

# **DRAFT** Consultant Report

**AI Labs Group S.L.**

## **DermaSensor Case Study for Legit.Health Plus**

**2026-01-27**

**Prepared By:**

Brendan O'Leary

### **Disclaimer**

The conclusions and recommendations provided in this Consultant Report are based upon ATW's evaluation of the data, results and conclusions resident in the source documents provided by the client. ATW cannot confirm that all relevant documents or information were made available, and a complete, independent verification of the results and conclusions reported in the source documents provided was not conducted. The recommendations and opinions expressed herein are advisory in nature, and the client retains all responsibility for any use or implementation.

This page is intentionally left blank.

# Contents

<b>1</b>	<b>DermaSensor: A case study in SLA market authorization</b>	<b>3</b>
1.1	Regulatory pathway and timeline . . . . .	3
1.2	Summary of clinical evidence . . . . .	3
1.3	The DERM-SUCCESS pivotal study . . . . .	4
1.3.1	Study design . . . . .	4
1.4	MRMC reader studies . . . . .	5
1.4.1	DERM-SUCCESS initial reader study . . . . .	5
1.4.2	DERM-SUCCESS pigmented lesion reader study . . . . .	5
1.4.3	DERM-ASSESS III reader study . . . . .	5
1.5	Post-market requirements and limitations . . . . .	5
1.5.1	FDA-mandated post-market surveillance . . . . .	5
1.5.2	Diversity limitations in the pivotal trial . . . . .	6
1.5.3	Post-market evidence generation . . . . .	6
<b>2</b>	<b>Indications for use</b>	<b>6</b>
<b>3</b>	<b>Client Questions</b>	<b>6</b>
3.1	Prospective versus retrospective study designs . . . . .	7
3.2	Hybrid study models . . . . .	8
3.3	Image capture device compatibility . . . . .	8
3.4	AI model validation with geographic data considerations . . . . .	9
<b>4</b>	<b>Summary Recommendations</b>	<b>11</b>
4.1	Retrospective Data Collection . . . . .	11
4.2	Standalone Performance Assessment . . . . .	12
4.3	MRMC Second-Read Study Design . . . . .	12
4.3.1	Telemedicine Considerations . . . . .	13
4.4	Precision/Repeatability Study . . . . .	14
4.5	Reference Standard . . . . .	15
4.6	Pre-Submission . . . . .	15

**References** **16**

**Attachments** **18**

DermaSensor De Novo Decision Summary (DEN230008) FDA Decision Summary 19

Ferris et al. (2025) DERM-SUCCESS FDA Pivotal Study: A Multi-Reader Mul... 52

Hartman et al. (2024) Multicenter prospective blinded melanoma detection... . . . 61

Jaklitsch et al. (2025) Prospective evaluation of an AI-enabled ESS device... . . . 69

Lee and Tung (2024) Comment on Use of an ESS and AI device in the asse... . . . 71

Merry et al. (2025) Primary Care Physician Use of Elastic Scattering S... . . . . 73

Motavaf et al. (2025) Current landscape of AI Regulation in Dermatology . . . . 84

Seiverling et al. (2025) Enhancing Diagnostic Precision in Primary Care: A ... . . 86

Venkatesh et al. (2024) Learnings from the first AI-enabled skin cancer de... . . . 93

# 1 DermaSensor: A case study in SLA market authorization

The DermaSensor De Novo authorization (DEN230008) represents the first AI-enabled skin cancer detection device cleared by FDA for use by primary care physicians. (Venkatesh et al. 2024, 1) This section examines DermaSensor's regulatory path, clinical validation strategy, and post-market requirements, drawing from the FDA decision summary, published clinical studies, and regulatory analysis. (*De Novo Classification Request for DermaSensor 2024b*)

## 1.1 Regulatory pathway and timeline

DermaSensor received Breakthrough Device Designation from FDA on May 11, 2021, enabling more intensive FDA engagement during development. The company submitted its De Novo request (DEN230008) on February 3, 2023, and FDA granted the De Novo on January 12, 2024, establishing 21 CFR 878.1830 and product code QZS. (*De Novo Classification Request for DermaSensor 2024a, 1*)

The De Novo indications for use are:

The DermaSensor device is indicated for use to evaluate skin lesions suggestive of melanoma, basal cell carcinoma, and/or squamous cell carcinoma in patients aged 40 and above to assist in the decision regarding referral of the patient to a dermatologist. The DermaSensor device should be used in conjunction with the totality of clinically relevant information from the clinical assessment, including visual analysis of the lesion, by physicians who are not dermatologists. The device should be used on lesions already assessed as suspicious for skin cancer and not as a screening tool. The device should not be used as the sole diagnostic criterion nor to confirm clinical diagnosis of skin cancer.

The device uses elastic scattering spectroscopy (ESS) to analyze optical properties of skin lesions rather than image-based analysis.

## 1.2 Summary of clinical evidence

Table 1 summarizes the clinical studies supporting DermaSensor's regulatory authorization and subsequent validation. The studies illustrate a strategy of prospective data collection followed by retrospective reader studies, enabling multiple analyses from shared datasets.

The DERM-SUCCESS pivotal study enrolled patients prospectively with histopathology ground truth, establishing a foundation of well-characterized cases. Three retrospective MRMC reader studies then used cases from DERM-SUCCESS and earlier data collections, maximizing the analytical value from each enrolled patient. The reader studies used either randomized crossover designs (comparing aided vs. unaided reads) or sequential read designs

**Table 1:** Clinical studies described in the FDA decision summary (DEN230008) as supporting the De Novo authorization. The FDA noted that eight studies were conducted during development, but only these four used the final version of the device and algorithm. MRMC = multi-reader multi-case.

Study (FDA name)	Type	Readers	Cases	Publication
DERM-SUCCESS	Prospective	30 PCPs	1,005 patients 1,579 lesions	Merry et al. 2025
DERM-SUCCESS initial reader study	MRMC	108 PCPs	50 lesions (25 malig, 25 benign)	Ferris et al. 2025
DERM-SUCCESS pigmented lesion reader study	MRMC	77 PCPs	136 lesions (69 malig, 67 benign)	—
DERM-ASSESS III reader study	MRMC	118 PCPs	100 lesions (50 malig, 50 benign)	Seiverling et al. 2025

(unaided assessment followed by device-aided revision). Notably, the pivotal study and all reader studies used board-certified primary care physicians, matching the intended use population.

### 1.3 The DERM-SUCCESS pivotal study

The DERM-SUCCESS study (DERM-Spectroscopy Utilization for Clinical Characterization of Evaluation of Suspicious Skin) served as the pivotal clinical trial supporting FDA authorization. ([Merry et al. 2025](#))

#### 1.3.1 Study design

The study was an international, multicenter, prospective, blinded clinical study conducted at 22 primary care centers (18 in the US, 4 in Australia) by 30 primary care physicians. The study enrolled 1,005 patients with 1,579 lesions suggestive of melanoma, BCC, or SCC. All enrolled lesions were biopsied, and pathology findings were validated by 2–5 central dermatopathologists.

Each enrolled lesion was scanned with the device, but investigators and subjects were blinded to the device results. Clinical management was provided per standard of care.

The co-primary endpoints were met: device sensitivity (95.5%) was superior to PCP sensitivity (83.0%), and sensitivity + specificity exceeded 1. (*De Novo Classification Request for DermaSensor 2024b*, 18) The relatively low specificity (20.7%) was clinically justified because the device is intended to aid referral decisions rather than rule out malignancy. A high sensitivity, low specificity profile is appropriate for a triage device. This specificity-sensitivity tradeoff has been a persistent challenge for AI-enabled dermatology devices, contributing to the discontinuation of earlier devices like MelaFind. ([Venkatesh et al. 2024](#), 2)

## 1.4 MRMC reader studies

In addition to the pivotal DERM-SUCCESS trial, three multi-reader multi-case (MRMC) studies demonstrated that the device improves PCP referral decisions.

### 1.4.1 DERM-SUCCESS initial reader study

This study used a randomized crossover design: 108 board-certified PCPs each assessed 50 lesion cases (25 malignant, 25 benign) twice, once with device output and once without, in random order. Device-aided sensitivity improved from 82.0% (unaided) to 91.4% (aided). (*De Novo Classification Request for DermaSensor 2024b*, 20; Ferris et al. 2025)

### 1.4.2 DERM-SUCCESS pigmented lesion reader study

This study used a sequential read design: 77 PCPs first assessed lesions unaided, then revised their decision after seeing device output. The study evaluated 136 lesion cases (69 malignant, 67 benign) enriched for pigmented lesions. Device-aided sensitivity improved from 80.5% to 86.3%. Melanoma sensitivity specifically improved from 68.8% to 75.4%. (*De Novo Classification Request for DermaSensor 2024b*, 21–22)

### 1.4.3 DERM-ASSESS III reader study

This study also used a sequential read design, evaluating 118 PCPs assessing 100 lesion cases (50 malignant, 50 benign) enriched for melanoma (68% of malignant cases). The cases were drawn from the DERM-ASSESS III prospective study conducted at dermatology centers. (Hartman et al. 2024) Device-aided sensitivity improved from 73.7% to 81.8%. Melanoma sensitivity improved from 70.2% to 79.1%. (*De Novo Classification Request for DermaSensor 2024b*, 22–23; Seiverling et al. 2025)

## 1.5 Post-market requirements and limitations

The DermaSensor authorization included notable post-market requirements and acknowledged limitations.

### 1.5.1 FDA-mandated post-market surveillance

A key special control for 21 CFR 878.1830 requires post-market surveillance unless FDA determines it is not necessary. (*De Novo Classification Request for DermaSensor 2024a*, 3) In DermaSensor's case, FDA specifically required post-market performance testing in underrepresented populations, with enrollment targets including Fitzpatrick skin phototypes IV, V, and VI. (*De Novo Classification Request for DermaSensor 2024a*, 4–5)

### 1.5.2 Diversity limitations in the pivotal trial

The DERM-SUCCESS trial enrolled a patient population with limited diversity: 97.1% of patients were White, and while 27.5% were from Fitzpatrick skin types IV–VI, only 12.7% were from the most pigmented categories (V and VI). (*De Novo Classification Request for DermaSensor 2024b*; Venkatesh et al. 2024, 2)

FDA determined that the pivotal data was insufficient to fully characterize device performance across all skin phototypes and required DermaSensor to conduct additional studies in Fitzpatrick IV–VI populations post-market. (*De Novo Classification Request for DermaSensor 2024a, 4*) This concern about algorithmic bias in AI dermatology devices has been raised more broadly in the literature. (Lee and Tung 2024, 122; Motavaf et al. 2025, 2)

### 1.5.3 Post-market evidence generation

Since authorization, investigator-initiated studies have continued to evaluate DermaSensor performance in clinical practice. A prospective study at three dermatology clinics evaluated the device on patient-identified lesions, demonstrating 100% sensitivity (22/22 malignancies detected) with an AUC of 0.79, consistent with the primary care validation results. (Jaklitsch et al. 2025)

## 2 Indications for use

To leverage the DermaSensor predicate, Legit.Health should adopt the following indications for use for Legit.Health Plus:

The Legit.Health Plus device is indicated for use to evaluate skin lesions suggestive of melanoma, basal cell carcinoma, and/or squamous cell carcinoma in patients aged 40 and above to assist in the decision regarding referral of the patient to a dermatologist. The Legit.Health Plus device should be used in conjunction with the totality of clinically relevant information from the clinical assessment, including visual analysis of the lesion, by physicians who are not dermatologists. The device should be used on lesions already assessed as suspicious for skin cancer and not as a screening tool. The device should not be used as the sole diagnostic criterion nor to confirm clinical diagnosis of skin cancer.

These indications are intentionally identical to DermaSensor's to maximize the strength of the predicate relationship for the 510(k) submission.

## 3 Client Questions

The following questions were raised during our engagement and are addressed in the context of the DermaSensor precedent. These answers assume that Legit.Health will be submitting

a pre-submission to the FDA about their proposed study design.

### 3.1 Prospective versus retrospective study designs

DermaSensor utilized active recruitment to build its dataset. Given that we aim to evaluate the same endpoints, does the FDA require us to replicate their prospective design, or would a retrospective MRMC study using an existing image library be sufficient for a 510(k) submission?

#### Response:

A retrospective study using an existing image library could be acceptable for Legit.Health's 510(k) submission, and there are strong arguments for this approach given Legit.Health Plus's technological characteristics.

The key distinction between Legit.Health Plus and DermaSensor is the data modality. DermaSensor uses a proprietary elastic scattering spectroscopy (ESS) sensor to analyze optical properties of skin lesions. Before DermaSensor conducted its prospective studies, no retrospective dataset of ESS readings existed, so prospective data collection was the only option. In contrast, Legit.Health Plus analyzes standard clinical images from commonly available capture devices such as smartphone cameras. Because these image modalities already exist and have been used in clinical practice and research for years, retrospective image datasets with histopathology-confirmed ground truth are available and could potentially serve as validation data.

For a retrospective study to provide meaningful validation, the following considerations are important:

1. **Well-established ground truth:** Histopathology confirmation for malignant lesions and appropriate follow-up or expert consensus for benign lesions
2. **Representative case mix:** Adequate representation across skin phototypes, lesion types, and anatomic locations relevant to the intended use population
3. **Independent validation data:** Separation of validation data from training data to avoid overfitting

Prospective data collection is generally considered stronger evidence, particularly for machine learning devices, because it addresses concerns about shortcut learning, dataset shift, selection bias, and, critically, data independence. When validation data is collected prospectively from sites that were not used for training, independence is inherent. With retrospective data, demonstrating this independence requires rigorous documentation and controls. FDA routinely accepts retrospective validation studies when the data quality, representativeness, and independence from training data are adequately established.

A retrospective validation study using a well-truthed image library is a viable approach for the clinical validation component of Legit.Health's 510(k). The pre-submission should address:

1. The source and quality of the retrospective dataset, including ground truth methodology
2. How the dataset represents the US intended use population (demographics, skin phototypes, lesion types)
3. Independence of validation data from training data
4. Any supplemental prospective data collection planned to address population representativeness gaps

## 3.2 Hybrid study models

If prospective recruitment is mandatory, is a hybrid model (partial prospective recruitment supplemented by retrospective analysis) viable, or must we mirror the predicate's methodology exactly?

### Response:

As discussed above, I do not believe that prospective recruitment will be mandatory in this case.

## 3.3 Image capture device compatibility

Recalling our past conversations, you mentioned that the specific device used for image capture (iPhone, Android, etc.) could be crucial for the FDA. We have already conducted usability studies in Europe. Do you believe this data is sufficient, or should the usability study be conducted in the US? If a US-based study is required, could we leverage the prospective clinical trial to collect this data (hardware/camera performance) simultaneously?

### Response:

The concern about the specific device used for image capture is not a usability concern. Machine learning algorithms can be susceptible to differences in image acquisition technologies. FDA's presumption is that they are susceptible until data demonstrates that they are not.

If the device will be labeled for use with the iPhone, significant validation data with the iPhone will be needed. If the device will be labeled for use with particular makes and models of Android phones, the same will be needed.

If retrospective datasets are available that address this, that's one approach. Another approach might be prospective collection of matched images of individual lesions from a variety of makes and models of phones. Showing high agreement between the algorithm's responses to images of the same lesions across different phones, lighting conditions, etc., can address FDA's questions about whether the machine learning technologies in play are susceptible to image acquisition conditions.

Usability is likely to be a much more trivial concern relative to this. I have not reviewed the European usability study, but it is likely that it is at least somewhat valuable. Prospective data collection using phones as discussed here could also help to address FDA's usability questions.

### 3.4 AI model validation with geographic data considerations

Our current AI performance testing was conducted primarily using European datasets. To ensure the FDA accepts the training and validation of our algorithm for the US market, do you believe we need to reinforce our validation with data from the US? Specifically, could we maintain our core training on European data but use the data collected in the US clinical study as part of the independent clinical validation of the models?

#### Response:

Yes, Legit.Health will almost certainly need to collect clinical validation data in the United States, ideally from three or more sites.

FDA has issued guidance on the acceptance of clinical data from studies conducted outside the United States. ([Acceptance of Clinical Data to Support Medical Device Applications and Submissions 2018](#)) While data collected outside the US can be used to support device authorizations, in practice it is often difficult to justify, and FDA can be reluctant to accept it. The guidance establishes that OUS data should be applicable to the US population and US medical practice.

The DermaSensor clinical validation studies (DERM-SUCCESS, DERM-ASSESS III) were conducted in the US and Australia.

The approach Legit.Health describes (training on European data and validating on US data) is scientifically sound. A model trained on European data that performs well on US validation data demonstrates geographic generalizability, and using US data exclusively for validation clearly establishes independence between training and validation datasets.

Two additional considerations are important regardless of where the image data originates:

1. **MRMC readers must be US clinicians:** The readers in any MRMC study must reflect the US intended user population. If Legit.Health Plus is intended for primary care physicians, the MRMC study should use US-based, board-certified primary care physicians. DermaSensor's reader studies used 77–118 US PCPs each. This requirement is independent of where the lesion images were collected.
2. **OUS ground truth requires justification:** If histopathology or other reference standard determinations were performed outside the US, Legit.Health must justify why this ground truth is applicable to US clinical practice. FDA guidance states that sponsors "should justify why non-U.S. data reflect what is expected for a U.S. population with respect to disease occurrence, characteristics, practice of medicine, and clinician

competency.” (*Clinical Performance Assessment 2022*, 14) Pathology practices and diagnostic criteria may vary between countries; Legit.Health should be prepared to demonstrate that the pathology methods and classifications used are consistent with US standards.

