Another study, using the Nevisense data, performed a similar assessment to compare individual dermatopathologist performance in pigmented lesions suspicious for melanoma, relative to the consensus gold diagnosis, and reported average individual dermatopathologist sensitivity of 84.9% and specificity of 98.1%. [39]

Section IV – Proposed Regulation of SLA

4.1 Regulatory Approach

FDA has cleared several Class II devices intended for lesion measurement and/or documentation of changes to lesion size or shape over time.

The purpose of today's discussion is to focus specifically on the ground truth and appropriate performance measures necessary for SLA intended for providing diagnostic information. Therefore, to support the discussion on the appropriate degree of accuracy and the ground truth to which accuracy is compared, the panel should be aware of the different use contexts that FDA anticipates for SLA devices:

- The users of the SLA device may be trained dermatologists, non-dermatology health care providers, or lay users.
- The device may be used to assess a single suspicious lesion, or it may be used for screening multiple or all skin lesions on the body as part of routine cancer screening.
- The device output may be used in conjunction with other clinical or historical aspects, as an adjunct to a provider, i.e., a piece of information to be considered with other sources of information, or the SLA device may serve as a standalone diagnostic device.

Each of these permutations may impact the regulation of the device and the clinical performance testing that may be needed to assess device accuracy. For the purposes of this panel meeting, the devices to be discussed will include any AI/ML-enabled devices that analyze data obtained from a suspicious skin lesion, regardless of the type of data collected (e.g., clinical photo, other imaging modality, measurement of physiological signals, etc.) or the device output that is intended to support diagnosis of skin lesions.

4.2 Performance Measure Benchmarks

Table 5 provides the proposed thresholds for minimal acceptable performance. We define the minimally acceptable performance threshold to be the lower margin of the 95% confidence interval of sensitivity and specificity. The proposed thresholds are predicated on balancing the benefits and risks of the devices with the public health need for assistive tools, considering the sensitivities reported by clinicians and the value added over the accuracy of healthcare providers.

Performance will be assessed when the device is operated by the intended user (dermatologist, non-dermatology health care provider, or lay user).

Accuracy of SLAs may be benchmarked relative to a pre-defined accuracy threshold or other performance metric based upon a selected ground truth. The potential means of obtaining ground truth are further discussed below.

Table 5. Proposed performance thresholds for SLA

BCC/SCC	Sensitivity 80%
	Specificity 80%
MM	Sensitivity 90**%**
	Specificity 70**%**

Notes for Table 5:
1. Performance will be assessed when the device is operated by the target user.
2. The proposed thresholds reflect literature review of sensitivity and specificity for MM and BCC/SCC by user group
3. Proposed performance goals may include pre-defined thresholds for sensitivity and specificity

4.2 Options for Ground Truth

Ground truth is defined as the diagnosis considered to be correct for a lesion. Histopathology has traditionally been the diagnostic benchmark for skin lesions and is therefore commonly utilized in clinical studies. Assessment of SLA accuracy in the literature has included alternatives to histopathology such as a clinical diagnosis made by a specialist (e.g. dermatologist), a consensus diagnosis by a panel of dermatologists, or a confirmed benign diagnosis as evidenced by long term follow-up over a period of months to years. [40-50] The alternatives are most often utilized for benign-appearing lesions for which a histopathologic diagnosis is not available because it would not undergo biopsy during the course of normal clinical practice. However, some studies have used the clinical diagnosis by a dermatologist or a panel of dermatologists as the ground truth even for suspicious-appearing lesions.

Ground truth may be derived from histological diagnosis (as consensus by a panel of core study dermatopathologists), or it may be the consensus of clinical observation by dermatologists. Other ground truth methods may be acceptable in some circumstances. This will be discussed with the panel in Question 1.

In addition to assessing sensitivity, in some studies comparing the diagnostic accuracy of AI versus dermatologists, the focus has been on the specificity of the AI algorithm by setting the sensitivity of the AI-guided device at the level of the sensitivity found for the dermatologists in the same study, which allows for a direct comparison of specificity of the AI algorithm versus the dermatologists at the same sensitivity. Given the differences in sensitivity and specificity among different users, different thresholds of diagnostic accuracy for SLAs may be needed for different user groups and different types of skin lesions. This will be addressed in Question 2.

In testing SLA performance, a benchmark must be set for the performance measure of accuracy (sensitivity and specificity). One benchmark proposed is predefined percent accuracy (for example, the values proposed in Table 5) relative to the ground truth. Another potential performance measure would be accuracy comparable to a panel of dermatologists who assess the same lesion images or dermoscopic images. A third approach, one used in other specialties for devices intended to provide clinical decision support, is to assess whether the device improves the performance of the provider, relative to their unaided diagnostic accuracy. Finally, a hybrid option may be acceptable as well, wherein clinically benign lesions are assessed without biopsy (e.g. by a panel of dermatologists with dermoscopic examination), whereas lesions suspicious for melanoma are assessed by biopsy. The panel will be asked which of these benchmarks should be used to defined sufficient accuracy for an SLA in clinical testing in Question 2.

Section V Benefit – Risk Discussion

It has been reported that PCPs assess and treat a large portion (> 50%) of dermatological conditions in practice. [28, 29] This includes primarily rashes but applies to skin lesions as well. With less experience evaluating skin lesions, it is expected that non-dermatologist healthcare providers may have greater reliance on SLA results when making the decision of whether to refer a patient for further evaluation and potential skin lesion biopsy.

While the ideal device would be 100% sensitive and specific, all diagnostic devices, particularly those that are AI/ML based, must balance sensitivity and specificity. Higher sensitivity will result in increased diagnosis but may be accompanied by more false positives and unnecessary biopsies. Greater specificity will reduce unnecessary biopsies but may result in false negatives and missed or delayed diagnosis and treatment. In the clinical setting, while optimizing both sensitivity and specificity is important, the risk of a false negative has the potential of more severe consequences (missed or delayed diagnosis and treatment) than a false positive (unneeded biopsy and patient apprehension), so sensitivity may be viewed as the more clinically important parameter, but this may be different with different users (e.g., for lay persons sensitivity maybe more important to be seen by a medical professional for further evaluation, while for a dermatologist specificity maybe more significant to ensure correct diagnosis). A balanced consideration of probable benefits and probable risks is an essential part of FDA's determination that there are reasonable assurances of medical device safety and effectiveness [51]. Table 6 provides a general benefit-risk assessment for SLA devices.