## 4 Summary Recommendations

DermaSensor's regulatory pathway involved multiple studies, including both a crossover design (DERM-SUCCESS initial reader study) and sequential second-read designs (DERM-SUCCESS pigmented lesion reader study and DERM-ASSESS III reader study). This complexity may reflect evolving decisions about indications for use during their development process. Legit.Health can take a more direct route by adopting DermaSensor's authorized indications from the outset.

The recommended clinical evidence strategy involves three studies:

1. Retrospective **standalone performance study** (algorithm performance independent of clinician interaction)
2. Retrospective **MRMC second-read study** (primary clinical evidence)
3. Prospective **precision/repeatability study** (to address multiple phone models and image acquisition variability)

### 4.1 Retrospective Data Collection

DermaSensor required prospective data collection because elastic scattering spectroscopy readings cannot be retrospectively assigned to lesions that were never measured with the device. In contrast, Legit.Health Plus can retrospectively assign algorithm outputs to images collected in the past. The primary challenge will be demonstrating that any retrospective image dataset is completely independent from images used to train, tune, or otherwise develop the device. If retrospective data is used, Legit.Health should be prepared to demonstrate:

- Controls to ensure the data used for testing is sequestered from the development process and from model developers
- A data access log tracking each time the data is accessed
- Clear separation between development and FDA clinical validation datasets, ideally with test data from sites different from those used in development

These recommendations are consistent with FDA's draft guidance on AI-enabled device software functions, which includes additional information on data management that will be valuable for Legit.Health's submission. ([Artificial Intelligence-Enabled Device Software Functions 2025, 19, 22](#))

This approach can address questions about the performance of the device and its impact on clinical decisions (the standalone performance study and the MRMC second-read study). If a well-truthed and independent dataset exists with multiple images of the same lesion from multiple phones or cameras, the precision/repeatability study could also use retrospective data collection. However, such a dataset is unlikely to exist. Therefore, Legit.Health should plan to collect precision/repeatability data prospectively (see Section 4.4).

## 4.2 Standalone Performance Assessment

The standalone performance study characterizes the algorithm’s intrinsic performance independent of clinician interaction. This study should include a large number of cases to provide precise estimates of sensitivity, specificity, and performance across subgroups. A subset of this dataset can then be used for the MRMC study, which evaluates the device’s impact on reader decisions.

The standalone study is designed to satisfy the special control requirement that “[t]esting must demonstrate at least 90% sensitivity of the device output for lesions with high metastatic potential, or an alternative clinical consideration must be provided to justify lower sensitivity. Clinical justification must be provided for the reported specificity.” This establishes an absolute performance threshold for the algorithm itself, independent of how it affects reader decisions.

A power analysis will be needed to determine study size, informed by anticipated device performance and the need for subgroup analyses across lesion types, skin phototypes, and anatomic locations. DermaSensor’s standalone performance was assessed in the DERM-SUCCESS pivotal study, which enrolled 1,005 participants with 1,579 lesions across 22 clinical sites. (*De Novo Classification Request for DermaSensor 2024b, 12*) Even so, this study did not provide sufficient data to support authorization for individuals under age 40, and FDA required post-market data collection for individuals with Fitzpatrick skin phototypes IV–VI due to underrepresentation in the pivotal study. (*De Novo Classification Request for DermaSensor 2024b, 4–5*) Legit.Health should plan for a comparable sample size to support the required subgroup analyses, and should consider whether enrichment with datasets for underrepresented populations may reduce the risk of similar limitations.

## 4.3 MRMC Second-Read Study Design

The MRMC study evaluates the device’s impact on clinician decision-making using a subset of the cases from the standalone performance dataset. This study is designed to satisfy the special control requirement that “[d]ata must demonstrate superior accuracy of device-aided users’ diagnostic characterization of the indicated lesions compared to the accuracy of unaided users.” Therefore, the co-primary endpoints should be sensitivity *improvement* (aided vs. unaided) and AUC improvement, not absolute thresholds, but demonstrated superiority over unaided performance. While FDA’s MRMC guidance was developed for radiology CADe devices, the scientific question is the same: does access to the device output improve clinician diagnostic accuracy? (*Clinical Performance Assessment 2022*)

To generate ROC curves and calculate AUC, the study must collect more than binary referral decisions. In the DermaSensor MRMC studies, clinicians provided three responses for each lesion: a binary diagnosis (malignant or benign), a binary management decision (refer or not), and a confidence rating on their management decision using a 1–10 scale. The ROC curves were then constructed by combining the referral decisions with the confidence levels, enabling AUC analysis at multiple thresholds. (*Ferris et al. 2025*) This approach demonstrated statistically significant improvement in diagnostic performance with device assistance: in

the DERM-SUCCESS initial reader study, AUC improved from 0.708 (unaided) to 0.762 (aided), and in the DERM-ASSESS III reader study, AUC improved from 0.630 to 0.671. (Ferris et al. 2025; Seiverling et al. 2025) Legit.Health should adopt a similar methodology, collecting confidence ratings both before and after the device output is displayed.

Specificity should be a secondary endpoint with a pre-specified noninferiority margin. Some reduction in specificity may be acceptable: DermaSensor's clinical utility data showed specificity decreased from 44.2% to 32.4% when using the device, but this was offset by the significant improvement in sensitivity and reduction in false negatives. (Ferris et al. 2025) The specificity endpoint must be carefully pre-specified, including any noninferiority margin.

A power analysis will be needed to determine study size, informed by prior estimates of the device's sensitivity and specificity. DermaSensor's pivotal MRMC study (DERM-SUCCESS initial reader study) used 108 primary care physicians evaluating 50 lesion cases. (Ferris et al. 2025)

The case set should be enriched to ensure sufficient pigmented lesions (melanoma, pigmented nevi) are present for meaningful subgroup analysis. DermaSensor conducted a separate MRMC study (DERM-SUCCESS pigmented lesion reader study) specifically to demonstrate performance on pigmented lesions; Legit.Health can address this by ensuring the primary MRMC dataset includes enough pigmented cases to power a pre-specified subgroup analysis.

Because Legit.Health Plus functions as a second-read device, a sequential reading design is appropriate. In clinical practice, a primary care provider will not photograph a skin lesion unless they already have clinical suspicion about it. The "first read" occurs in the examination room when the clinician visually assesses the lesion and decides it warrants further evaluation. Only then does the clinician capture an image and submit it to Legit.Health Plus for a "second read." The device output informs the clinician's referral decision, but it does not replace or precede the initial clinical assessment.

FDA guidance describes this sequential design as follows: "A conventional reading by the readers without the CADe device (reader alone)" followed by "a second-read by the readers in which the CADe output is displayed immediately after conducting a conventional interpretation." (*Clinical Performance Assessment 2022*, 11) Because Legit.Health is pursuing indications for use that match DermaSensor's (evaluation of lesions "already assessed as suspicious for skin cancer"), this sequential design is appropriate.

### 4.3.1 Telemedicine Considerations

The sequential design described above assumes the intended use case where a clinician has physically examined the patient. This clinical workflow differs from a telemedicine scenario in which a dermatologist receives images without having examined the patient. In that use case, the physician's first encounter with the lesion would be the image itself. Depending on how the device is integrated into the workflow, the device output could be available concurrently with or even before the physician's assessment, or it could still function as a second-read tool if the physician reviews the image before seeing the device output. The appropriate study design would depend on the intended clinical workflow for that use case.

## 4.4 Precision/Repeatability Study

A separate, smaller prospective study should be conducted in a dermatology setting to assess device precision across image acquisition systems. Dermatologists would capture images of skin lesions using multiple smartphone models (e.g., the ten most popular devices in the US market).

This study addresses a unique challenge for image-based AI devices that did not apply to DermaSensor. Machine learning algorithms can be susceptible to differences in image acquisition technologies, and FDA's presumption is that they are susceptible until data demonstrates otherwise.

The goal of this study is to gather extensive data on a relatively small number of lesions. Each lesion should be photographed with every device under multiple conditions, generating many images per lesion. The lesion set should represent the full spectrum of lesion types the device will encounter in clinical use (melanoma, BCC, SCC, and benign lesions across skin phototypes and anatomic locations), but a large total number of lesions is not required. The focus is on demonstrating consistency of device output across acquisition systems, not on establishing diagnostic accuracy (which is addressed by the MRMC study).

A dermatology setting is appropriate for this study because dermatology clinics are naturally enriched with the malignant and suspicious lesion types that would be difficult to efficiently capture in primary care. To ensure benign lesions are adequately represented, dermatologists can also capture images of non-suspicious moles and other benign findings during the study.

Study design considerations:

- **Matched image capture:** Each lesion photographed with all devices under similar conditions
- **Primary endpoint:** Agreement in algorithm output across acquisition systems
- **Secondary endpoints:** Sensitivity and specificity by phone model, agreement in underlying algorithm score using an appropriate metric such as ICC
- **Environmental factors:** Consider also varying lighting conditions to further demonstrate robustness<sup>1</sup>

This study design can form the basis for a predetermined change control plan (PCCP) that would enable Legit.Health to add additional image acquisition systems over time without requiring a new 510(k) submission for each device, provided the new systems demonstrate comparable performance using the same methodology and do not require additional training or tuning of the algorithm.

---

<sup>1</sup>Only one variable should be changed at a time. If, for example, you intend to demonstrate that the technology works for 10 acquisition systems under both bright and dim conditions, you might need 20 images of each lesion to do this (one for each acquisition system and lighting condition).

## 4.5 Reference Standard

The reference standard methodology used in DermaSensor’s clinical studies provides a model for Legit.Health’s validation approach. In the DERM-SUCCESS pivotal study, all enrolled lesions were biopsied, and device output was compared to histopathology findings validated by two to five central study dermatopathologists (depending on histological severity and discordance). (*De Novo Classification Request for DermaSensor 2024b, 12*) This design eliminated verification bias by ensuring 100% of lesions had pathology-confirmed ground truth.

For the retrospective MRMC reader studies, lesion cases were drawn from these prospectively collected, histopathology-confirmed datasets. The ground truth for each case was established before the reader study, and readers in the MRMC studies were not involved in the original truthing process.

Any retrospective image dataset used for Legit.Health’s FDA clinical validation should have comparable reference standard characteristics:

- **Histopathology confirmation** for all malignant lesions (melanoma, BCC, SCC)
- **Central pathology review** with a defined consensus process for discordant cases
- **Separation of truthers from readers:** Clinicians who established ground truth should not participate in the MRMC reader study (*Computer-Assisted Detection Devices Applied to Radiology Images and Radiology Device Data - Premarket Notification [510(k)] Submissions 2022*)

For benign lesions that were not biopsied, an expert consensus panel or clinical follow-up may be acceptable, but this approach should be discussed with FDA in the pre-submission.

## 4.6 Pre-Submission

A pre-submission to FDA is essential to confirm these study designs before data collection begins. Key topics to address in the pre-submission include:

1. Acceptability of retrospective data and the independence of FDA clinical validation data from training data
2. Proposed MRMC study design, including reader population, case mix, and endpoints
3. Statistical analysis plan, including power calculations and noninferiority margins for specificity
4. Approach to demonstrating device performance across image acquisition systems
5. Representativeness of the study population for the US intended use population (demographics, skin phototypes, lesion types)

## References

- Acceptance of Clinical Data to Support Medical Device Applications and Submissions: Frequently Asked Questions.* 2018. Final Guidance Document. Docket Number FDA-2013-N-0080. U.S. Food and Drug Administration. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/acceptance-clinical-data-support-medical-device-applications-and-submissions-frequently-asked>.
- Artificial Intelligence-Enabled Device Software Functions: Lifecycle Management and Marketing Submission Recommendations.* 2025. Draft Guidance Document. Docket Number FDA-2024-D-4488. U.S. Food and Drug Administration. <https://www.regulations.gov/document/FDA-2024-D-4488-0002>.
- Clinical Performance Assessment: Considerations for Computer-Assisted Detection Devices Applied to Radiology Images and Radiology Device Data in Premarket Notification (510(k)) Submissions.* 2022. Final Guidance Document. Docket Number FDA-2009-D-0503. U.S. Food and Drug Administration. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-performance-assessment-considerations-computer-assisted-detection-devices-applied-radiology>.
- Computer-Assisted Detection Devices Applied to Radiology Images and Radiology Device Data - Premarket Notification [510(k)] Submissions.* 2022. Final Guidance Document. Docket Number FDA-2009-D-0503. U.S. Food and Drug Administration. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/computer-assisted-detection-devices-applied-radiology-images-and-radiology-device-data-premarket>.
- De Novo Classification Request for DermaSensor: Authorization Letter.* 2024a. De Novo Authorization Letter. No. DEN230008. U.S. Food and Drug Administration. [https://www.accessdata.fda.gov/cdrh\\_docs/pdf23/DEN230008.pdf](https://www.accessdata.fda.gov/cdrh_docs/pdf23/DEN230008.pdf).
- De Novo Classification Request for DermaSensor: Decision Summary.* 2024b. De Novo Decision Summary. No. DEN230008. U.S. Food and Drug Administration. [https://www.accessdata.fda.gov/cdrh\\_docs/reviews/DEN230008.pdf](https://www.accessdata.fda.gov/cdrh_docs/reviews/DEN230008.pdf).
- Ferris, Laura K, Erik Jaklitsch, Elizabeth V Seiverling, et al. 2025. “DERM-SUCCESS FDA Pivotal Study: A Multi-Reader Multi-Case Evaluation of Primary Care Physicians’ Skin Cancer Detection Using AI-Enabled Elastic Scattering Spectroscopy.” *Journal of Primary Care & Community Health* 16: 21501319251342106. <https://doi.org/10.1177/21501319251342106>.
- Hartman, Rebecca I, Nicole Trepanowski, Michael S Chang, et al. 2024. “Multicenter prospective blinded melanoma detection study with a handheld elastic scattering spectroscopy device.” *JAAD International* 15: 24–31. <https://doi.org/10.1016/j.jdin.2023.10.011>.

- Jaklitsch, Erik, Selina Chang, Sabrina Bruno, Nicholas D'Angelo, Joe K Tung, and Laura K Ferris. 2025. "Prospective evaluation of an AI-enabled elastic scattering spectroscopy device for triage of patient-identified skin lesions in dermatology clinics." *JAAD International* 23: 27–28. <https://doi.org/10.1016/j.jdin.2025.07.007>.
- Lee, Tai-Lin Irene, and Joe K Tung. 2024. "Comment on 'Use of an elastic-scattering spectroscopy and artificial intelligence device in the assessment of lesions suggestive of skin cancer: A comparative effectiveness study'." *JAAD International* 17: 122–23. <https://doi.org/10.1016/j.jdin.2024.06.005>.
- Merry, Stephen P, Ivana T Croghan, Kimberly A Dukes, et al. 2025. "Primary Care Physician Use of Elastic Scattering Spectroscopy on Skin Lesions Suggestive of Skin Cancer." *Journal of Primary Care & Community Health* 16: 21501319251344423. <https://doi.org/10.1177/21501319251344423>.
- Motavaf, Mateen, Danyal Raza, William Frye, and Melodi Javid Whitley. 2025. "Current landscape of AI Regulation in Dermatology." *Journal of the American Academy of Dermatology*, ahead of print. <https://doi.org/10.1016/j.jaad.2025.11.035>.
- Seiverling, Elizabeth V, Asghar Shah, Martin A Weinstock, Jane M Grant-Kels, Nathan P Falk, and Daniel M Siegel. 2025. "Enhancing Diagnostic Precision in Primary Care: A Multireader Multicase (MRMC) Study of an AI-Powered Handheld Elastic Scattering Spectroscopy Device for Informed Referral Decisions in Melanoma Evaluation." *The Journal of Clinical and Aesthetic Dermatology* 18 (10): 59–65.
- Venkatesh, Kaushik P, Kushal T Kadakia, and Stephen Gilbert. 2024. "Learnings from the first AI-enabled skin cancer device for primary care authorized by FDA." *Npj Digital Medicine* 7: 156. <https://doi.org/10.1038/s41746-024-01161-1>.

## Attachments

The following pages contain copies of the references cited in this report.

## DermaSensor De Novo Decision Summary (DEN230008)

### DE NOVO CLASSIFICATION REQUEST FOR DERMASENSOR

#### REGULATORY INFORMATION

FDA identifies this generic type of device as:

**Software-aided adjunctive diagnostic device for use by physicians on lesions suspicious for skin cancer.** A software-aided adjunctive diagnostic device for use by physicians on lesions suspicious for skin cancer is a prescription device that uses a software algorithm to analyze optical or other physical properties of a skin lesion and returns a classification of the skin lesion. The device is intended for use by a physician not trained in the clinical diagnosis and management of skin cancer as an adjunctive second-read device following identification of a suspicious skin lesion. It is not for use as a standalone diagnostic and is not for use to confirm a clinical diagnosis.

**NEW REGULATION NUMBER:** 21 CFR 878.1830

**CLASSIFICATION:** Class II

**PRODUCT CODE:** QZS

#### BACKGROUND

**DEVICE NAME:** DermaSensor

**SUBMISSION NUMBER:** DEN230008

**DATE DE NOVO RECEIVED:** February 2, 2023

**SPONSOR INFORMATION:** DermaSensor, Inc.  
80 SW 8<sup>th</sup> Street #2000  
Miami, Florida 33130 USA

#### INDICATIONS FOR USE

The DermaSensor device is indicated for use to evaluate skin lesions suggestive of melanoma, basal cell carcinoma, and/or squamous cell carcinoma in patients aged 40 and above to assist in the decision regarding referral of the patient to a dermatologist. The DermaSensor device should be used in conjunction with the totality of clinically relevant information from the clinical assessment, including visual analysis of the lesion, by physicians who are not dermatologists. The device should be used on lesions already assessed as suspicious for skin cancer and not as a screening tool. The device should not be used as the sole diagnostic criterion nor to confirm clinical diagnosis of skin cancer.