Table 6. Benefits and Risks Associated with SLA

Benefits of SLA Diagnostic Testing	Risks of SLA Diagnostic Testing
Greater access to testing by reducing barriers to healthcare access	**Increased use of healthcare resources and more skin lesion biopsies** due to false positive results
Earlier testing to improve outcomes in skin cancer, especially melanoma	**Delay in diagnosis** due to false negative results
Enhanced assessment as an additional tool aiding clinical decisions especially with borderline lesions	**Poor Positive Predictive Value** when skin cancer has low prevalence in a given population

Section VI – Panel Questions

Question 1: Ground Truth

In clinical trials for diagnostic devices, accuracy is assessed by comparing the device output to the ground truth. For skin lesions, particularly when ruling out malignancy, clinical practice has traditionally relied on histology for ground truth. FDA has requested that histological diagnosis (core specimen processing with a consensus diagnosis from an expert dermatopathologist panel) be used for ground truth because it provides the greatest certainty in the diagnosis. Device developers, however, cite concerns, both practical and ethical, in requiring biopsy of all lesions, particularly those that appear benign. They have proposed alternate means of defining ground truth, including consensus opinion of experts (of visual or dermoscopic examination of the lesion(s)), opinion of one expert (visual or dermoscopic examination), or other methods.

Should histological diagnosis be required for obtaining ground truth diagnosis in all lesions (both suspected malignant and benign) during SLA clinical trials? 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 would be acceptable?

Question 2: Performance Thresholds

A. Some SLA devices may be used as one of multiple sources of information for clinical decision making, meaning that the output will be an adjunctive tool, to be used by a provider in concert with clinical and historical information, in reaching a management decision. The provider may be a dermatologist or a non-dermatologist health care provider. Table 5 above provides proposed performance thresholds for sensitivity and specificity for melanoma, BCC, and SCC. <u>For SLAs intended to be used as one of multiple sources of information by a non-dermatologist healthcare provider or dermatologist</u>:
   i. Should the performance thresholds of SLA devices intended to provide adjunctive information for clinical decision making be a pre-defined sensitivity and specificity across all SLAs e.g., Table 5, or should performance be compared to another metric, such as the performance of the study dermatologists without use of the SLA? Or for providing adjunctive information, can performance be assessed by whether having the SLA improves the accuracy of the study dermatologists?
   ii. If preset thresholds are preferable, are the proposed thresholds for sensitivity and specificity in Table 5 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, and SCC)? If so, what performance thresholds do you recommend for each?

B. Other SLA may be used as standalone devices, meaning that the output will be relied upon at face value to guide management. **Devices that will provide results to the lay user without provider input will always be standalone**. <u>For SLA intended to be used as standalone devices</u>:

i. Should the performance thresholds of SLA devices intended for standalone use be pre-defined sensitivity and specificity across all SLAs, or should performance be compared to another metric, such as the performance of the study dermatologists?
ii. If preset thresholds are preferable, are the proposed thresholds for sensitivity and specificity in Table 5 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 (melanoma, BCC, and SCC)? If so, what performance thresholds do you recommend for each?

Question 3: Performance in US Population

Panelists should consider whether these SLA devices must be able to analyze skin lesions with an acceptable sensitivity and specificity in all patients prior to FDA clearance, or whether proof of performance data in higher-prevalence populations (e.g., non-Hispanic white individuals) can be provided to allow these high-prevalence populations access to this technology, followed by clinical studies in lower prevalence populations. The potential benefit of a stepwise approach is that it may allow for earlier access to this technology for populations at high-risk, but it may increase the risk of false positive and false negative results in lower prevalence populations in whom the device has not been adequately trained and tested. However, requiring SLA to be tested in patients with lower incidence before entering the market could delay the time to market due to extended enrollment times for statistically relevant numbers of darker skin individuals with skin cancer.

Should FDA allow SLAs to be marketed based on study data from a limited US demographic (e.g., in higher incidence populations) with subsequent data collection in lower incidence populations to expand the indications for use? Or, should the FDA require study data from across all US demographics, regardless of specific cancer incidence?

Although the previous questions have focused on skin cancer, SLA may also be used for other lesions that have similar prevalence across all US demographics but look different in different Fitzpatrick skin types. To ensure generalizability across the entire US population, should FDA require SLAs indicated for use beyond cancerous lesions be tested in a representative US population?

Section VII Appendices

Appendix A: Fitzpatrick Classification of Skin Types

Dr. Thomas Fitzpatrick developed the concept of skin phototypes while developing ultraviolet (UV) therapy protocols. It was a departure from prior descriptions of skin by the visible color and instead emphasized the response to UV exposure. Generally, individuals who appear to have light skin and hair will burn easily and tan poorly when exposed to UV without sun protective measures.[52, 53] Conversely, generally individuals with dark skin will tan and not burn. However, some individuals who appear to be light skinned will tan well, and some individuals with dark hair and eyes and olive-complected skin will burn more than anticipated.

The Fitzpatrick skin type, or phototype, is used to describe the risk of sunburn and correlates well with risk of skin cancer. The three most common skin cancers (melanoma, BCC, and SCC) are more prevalent in Fitzpatrick I and II individuals and are relatively less prevalent in Fitzpatrick V and VI individuals. However, as stated in Section II, people of color are more likely to develop melanoma in areas that are not sun exposed, such as the sole or under the nail. Due in part to lower expected risk and screening, these melanomas are often detected late.

Fitzpatrick Skin Type	Skin Color	Reaction to Sun Exposure
I	Pale white	Always burns, never tans
II	White	Usually burns, sometimes tans
III	White	Sometimes burns mildly, tans average
IV	Olive	Rarely burns, tans well
V	Brown	Very rarely burns, tans very easily
VI	Dark brown	Never burns, tans very easily

Appendix B: ABCD (E, F, U) Criteria for Melanoma

In 1985, dermatologists from New York University first devised the acronym ABCD (asymmetry, border irregularity, color variegation, diameter >6 mm) to educate primary care clinicians and laypeople on the identification of early melanoma. Subsequent studies have expanded information on its usefulness. [54-63]

Asymmetry: if a lesion is bisected, one half is not identical to the other half

Border irregularities

Color variegation: presence of multiple shades of brown, or presence of any blue, black, gray, or white

Diameter: ≥6 mm

In 2004, the criteria were enhanced with the addition of "E" (evolution) to incorporate the fundamental concept of change, including a modification over time of a pre-existing nevus or the development of a new lesion, especially in individuals older than 40 years.

Since its introduction, additional mnemonic letters added include:

Evolution: change in size, shape, or color, or a new lesion

Funny looking: lesion that looks “wrong”

Ugly duckling: a pigmented lesion that is obviously different from the others in a given individual

These criteria apply most commonly to the superficial spreading subtype and are less applicable to nodular and desmoplastic melanoma subtypes. Moreover, melanomas in children and adolescents often lack the conventional ABCDE criteria as do amelanotic melanomas. The diagnostic accuracy of the mnemonic has been assessed in a few studies, all having methodologic limitations [64, 65]. The sensitivity and specificity of the ABCDE criteria vary when they are used individually or in combination, and the risks of over- and under-referral must be balanced accordingly. The use of a single criterion is sensitive but not specific, meaning that many benign lesions would be biopsied or referred, whereas using more than one criterion for referral is more specific but increases the chances of missing malignant lesions.

In a retrospective study of 1140 lesions including 460 melanomas, the sensitivity in identifying a lesion as a melanoma was 97 percent when using a single criterion and 43 percent when using all five criteria jointly. By contrast, specificity was 36 percent for a single criterion and 100 percent for all five criteria. [19]

Appendix C: Skin Lesion Visualization Devices

Class I devices

- **White Light Dermatoscopes and cameras:**

Dermatoscopes provide magnified and illuminated images of skin. They are different from skin lesion analyzer devices as they do not provide any image processing or analyzing features.

Considering the advances in Charge-coupled device (CCD) and Complementary metal–oxide–semiconductor (CMOS) imaging technologies, usually quality and resolution of digital cameras for white light imaging (even in the smart phones) are very good.

Therefore, using these devices for white light imaging of skin (without any analysis) is considered low risk. They are classified as class I devices under the following regulations:

Regulation: 21 CFR 880.6350, Battery-powered medical examination light

Identification: A battery-powered medical examination light is a battery-powered device intended for medical purposes that is used to illuminate body surfaces and cavities during a medical examination.

- Regulation: 21 CFR 880.6320, AC-powered medical examination light

  Identification. An AC-powered medical examination light is an AC-powered device intended for medical purposes that is used to illuminate body surfaces and cavities during a medical examination.

- Regulation: 21 CFR 878.4160, surgical camera and accessories

  Identification. A surgical camera and accessories is a device intended to be used to record operative procedures.

- **Image Archiving:**

These devices are intended to store and archive patient information and images[1]. They could also be a "software as medical device" (SaMD)[2]. They are classified as class I devices under the following regulations,

Regulation: 21 CFR 892.2020 Medical Image Communications Device,
Identification: A medical image communications device provides electronic transfer of medical image data between medical devices. It may include a physical communications medium, modems, interfaces, and a communications protocol.

Class II devices:

- **Skin Imaging to Demonstrate the Relative Location of Blood, Collagen and Pigments:**

These devices are dermatoscopes and skin imaging devices that use multispectral images and provides analyzed images showing the relative location of blood, collagen, and pigments. These devices are different from skin lesion analyzer devices as: they only provide additional imaging information as an adjunct tool to a healthcare provider, they do not correlate the imaging

information to a disease, and they do not provide any interpretation for a disease diagnosis or disease state. These devices are classified as class II devices requiring 510(k).

- **Digital photography**

Significant advances in the last two decades in CCD and CMOS imagers made high quality and high-resolution digital cameras available at low cost. These devices capture white light images and help dermatologists to visualize and document skin texture, color and abnormalities.

- **Total Body Photography**

Total body digital photography uses digital cameras to quickly image the entire body skin. These devices help to document and monitor the entire body skin surface and its abnormalities.

- **Dermoscopy (Epiluminescence Microscopy)**

Dermoscopes, also known as epiluminescence microscopes, are handheld skin examination microscopes placed on a skin lesion that allow magnified and illuminated views of the skin . They help provide in vivo evaluation of colors and microstructures of the skin not visible to the naked eye.

- **Skin lesion measurement tools**

Unautomated and manual tools like rulers to measure the size of skin lesions.

- **Reflectance Confocal Microscopy (RCM)**

Reflectance confocal microscopy visualizes and captures high resolution skin tissue images. This technique scans the skin with a low power laser beam and captures the reflected light from the microscope focal point through a confocal pin hole filter. Capturing multiple two-dimensional images at different depths enables the reconstruction of three-dimensional structures (a process known as optical sectioning).

The Cochrane group reviewed the literature reporting accuracy of RCM in lesions suspicious for skin cancer compared with a reference standard of either histological confirmation or clinical follow-up. [63],The review included 10 studies including data for BCC or SCC. Meta-analysis demonstrated RCM to have higher specificity than sensitivity when assessing all lesions (sensitivity 76%; specificity 95%), but to be highly sensitive in assessment of BCC (sensitivity 94%; specificity 85%). Summary sensitivity and specificity could not be calculated for SCC due to paucity of adequate studies. A second meta-analysis that included 18 publications estimated use of RCM for melanoma to have 92% sensitivity (95% CI 87 to 95%) with 72% specificity (95% 62 to 81). [66, 67]

- **Multispectral and Hyperspectral Imaging**

These devices use reflectance microscopic or macroscopic imaging with different illumination wavelengths to demonstrate different properties of the tissue, including blood, collagen and pigments of the unstained skin or tissue.