### LIMITATIONS

The performance of the device has not been specifically evaluated in patients with increased risk for skin cancer, e.g., inherited or drug-induced photosensitivity; genetic predisposition to melanoma or basal cell carcinoma (BCC); immune compromise; or other medical conditions that increase the risk of skin cancer or its metastasis.

The device is intended to assist in clinical decisions related only to the skin malignancies melanoma (including severely atypical nevi), squamous cell carcinoma (SCC), and BCC. It has been tested on each of these three common skin cancer types but has not been tested on rare skin cancer types; thus, it should not be used for lesions that are suggestive of malignancies other than melanoma, BCC and/or SCC.

The device is intended for use on primary lesions only and has not been tested on lesions that are previously biopsied, recurrent, or metastatic; on scars, tattoos, sunburned skin, or within a hairy area; or which are located on palms, soles, mucosal surfaces, genitals, ears, within 1 cm of the eye, or under nails.

Consistent with the lower prevalence of skin cancer in Fitzpatrick skin phototypes IV-VI, less data is available for sensitivity of the DermaSensor device for melanoma in these patients. The decision to refer patients with suspicious pigmented lesions in this group should be primarily based on clinical concern.

The sale, distribution, and use of the DermaSensor are restricted to prescription use in accordance with 21 CFR 801.109.

The device is not intended to be used as a stand-alone diagnostic device.

The device is not intended to replace biopsy.

The device is not intended to replace clinical decision making.

PLEASE REFER TO THE LABELING FOR A MORE COMPLETE LIST OF CONTRAINDICATIONS, WARNINGS AND PRECAUTIONS.

### DEVICE DESCRIPTION

The DermaSensor device (hereinafter referred to as 'DermaSensor', or the 'DermaSensor device'; **Figure 1**) utilizes optical spectroscopy and an artificial intelligence/machine learning (AI/ML) based software algorithm to analyze an intact skin lesion to which the device is non-invasively applied.

**Figure 1: DermaSensor Device**

The device is a combination of a handheld unit and a base unit. The handheld unit contains a xenon arc lamp and a fiber-optic probe tip which together transmit broadband white light to a lesion surface. Samples of the backscattered light from the tissue are collected by an adjacent detection optical fiber, also within the probe tip, and are conveyed to a microspectrometer, yielding Elastic Scattering Spectroscopy (ESS) spectral recordings. The handheld unit is operated using a touch-screen interface with step-by-step guidance. The small fiber-optic tip is the only component that contacts the patient. The handheld unit remains in the base when not actively being applied to a lesion and its battery is recharged by the base's wireless charging mechanism. The base unit also contains calibration material that is accessible to the handheld unit.

In the DermaSensor device, analysis of the optical recordings of backscattered light over the range of wavelengths is carried out using a proprietary ML-derived classifier algorithm. The spectrum of scattered intensity vs. wavelength is a pattern, which is analyzed by a proprietary classifier algorithm in the device's built-in microprocessor to assess for the potential presence of melanoma, squamous cell carcinoma, or basal cell carcinoma. An internal microprocessor and classifier algorithm analyze the recording and provide results to the user at the point of care. Results are provided as "Monitor," for a negative result, or "Investigate Further" for a positive result. For positive output ("Investigate Further") the DermaSensor additionally displays a Similarity Score of 1-10, with higher scores representing greater similarity to signals seen in malignant lesions.

#### **SUMMARY OF NONCLINICAL/BENCH STUDIES**

##### **BIOCOMPATIBILITY/MATERIALS**

DermaSensor does not contain substances that are intended to be introduced into the human body nor absorbed by or locally dispersed in the human body. The patient contacting materials are comprised of a stainless-steel tip encapsulating silica core fiber optics and synthetic polymer. The device tip is a surface-contacting device intended to contact intact skin for less than 3 minutes, and the user remains gloved during operation; therefore, meeting the requirements of <24 hours for only limited (e.g., transient) use. The risk assessment of the materials, manufacturing process, nature of contact and duration determined that only cytotoxicity, sensitization, and irritation testing were applicable to be performed. These tests were conducted according to the following applicable standards and guidance:

- ISO 10993-1 Fifth edition 2018-08 - *Biological evaluation of medical devices -- Part 1: Evaluation and testing within a risk management process.*
- ISO 10993-5: 2009/revised 2014, *Biological evaluation of medical devices – Part 5: Tests for in vitro cytotoxicity.*
- ISO 10993-10: 2010/revised 2014, *Biological evaluation of medical devices – Part 10: Tests for irritation and skin sensitization.*
- ISO 10993-12: 2012, *Biological evaluation of medical devices – Part 12: Sample preparation and reference materials.*
- FDA guidance: Use of International Standard ISO 10993-1: 2016, "*Biological evaluation of medical devices - Part 1: Evaluation and testing within a risk management process*" - *Guidance for Industry and Food and Drug Administration Staff*

The results are provided in **Table 1**. The device passed the tests and the results demonstrate mitigation of the risk of adverse tissue reactions and infections for the patient.

**Table 1: Summary of biocompatibility test methods and results**

Endpoint	Test Method	Findings
Cytotoxicity	MEM Elution Assay Method (ISO 10993-5:2009)	No evidence of cell lysis or toxicity
Sensitization	Guinea Pig Maximization Sensitization Test (ISO 10993-10:2010)	Test article met the requirements at 24, 48 and 72 hrs.
Irritation	Intracutaneous Irritation Test (ISO 10993-10:2010)	No erythema or edema observed at 24, 48 or 72 hrs.

#### **CLEANING AND DISINFECTION**

Cleaning and disinfection validations were conducted to validate the cleaning and disinfecting instructions for the reusable device components of the DermaSensor device. Cleaning and disinfection validation was performed in accordance with FDA guidance *Reprocessing Medical Devices in Health Care Settings: Validation Methods and Labeling – Guidance for Industry and Food and Drug Administration Staff* (March 17, 2015), AAMI TIR12:2010, and AAMI TIR30:2011 (R2016). The cleaning procedure per the instructions for use was performed and the bio load reduction and cytotoxicity (MEM Elution) were evaluated. All testing and results were considered to be adequate and met the above standards.

**SHELF LIFE/STERILITY**

The DermaSensor device is provided non-sterile and is not intended for sterilization by the end user. The use life of the DermaSensor device is a minimum of three years. Product use-life testing was performed for the device and the results were considered to be adequate.

**ELECTROMAGNETIC COMPATIBILITY (EMC) & ELECTRICAL SAFETY**

The following Electrical Safety and Electromagnetic Compatibility testing has been performed:

- ETSI EN 301 489 -17 V3.2.2 (2019-12) and V3.2.3 (2019-11); ElectroMagnetic Compatibility (EMC) standard for radio equipment and services; Part 17: Specific conditions for Broadband Data Transmission Systems; Harmonized Standard for ElectroMagnetic Compatibility
- IEC 60601-1-2:2014 Medical Electrical Equipment - Part 1-2: General requirements for basic safety and essential performance - Collateral Standard: Electromagnetic disturbances - Requirements and tests
- IEC 60601-1:2005 + A1:2012 Medical electrical equipment - Part 1: General requirements for basic safety and essential performance
- IEC 62133-2:2017 Secondary cells and batteries containing alkaline or other non-acid electrolytes - requirements for portable sealed secondary cells, and for batteries made from them, for use in portable applications-Part 2: Lithium systems
- IEC 62471:2006 Photobiological safety of lamps and lamp systems. According to the standard, the pulsed light lamp has been considered as a non-general lighting lamp (non-GLS)
- IEC 60601-1-6: 2010 + A1:2013 Medical electrical equipment - Part 1-6: General requirements for basic safety and essential performance - Collateral standard: Usability
- IEC 60601-2-57:2011 Medical electrical equipment - Part 2-57: Particular requirements for the basic safety and essential performance of non-laser light source equipment intended for therapeutic, diagnostic, monitoring and cosmetic/aesthetic use

DermaSensor passed all relevant portions of the testing, and the testing and results were considered to be adequate.

### SOFTWARE/CYBERSECURITY

All components of the device are controlled/monitored by software, which is responsible for the functionality, user interface, safety checks and performance accuracy. The software was developed and tested according to the following FDA guidance documents and standards:

- FDA guidance, *Guidance for the Content of Premarket Submission for Software Contained in Medical Devices* (May 11, 2005)
- FDA guidance, *Content of Premarket Submissions for Management of Cybersecurity in Medical Devices* (October 2, 2014)
- FDA guidance, *Cybersecurity in Medical Devices: Quality System Considerations and Content of Premarket Submissions* (September 2023)
- IEC 62304: 2006 *Medical device software - Software life-cycle processes*
- ISO 14971:2019 *Medical devices - Application of risk management to medical devices*

Software documentation and testing, including cybersecurity information, demonstrates that the software will operate in a manner described in the specifications. The hazard analysis characterized software and cybersecurity risks, including device malfunction, measurement-related errors, sensor, cable and other hardware failures, and unauthorized access by malicious end users. The submission describes verification and validation testing to address the potential hazards with satisfactory results. The cybersecurity documentation included a cybersecurity hazard analysis and mitigation information, an upgrade and maintenance plan, other information for safeguarding the device during manufacturing and upon commercial distribution, and warning and precaution information in the product labeling.

Overall, the software documentation contains sufficient detail to provide reasonable assurance that the software will operate in a manner described in the specifications. Software security information demonstrated the device is protected from cyber vulnerability threats. All testing and results were considered to be adequate and met the above standards.

#### Algorithm Development

Initial work on the ESS technology identified the optimal wavelength range and spectral features relevant to skin cancer detection. Clinical development began with a blinded, multicenter clinical study through an initial ML retrospective algorithm, for which the algorithm was developed and tested on the same data set to evaluate the ability of ESS to differentiate malignant from benign melanocytic lesions. A subsequent blinded, multicenter clinical study investigated the performance of an algorithm using ESS and ML for the most common types of skin cancers (Rodriguez-Diaz et al. *Photochem, Photobiol.*

2019;95(6):1441-1445). The training dataset contained 950 lesions contributing ~4,200 spectra while the testing dataset included 361 lesions, for a total of 1,311 skin lesions. The training and testing datasets were independent and were chosen randomly. This study dataset was later used entirely for algorithm training.

The algorithm was trained in the spectral range of 360 to 810 nm, selected due to the low signal-to-noise ratio in this range. Forty-seven wavelengths were chosen within this range to reduce dimensionality. Tuning sensitivity estimates were 96.5% (221/229). Specificity for physician-biopsied lesions was 20.6% (28/136). The estimated specificity for physician-selected un-biopsied lesions was 33.0% (128/388). When pooling lesions across the tuning datasets, there was high sensitivity (i.e., at or above 90%) for each of the three individual cancers. The high sensitivity observed in tuning provided the basis for further clinical studies. The associated specificity of 20.6% was clinically accepted because it was paired with the high sensitivity and was supported by the greater benefit of identifying skin cancer than the risk of excess referrals.

The patient demographics for the lesions used to train and tune the dataset are summarized in **Table 2**. The training and tuning data included a majority of older patients (age 60-79), Fitzpatrick phototype I-III, and individuals who self-identified as White. This pattern reflects the greater incidence of skin cancer in these populations, particularly considering that lesions on the palms, soles, and under the nails were specifically excluded from the study.

**Table 2: Summary of patient demographics in algorithm training and tuning datasets**

	Training Dataset		Tuning Dataset	
	Sample Size	Percent (%)	Sample Size	Percent (%)
<b>Age</b>				
20-39	81	7.6%	62	14.2%
40-59	87	8.2%	87	19.9%
60-79	604	56.6%	216	49.3%
80+	45	4.2%	73	16.7%
Unknown	250	23.4%	0	0.0%
<b>Fitzpatrick Phototype</b>				
I	67	6.3%	64	14.6%
II	332	31.1%	160	36.5%
III	244	22.9%	91	20.8%
IV	46	4.3%	50	11.4%
V	8	0.7%	54	12.3%
VI	3	0.3%	16	3.7%
Unknown	367	34.4%	3	0.7%
<b>Ethnicity/Race</b>				
White	784	73.5%	402	91.8%
Black or African American	5	0.5%	15	3.4%
Asian	9	0.8%	5	1.1%

Multiracial	2	0.2%	1	0.2%
Native Hawaiian or Other	17	1.6%	15	3.4%
Unknown	250	23.4%	0	0.0%
<b>Gender</b>	<b>Sample Size</b>	<b>Percent (%)</b>	<b>Sample Size</b>	<b>Percent (%)</b>
Male	424	39.7%	169	38.6%
Female	265	24.8%	195	44.5%
Unknown	378	35.4%	74	16.9%

The anatomical locations of the lesions used to train and tune the dataset are summarized in **Table 3**. Lesions on the head and neck constituted the majority of lesions in the training and tuning data set. This pattern is consistent with the relatively higher rate of skin cancer on the head and neck, particularly of SCC and BCC, due to greater sun exposure.

**Table 3: Summary of anatomical location of lesions in algorithm training and tuning datasets**

Body site	Training Dataset		Tuning Dataset	
	Sample Size	Percent (%)	Sample Size	Percent (%)
Sun-exposed				
Head/neck	136	5.9%	192	15.6%
Chest/Upper back	75	3.3%	140	11.4%
Forearms/hands	80	3.5%	120	9.7%
Lower legs	33	1.4%	44	3.6%
Other	25	1.1%	41	3.3%
Unexposed				
Buttocks/hips	2	0.1%	5	0.4%
Upper legs	25	1.1%	40	3.2%
Lower trunk	16	0.7%	39	3.2%
Other	1	0.0%	2	0.2%
Unknown	1900	82.9%	610	49.5%

The diagnoses for the lesions used to train and tune the dataset are summarized in **Table 4**. The greatest number of lesions were BCC, followed by SCC, and then melanoma (including highly atypical nevi). This pattern is consistent with the relatively higher rate of BCC than SCC than atypical pigmented lesions in the US population. Additional lesions were included as mimics of melanoma, SCC, and BCC to support sensitivity and specificity assessments.

**Table 4: Summary of lesion diagnoses in algorithm training and tuning datasets**

	Training Dataset	Tuning Dataset

Diagnosis Distribution	Sample Size	Percent (%)	Sample Size	Percent (%)
Melanoma and severely dysplastic nevi	187	8.2%	38	3.1%
SCC	207	9.0%	83	6.7%
BCC	294	12.8%	107	8.7%
Mixed Malignant	2	0.1%	1	0.1%
Melanoma mimic				
SK (brown)	158	6.9%	209	17.0%
Benign Nevi	631	27.5%	367	29.8%
Spitz Nevi	0	0.0%	0	0.0%
Lentigo	98	4.3%	41	3.3%
Angioma	15	0.7%	12	1.0%
Dermatofibroma	4	0.2%	9	0.7%
Tattoo/foreign objects	0	0.0%	0	0.0%
SCC Mimics				
AK	172	7.5%	145	11.8%
Wart	14	0.6%	4	0.3%
Dermatitis	0	0.0%	0	0.0%
SK (nonbrown)	75	3.3%	104	8.4%
BCC Mimics				
Acne/rosacea	0	0.0%	0	0.0%
Sebaceous hyperplasia/cyst	3	0.1%	4	0.3%
Other benign tumors	0	0.0%	0	0.0%
Insect bite	0	0.0%	0	0.0%
Xanthoma/ xanthelasma	0	0.0%	0	0.0%
Other malignant lesions	6	0.3%	1	0.1%
Other benign lesions	427	18.6%	109	8.8%

**NON-CLINICAL PERFORMANCE TESTING – BENCH**

Bench testing was conducted to demonstrate that the DermaSensor device performs as expected under the anticipated conditions of use. **Table 5** describes the bench testing conducted to demonstrate the device performance characteristics.

**Table 5: Summary of testing and description/acceptance criteria**

Test	Description/Acceptance Criteria
Dimensional	The tolerance stack for 1) optical design components and the 2) wireless charging circuit is acceptable.
Battery Performance	Testing was conducted to IEC 62133, IEC 62471, ETSI EN 301 489-17, IEC 60601-1-2, IEC 60601-1, IEC 60601-2-57, and 60601-1-6 and passed. All batteries tested surpassed the

	use life target and the review of data showed the batteries were not adversely impacted by the stress of the testing.
Lamp Reliability	All the test articles registered sample counts and demonstrated stability within the range specified in the acceptance criteria, indicating that the lamp life is stable over the simulated use life days.
Optical and Lamp Performance	Lamp engine testing was conducted to IEC 62471 and found to be exempt. Testing was conducted to ETSI EN 301 489-17, IEC 60601-1-2, IEC 60601-1, IEC 60601-2-57, and 60601-1-6 and passed. The optical configuration provided the maximum captured light from the lamp back to the microspectrometer.
WiFi connectivity	Testing was conducted to ETSI EN 301 489-17, IEC 60601-1-2, IEC 60601-1, IEC 60601-2-57, and 60601-1-6 and passed. IP address assigned demonstrating that the device connection is successful.
Calibrator Configuration	The peak signal of a calibration recording achieved the necessary dynamic range and demonstrated that the calibration is successful.
Wireless Charging	Testing was conducted to ETSI EN 301 489-17, IEC 60601-1-2, IEC 60601-1, IEC 60601-2-57, and 60601-1-6 and passed. The Handheld Unit charged while placed in the Base at 120v and 240v. The Handheld Unit remained functional while placed in and removed from the Base.
Label Validation	All text content is legible.
Estimation of Useful Life	The device functioned beyond useful life and simulated use days.
PCBA functional verification	Testing was conducted to ETSI EN 301 489-17, IEC 60601-1-2, IEC 60601-1, IEC 60601-2-57, and 60601-1-6 and passed. Testing of the device operations performed as intended indicating the PCBAs function as intended.
System Packaging and Distribution	Testing was conducted to ASTM D4169 and passed.
Manual Cleaning, Disinfection, Sensitization, Cytotoxicity, Irritation	Testing was conducted to AAMI TIR12, AAMI TIR30, EN ISO 10993-1, EN ISO 10993-5, EN ISO 10993-10, EN ISO 10993-12 and results passed.

### **HUMAN FACTORS/USABILITY**

Human factors testing was successfully completed with formative and summative validation studies, summarized below (**Table 6**).

**Table 6: Summary of Human Factors Testing**

Study	Summary
Human Factors Validation Testing	<p>After formative testing, summative human factors validation testing evaluated user tasks through simulated-use to assess the use-related hazards of the device. Key attributes of the intended users, uses, and use environments were considered and sufficiently simulated to demonstrate safe and effective operation of the device under actual conditions of use.</p> <p>Two (2) user groups participated in the testing: 15 primary care physicians (PCPs) representing a range of clinical experiences working in primary care, and 15 mid-level practitioners that included Physician Assistants (PAs), Nurse Practitioners (NPs), and Registered Nurses (RNs). Participant performance during the simulated use scenarios, responses to knowledge/comprehension questions, and any participant comments on performed user tasks and the device-user interfaces were reviewed.</p> <p>All 30 participants were able to identify the indications for use and all contraindications for use, and that this device should not be used for the sole diagnosis of skin cancer. Use difficulties, close calls, and use errors were analyzed to determine their likely root cause(s) and were adequately mitigated. The results demonstrated that participating PCPs were able to independently and appropriately take recordings from mock lesions with acceptable levels of residual risk.</p>

**REPEATABILITY/REPRODUCIBILITY**

A repeatability study was conducted across two study sites, each with three PCP investigators. At least two investigators completed the full recording process four times for each enrolled lesion, twice with one device (Device A) and twice with a second device (Device B). Analyses included Intraclass Correlation (ICC) and Fleiss’s Kappa (reproducibility with multiple raters); Positive Percent Agreement (PPA) between two runs and Negative Percent Agreement (NPA) for repeatability; and Average Positive Agreement (APA) between three physicians with Device A and Average Negative Agreement (ANA) between three physicians with Device A for reproducibility.

Secondary endpoints included a repeat of the primary analysis, separately for all malignant and all high-risk lesions using biopsy results as the gold standard. The malignant lesion analysis included all malignant lesions but did not include highly atypical melanocytic nevi since this pathology was not considered malignant and since there is a high rate of discordance for it among dermatopathologists.

The device agreement between device scans (i.e., repeatability) for all 64 lesions was 84.2% (95% CI 78.2-89.1%) and between physicians (i.e., reproducibility) was 87.8% (95% CI: 82.2-92.1%). For lesions that were high risk, the agreement was 97.1% (95% CI: 89.9-99.6%) for repeatability and 97.1% (95% CI: 89.9- 99.6%) for reproducibility. The PPA and NPA for repeatability were 94.8% (95% CI: 92.1-97.1%) and 46.0% (95% CI:

32.1-60.0%) respectively. The APA and ANA for reproducibility were 91.6% (95% CI: 86.8-95.7%) and 44.8% (95% CI: 20.8-63.3%) respectively.

The device agreement of the spectral score (1-10) between device scans (i.e., repeatability) was assessed. The repeatability percentage of scores 1-8 is below 37.5%, score 9 is 50.0%, and score 10 is 62.5%. The device agreement of the underlying continuous model output between device scans (repeatability) was also assessed.

Overall, the device shows low repeatability and reproducibility of negative binary decisions (NPA and ANA), low repeatability of the reported spectral similarity score, and high variability of the underlying continuous model output. The device shows repeatability and reproducibility of positive binary decisions (PPA and APA) greater than 90%. The repeatability and reproducibility information was considered acceptable within the context of the clinical performance testing and is included in the labeling.

### **CLINICAL PERFORMANCE TESTING**

Eight clinical studies were conducted to assess performance of the DermaSensor device and algorithm during development. The final version of the device and algorithm were tested with lesions from the intended lesion types in four of these studies. The results of those four studies provide the basis for determination of safety and effectiveness of the device when used as intended to evaluate skin lesions that a primary care physician has identified as suggestive of melanoma, squamous cell carcinoma (SCC), and/or basal cell carcinoma (BCC).

#### **Standalone Algorithm Performance Evaluation**

The standalone performance of the device was assessed in the DERM-SUCCESS pivotal clinical study. The study was designed to identify the sensitivity and specificity of the device output compared to ground truth diagnosis (biopsy). The study also assessed the sensitivity and specificity of primary care physicians (PCPs) on the same lesion set. The study was an international, multicenter, prospective, blinded clinical study conducted at 22 study sites including 18 locations in the United States, with a range of sun exposure patterns (Arizona, California, Florida, Kansas, Minnesota, Pennsylvania, Rhode Island, Tennessee, Texas, Utah, Virginia), and four in Australia. The investigators were 30 primary care physicians, and the study enrolled a total of 1,005 participants with 1,579 lesions. The study enrolled participants with skin lesions suggestive of melanoma, BCC, and/or SCC. Each enrolled lesion was scanned with the device. Investigators and subjects were blinded to the device results, and clinical management was provided per standard of care. All enrolled lesions were biopsied. DermaSensor device output (binary classification of a negative result, "Monitor," or a positive result, "Investigate Further") was compared to pathology findings that were validated by two to five central study dermatopathologists depending on the histological severity and discordance for the diagnoses.

#### **Patient inclusion/exclusion**

The study enrolled men or women of any ethnic group aged 22 and older with a primary skin lesion(s) suggestive of melanoma, BCC, or SCC. Participants were recruited in

primary care offices where physicians performed their own clinical assessments, used the DermaSensor device, and performed biopsies of the lesions.

Patients were excluded from enrollment in the study if they met any of the following exclusion criteria:

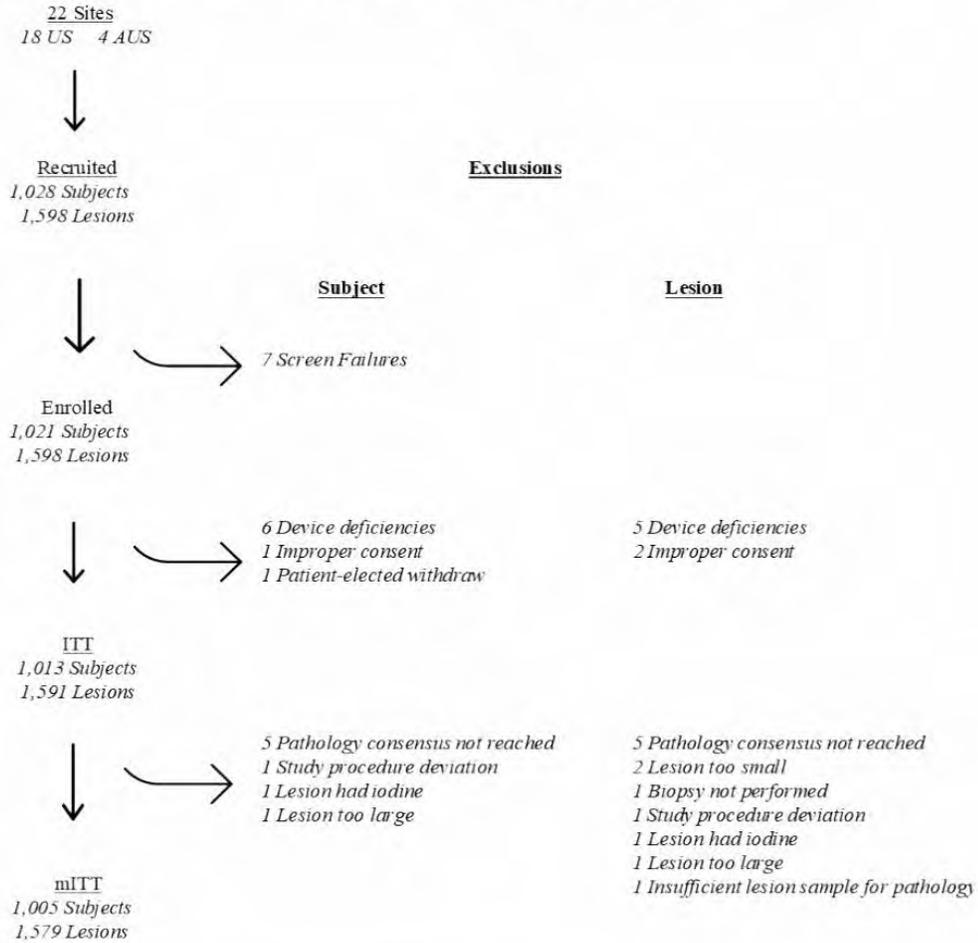
- Lesion < 2.5mm in diameter or > 15mm in diameter
- Lesion surface not accessible (e.g. inside ears, under nails, completely covered by a crust or scale)
- Lesion on area of crust, psoriasis, eczema, or similar skin condition
- Lesion has erosion and/or ulceration with no area >2.5mm intact
- Lesion has foreign matter (e.g. tattoo, splinter, dermoscopy oils, or other medicated or non-medicated topical solutions)
- Lesion in which the device tip cannot be placed entirely within the border of the targeted area
- Lesion located on acral skin (e.g. sole or palms)
- Lesion located within 1 cm of the eye
- Lesion on or adjacent to scars, areas previously biopsied, or areas subjected to any past surgical intervention
- Lesion located on mucosal surfaces (e.g. genitals, lips)
- Lesion located on acute sunburn
- Six (6) or more lesions suggestive of melanoma, basal cell carcinoma, and/or squamous cell carcinoma requiring biopsy to assess risk of malignancy

See **Table 7** and **Figure 2**, for additional summary information regarding the number of participants and inclusion/exclusion of participants in this study.

**Table 7: Participant accountability**

	<b>Participants n (%)</b>	<b>Lesions n (%)</b>
Enrolled	1,021	1,598
ITT Safety Population	1,013 (99.2%)	1,591 (99.6%)
<i>Excluded from ITT Population</i>	8 (0.8%)	7 (0.4%)
mITT Effectiveness Population	1,005 (98.4%)	1,579 (98.8%)
<i>Excluded from mITT Population</i>	8 (0.8%)	12 (0.8%)

**Figure 2: Inclusion/Exclusion diagram for all recruited participants and lesions**



**Table 8** summarizes the demographics of participants in the study. The participants in the DERM-SUCCESS pivotal study included a majority of older patients (age 60-79), Fitzpatrick phototype I-III, and individuals who self-identified as White. This pattern reflects the greater incidence of skin cancer in these populations, particularly considering that lesions on the palms, soles, and under the nails were specifically excluded from the study.

**Table 8: Participant characteristics in DERM-SUCCESS pivotal study (mITT, n=1,005)**

Characteristics	n (%)
<b>Sex</b>	
Male	487 (48.5%)

Female	518 (51.5%)
<b>Age</b>	
Mean (STD)	58.5 (15.1)
Median (Q1-Q3)	60.0 (49.0-69.0)
Min-Max	22-95
<b>Ethnicity</b>	
Hispanic or Latino	79 (7.9%)
Not Hispanic or Latino	913 (90.8%)
Unknown	13 (1.3%)
<b>Race</b>	
White	976 (97.1%)
Native Hawaiian or Other Pacific Islander	3 (0.3%)
Asian	9 (0.9%)
Black or African American	7 (0.7%)
Other/Multiracial	10 (1.0%)
<b>Fitzpatrick Phototype</b>	
I - Always burns, never tans	99 (9.9%)
II - Always burns, tans minimally	278 (27.7%)
III - Sometimes mild burn, tans uniformly	352 (35.0%)
IV - Burns minimally, always tans well	148 (14.7%)
V - Very rarely burns, tans very easily	110 (10.9%)
VI - Never burns	18 (1.8%)
<b>Risk Factors</b>	
Ultraviolet light exposure (natural or tanning bed)	455 (45.3%)
Weakened immune system	32 (3.2%)
Family history of skin cancer	332 (33.0%)
Xeroderma Pigmentosum	1 (0.1%)
Fair skin; freckling; light hair	362 (36.0%)
Many moles and/or dysplastic nevi	331 (32.9%)
Personal history of skin cancer	233 (23.2%)
New or changing lesion(s)	725 (72.1%)
None of the above	29 (2.9%)
<b>Number of Eligible Lesions per subject Enrolled</b>	
1	657 (65.4%)
2	207 (20.6%)
3	78 (7.8%)
4	41 (4.1%)
5	22 (2.2%)

**Table 9** summarizes the lesion characteristics of patients in the study. In the DERM-SUCCESS pivotal study, lesions on the trunk constituted the majority of lesions, followed by lesions on the arms and head.

**Table 9: Lesion characteristics in DERM-SUCCESS pivotal study (mITT, n=1,579)**

Anatomic Location	n	% Anatomic Location	% Total
<b>Head</b>	<b>247</b>		<b>15.6%</b>
Scalp	50	20.2%	3.2%
Forehead	59	23.9%	3.7%
Cheek or nose	85	34.4%	5.4%
Chin	8	3.2%	0.5%
Other	45	18.2%	2.8%
<b>Arm</b>	<b>299</b>		<b>18.9%</b>
Upper Arm	108	36.1%	6.8%
Elbow	11	3.7%	0.7%
Forearm	121	40.5%	7.7%
Hand	26	8.7%	1.6%
Other	33	11.0%	2.1%
<b>Leg</b>	<b>207</b>		<b>13.1%</b>
Upper Leg	66	31.9%	4.2%
Lower Leg	116	56.0%	7.3%
Knee	10	4.8%	0.6%
Foot	7	3.4%	0.4%
Other	8	3.9%	0.5%
<b>Trunk</b>	<b>826</b>		<b>52.3%</b>
Neck	63	7.6%	4.0%
Chest	128	15.5%	8.1%
Upper Back	361	43.7%	22.9%
Lower Back	149	18.0%	9.4%
Abdomen	64	7.7%	4.1%
Pubic, Inguinal	5	0.6%	0.3%
Buttocks	9	1.1%	0.6%
Other	47	5.7%	3.0%

**Table 10** summarizes the diagnoses of the lesions included in the study according to dermatopathology. The final pathology results were “high risk lesions” (high risk melanocytic lesions (melanoma and highly atypical nevi), SCC, and BCC). Consistent with the known occurrence rate, BCC was the most common lesion, with a slightly smaller number of SCC, and approximately one half the number of high risk melanocytic lesions.

**Table 10: Final dermatopathology diagnoses in DERM-SUCCESS pivotal study (mITT, n=1,579)**

Parent Pathology or Diagnosis	n (%)
<b>High Risk Lesions</b>	<b>224 (14.2%)</b>
Melanoma and severely atypical nevi	48 (3.0%)
Squamous Cell Carcinoma (SCC)	86 (5.4%)
Basal Cell Carcinoma (BCC)	90 (5.7%)
<b>Low Risk Lesions</b>	<b>1355 (85.8%)</b>
Benign melanocytic nevus	500 (36.9%)
Seborrheic keratosis	490 (36.2%)
Actinic Keratosis (AK)	71 (5.2%)
Lentigo	65 (4.8%)
Other	53 (3.9%)
Verruca / Wart	48 (3.5%)
Dermatofibroma	42 (3.1%)
Lichenoid keratosis	29 (2.1%)
Angioma or vascular lesion	20 (1.5%)
Categories with <20 lesions	37 (2.3%)
<i>Sebaceous hyperplasia</i>	<i>9 (0.7%)</i>
<i>Acrochordon</i>	<i>8 (0.6%)</i>
<i>Scar</i>	<i>6 (0.4%)</i>
<i>Epidermal cyst</i>	<i>5 (0.4%)</i>
<i>Folliculitis</i>	<i>3 (0.2%)</i>
<i>Angiofibroma</i>	<i>2 (0.1%)</i>
<i>Rosacea</i>	<i>1 (0.1%)</i>
<i>Excoriation</i>	<i>1 (0.1%)</i>
<i>Solar elastosis</i>	<i>1 (0.1%)</i>
<i>Angiokeratoma</i>	<i>1 (0.1%)</i>

Study endpoints

The co-primary endpoints of the DERM-SUCCESS study were: DermaSensor sensitivity compared to that of the study physicians (primary care physicians); and a sensitivity + specificity > 1. The endpoint of sensitivity compared to the primary care physicians was selected to determine if the device provides superior sensitivity compared to primary care physicians. The second co-primary endpoint was selected to confirm that the performance was non-random. The secondary endpoint of the DERM-SUCCESS study was DermaSensor sensitivity compared to a performance goal of 90%. This endpoint was selected to compare the sensitivity of the DermaSensor device to that of the current gold standard for skin lesion evaluation, i.e., dermatologist evaluation, which literature review indicated is approximately 90%.