Photoplethysmography imaging is another technology that can be used for measuring relative blood flow in microcirculation and assesses reflected light due to the variations caused by motion of the red blood cells and pulsation in the vasculature.

- **Optical Coherence Tomography (OCT)**

Optical Coherence Tomography Systems use low-coherent light sources to reconstruct the reflected light from different depths of the tissue. They can be used in the evaluation of in-vivo or ex-vivo human tissue microstructure, by providing two or three-dimensional real-time depth visualization. Depending on the properties of the light source (superluminescent diodes, ultrashort pulsed lasers, or supercontinuum lasers), optical coherence tomography can provide images in sub-micrometer resolution.

A 2018 systematic review assessed reported performance of OCT in skin cancer, including MM, BCC, and SCC compared to a reference standard of histological confirmation or clinical follow-up, identifying five studies with 529 cutaneous lesions (282 malignant lesions). [68] Meta-analysis estimated the sensitivity of OCT for identification of BCC at 95% (95% CI 91-97%) and specificity at 77% (95% CI 69-83%). Insufficient data was available for SCC and MM.

- **High Frequency Ultrasound**

High-frequency ultrasound (HFUS) provides images of superficial cutaneous structures and color doppler high-frequency ultrasound enables visualization of blood flow. HFUS has lower resolution than OCT but has higher scan depth.

A meta-analysis of studies evaluating the accuracy of HFUS (20 MHz or more) to assess lesions suspicious for melanoma, BCC, or SCC, compared to reference standard of histological confirmation or clinical follow-up, identified 29 datasets: 20 for melanoma, nine for BCC, and none for SCC. [69] Most studies were hampered by inadequate visualization. Derived sensitivities for US evaluations were 83% (95% CI 75-90%) with variable specificities ranging from 33% to 73%. It was not possible to estimate accuracy for BCC due to high variability.

- **Electrical Impedance**

These devices measure electrical impedance of skin lesions in real-time and provide a scoring output based on the electrical differences between the normal and abnormal tissues.

- **Raman Spectroscopy**

Raman spectroscopy uses a laser as an illumination source and detects the inelastic scattering of photons (Raman scattering). Most light scattered by skin or tissue is at the same frequency as the excitation source (Rayleigh or elastic scattering). A small amount of the scattered light shifts in energy from the laser frequency because of interactions between the incident electromagnetic waves and the vibrational energy levels of the molecules in the sample. Plotting the intensity of the shifted light against the frequency (wavelength) produces a Raman spectrum of the sample. Raman spectroscopy may be used as a complementary assessment in addition to other image-based technologies.

- **Smartphone Apps**

Currently there are many smartphone apps developed for skin lesion assessment, which can be categorized into 4 types based on their functions: teledermatology, informational, educational, and interactive. Among them, the interactive apps provide lesion assessment through a computerized algorithm without professional input. While the output varies from risk classification to lesion type identification, the ultimate goal of these apps is usually early

detection of melanoma. The majority of the apps are intended to be used by lay persons. Currently, the interactive apps with lesion assessment function have not been cleared or approved by the FDA, but some are being marketed in Europe, Australia or New Zealand with CE mark.

Appendix D: Approved Devices MelaFind (P090012) and Nevisense (P150046)

Excerpts from the PMA P090012 FDA Summary of Safety and Effectiveness Data for MelaFind is provided below or reference on prior clinical study designs.

The complete SSED of MelaFind can be accessed at
https://www.accessdata.fda.gov/cdrh_docs/pdf9/p090012b.pdf

SUMMARY OF PRIMARY CLINICAL STUDY (PROTOCOL 20061)

The applicant performed a clinical study to establish a reasonable assurance of safety and effectiveness of detecting malignant melanoma and high-grade lesions with MelaFind for use on clinically atypical cutaneous pigmented lesions with one or more clinical or historical characteristics of melanoma, excluding those with a clinical diagnosis of melanoma or likely melanoma in the US. Data from this clinical study were the basis for the PMA approval decision. A summary of the clinical study is presented below.

MelaFind's classifier algorithm was developed and tested in six clinical studies, which enrolled a total of 9439 images of 9078 lesions from 6931 patients, including 630 melanomas, at 40 clinical study sites in the United States and abroad over seven years. Five of these – Protocols 20011, 20012, RCP2007-05, 20031-A, and 20031-B – were prepivotal clinical studies to develop the automatic MelaFind image analysis algorithm. The last, Protocol 20061, was the pivotal trial to evaluate the safety and effectiveness of MelaFind. Adjunctively, two web-based (electronic) physician reader studies were performed based upon stored images and case histories collected by live assessment in pivotal study (Protocol 20061). The effect of electronic lesion assessment compared to live lesion assessment was not evaluated in studies conducted with this device.

A. Study Design

Patients were treated between Jan 31, 2007 and July 7, 2008. The database for this PMA reflected data collected through July 7, 2008 and included 1383 patients having 1831 pigmented skin lesions (PSLs). Of the 1831 lesions enrolled, 1632 lesions considered to be eligible and evaluable for analysis. There were 7 investigational sites.

Protocol 20061 was a prospective, multi-center, blinded clinical study. Examining dermatologists were blinded to the MelaFind results, dermatopathologists were blinded to both the dermatological diagnoses and MelaFind results, and MelaFind was blinded to both dermatological and histological diagnoses. Enrollment was to proceed until at least 93 eligible and evaluable dermato-histologically confirmed melanomas were enrolled among lesions receiving dermatological diagnosis of either 'melanoma cannot be ruled out' or 'not melanoma', with a minimum total number of lesions of 1200. Sensitivity and specificity as primary endpoints were determined to be appropriate metrics for evaluating safety and effectiveness of MelaFind to correctly identify malignant melanoma. The sponsor used exact method to calculate the sample

size based on one-sided alpha=0.05 and used the "mid-P exact method" to compute a one-sided 95% confidence interval (CI) on sensitivity for their statistical analysis.

Procedures:

Three high resolution digital photographs of the lesions were obtained – two clinical views (from 21 inches and 8 inches away) and a dermoscopic image – using standard cameras. Enrolled lesions were to be assessed by MelaFind as 1 (positive) or 0 (negative), by dermatologists based upon clinical and dermatoscopic (if performed) prebiopsy lesion categorization as definite melanoma (100% likely melanoma), melanoma cannot be ruled-out (likely, 67-99%; possible, 34-66%; and unlikely, 1-33%), and not melanoma (0% likely) and by dermatopathologists using histologic diagnosis of biopsy specimens.

All study lesions were biopsied; no study lesions were followed to assess lesion change with time (evolution). Biopsies were reviewed by at least two central dermatopathologists; the positive class of lesions consisted of melanomas (in situ and invasive), and high-grade lesions (high grade dysplastic nevi, atypical melanocytic proliferation/hyperplasia). Breslow thicknesses of invasive melanomas were recorded.

Clinical Endpoints

With regards to safety, effectiveness, and success/failure criteria, the sponsor met the following primary endpoints:

Primary Aim 1: To demonstrate that MelaFind's sensitivity to malignant melanoma, among lesions with dermatological diagnoses of "Melanoma cannot be ruled out" or "Not melanoma", is at least 95% at a 95% confidence level.

Analysis of Primary Aim 1

Primary Aim 1 is met with 90% Confidence Interval (CI) where 1-sided 95% Lower confidence bound is greater than 95% but not met with 95% CI (2-sided 95% LCB).

	Sensitivity	**95% CI**		**90% CI**	
MelaFind	98.3%	94.1%	99.7%	95.0%	99.5%

Primary Aim 2: To demonstrate that, along with this high level of sensitivity, the specificity of MelaFind for lesions that are not malignant melanoma, among lesions with dermatological diagnoses of "Melanoma cannot be ruled out" or "Not melanoma," is superior to the specificity of study dermatologists.

Analysis of Primary Aim 2:

Specificity (10.6%) is superior to the study dermatologists (5.5%)

	Specificity	95% CI	
MelaFind	10.6%	9.7%	13.2%
Study dermatologist	5.5%	4.5%	7.3%
Difference	5.1%	3.3%	7.7%

Summary of Diagnostic Performance of MelaFind

Assessment including all lesions based upon dermatopathology	MelaFind
True Positive (TP)	172
False Negative (FN)	3
True Negative (TN)	157
False Positive (FP)	1300
Sensitivity*: P20061 lesion cohort	(172/175) 98.3%**
Specificity*: P20061 lesion cohort	(157/1457) 10.8%**
* This does not represent true (per subject) sensitivity or specificity.	
**The sensitivity and specificity are based on all eligible and evaluable lesions selected for biopsy by the investigating dermatologist and does not include all possible candidate lesions per subject.	

Population Demographics and Baseline Parameters

Demographics		All Enrolled Subject Population			
		Atypical			All Pigmented Lesions
		Melanoma (F2)	Melanoma Cannot Be Ruled Out (F3)	Not melanoma (F4)	All Populations***
N, Lesions from Patients Enrolled		25	1702	103	1831
Gender*	Female	11	920	61	993
	Male	14	782	42	838
Age*	<21 years	0	102	5	107
	21 – 55 years	13	1082	46	1142
	>55 years	12	518	52	582
Anatomic Location*	Face	1	51	7	59
	Posterior Torso	7	745	34	787
	Anterior Torso	0	358	18	376
	Extremity: Arm/ Leg	16	491	30	537
	Neck	0	32	8	40
	Scalp	1	25	6	32
Fitzpatrick Skin Type*	I	3	112	2	117
	II	11	897	73	981
	III	8	610	24	642
	IV	3	75	3	82
	V	0	5	1	6
	VI	0	3	0	3
Geographic Sites, Patients***	US – sun belt	22	1349	88	1459
	US – non-sun belt	3	353	15	372
	Non - US	0	0	0	0
* Table presents total lesion counts. Patients who contributed more than one lesion to the study are represented in more than one population when those lesions occur in more than one population					
** Geographic sites were tabulated based on clinical study site. US – sun belt included Alabama, California, Florida, and North Carolina. US – non-sun belt included Pennsylvania and Illinois. All					

SUMMARY OF SUPPLEMENTAL CLINICAL INFORMATION

SUMMARY OF READER STUDY

A prospective, randomized, and investigator blinded web-based reader study under Protocol 20063 was conducted electronically. The study was intended to assess and compare the biopsy

sensitivity and specificity of MelaFind®, to the average biopsy/referral sensitivity and specificity of expert and general dermatologists as well as primary care physicians who did not participate in the pivotal clinical study. In this reader study, randomly selected lesions were evaluated by physicians who recorded their biopsy decisions. One-hundred-thirty lesions (65 melanomas and 65 non-melanomas) were selected randomly from the database, maintaining the prevalence of non-melanoma lesion types observed. Non-melanomas were matched by age and anatomic site to melanomas. All lesion images underwent review by the principal investigator for image quality. Physicians (pigmented skin lesion experts, general dermatologists, and primary care physicians) were recruited until at least 40 participants from each category completed the study. Physician heterogeneity was assessed using kappa statistics.

Procedure:155 physicians completed the study. Reader study physicians reviewed three high resolution digital images taken from standard cameras – clinical images from 21 and 8 inches away from the lesion, and a dermoscopic image. In addition, twenty-four items of information were provided, including clinical history, risk factors for melanoma, and the results of physical examination findings by the investigating physician.

Primary Objective. The primary objective of Protocol 20063, the adjunctive reader study, was to test the hypothesis that MelaFind sensitivity to identify melanoma was at least as good as that of investigators using photographs and histories of the same lesions collected by pivotal study investigators by live assessment. MelaFind sensitivity was 97%, which was statistically significantly superior to that of 110 dermatologists who, on the average, missed (i.e., elected not to biopsy) 28% of melanomas in this electronic study (p-value < 0.0001).