Results of DERM-SUCCESS pivotal study

The DERM-SUCCESS pivotal study demonstrated an overall sensitivity of 95.5% when assessed over all ages and Fitzpatrick skin phototypes for identification of melanoma, SCC,

or BCC. The first co-primary endpoint was met, as the device sensitivity for high-risk lesions, defined as the three target malignancies of melanoma (including highly atypical nevi), SCC, and BCC, across all ages and Fitzpatrick skin phototypes was 95.5%, which was superior to the sensitivity of PCPs for the same population (83.0%, p-value <0.0001) as shown in **Table 11**. Specificity for all high-risk lesions was 20.7% (18.5%-23.1%), lower than the PCP specificity of 54.2% (50.7-57.6%). The second co-primary endpoint was met with sensitivity (95.5%) + specificity (20.7%) > 1. The secondary endpoint was met, with overall sensitivity >90% (<0.001). The specificity of 20.7% was deemed clinically acceptable when paired with the high overall sensitivity (95.5%) due to the significantly greater risk of a false negative, particularly for melanoma, than a false positive. The low specificity is also mitigated by the intended use of the device as one component/adjunctive information in the decision of the provider.

**Table 11: Summary of DERM-SUCCESS study sensitivity and specificity results**

Sensitivity (%) and 95% Confidence Interval <sup>1</sup>				
	# Lesions	Device	PCP	P-value <sup>2</sup>
All malignant lesions - sensitivity	224	95.5% (91.7%-97.6%)	83.0% (77.7%-87.3%)	<0.001
Benign lesions - specificity	1355	20.7% (18.5%-23.1%)	54.2% (50.7-57.6%)	N/A

<sup>1</sup> 95% Confidence Interval calculated accounting for within-subject correlations using Wilson method

<sup>2</sup> P-value obtained using the method of moments for clustered paired data

The performance of the DermaSensor device across age groupings is presented in **Table 12**.

**Table 12: Device performance across age groupings**

Age Group	Device Sensitivity	Device Specificity	PCP Unaided Sensitivity	PCP Aided Sensitivity	Difference
40-49*	100% (9/9)	22.9% (39/170)	83.1%*	91.8%*	8.7%*
50-59	95.6% (43/45)	24.1% (72/299)	77.7-83.6%	90.3-94.4%	6.8-16.8%
60-69	93.9% (62/66)	22.0% (89/405)	72.9-84.9%	82.0-91.0%	6.2-9.2%
70-79	98.3% (59/60)	12.6% (27/214)	84.6-90.7%	87.4-98.6%	2.8-7.9%
80-89	97.3% (36/37)	10.8% (7/67)	79.6-93.9%	98.9-99.1%	5.0-19.4%

\* Only DERM-SUCCESS Pigmented Reader Study results are available.

The area under the receiver operating characteristic (ROC) curve (AUROC) for the device was 0.7896. This was greater than the AUROC for PCPs, which was 0.7404 for all lesions. Device AUROC was also greater than the AUROC for the PCPs when limited to study lesions for which they reported low confidence in their clinical diagnosis, which was 0.5555. When using the ROC curves to interpolate the PCPs' specificity at the same level of sensitivity as the device (i.e., 95.5%), the corresponding specificity of the device is 20.7% while the PCP specificity is 9.0%.

#### Spectral similarity score evaluation

For positive output (“Investigate Further”) the DermaSensor displays a Similarity Score of 1-10, with higher scores representing greater similarity to signals seen in malignant lesions. Standalone performance testing demonstrated that higher scores were associated with a higher positive predictive value (PPV), as calculated relative to the frequency of the score among all spectral scores displayed (total 1,288 scores; **Table 13**). The PPVs for each score, when assessed alone, are impacted by prevalence (frequency) of each score within the lesions of the pivotal study, which varied from 4.8% to 16.5%. Prevalence of skin cancer in lesions assessed for skin cancer may be different in real world use, and this would change the PPV. Therefore, this assessment was intended to show a correlation between higher scores and greater likelihood of malignancy, and not to predict PPV of the score in clinical use outside of the DERM-SUCCESS pivotal study.

**Table 13: Summary of positive predictive values by spectral similarity score output**

Spectral Score	Positive Total	PPV (Frequency)	Lower and upper bounds
1	13 / 202	6.4% (15.7%)	3.7 - 10.9%
2	7 / 213	3.3% (16.5%)	1.5 - 7%
3	15 / 177	8.5% (13.7%)	4.9 - 14.2%
4	14 / 132	10.6% (10.2%)	6.2 - 17.5%
5	12 / 103	11.7% (8%)	6.6 - 19.7%
6	25 / 104	24% (8.1%)	16.7 - 33.2%
7	33 / 117	28.2% (9.1%)	20.5 - 37.5%
8	26 / 101	25.7% (7.8%)	18.1 - 35.2%
9	31 / 77	40.3% (6%)	30.3 - 51.1%
10	38 / 62	61.3% (4.8%)	48.3 - 72.9%

#### Standalone Algorithm Performance Testing Conclusions

The DERM-SUCCESS pivotal study demonstrated an overall sensitivity for malignancy of 95.5%. The first co-primary endpoint was met, as the device sensitivity for all malignancies was superior to the sensitivity of PCPs, 83.0%, for the same population. Specificity for benign lesions was 20.7%, meeting the second co-primary endpoint of sensitivity + specificity > 1. The secondary endpoint was met, with overall sensitivity >90%. A correlation was observed between the magnitude of the similarity score and positive predictive value, thereby meeting the second co-primary endpoint.

#### Safety

No adverse events were observed during the DERM-SUCCESS pivotal study. False negative and false positive device outcomes were not considered adverse events as part of the study due to the non-interventional design. However, false negatives would be considered adverse events during real world use of the device.

#### Adjunctive Use Performance Testing

Three studies were completed to assess performance of the device as an aid to PCPs. The first study was conducted with all lesions suggestive of skin cancer, per the indications for

use, to demonstrate superior performance of the device-aided PCP compared to unaided PCPs. However, melanoma poses the highest risk of the three lesion types assessed due to a higher metastatic potential, higher morbidity, and higher mortality compared with SCC or BCC. Therefore, two additional reader studies were performed to ensure the device does not negatively impact the clinical decision of readers specifically in case sets enriched for pigmented lesions. Both studies were performed in a sequential read design to mimic the intended use of the device as a second read device.

DERM-SUCCESS initial reader study: "A Multi-Reader, Multi-Case (MRMC) Companion Study to the DERM-SUCCESS Clinical Study"

The DERM-SUCCESS initial reader study was intended to assess superiority of device-aided PCPs over unaided PCPs across all skin cancers (melanoma, SCC, and BCC). The study included 108 readers trained in internal medicine or family practice who evaluated 25 malignant (20% melanoma, 40% SCC, 40% BCC) and 25 benign lesions. Readers were shown images of the lesions either with or without the device output in random order. All lesions were viewed twice during the study, once without device output and once with device output. After each lesion, the readers were asked to provide a management decision of "Refer" or "Do not refer." The co-primary endpoints were aided sensitivity superiority to unaided sensitivity, and aided sensitivity + specificity >1. The study demonstrated that:

- device-aided PCPs had a higher sensitivity of 91.4% (95% CI: 85.7-97.1%) compared with unaided PCP sensitivity 82.0% (95% CI: 76.4-87.6%; p=0.0027); and
- sensitivity + specificity of PCPs aided by the device was greater than 1 (Odds Ratio 6.8, CI: 4.7 to 9.8, p<0.001).

The co-primary endpoints were therefore met.

Additional analyses showed that:

- device-aided PCPs had a lower specificity of 32.4% (95% CI: 20.7-44.1%) compared with unaided PCP specificity 44.2% (95% CI: 36.0-52.4%); and
- when assessed for overall performance of aided vs. unaided reader performance, the area under the receiver operating characteristic curve (AUROC) increased 5.4% from 0.708 to 0.762.

The data in this study, in which DermaSensor output was provided concurrently with the image of the lesion, demonstrated that PCPs performed better when viewing images with the device output than when viewing the image alone. However, the increased sensitivity was associated with reduced specificity. The specificity of the device-aided PCP was higher in this study than the specificity of the device alone, as quantified in the standalone performance testing, indicating that adjunctive use by PCPs may, to some extent, mitigate the risk associated with the low specificity of the device alone. Nonetheless, use of the device is still expected to result in higher referral of both true positives and false positives. The demonstrated increase in AUROC, which reflects both sensitivity and specificity, indicates that overall accuracy was greater in aided reads than in unaided reads.

DERM-SUCCESS pigmented lesion reader study: “A Multi-Reader, Multi-Case (MRMC) Study for Evaluating the Impact of the DermaSensor Device on Primary Care Physicians' Assessment of Pigmented Lesions”

The DERM-SUCCESS pigmented lesion reader study was intended to demonstrate non-inferiority of PCPs aided by the device compared to unaided PCPs for pigmented lesions suggestive of skin cancer. Seventy-seven PCP readers were shown an image of a lesion and entered an initial management decision of “Refer” or “Do not refer.” The device output was then revealed, and the readers entered a second management decision. The lesions included 69 malignant lesions (36% melanoma, 26% SCC, 36% BCC) and 67 benign lesions (59 pigmented, 10 non-pigmented). The primary endpoint was aided AUROC non-inferiority for all skin cancers. When the device result was shown the aided PCP AUROC was found to be greater than the unaided PCP AUROC by 1.5%, with statistically significant non-inferiority ( $p < 0.001$ ). The primary endpoint was therefore met.

Secondary endpoints were aided reader sensitivity non-inferiority for all skin cancers, aided reader sensitivity non-inferiority for melanoma, and specificity non-inferiority margin of 20%. Analysis of secondary endpoints showed:

- aided PCP sensitivity for all skin cancers increased from 80.5% (95% CI: 75.1-85.9%) to 86.3% (95% CI: 65.3-85.6%) with statistically significant superiority ( $p < 0.001$ );
- aided PCP sensitivity for melanoma increased from 68.8% (95% CI: 59.6-78.1%) to 75.4% (95% CI: 65.3-85.6%) with statistically significant non-inferiority ( $p < 0.001$ ); and
- aided PCP specificity for all skin cancers decreased from 45.7% (95% CI: 39.2-52.2%) to 35.1% (95% CI: 27.4-42.9%) with statistically significant non-inferiority ( $p = 0.003$ ).

This second-read design study demonstrated that PCPs improved their initial management decision when provided with the device output and were able to revise the management decision. Use of the device as a second read improved the sensitivity of the study PCPs for skin cancers overall and did not decrease sensitivity for lesions with the highest metastatic potential (i.e., melanoma). However, the increased sensitivity was associated with reduced specificity, which may result in higher referral of both true positives and false positives. AUROC, which reflects both sensitivity and specificity, indicated that the device led to improvement in overall accuracy based on the point-estimate; the statistically significant non-inferiority of AUROC supports that in a study enriched with pigmented lesions, PCP accuracy does not decrease for lesions with high metastatic potential, mitigating the risk of false negatives in this high-risk patient population.

The results are summarized in **Table 14**.

**Table 14: PCP Management AUROC and Sensitivity Without and With Device**

Lesion Assessments	Estimate (95% Confidence Interval)		P-value
	Without Device	With Device	

AUROC	n=20,944	0.702 (0.653-0.751)	0.717 (0.661-0.773)	<0.001
Sensitivity	n=10,626	80.5% (75.1-85.9%)	86.3% (81.1-91.5%)	<0.001*
Sensitivity for Melanoma	n=3,850	68.8% (59.6-78.1%)	75.4% (65.3-85.6%)	0.029*
Specificity	n=10,318	45.7% (39.2-52.2%)	35.1% (27.4-42.9%)	0.003

\*p-value for superiority, otherwise p-value is for non-inferiority

DERM-ASSESS III reader study: "A Multi-Reader, Multi-Case (MRMC) Study for Evaluating the Impact of the DermaSensor Device on Primary Care Physicians' Assessment of Lesions Suggestive of Melanoma"

The DERM-ASSESS III reader study for lesions suggestive of melanoma was intended to demonstrate non-inferiority of PCPs aided by the device compared to unaided PCPs, for lesions suggestive of melanoma. The readers, 118 PCPs, were shown an image of a lesion and entered an initial management decision of "Refer" or "Do not refer." The device output was then revealed, and the readers entered a second management decision. The lesions included 50 malignant lesions (68% melanoma, 16% SCC, 16% BCC) and 50 benign lesions. The primary endpoint was aided AUROC non-inferiority for all skin cancers. When the device result was shown the aided read AUROC was found to be greater than the unaided AUROC by 4.1%, with statistically significant non-inferiority ( $p < 0.001$ ) and superiority ( $p = 0.036$ ). The primary endpoint was therefore met. Secondary endpoints aided reader sensitivity non-inferiority for all skin cancers, aided sensitivity non-inferiority for melanoma, and specificity non-inferiority margin of 20%, were also met. The results are summarized in **Table 15**.

This study also demonstrated that use of the device as a second read source of information improved the sensitivity of the PCPs in the study. The increase in sensitivity was associated with lower specificity. This is anticipated to result in higher referral of both true positives and false positives. AUROC, which reflects both sensitivity and specificity, was improved when the device output was provided, indicating that the device led to improvement in overall accuracy based on the point-estimate; the statistically significant non-inferiority of AUROC supports that in a study enriched with pigmented lesions, PCP accuracy does not decrease for lesions with high metastatic potential, mitigating the risk of false negatives in this high-risk patient population.

**Table 15: PCP Management AUROC and Sensitivity Without and With Device**

	Lesion Assessments	Estimate (95% Confidence Interval)		P-value
		Without Device	With Device	
AUROC	n=11,800	0.630 (0.582-0.678)	0.671 (0.611-0.732)	0.036*
Sensitivity	n=5,900	73.7% (67.7-79.6%)	81.8% (76.0-87.6%)	<0.001*

Sensitivity for Melanoma	n=3,850	70.2% (62.9-77.6%)	79.1% (72.4-85.7%)	<0.001*
Specificity	n=5,900	44.2% (38.1-50.2%)	38.6% (30.9-46.2%)	<0.001

\*p-value for superiority, otherwise p-value is for non-inferiority

### Clinical Performance Results Analysis by Skin Cancer Type

Melanoma, SCC, and BCC are all skin cancers, but they are three distinct diseases, with different natural histories, risk of metastasis, urgency of diagnosis and treatment, and mortality rates. Device standalone performance was assessed for each lesion type and is summarized below (**Table 16**). Device standalone performance and the performance of the device in aiding PCP performance was further assessed by patient age and Fitzpatrick skin phototypes; this is also summarized below by each lesion type.

**Table 16 (a, b): Standalone device sensitivity and specificity by skin cancer subtype (based on binary output correlation with final histological diagnosis)**

#### **a. All participants**

Final Diagnosis	n	True Positives or True Negatives <sup>1</sup>	Sensitivity or Specificity (%) <sup>2</sup>
<b>High risk lesions</b>	224	214	95.5% (91.7%-97.6%)
Melanoma and severely atypical nevi	48	42	87.5% (76.4%-93.8%)
<i>Melanoma</i>	29	24	82.8% (64.2%-94.2%)
<i>Highly Atypical Nevus</i>	19	18	94.7% (91.9%-96.6%)
Squamous Cell Carcinoma (SCC)	86	84	97.7% (91.1%-99.4%)
Basal Cell Carcinoma (BCC)	90	88	97.8% (91.3%-99.5%)
<b>Low risk lesions</b>	1355	281	20.7% (18.5%-23.1%)

#### **b. Participants age 40+ (Intended use population)**

Final Diagnosis	n	True Positives or True Negatives <sup>1</sup>	Sensitivity or Specificity (%) <sup>2</sup>
<b>High risk lesions</b>	217	209	96.3% (92.9-98.4%)
Melanoma and severely atypical nevi	41	37	90.2% (76.9-97.3%)
Squamous Cell Carcinoma (SCC)	86	84	97.7% (91.1%-99.4%)
Basal Cell Carcinoma (BCC)	90	88	97.8% (91.3%-99.5%)
<b>Low risk lesions</b>	1155	234	20.3% (18.0-22.7%)

<sup>1</sup>True positives for high risk lesions and true negatives for low risk lesions

<sup>2</sup>95% CI calculated accounting for within-subject correlations using Wilson method.

Melanoma (including severely/highly atypical nevi):

Standalone device sensitivity for melanoma was highest in Fitzpatrick skin phototypes I and II (100% and 91.7%, respectively for the full population; 100% and 90.9%, for age 40+). The study included Fitzpatrick skin phototypes VI, V, and VI with a relatively small number of lesions per phototype (5, 6, and 1), consistent with the lower incidence of melanoma in Fitzpatrick skin phototypes IV-VI. Standalone device sensitivity was high in participants aged 40-59 (92.9%) and 60-79 (91.7%) but relatively lower in participants under age 40 (71.4%) and participants aged over 80 (66.7%).

When used as a second read device for adjunctive use in clinical decision support, the DermaSensor device improved performance of PCPs for participants ages 40 and older and with Fitzpatrick skin phototypes I-VI. **Table 17** and **Table 18** below summarize device standalone sensitivity and specificity and degree of effect on readers by subgroup by Fitzpatrick skin phototype and age group. We note that melanoma patients with Fitzpatrick skin phototypes IV-VI, were less represented in the studies consistent with their lower incidence in the US population. To assess safety in this patient subgroup, the performance of the device was compared to the PCP performance for the 12 melanomas enrolled in this subgroup in the DERM-SUCCESS standalone performance test study. The unaided PCP melanoma sensitivity was 66.7% (8/12), compared to the device sensitivity of 83.3% (10/12) for the same lesions. Thus, the data are suggestive of a margin of safety between the device sensitivity and PCP sensitivity for melanoma patients in this subgroup.