Protocol 20063: ANOVA Results for Biopsy/Referral Sensitivity and Specificity with 95% CI, n = 100:

	Sensitivity	Std. Dev.	CI	Specificity	Std. Dev.	CI
All Derms	0.72	0.03	(0.66, 0.78)	0.51	0.04	(0.43, 0.58)
MelaFind	0.97	0.15	(0.90, 0.99)	0.09	.19	(0.04, 0.19)
Difference	0.25	0.03	(0.18, 0.32)	-0.41	0.05	(-0.51, -0.31)

PANEL MEETING RECOMMENDATION AND FDA'S POST-PANEL ACTION

- Panel Meeting Recommendation

At an advisory committee meeting held on November 18, 2010, the General and Plastic Surgery Devices Panel raised concern regarding MelaFind use by non-dermatologists. The indications for use defined the operator of MelaFind to be at the physician level whereas the pivotal study only used study investigators that were board certified dermatologists. In addition, panelists were concerned that without training in the proper use of MelaFind, operator's would not select the appropriate lesions for MelaFind use and would not correctly use the device to guide their clinical decision to biopsy in order to rule-out melanoma in accordance to the indications and instructions for use. Other concerns involved having a MelaFind negative reading influence a decision to not biopsy a lesion with clinical suspicion

of melanoma, which may potentially result in a false negative diagnosis and delay in care; and the guidance that should be provided to users based on the high number of melanomas confirmed by dermatopathology among the nonevaulables.

At the meeting, the Panel voted 10 votes yes and six votes no that there is reasonable assurance the device is safe, and eight votes yes and six votes no that there is reasonable assurance that the device is effective, and eight votes yes, seven votes no and one absente[d] that the benefits of the device do outweigh the risks in patients who meet the criteria specified in the proposed indication.

B. FDA's Post-Panel Action

The sponsor provided a revised indication for use that defines MelaFind use by physicians trained in the clinical diagnosis and management of skin cancer (i.e. dermatologists) who have also successfully completed a training program in the appropriate use of MelaFind. This addresses the concern regarding the use of MelaFind by a non-dermatologist and the potential concerns regarding the appropriate use of MelaFind according to the indications and instructions for use. In addition, the revised indications for use and labeling also states that MelaFind should not be used to confirm a clinical diagnosis of melanoma and that it is one element of the overall clinical assessment. Also, MelaFind negative lesions should be based on the remainder of the entire clinical context and lesions that are "non-evaluable" by MelaFind should be carefully re-evaluated for biopsy. These indications and labeling were found acceptable to address these outstanding concerns since lesions that are clinically diagnosed to be suspicious for melanoma will not be evaluated by MelaFind and a MelaFind negative reading is only part of the assessment for a clinical decision to biopsy and will not replace clinical judgement. In addition, non-evaluable lesions will now be re-evaluated for biopsy which is supported by the clinical data.

CDRH DECISION

CDRH issued an approval order on November 1, 2011. The final conditions of approval cited in the approval order are described below.

The sponsor must conduct a post approval study that will evaluate whether MelaFind increases the sensitivity of physicians in diagnosing melanomas and high-grade lesions, while the false positive rate of physicians is not substantially elevated.

The study will be a multi-center, single arm, observational, prospective study to gather data on relative sensitivity, among other study endpoints. Data to be collected includes: relative sensitivity comparing physicians' performance before and after using MelaFind as the primary study endpoint; real-world use of MelaFind, i.e., the patient characteristics including age, gender, race/ethnicity, and Fitzpatrick Skin Type, the number of lesions that were examined by MelaFind, the proportion of lesions that meet the labeled Indications For Use among all the lesions examined by MelaFind, the proportions of positive and negative findings of MelaFind among all of the lesions examined, the proportion of lesions that are un-evaluable by MelaFind, the proportion of lesions that are found to be un-evaluable for each user of MelaFind, the number of attempts with MelaFind that were performed for each lesion before a definitive reading resulted or the lesion was declared un-evaluable, and the impact of MelaFind use on the per physician biopsy rate for pigmented lesions; and an evaluation of safety and effectiveness of MelaFind, i.e., the proportion of biopsy from the lesions that MelaFind identifies as positive and the results of those biopsies, the proportion of biopsy among the "unreadable" lesions and the results of those biopsies, the proportion of biopsy

from the lesions that MelaFind identifies as negative and the results of those biopsies, and the proportion of the biopsied lesions (from each of the above – MelaFind positive, MelaFind negative, and un-evaluable) returned as melanoma on pathology. This study must enroll 78 patients with one or more eligible and evaluable histologically-confirmed melanoma and/or high-grade lesion based on the null hypothesis that the relative sensitivity is less than or equal to 1.1. The study power will be at least 85%.

Patients with lesions evaluated with MelaFind during the enrollment period, but not biopsied at that time, will be followed at 1 year ±3 months and 2 years ±3 months. At least 50% of the study sites will be new (i.e., they did not participate in the MelaFind pivotal study). The study sites will include a mix of academic centers and private practices.