**Table 17 (a,b): Summary of Device Melanoma Performance and PCP Melanoma Performance Based on Fitzpatrick Skin Phototypes**

**a. All participants**

Fitzpatrick Phototype	DERM-SUCCESS		DERM-SUCCESS			
	Standalone device performance		Pigmented lesion reader study – sensitivity			
	Sensitivity	Specificity	No. Lesions	PCP Unaided	PCP Aided	% Change for Aided
<b>Overall</b>	87.5% (42/48)	23.7% (216/913)	25	68.8%	75.4%	6.6%
I	100.0% (10/10)	19.3% (16/83)	4	74.0%	81.2%	7.1%
II	91.7% (11/12)	24.4% (50/205)	9	80.7%	86.1%	5.5%
III	78.6% (11/14)	21.4% (65/304)	7	59.6%	75.1%	15.6%
IV	80.0% (4/5)	26.2% (43/164)	2	72.7%	57.8%	-14.9%
V	83.3% (5/6)	26.1% (37/142)	3	45.5%	48.1%	2.6%
VI	100.0% (1/1)	33.3% (5/15)	0	N/A	N/A	N/A

**b. Participants age 40+ (Intended use population)**

Fitzpatrick Phototype	DERM-SUCCESS		DERM-SUCCESS			
	Standalone device performance		Pigmented lesion reader study – sensitivity			
	Sensitivity	Specificity	No. Lesions	PCP Unaided	PCP Aided	% Change for Aided
<b>Overall</b>	90.2% (37/41)	23.0% (171/742)	21	69.9%	78.4%	8.5%
I	100% (9/9)	20.9% (14/67)	3	72.7%	82.7%	10.0%

II	90.9% (10/11)	25.3% (45/178)	8	79.4%	86.4%	7.0%
III	81.8% (9/11)	20.0% (48/240)	6	59.1%	76.4%	17.3%
IV	80.0% (4/5)	26.2% (43/164)	2	72.7%	57.8%	-14.9%
V	100% (4/4)	26.7% (31/116)	2	57.1%	66.2%	9.1%
VI	100.0% (1/1)	33.3% (5/15)	0	N/A	N/A	N/A

**Table 18: Summary of Device Melanoma Performance and PCP Melanoma Performance Based on Age Group**

Age Group	DERM-SUCCESS		DERM-SUCCESS			
	Standalone device performance		Pigmented Lesion reader study – sensitivity			
	Device Sensitivity	Device Specificity	No. Lesions	PCP Unaided	PCP Aided	% Change for Aided
Age 20-40	71.4% (5/7)	25.3% (46/182)	4	63.3%	60.1%	-3.20%
Age 41-60	92.9% (13/14)	27.4% (57/208)	8	75.8%	79.2%	3.40%
Age 61-80	91.7% (22/24)	28.3% (30/106)	13	66.2%	77.8%	11.60%
Age 81+	66.7% (2/3)	0.0% (0/4)	0	NA	NA	NA
<b>Overall</b>	87.5% (42/48)	23.7% (216/913)	25	68.8%	75.4%	6.6%
<b>Ages 40+</b>	90.2% (37/41)	23.0% (171/742)	21	69.9%	78.4%	8.5%

The DERM-SUCCESS pigmented lesion reader study was analyzed to determine whether the improvement in sensitivity in the context of reduced specificity had a net positive change in referral decision. The analysis demonstrated that the ratio of readers’ management decisions that were correctly changed from unaided “Do not refer” to aided “Refer” (resulting in correction of a potential missed diagnosis), compared to management decisions that were incorrectly changed from unaided “Refer” to aided “Do not refer” (resulting in a missed diagnosis) was 248:49, indicating that the device was more likely to correct user management decision than to harm management decision by a factor of 5, indicating an overall benefit as an aid in second reading of pigmented lesions suspicious for melanoma (Table 19).

**Table 19: Summary of converted reader decisions.**

Total melanoma reads 1925	Unaided reads	Aided reads converted
	False negative 600	After device true positive 248 (43% conversion rate)
	True positive 1325	After device false negative 49 (3.7% conversion rate)

Number of reads = number of melanomas x number of readers

**Squamous Cell Carcinoma (SCC):**

Overall standalone device sensitivity for SCC detection was 97.7%, with specificity of 17.3% for non-melanocytic lesions. There was insufficient support (number of lesions studied) for participants aged below 40 to confirm that benefit outweighed risk. The table

below summarizes performance by clinically-relevant subgroups from the DERM-SUCCESS standalone performance study and DERM-SUCCESS pigmented lesion reader study for SCC lesions.

When used as a second read device for adjunctive use in clinical decision support, the DermaSensor device improved performance of PCPs for participants aged 40 and over. **Table 20** and **Table 21** below summarize device standalone sensitivity and specificity by Fitzpatrick skin phototype and age group.

**Table 20: Summary of Device SCC Performance for All Patients Based on Fitzpatrick Skin Phototypes**

Fitzpatrick Phototype	DERM-SUCCESS Standalone device performance	
	Sensitivity	Specificity
I	90.0% (9/10)	16.2% (11/68)
II	100% (28/28)	19.5% (46/236)
III	100% (19/19)	10.8% (29/269)
IV	100% (18/18)	20.4% (31/152)
V	90.0% (9/10)	23.3% (27/116)
VI	100% (1/1)	28.6% (4/14)

**Table 21: Summary of Device SCC Performance and PCP SCC Performance for All Patients Based on Age Group**

Age (years)	DERM-SUCCESS Standalone device performance		DERM-SUCCESS Pigmented Lesion reader study – sensitivity		
	Sensitivity	Specificity	Unaided Reader	Aided Reader	% Change for Aided
20-40	100% (1/1)	22.0% (9/41)	84.40%	93.50%	9.10%
41-60	100% (10/10)	21.1% (57/270)	88.30%	95.70%	7.40%
61-80	96.7% (58/60)	15.4% (77/500)	76.80%	81.70%	4.90%
81+	100% (15/15)	11.4% (5/44)	94.50%	99.70%	5.20%

<sup>1</sup>A total of 77 PCP readers performed evaluations of these cases.

#### Basal Cell Carcinoma (BCC):

Overall standalone device sensitivity for BCC detection was 97.8% with specificity of 17.3% for non-melanocytic lesions. **Table 22** and **Table 23** below summarize performance by clinically-relevant subgroups from the DERM-SUCCESS clinical study and DERM-SUCCESS pigmented reader study for BCC lesions.

When used as a second read device for adjunctive use in clinical decision support, the DermaSensor device improved performance of PCPs for patients aged 40 and older. The

tabulated results below summarize device standalone sensitivity and specificity and degree of effect on readers by subgroup by Fitzpatrick skin phototype and age group.

**Table 22: Summary of Device BCC Performance for All Patients Based on Fitzpatrick Skin Phototypes**

Fitzpatrick Phototype	DERM-SUCCESS Standalone device performance	
	Sensitivity	Specificity
I	100% (7/7)	16.2% (11/68)
II	97.4% (37/38)	19.5% (46/236)
III	100% (35/35)	10.8% (29/269)
IV	100% (2/2)	20.4% (31/152)
V	85.7% (6/7)	23.3% (27/116)
VI	100% (1/1)	28.6% (4/14)

**Table 23: Summary of Device BCC Performance and PCP BCC Performance for All Patients Based on Age Group**

Age (years)	DERM-SUCCESS Standalone device performance		DERM-SUCCESS Pigmented Lesion reader study – sensitivity		
	Sensitivity	Specificity	Unaided Reader	Aided Reader	% Change for Aided
20-40	100% (1/1)	22% (9/41)	NA	NA	NA
41-60	94.4% (34/36)	21.1% (57/270)	87.20%	92.60%	5.40%
61-80	100% (39/39)	15.4% (77/500)	92.60%	97.80%	5.20%
81+	100% (14/14)	11.4% (5/44)	92.90%	97.40%	4.50%

#### Summary of Clinical Information

Overall, the data presented demonstrate increased sensitivity of melanoma, SCC, and BCC detection when PCPs use the DermaSensor device in the intended use population as a second-read device for adjunctive use, and that adjunctive use, to some extent, mitigates the risk of increased referrals for benign lesions associated with a relatively low specificity of the device.

#### POSTMARKET SURVEILLANCE

In order to satisfy special control (1) below, DermaSensor must collect and report post-market surveillance data acquired under anticipated conditions of use to demonstrate that the device performs as intended when used to analyze data from the intended patient population. Specifically, the sponsor must conduct post-market clinical validation performance testing of the DermaSensor device in patients from demographic groups representative of the U.S. population, to include populations for whom uncertainty remains in the data provided due to lower representation in the mITT population (i.e., Fitzpatrick skin phototypes with lower prevalence of skin cancer).

FDA expects that the post-market clinical validation performance testing will include a statistically justified study sample size to confirm that performance of the device in post-market use is not inferior to the performance observed in the pre-market study for the studied subgroup(s) or the overall population. The study should enroll a representative range of subjects with overrepresentation of patients who have Fitzpatrick skin phototypes with lower prevalence of skin cancer (e.g., Fitzpatrick skin phototypes IV, V, and VI). The study should record age, sex, socioeconomic status, race, and risk factors for skin cancer development and progression, and assess their impact on the device's effectiveness as measured by sensitivity, specificity, and/or improvement of the user's management decision.

### LABELING

Device labeling includes the product labels on the device and instructions for use labeling. The instructions for use include a description of the device components and outputs, device technical parameters, instructions for use of the device, electromagnetic compatibility information, shelf life, and device reprocessing and disposal instructions. The instructions for use also contain relevant information about the patients in the algorithm development and training datasets, and findings from the standalone and adjunctive use clinical studies with the performance characteristics of the device when used as intended. The document also states the shelf life for any sterile components as well as the necessary measures to properly dispose of any single use items and clean the reusable components of the device. The sale, distribution, and use of DermaSensor device are restricted to prescription use in accordance with 21 CFR 801.109.

Labeling also includes the following:

- A description of performance measures, including sensitivity and specificity, and statistical confidence intervals, as well as performance of the device for all clinically relevant subgroups within the intended use population
- A statement that the device is not intended for use as a standalone diagnostic.
- A statement that performance of the device has not been specifically evaluated in patients with increased risk for skin cancer, e.g., inherited or drug-induced photosensitivity; genetic predisposition to melanoma or basal cell carcinoma (BCC); immune compromise; or other medical conditions that increase the risk of skin cancer or its metastasis.
- A statement that the device should not be used for lesions that are suggestive of malignancies other than melanoma, BCC and/or SCC.
- A statement that the device is intended for use on primary lesions only and has not been tested on lesions that are previously biopsied, recurrent, or metastatic; on scars, tattoos, sunburned skin, or within a hairy area; or which are located on palms, soles, mucosal surfaces, genitals, ears, within 1 cm of the eye, or under nails.
- A statement that, consistent with the lower prevalence of skin cancer in Fitzpatrick skin phototypes IV-VI, less data is available for sensitivity of the DermaSensor device for melanoma in these patients. The decision to refer patients with suspicious pigmented lesions in this group should be primarily based on clinical concern.

Labeling will be updated in accordance with data collected via post-market surveillance to provide updated clinical performance data and effectiveness data of the device.

**RISKS TO HEALTH**

The table below identifies the risks to health that may be associated with use of a software-aided adjunctive diagnostic device for use by physicians on lesions suspicious for skin cancer and the measures necessary to mitigate these risks.

<b>Identified Risks to Health</b>	<b>Mitigation Measures</b>
False negative results, leading to failure to treat cancer and cancer progression, or false positive results, leading to unnecessary referrals and/or medical procedures	Clinical performance testing Postmarket surveillance Non-clinical performance testing Labeling
False results or failure to generate a result due to use error or improper device use	Precision testing Human factors testing Labeling
False results or failure to generate a result due to device failure or malfunction	Non-clinical performance testing Precision testing Software verification, validation, and hazard analysis Labeling
Electrical, thermal, mechanical, or light exposure-related injury	Electrical, mechanical, and thermal safety testing Software verification, validation, and hazard analysis Labeling
Interference with other devices	Electromagnetic compatibility testing
Adverse tissue reaction	Biocompatibility evaluation
Infection and cross contamination	Cleaning and disinfection validation Labeling

**SPECIAL CONTROLS**

In combination with the general controls of the FD&C Act, the software-aided adjunctive diagnostic device for use by physicians on lesions suspicious for skin cancer are subject to the following special controls:

- (1) Data obtained from premarket clinical performance validation testing and post-market surveillance acquired under anticipated conditions of use must demonstrate that the device performs as intended in the intended patient population, unless FDA determines based on the totality of the information provided for premarket review that data from post-market surveillance is not required.
  - (i) Data must demonstrate superior accuracy of device-aided users' diagnostic characterization of the indicated lesions compared to the accuracy of un-aided users.
  - (ii) Clinical testing must evaluate patients across a range of skin phototypes, risk factors, and anatomic areas that represents the intended use population.
  - (iii) Standalone device performance testing must demonstrate the accuracy of the device output relative to ground truth, including the following:

- (A) Testing must demonstrate at least 90% sensitivity of the device output for lesions with high metastatic potential, or an alternative clinical consideration must be provided to justify lower sensitivity. Clinical justification must be provided for the reported specificity.
  - (B) Lesions must be selected by representative users and include a justified quantity and range of mimic lesions per diagnosis.
  - (C) Justification must be provided for the determination of ground truth.
  - (D) Testing must include a representative range of individuals with diverse risk factors (including age, body site, and skin phototype, and other clinical factors), and analysis of standalone performance must include subgroup analysis by relevant risk factors.
- (2) Non-clinical performance testing must demonstrate that the device performs as intended under anticipated conditions of use, including compatibility testing of the device software with specific signal or image acquisition hardware. Testing must include a description of compatible hardware and processes, pre-specified compatibility testing protocols and dataset(s).
- (3) Performance testing must demonstrate device precision, including repeatability and reproducibility of device performance, across operators and challenging use conditions.
- (4) Performance testing must demonstrate the electrical safety, mechanical safety, thermal safety, and electromagnetic compatibility of any electrical components of the device.
- (5) Performance testing must validate reprocessing instructions for reusable components of the device.
- (6) The patient-contacting components of the device must be demonstrated to be biocompatible.
- (7) Software verification, validation, and hazard analysis must be performed.
- (8) Human factors assessment must demonstrate that the device can be safely and correctly used by intended users.
- (9) Labeling must include:
- (i) A summary of standalone and clinical performance testing conducted with the device. The summary must describe performance measures, including sensitivity and specificity, and statistical confidence intervals, as well as performance of the device for all clinically relevant subgroups within the intended use population;
  - (ii) A description of the patient population that was used in development or training of the device algorithm;
  - (iii) Device limitations or subpopulations for which the device may not perform as expected or for whom the device has not been validated;

- (iv) Information for interpretation of the device outputs detailing the risks associated with misinterpretation of the device outputs;
- (v) Warnings to avoid unsafe exposure to any energy-emitting components of the device, e.g., excluding lesions close to the eye;
- (vi) A statement that the device is not intended for use as a standalone diagnostic; and
- (vii) Instructions for device maintenance and validated methods and instructions for reprocessing of any reusable components.

### **BENEFIT-RISK DETERMINATION**

#### Risks and Other Factors

The risks of the device are based on nonclinical laboratory studies as well as data collected in the clinical studies described above.

No device or procedure related adverse events (AEs), serious adverse events (SAEs), or unanticipated adverse device effects (UADEs) were observed in the DERM-SUCCESS pivotal study.

The main risks of the DermaSensor device are false negative output and false positive output. False negative output may cause a provider to monitor rather than refer a lesion, which could result in delayed diagnosis of melanoma, SCC, or BCC. False positive output may result in referral of a benign lesion for dermatology evaluation. This may result in patient anxiety and potentially increase the probability of biopsy. These risks are mitigated by the high standalone device sensitivity in the intended use population (>90%); by the PCP's use of the device output as a second read and in addition to all available clinical information and examination; and by the final management decision by the dermatologist to whom the patient is referred.

#### Benefits

The probable benefits of the device are based on nonclinical laboratory studies as well as data collected in the clinical studies described above.

The main benefit of the DermaSensor device is the demonstrated ability to improve the management decision of PCPs in the intended use population. In participants aged 40 and older with melanoma (including severely/highly atypical nevi), SCC, or BCC lesion(s), this benefit was indicated by high standalone device sensitivity (>90%), in a pivotal study that included a large number of participants and lesions (>1000 each) across 22 clinical sites and Fitzpatrick phototypes I-VI. Three reader studies demonstrated that use of the DermaSensor resulted in superiority of the aided PCP sensitivity over the unaided sensitivity of the same PCP and that the device led to a greater increase in true positive aided referral decisions than false negative aided referral decisions.

The clinical performance testing, including standalone device sensitivity and demonstrated aid in reader studies, supports the benefit of the DermaSensor device as a second-read adjunctive aid to

primary care physicians in correct referral decisions for individuals aged 40 and over with melanoma (including highly atypical nevi), SCC, and BCC.

#### Uncertainty

The performance of the device has not been specifically evaluated in patients with increased risk for skin cancer, e.g., inherited or drug-induced photosensitivity; genetic predisposition to melanoma or BCC; immune compromise; or other medical conditions that increase the risk of skin cancer or its metastasis. Device-aided user performance may differ compared with the premarket studies provided, due to, for example, differences in reliance on the device in the real world compared to in the retrospective studies provided. Sensitivity of the DermaSensor device in participants under the age of 40 was less than 90%. Algorithm training and clinical testing had limited representation of patients with Fitzpatrick skin phototypes IV, V, and VI, reflective of the relative incidence of skin cancer in these patient subpopulations. This uncertainty is mitigated by the post-market surveillance special control and labeling which includes a precaution stating that limited data is available regarding the safety and effectiveness of the DermaSensor device for patients with Fitzpatrick skin phototypes IV, V, and VI and those at increased risk for skin cancer, e.g., inherited or drug-induced photosensitivity; genetic predisposition to melanoma or BCC; immune compromise; or other medical conditions that increase the risk of skin cancer or its metastasis.