Section VIII References

1. Yanes, D., et al., *Predicting Negative Outcomes of Cryolipolysis in Patients With and Without Raynaud Disease.* Dermatol Surg, 2021. **47**(5): p. 675-677.
2. Nikolis, A., *Response to: Shockwave Therapy for the Prevention of Paradoxical Adipose Hyperplasia After Cryolipolysis: Myth or Reality?* Aesthet Surg J, 2021. **41**(8): p. Np1139-np1140.
3. National Cancer Institute, S., Epidemiology, and End Results Program (SEER). *Cancer Stat Facts: Melanoma of the Skin*.
4. National Cancer Institute, S., Epidemiology, and End Results Program (SEER). *Cancer Stat Facts: Common Cancer Sites*. 2021 January 9.
5. Guy, G.P., Jr., et al., *Vital signs: melanoma incidence and mortality trends and projections - United States, 1982-2030.* MMWR Morb Mortal Wkly Rep, 2015. **64**(21): p. 591-6.
6. Ekwueme, D.U., et al., *The health burden and economic costs of cutaneous melanoma mortality by race/ethnicity-United States, 2000 to 2006.* J Am Acad Dermatol, 2011. **65**(5 Suppl 1): p. S133-43.
7. Culp, M.B. and N.B. Lunsford *Melanoma Among Non-Hispanic Black Americans*. Preventing chronic disease, 2019. **16**, E79 DOI: 10.5888/pcd16.180640.
8. Samarasinghe, V. and V. Madan, *Nonmelanoma skin cancer.* J Cutan Aesthet Surg, 2012. **5**(1): p. 3-10.
9. Miller, D.L. and M.A. Weinstock, *Nonmelanoma skin cancer in the United States: Incidence.* Journal of the American Academy of Dermatology, 1994. **30**(5, Part 1): p. 774-778.
10. Jerant, A.F., et al., *Early detection and treatment of skin cancer.* Am Fam Physician, 2000. **62**(2): p. 357-68, 375-6, 381-2.
11. Alam, M., et al., *Guidelines of care for the management of cutaneous squamous cell carcinoma.* Journal of the American Academy of Dermatology, 2018. **78**(3): p. 560-578.
12. Euvrard, S., J. Kanitakis, and A. Claudy, *Skin cancers after organ transplantation.* N Engl J Med, 2003. **348**(17): p. 1681-91.
13. Waldman, A. and C. Schmults, *Cutaneous Squamous Cell Carcinoma.* Hematol Oncol Clin North Am, 2019. **33**(1): p. 1-12.
14. Smith, K.J., S. Hamza, and H. Skelton, *Histologic features in primary cutaneous squamous cell carcinomas in immunocompromised patients focusing on organ transplant patients.* Dermatol Surg, 2004. **30**(4 Pt 2): p. 634-41.
15. Pampena, R., et al., *A meta-analysis of nevus-associated melanoma: Prevalence and practical implications.* J Am Acad Dermatol, 2017. **77**(5): p. 938-945 e4.
16. *MelaFind: Summary of Safety and Effectiveness Data (SSED)*
17. *Nevisense: Summary of Safety and Effectiveness Data (SSED)*.
18. Kim, Y.-h., A. Kobic, and N.Y. Vidal, *Distribution of race and Fitzpatrick skin types in data sets for deep learning in dermatology: A systematic review.* Journal of the American Academy of Dermatology, 2021.
19. Thomas, L., et al., *Semiological value of ABCDE criteria in the diagnosis of cutaneous pigmented tumors.* Dermatology, 1998. **197**(1): p. 11-7.
20. Scope, A., et al., *The "Ugly Duckling" Sign: Agreement Between Observers.* Archives of Dermatology, 2008. **144**(1): p. 58-64.
21. Tsang, M.W. and J.S. Resneck, Jr., *Even patients with changing moles face long dermatology appointment wait-times: a study of simulated patient calls to dermatologists.* J Am Acad Dermatol, 2006. **55**(1): p. 54-8.
22. Awadalla, F., et al., *Dermatologic disease in family medicine.* Fam Med, 2008. **40**(7): p. 507-11.
23. Dinnes, J., et al., *Visual inspection for diagnosing cutaneous melanoma in adults.* Cochrane Database Syst Rev, 2018. **12**: p. CD013194.

24. Dinnes, J., et al., *Dermoscopy, with and without visual inspection, for diagnosing melanoma in adults.* Cochrane Database of Systematic Reviews, 2018(12).
25. Chuchu, N., et al., *Teledermatology for diagnosing skin cancer in adults.* Cochrane Database Syst Rev, 2018. **12**: p. CD013193.
26. Dinnes, J., et al., *Dermoscopy, with and without visual inspection, for diagnosing melanoma in adults.* Cochrane Database Syst Rev, 2018. **12**: p. CD011902.
27. Dinnes, J., et al., *Visual inspection and dermoscopy, alone or in combination, for diagnosing keratinocyte skin cancers in adults.* Cochrane Database Syst Rev, 2018. **12**: p. CD011901.
28. Ramsay, D.L. and P.E. Weary, *Primary care in dermatology: whose role should it be?* J Am Acad Dermatol, 1996. **35**(6): p. 1005-8.
29. Barnett, M.L., Z. Song, and B.E. Landon, *Trends in physician referrals in the United States, 1999-2009.* Arch Intern Med, 2012. **172**(2): p. 163-70.
30. Anderson, A.J., J.M. Kirkwood, and L.K. Ferris, *A pragmatic approach to melanoma screening in collaboration with primary care providers.* Cutis, 2016. **97**(6): p. 382-3.
31. Burton, R.C., et al., *General practitioner screening for melanoma: sensitivity, specificity, and effect of training.* J Med Screen, 1998. **5**(3): p. 156-61.
32. Curiel-Lewandrowski, C. and S.M. Swetter, *Lack of harms from community-based melanoma screening by primary care providers.* Cancer, 2016. **122**(20): p. 3102-3105.
33. Federman, D.G., J. Concato, and R.S. Kirsner, *Comparison of dermatologic diagnoses by primary care practitioners and dermatologists. A review of the literature.* Arch Fam Med, 1999. **8**(2): p. 170-2.
34. Madeja, J., G. Kelsberg, and S. Safranek, *Does screening by primary care providers effectively detect melanoma and other skin cancers?* J Fam Pract, 2020. **69**(2): p. E10-E12.
35. Vestergaard, T., et al., *Introducing teledermoscopy of possible skin cancers in general practice in Southern Denmark.* Fam Pract, 2020. **37**(4): p. 513-518.
36. Madeja, J., G. Kelsberg, and S. Safranek, *Does screening by primary care providers effectively detect melanoma and other skin cancers?* J Fam Pract. , 2020. **69**(2): p. E10-E12.
37. Youl, P.H., et al., *Diagnosing skin cancer in primary care: how do mainstream general practitioners compare with primary care skin cancer clinic doctors?* Med J Aust. , 2007. **187**(4): p. 215-20.
38. Braun, R.P., et al., *Agreement of Dermatopathologists in the Evaluation of Clinically Difficult Melanocytic Lesions: How Golden Is the 'Gold Standard'?* Dermatology, 2012. **224**(1): p. 51-58.
39. Braun, R.P., et al., *Electrical Impedance Spectroscopy in Skin Cancer Diagnosis.* Dermatol Clin, 2017. **35**(4): p. 489-493.
40. Tschandl, P., et al., *Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study.* Lancet Oncol, 2019. **20**(7): p. 938-947.
41. Sun, M.D., et al., *Accuracy of commercially available smartphone applications for the detection of melanoma.* Br J Dermatol, 2021.
42. Singh, N. and S.K. Gupta, *Recent advancement in the early detection of melanoma using computerized tools: An image analysis perspective.* Skin Res Technol, 2019. **25**(2): p. 129-141.
43. Maron, R.C., et al., *Systematic outperformance of 112 dermatologists in multiclass skin cancer image classification by convolutional neural networks.* Eur J Cancer, 2019. **119**: p. 57-65.
44. Marchetti, M.A., et al., *Results of the 2016 International Skin Imaging Collaboration International Symposium on Biomedical Imaging challenge: Comparison of the accuracy of computer algorithms to dermatologists for the diagnosis of melanoma from dermoscopic images.* J Am Acad Dermatol, 2018. **78**(2): p. 270-277 e1.
45. Malhi, I.S. and Z.Z.N. Yiu, *Algorithm-based smartphone apps to assess risk of skin cancer in adults: critical appraisal of a systematic review.* Br J Dermatol, 2021. **184**(4): p. 638-639.

46. Haenssle, H.A., et al., *Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists.* Ann Oncol, 2018. **29**(8): p. 1836-1842.
47. Freeman, K., et al., *Algorithm based smartphone apps to assess risk of skin cancer in adults: systematic review of diagnostic accuracy studies.* BMJ, 2020. **368**: p. m127.
48. Chuchu, N., et al., *Smartphone applications for triaging adults with skin lesions that are suspicious for melanoma.* Cochrane Database Syst Rev, 2018. **12**: p. CD013192.
49. Brinker, T.J., et al., *Comparing artificial intelligence algorithms to 157 German dermatologists: the melanoma classification benchmark.* Eur J Cancer, 2019. **111**: p. 30-37.
50. Brinker, T.J., et al., *A convolutional neural network trained with dermoscopic images performed on par with 145 dermatologists in a clinical melanoma image classification task.* Eur J Cancer, 2019. **111**: p. 148-154.
51. FDA, *Factors to Consider When Making Benefit-Risk Determinations in Medical Device Premarket Approval and De Novo Classifications* HHS, Editor. 2019.
52. Oakley, A. *Fitzpatrick skin phototype*.
53. Gupta, V. and V.K. Sharma, *Skin typing: Fitzpatrick grading and others.* Clin Dermatol, 2019. **37**(5): p. 430-436.
54. Abbasi, N.R., et al., *Early diagnosis of cutaneous melanoma: revisiting the ABCD criteria.* JAMA, 2004. **292**(22): p. 2771-6.
55. Fox, G.N., *ABCD-EFG for diagnosis of melanoma.* Clin Exp Dermatol, 2005. **30**(6): p. 707.
56. Hazen, B.P., et al., *The clinical diagnosis of early malignant melanoma: expansion of the ABCD criteria to improve diagnostic sensitivity.* Dermatol Online J, 1999. **5**(2): p. 3.
57. Lee, H. and K. Kwon, *A mathematical analysis of the ABCD criteria for diagnosing malignant melanoma.* Phys Med Biol, 2017. **62**(5): p. 1865-1884.
58. Friedman, R.J., D.S. Rigel, and A.W. Kopf, *Early detection of malignant melanoma: The role of physician examination and self-examination of the skin.* CA: A Cancer Journal for Clinicians, 1985. **35**(3): p. 130-151.
59. Cohn, B.A., *In consideration of the E in the melanoma ABCDE mnemonic.* Arch Dermatol, 2006. **142**(4): p. 528-9; author reply 529.
60. De Giorgi, V., et al., *Skin self-examination and the ABCDE rule in the early diagnosis of melanoma: is the game over?* Br J Dermatol, 2013. **168**(6): p. 1370-1.
61. Duarte, A.F., et al., *Clinical ABCDE rule for early melanoma detection.* Eur J Dermatol, 2021. **31**(6): p. 771-778.
62. Robinson, J.K. and R. Turrisi, *Skills training to learn discrimination of ABCDE criteria by those at risk of developing melanoma.* Arch Dermatol, 2006. **142**(4): p. 447-52.
63. Titus, L., *Skin self-examination and the ABCDE rule in the early diagnosis of melanoma: is the game over? Reply from author.* Br J Dermatol, 2013. **168**(6): p. 1371-2.
64. Whited, J.D. and J.M. Grichnik, *Does This Patient Have a Mole or a Melanoma?* JAMA, 1998. **279**(9): p. 696-701.
65. Harrington, E., et al., *Diagnosing malignant melanoma in ambulatory care: a systematic review of clinical prediction rules.* BMJ Open, 2017. **7**(3): p. e014096.
66. Dinnes, J., et al., *Reflectance confocal microscopy for diagnosing keratinocyte skin cancers in adults.* Cochrane Database Syst Rev, 2018. **12**: p. CD013191.
67. Dinnes, J., et al., *Reflectance confocal microscopy for diagnosing cutaneous melanoma in adults.* Cochrane Database Syst Rev, 2018. **12**: p. CD013190.
68. Ferrante di Ruffano, L., et al., *Optical coherence tomography for diagnosing skin cancer in adults.* Cochrane Database Syst Rev, 2018. **12**: p. CD013189.
69. Dinnes, J., et al., *High-frequency ultrasound for diagnosing skin cancer in adults.* Cochrane Database Syst Rev, 2018. **12**: p. CD013188.