#### Patient Perspectives

This submission did not include specific information on patient perspectives for this device.

#### Benefit/Risk Conclusion

In conclusion, given the available information above, the data support that for the following indication statement:

The DermaSensor device is indicated for use to evaluate skin lesions suggestive of melanoma, basal cell carcinoma, and/or squamous cell carcinoma in patients aged 40 and above to assist in the decision regarding referral of the patient to a dermatologist. The DermaSensor device should be used in conjunction with the totality of clinically relevant information from the clinical assessment, including visual analysis of the lesion, by physicians who are not dermatologists. The device should be used on lesions already assessed as suspicious for skin cancer and not as a screening tool. The device should not be used as the sole diagnostic criterion nor to confirm clinical diagnosis of skin cancer.

The probable benefits outweigh the probable risks for the DermaSensor device. The device provides benefits, and the risks can be mitigated by the use of general controls and the identified special controls.

#### CONCLUSION

The De Novo request for the DermaSensor is granted, and the device is classified as follows:

Product Code: QZS

Device Type: Software-aided adjunctive diagnostic device for use by physicians on lesions suspicious for skin cancer.

Regulation Number: 21 CFR 878.1830

Class: II

## Ferris et al. (2025)

Original Research

## DERM-SUCCESS FDA Pivotal Study: A Multi-Reader Multi-Case Evaluation of Primary Care Physicians' Skin Cancer Detection Using AI-Enabled Elastic Scattering Spectroscopy

Journal of Primary Care & Community Health  
 Volume 16: 1–9  
 © The Author(s) 2025  
 Article reuse guidelines:  
[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)  
 DOI: 10.1177/21501319251342106  
[journals.sagepub.com/home/jpc](https://journals.sagepub.com/home/jpc)  


Laura K. Ferris<sup>1</sup>, Erik Jaklitsch<sup>2</sup>, Elizabeth V. Seiverling<sup>2,3</sup>, Thomas Agresta<sup>4</sup>, Peggy Cyr<sup>5</sup>, Laurie Caines<sup>6</sup>, Na Wang<sup>7</sup>, and Daniel M. Siegel<sup>8</sup>

### Abstract

**Background:** Elastic Scattering Spectroscopy (ESS), an optical tissue sampling technique, distinguishes between benign and malignant tissue in vivo without the need to perform a surgical biopsy. A handheld device that employs ESS enabled with an artificial intelligence algorithm was developed as an objective tool to aid primary care physicians (PCPs) in their management of lesions suspicious for skin cancer. The aim of this study was to assess and compare the diagnostic and management performance of PCPs with and without the use of the ESS device in detecting skin cancer.

**Methods:** In this clinical utility study, 108 PCPs evaluated 100 skin lesion cases (50 aided with the device output and 50 unaided by the device). For each case, PCPs provided a diagnosis, management decision, and level of confidence in that decision initially without, and then subsequently with, device output. Sensitivity, specificity, AUC, and confidence in their assessment prior to, and then with, device output were compared.

**Results:** With visual assessment assisted by device output, diagnostic sensitivity increased significantly from 71.1% to 81.7% ( $P=.0085$ ) and referral sensitivity increased significantly from 82.0% to 91.4% ( $P=.0027$ ) compared to visual assessment only. Device-aided diagnostic specificity decreased from 60.9% to 54.7% ( $P=.1896$ ), and referral specificity decreased from 44.2% to 32.4% ( $P=.0256$ ). Overall management performance (ie, AUC) also increased from 0.708 to 0.762, and increased from 0.567 to 0.682 for lesions which physicians reported low confidence in their unassisted management decision. Physicians reporting high confidence in their management assessment increased from 36.8% to 53.4%.

**Conclusion:** Use of the ESS device output by PCPs significantly improved their diagnostic and management sensitivities as well as their overall management performance. The findings suggest the ESS device can improve PCP skin cancer detection and confidence in their skin lesion evaluation and management.

### Keywords

melanoma, squamous cell carcinoma, basal cell carcinoma, artificial intelligence, machine learning, skin cancer, technology, diagnostic tools, elastic scattering spectroscopy

Received: January 2, 2025; revised: April 8, 2025; accepted: April 15, 2025

### Introduction

Skin cancer is the most common malignancy in the U.S.<sup>1</sup> The most common skin cancers are basal cell carcinomas (BCC) and squamous cell carcinomas (SCC), which arise from keratinocytes in the epidermis and are collectively referred to as keratinocyte carcinomas (KCs); approximately

5.4 million are diagnosed in the U.S. each year.<sup>2,3</sup> Melanoma, the next most common type of skin cancer, has lower incidence but higher risk of mortality than KCs; in 2023, an estimated 186 680 new cases are expected to be diagnosed in the U.S.<sup>1,2</sup> In the U.S., approximately 7990 people die from melanoma and 8000 die from KCs each year.<sup>1,4,5</sup> Most skin



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

cancers are treated surgically, and treatment of more advanced melanoma and KCs often result in significant surgical morbidity to sensitive areas.

Access to a dermatologist for diagnosis and management of cutaneous malignancies is challenging. Patients experience long waits for dermatology appointments, with 40% of Americans facing scarce dermatological services, and only 15% of high-risk Americans have ever been screened for skin cancer.<sup>6-8</sup> This underscores the pivotal role of primary care physicians (PCPs) in promptly detecting cutaneous malignancies. However, differentiating malignant from benign skin lesions can be challenging for physicians, particularly non-dermatologists. The “ABCDE” (asymmetry, border irregularity, color changes, diameter, evolution) criteria for the evaluation of a melanocytic growth are helpful for detecting superficial spreading melanomas; however, these criteria result in variable performance and do not aid in detection of KCs.<sup>9</sup>

To address this need, various non-invasive approaches have been developed to aid clinicians, including imaging algorithms, electrical impedance spectroscopy, radiation-based photometric analysis, optical coherence tomography, and dermoscopy. For many of these technologies, researchers have developed algorithms using artificial intelligence that are applied to the lesion data to provide an automated cancer risk assessment.<sup>10-15</sup> While promising, many face challenges to their widespread implementation, including narrow clinical scope (ie, only 1 type of skin cancer assessed), high cost, slow speed, and/or complicated output requiring extensive training.<sup>14,15</sup> Algorithm-based skin lesion imaging devices have shown potential to improve physician cancer detection, but further testing and improvements are needed, as commercially-available melanoma smartphone applications have a mean sensitivity of only 28%. Until recently, no skin cancer detection devices nor AI-based algorithms have been approved by the US Food and Drug Administration (FDA) for use by PCPs nor patients.<sup>15,16</sup>

Elastic Scattering Spectroscopy (ESS), an optical tissue sampling technique, captures sub-cellular information to distinguish between benign and malignant tissue in vivo

without the need to remove tissue.<sup>17</sup> A handheld device employing ESS and an AI-powered algorithm was developed as an objective tool for PCPs to aid in their management of lesions suspicious for skin cancer. The device has shown promise in several clinical studies to date.<sup>18-22</sup> The algorithm was developed to be highly sensitive in discerning malignant skin lesions (melanoma, BCC, and SCC) from benign skin lesions of concern to PCPs, the sensitivity across studies ranged from 90% to 96% and specificity ranged from 21% to 61%, depending on lesion selection criteria. This is the first technology to be cleared by the FDA to provide PCPs with an immediate, objective report of the skin cancer risk of a lesion; it is also the first technology to ever be cleared or approved for any user for all 3 common types of skin cancer.

In this study, PCPs actively practicing in the U.S. were presented with lesion cases (lesion images with patient and lesion clinical information) enrolled during the DERM-SUCCESS clinical validation study.<sup>23</sup> The purpose of the study was to evaluate the impact of the device results on PCPs decision-making regarding skin lesion evaluations. The DERM-SUCCESS study analyzed device performance on 1579 lesions from 1005 enrolled patients across 22 primary care study sites, and this companion study evaluated the device’s ability to impact and improve upon PCP skin cancer diagnostic and management performance compared to current practice without use of the device.

## Methods

### Study Overview

This multi-reader multi-case (MRMC) study was designed to assess and compare the sensitivity and specificity of PCPs in their clinical assessment of lesions suggestive of skin cancer with and without knowledge of the device result. This study was reviewed and approved by an ethical approval granting committee and an IRB-approved written informed consent was used and obtained for all study participants prior to participation in the DERM-SUCCESS clinical study.

<sup>1</sup>Department of Dermatology, University of North Carolina, Chapel Hill, USA

<sup>2</sup>University of Pittsburgh School of Medicine, PA, USA

<sup>3</sup>Tufts University School of Medicine, Boston, MA, USA

<sup>4</sup>University of Connecticut School of Medicine, Farmington, USA

<sup>5</sup>Department of Family Medicine, Maine Medical Center, Portland, USA

<sup>6</sup>UConn Health, Farmington, CT, USA

<sup>7</sup>Biostatistics and Epidemiology Data Analytics Center, Boston University, MA, USA

<sup>8</sup>SUNY Downstate Health Sciences University, Brooklyn, USA

#### Corresponding Author:

Laura K. Ferris, Department of Dermatology, University of North Carolina, 410 Market Street Suite 400, Chapel Hill, NC 27516, USA.

Email: ferrisk@email.unc.edu

### Device Information

The ESS device (DermaSensor™ device, manufactured by DermaSensor Inc., Miami, FL, USA) has been described previously as cleared by the FDA for use by non-dermatology expert physicians in January 2024.<sup>21,23</sup> Briefly, this device emits light pulses dispersed over isolated areas of the lesion, with backscattered optical reflectance representing its architecture (eg, nuclear and chromatic characteristics). The ESS device was trained and validated with >20,000 spectral scans from >4500 lesions including histologically-confirmed melanoma, BCC, SCC, and benign lesions as well as unbiopsied benign lesions diagnosed by board-certified dermatologists from 6 clinical studies conducted in primary care and dermatology settings.<sup>23-27</sup>

In accordance with FDA requirements, the AI algorithm was locked prospectively; thus, algorithm development occurred with neither knowledge of the device data nor device results. The device classifies lesions as “Investigate Further” or “Monitor” based on similarity of their scanned properties to malignant or benign lesions. For “Investigate Further” lesions, a 1 to 10 score reflects the degree of spectral similarity to malignant lesions from prior studies. The likelihood of malignancy (positive predictive value) in the DERM-SUCCESS study ranged from 6% for a spectral score of “1” to 61% for a score of “10.”

### Study Design

This study was conducted in 2 phases. In each phase, PCP study participants reviewed high-resolution digital images and clinical information for 50 skin lesion cases, which included 25 malignant cases, independently and in random order. These prospectively-collected lesion cases were a randomly selected subset of lesions from the DERM-SUCCESS clinical study and were all biopsied as suspicious for melanoma, BCC, and/or SCC by primary care physicians. All were subsequently diagnosed by 2 to 5 dermatopathologists depending on pathology type and discordance; thus, all lesions are representative of the pathologies of common primary care lesions suspicious for skin cancer. This selection process also ensured the device performance was reflective of that in the clinical study. Additionally, a standardized photograph protocol was implemented in the clinical study where captured images were reviewed by a panel of physicians (including both dermatologists and PCPs) for image quality and consistency prior to inclusion.

Physicians were recruited for participation via a survey platform (Sermo, Inc., United States) and were compensated for their time per the platform’s standards. In the first phase, each physician evaluated the 50 skin lesion cases based on their assessment alone, without the device output. In the second phase, following a break of at least 2 h, physicians evaluated the same lesions in a different randomized

order, based on the same clinical information provided for each case but with the addition of the device output. For both phases and every lesion case, physicians completed a questionnaire about their diagnosis (malignant or benign), their recommended management decision (further evaluate or not), and their confidence level on their management decision (1-10 scale, where 1=no confidence and 10=high confidence). Physicians were blinded to the proportion of benign versus malignant lesions included in the study. A power calculation predicted 93% power with 100 readers for 25 malignant lesions, assuming a 10% increase in diagnostic sensitivity (a measurement of PCPs’ ability to detect malignant lesions).

One hundred and eighteen (118) board-certified PCPs completed at least 1 lesion case in the study, and 108 PCPs completed all 100 lesion case assessments and were included in effectiveness analysis. Physician referral sensitivity, a measurement of PCPs’ ability to appropriately refer malignant cases, was a primary endpoint to assess improvements in skin cancer management decisions. To ensure PCP performance with the device was better than random chance, a second co-primary endpoint was evaluated (whether sensitivity + specificity > 1).

### Statistical Analyses

Co-primary endpoint 1 was referral sensitivity of the physicians using biopsy results as the reference standard, which was defined as the number of true positive cancers that were referred for further evaluation and tested at a 1-sided .025 level of significance. The second co-primary endpoint was referral sensitivity and specificity of the physicians with knowledge of device result using biopsy results as the reference standard, tested at a 1-sided .025 level of significance.

Additionally, physicians’ diagnostic sensitivity was analyzed with and without device result knowledge. The hypotheses compared the true diagnostic sensitivity for physicians with and without device result knowledge. The Area Under the Curve (AUC) of the Receiver Operating Curve (ROC) was calculated based on physicians’ referral decisions and confidence levels, using histopathology diagnosis as the reference. PCP referral and diagnostic specificity with and without device result knowledge were evaluated, including a confidence interval of the difference between the 2 specificities, calculated using the Obuchowski and Rockette ANOVA approach. All analyses were conducted using SAS Version 9.4 and using R.

### Results

Of the 108 U.S. board certified physicians included, 48.1% (n=52) were Internal Medicine physicians and 51.9% (n=56) were Family Medicine physicians (Supplemental Table S1). There were 65.7% (n=71) male and 34.3%

**Table 1.** Management and Diagnostic Sensitivity With and Without the Device.<sup>c</sup>

Outcome	Lesion assessments <sup>a</sup>	Estimate (%) and 95% confidence interval <sup>b</sup>		P-value <sup>b</sup>
		Without device	With device	
Management sensitivity (co-primary endpoint)	n = 5400	82.0% (76.4%-87.6%)	91.4% (85.7%-97.1%)	.0027
Diagnostic sensitivity	n = 5400	71.1% (63.4%-78.8%)	81.7% (72.4%-90.9%)	.0085

<sup>a</sup>For sensitivity analyses, only positive (malignant) lesion cases are included; for specificity analyses, only negative (benign) lesion cases are included.

<sup>b</sup>MRMC analysis of variance (ANOVA) method of Obuchowski and Rockette.<sup>28</sup>

<sup>c</sup>The authors recognize that there is debate about the definition of melanoma and high-risk melanocytic lesions. The FDA required that severely dysplastic nevi should be considered positive for referral to dermatology. Thus, the DERM-SUCCESS clinical study and this study grouped highly dysplastic nevi, atypical junctional melanocytic lesions, melanoma in situ, and invasive melanoma as all "high risk" lesions.

(n=37) female physicians. Years in practice varied with most physicians reporting 21+ years in practice (31.5%), followed by 1 to 5 years (19.4%), 11 to 15 years (17.6%), 6 to 10 years (17.6%), and 16 to 20 years (13.9%). The majority of physicians reported practicing in various types of non-academic settings, with 13.9% (n=15) practicing in an academic center. Among this group, 40.7% (n=44) reported referring patients with skin lesions to a dermatologist "most of the time" and 49.1% (n=53) reported referring "sometimes"; no physicians in the study reported doing so "never" or "always." The majority, 56.5% (n=61), rated themselves as having "intermediate" competency in skin lesion assessment.

Patient cases (n=50) had a median age of 59 years, were 100% of White race, and the majority had Fitzpatrick skin type of 2 (36.0%, n=18) or 3 (32.0%, n=16; Supplemental Table S2). Half of lesions included were characterized as high risk (50%, n=25), with the same clinical study proportions for basal cell carcinoma (40%, n=10), squamous cell carcinoma (36.0%, n=9), melanoma (16%, n=4), and severely dysplastic nevi (8.0%, n=2); half were low risk (50%, n=25), which included benign melanocytic nevi (36.0%, n=9), seborrheic keratosis (36.0%, n=9), and benign other (28%, n=7; see footnote; Supplemental Table S3). Other characteristics of the patients and lesions are shown in Supplemental Tables S2 and S3, including risk factors such as new/changing lesions, ultraviolet light exposure, and family history. Patient characteristics in the lesion cases sample set were very similar to those of the DERM-SUCCESS clinical study.<sup>23</sup> In both studies, there were no clinically meaningful differences between the selected lesion cases and the clinical study cases for any lesion characteristics, including location, pigmentation, texture, or size.

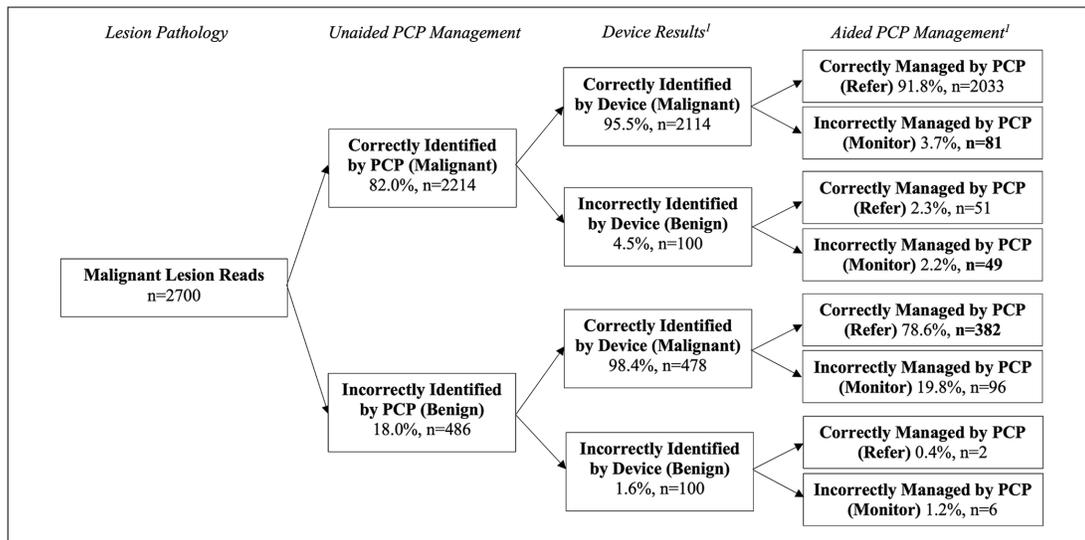
In accordance with the study protocol, the prevalence of malignant lesions in this study was 50%, which was higher than the 14.2% prevalence (224 malignant lesions of 1579 total lesions) in the clinical study (Supplemental Table S3). The device sensitivity for the 25 malignant lesions in this study was 96.0% (compared to 95.5% in the clinical study),

and the device specificity for the 25 benign lesions was 20.0% (compared to 20.7% in the clinical study). Within those 2 groups of 25 cases each, the lesion pathology breakdown was also similar to the clinical study pathology.

The 25 malignant and 25 benign lesions were assessed by each of the 108 PCPs, yielding 2700 lesion assessments of each type without the device result and 2700 of each type with the device result, totaling 5400 assessments for malignant lesions and 5400 for benign lesions. The management sensitivity of physicians with the device result was significantly higher than that without the device result (91.4% vs 82.0%,  $P=.0027$ ); thus, the physicians' false negative rate decreased by more than half, from 18.0% to 8.6% (Table 1). The diagnostic sensitivity for PCPs with the device result of 81.7% (95% CI: 72.4%-90.9%) was also significantly higher compared to their sensitivity without the device result of 71.1% (95% CI: 63.4%-78.8%;  $P=.0085$ ; Table 1).

For benign lesions, the associated management specificity was 44.2% (95% CI: 36.0%-52.4%) without the device result and 32.4% (95% CI: 20.7%-44.1%) with the device result ( $P=.0256$ ). Statistical significance testing for specificity was not planned but using the significance testing that was used for sensitivity (ie, a 1-sided .025 level of significance), the specificity decrease was not significant. PCP diagnostic specificity for benign lesions was not significantly different without the device output 60.9% (95% CI: 52.5%-69.3%) as compared to specificity with the device output 54.7% (95% CI: 42.3%-67.1%;  $P=.1896$ ). PCPs incorrectly referred an additional 319 of 2700 (11.8%) benign lesion cases when aided by the device output, but correctly referred an additional 254 of 2700 (9.4%) malignant lesion cases when aided by the device output. For benign lesions that PCPs had referred unaided, 21.2% of these were not referred when aided by the device output.

The referral sensitivity of the PCPs with knowledge of the device result yielded a sensitivity of 91.4% (95% CI: 85.7%-97.1%) and a specificity of 32.4% (95% CI: 20.7%-44.1%). The PCPs with the device result were significantly better at detecting malignant lesions than expected by chance, with an odds ratio of 6.8 (95% CI: 4.7-9.8;

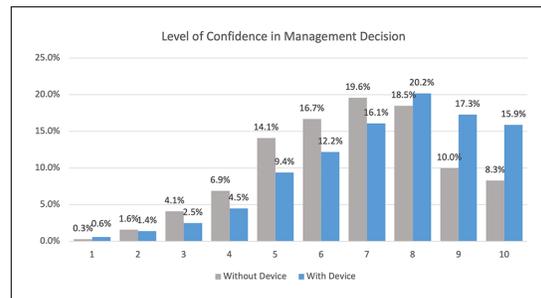


**Figure 1.** Flow chart of PCP decision-making for lesion management and impact of device results. Bolded values used to calculate ratio of increased detection. <sup>1</sup>Indicates percent of initial PCP management decision total.

$P < .0001$ ). The 382 device true positive results changed PCP reads from incorrectly managed to correctly managed malignancies. By comparison, 81 device true positive results and 49 device false negative results (130 in total) changed PCP reads from correctly managed to incorrectly managed malignancies (Figure 1). Thus, device use impact on skin cancer management has a net  $2.9\times$  ratio of increased detection (382 positively impacted vs 130 negatively impacted), supporting the device’s benefit/risk profile.

High confidence in the physician management decision (defined as a rating of 8-10) was reported for 43.2% of malignant lesions without the device result and 63.3% of malignant lesions with the device result. For all lesion assessments combined ( $n=5400$ ), the proportion of assessments made with a high level of confidence in the physician management decision was 36.8% without the device result and 53.4% with the device result (Figure 2).

PCP management sensitivity for malignant lesions with device output information was higher for low-confidence referral decisions (81.9% [95% CI: 70.4%-93.3%]) when compared to corresponding decisions without the device output (76.8% [95% CI: 67.2%-86.4%]). For high-confidence referral decisions, PCP management sensitivity for malignant lesions with device output was also higher (95.6% [95% CI: 91.7%-99.4%]) as compared to corresponding decisions without device output (88.7% [95% CI: 83.4%-94.0%]) as shown in Table 2. In addition, PCPs completed 58% more high confidence (ie, 8-10) referrals of



**Figure 2.** Level of Confidence in management decision ( $n=10800$ ).

malignant lesions with device output ( $n=1635$ ) than they completed without device output ( $n=1034$ ).

The PCPs’ observed AUC performance was 0.708 without device output and improved to 0.762 with device use (Figure 3a, Supplemental Figure S1A). The device stand-alone AUROC was 0.780 in the clinical study.<sup>18</sup> As mentioned previously, the PCP management performance with use of the device output was 91.4% (95% CI: 85.7%-97.1%) for sensitivity (Figure 3b) and 32.4% (95% CI: 20.7%-44.1%) for specificity. Fixing the PCP sensitivity of 91.4% on the ROC curve without use of the device, the interpolated PCP specificity level was 23.2%. Therefore, for PCPs unaided by the device result, interpolated specificity would be approximately 23.2% instead of 32.4% to achieve the

**Table 2.** Management Sensitivity and Specificity According to Confidence Level<sup>a</sup> (n = 10800).  
A: Low/High Confidence.<sup>b</sup>

Level of confidence	#Malignant lesions	#Physician positive	Sensitivity (95% CI) <sup>a</sup>	#Benign lesions	#Physician negative	Specificity (95% CI) <sup>a</sup>
Without device			n = 2700			n = 2700
Low	630	468	74.3% (67.6%-81.0%)	829	331	39.9% (32.7%-47.1%)
High	2070	1746	84.3% (78.4%-90.3%)	1871	863	46.1% (36.5%-55.7%)
With device			n = 2700			n = 2700
Low	361	296	82.0% (75.7%-88.3%)	630	142	22.5% (15.6%-29.5%)
High	2339	2172	92.9% (87.6%-98.2%)	2070	733	35.4% (21.1%-49.7%)

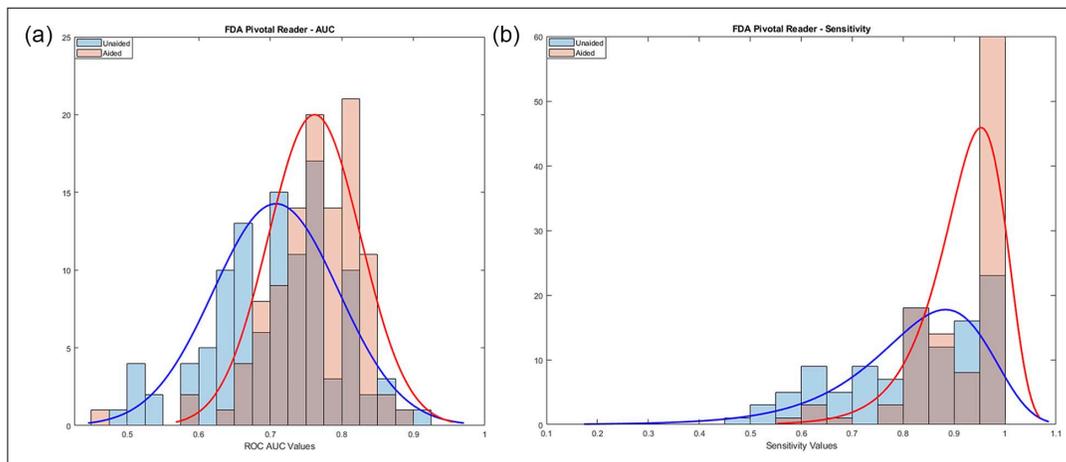
B: Low/Medium/High Confidence.<sup>c</sup>

Level of confidence	#Malignant lesions	#Physician positive	Sensitivity (95% CI) <sup>a</sup>	#Benign lesions	#Physician negative	Specificity (95% CI) <sup>a</sup>
Without device			n = 2700			n = 2700
Low	319	245	76.8% (67.2%-86.4%)	379	136	35.9% (26.4%-45.3%)
Medium	1215	935	77.0% (71.0%-82.9%)	1502	662	44.1% (36.2%-51.9%)
High	1166	1034	88.7% (83.4%-94.0%)	819	396	48.4% (36.6%-60.1%)
With device			n = 2700			n = 2700
Low	182	149	81.9% (70.4%-93.3%)	301	67	22.3% (13.0%-31.6%)
Medium	807	684	84.8% (77.3%-92.3%)	1229	379	30.8% (21.2%-40.5%)
High	1711	1635	95.6% (91.7%-99.4%)	1170	429	36.7% (21.1%-52.2%)

<sup>a</sup>Obuchowski and Rockette method for 95% confidence interval.

<sup>b</sup>Two-level classified as: low (1-5) and high (6-10).

<sup>c</sup>Three-level classified as: low (1-4), medium (5-7), and high (8-10).



**Figure 3.** (a) Physician management ROC curves unaided and aided by device and (b) physician sensitivity values unaided and aided by device. The Area Under the Curve (AUC) of the Receiver Operating Curve (ROC) was calculated based on physicians' referral decisions, using histopathology diagnosis as the reference.

same level of skin cancer detection that was achieved with use of the device, a difference of 9.2%. Additional ROC and AUC analyses were performed for lesions in which PCPs

had low confidence in their clinical assessment when unaided by the device output. For such low confidence lesions, the PCPs' observed AUC was 0.567 without the

device output and was 0.682 with the device output (Supplemental Figure S1B).

The DERM-SUCCESS clinical validation study and this companion clinical utility study were performed prior to FDA clearance, which resulted in device labeling for patients 40 years and older. Subanalysis of our study for patients above age 40 years (22 malignant, 22 benign) found the management sensitivity of physicians with the device result (95.9%, 95% CI: 94.5%-97.2%) was significantly higher than it was without the device result (83.6%, 95% CI: 81.0%-86.1%). Management specificity also decreased with device output (32.6%, 95% CI: 29.5%-35.9%) as compared to without (41.5%, 95% CI: 37.8%-45.1%). Among this subgroup, the PCPs' observed AUC performance improved to 0.800 (95% CI: 0.787%-0.812%) with device use from 0.704 (95% CI: 0.686%-0.722%) without device output. PCP diagnostic sensitivity in patients over 40 also increased when aided by the device (88.3%, 95% CI: 86.1%-90.4%) as compared to without it (73.3%, 95% CI: 70.2%-76.4%), and diagnostic specificity minimally decreased with device output (53.1%, 95% CI: 49.5%-56.7%) as compared to without it (58.1%, 95% CI: 54.8%-61.4%).

At the end of the study, physician participants were asked a series of 7 questions relating to device applicability in clinical practice. Only 1 of the 108 PCPs (<1%) reported no benefits from the device. Among the other PCPs, the most commonly-reported benefits of the ESS device were: "providing an immediate, objective result to inform your management of suspicious lesions" (82%), "detecting more skin cancer" (82%), and "providing you with greater confidence in your clinical assessments and management decisions" (81%). Also, 98% of PCPs agreed (81% strongly agreed and 17% agreed) that skin cancer is a disease deserving of better surveillance in the primary care setting. The PCPs' most commonly-reported estimate of their own sensitivity for correctly managing malignant lesions was 71% to 80%.

## Discussion

We show that the availability of the AI-enabled, non-invasive handheld ESS device output significantly increased the sensitivity of PCPs from 82.0%, when using only standard of care visual inspection and patient history, to 91.4% when they were also provided with the device result ( $P=.0027$ ). Thus, availability of the device output halved the physicians' false negative rate from 18.0% to 8.6%. When comparing to the physicians' clinical performance in the DERM-SUCCESS clinical study,<sup>23</sup> the unaided, standard of care sensitivity of the physicians for both studies was similar, that is, 83.0% for the clinical study and 82.0% for this utility study. When comparing to the device's performance in the clinical study, the PCPs' aided sensitivity of 91.4% was lower but similar to that of the device algorithm's standalone sensitivity of 96.0%.

The increases in sensitivity demonstrate the potential for a direct benefit to PCPs' ability to correctly manage high risk, potentially cancerous lesions when aided by the device result. Increasing skin cancer detection by 9.4% (ie, 91.4% compared to 82.0%) is highly impactful, given that 5.6 million cutaneous malignancies are diagnosed in the U.S. each year. Notably, while consensus among dermatologists and dermatopathologists has yet to be reached,<sup>29</sup> any significant increase in detection must be weighed against concerns for the overdiagnosis of malignancy. For this study, 254 additional malignant lesions were correctly managed, while 319 additional benign lesions were incorrectly referred with use of the device output.

To our knowledge, this is the first prospective clinical utility study completed for FDA clearance for any kind of skin cancer detection device for PCPs. Our results demonstrate a significant improvement in both diagnostic and management sensitivity of physicians with device result availability, while maintaining acceptable levels of specificity for all patients and for those patients 40 years and above, suggesting that the integration of this AI-powered device into their practice is beneficial to PCP skin cancer evaluations. The significant 10.6% improvement in diagnostic sensitivity (15.0% for patients 40 years and above), coupled with an observed drop of 6.2% in specificity (5.0% for patients 40 years and above) that was not significant, further supports that use of the device may enable physicians to evaluate skin lesions more confidently and correctly. We also found increasing physician performance as physician confidence increases, both with and without device use. Furthermore, device use increased physician confidence in management decisions and, notably, physician sensitivity was consistently higher with device use than without device use across different levels of physician confidence.

The observed physician AUC of 0.708 without device use was lower than the AUC of 0.762 with device use, and both were lower than standalone device AUC of 0.780. Physician management performance with use of the device was 91.4% sensitivity and 32.4% specificity. This difference in AUC with versus without the use of the device result was even more pronounced for lesions in which PCP management confidence was low (0.567 vs 0.682, respectively). Further, physicians reported increased confidence in their assessments with the availability of the device result. Results demonstrated consistent benefits of device use in terms of improved sensitivity across a wide range of physician, patient, and lesion characteristics.

Generally, PCPs are aware that they have room for improvement in their detection of skin cancer, as their most commonly self-reported estimate of sensitivity for correctly managing malignant lesions was 71% to 80%. Collectively, responses by the physician study participants show their strong agreement on the device benefits in clinical practice

and the need for improved skin cancer surveillance in primary care, which demonstrate (1) a strong unmet need for augmentation of PCPs' clinical assessments and management decisions in this context, and (2) that use of the ESS device could help meet that need for nearly all physicians that participated in this study.

This study has a few limitations of note. Due to ethical considerations since the device was experimental during the clinical studies, the DERM-SUCCESS Part 1 clinical validation study exclusively evaluated device performance compared to PCPs assessment and did not provide a device result to impact PCP lesion management. Thus, this companion reader study evaluated the impact of the same device results on a randomized subset of the same PCP-selected lesions, but it was not the same PCPs performing in-person lesion evaluations. Therefore, although patient and lesion clinical information and lesion tactile information (eg, smooth or rough, elevated or flat) was provided, the individual clinician's own tactile evaluation of the lesions was not possible; while this does not mimic true clinical practice, this is consistent with care provided via telemedicine. Also, recall bias could impact re-evaluation of lesions, and all lesions were from Caucasian patients with lighter skin types; relative performance with and without aid of the device output in darker skin types was not assessed. Inherent to the study design and endpoints, the ratio of benign to malignant lesions among the test pool was not reflective of primary clinical practice where the large majority of lesions that are encountered are benign, though this does not impact the reported performance results (ie, sensitivity, specificity, and AUC). However, the incorporation of KCs among the test pool is a strength that more accurately represents the distribution of malignant lesions encountered in primary care.

In conclusion, this study demonstrated that, when added to clinical information and lesion tactile information, PCP use of the AI-enabled ESS device was able to significantly improve skin cancer diagnosis and management sensitivity, as well as physician overall diagnostic accuracy (ie, AUC). The findings demonstrate that utilization of the ESS device output may improve physician performance in the management of suspicious lesions and referral of skin cancer to ensure timely diagnosis. Future investigations may provide further evidence of device performance and/or impact on physician diagnosis and management on varied atypical lesion and skin types in clinical practice settings with variable pre-test probability (patient-identified lesions, dermatology vs primary care).

#### Acknowledgments

The study team would like to take this time to commend Joseph Massaro, PhD (deceased) for writing the initial study statistical plan and for his contributions to the study design. We would like to thank Kim Ann Dukes, PhD, Margaret Shea, Kanisha Mittal,

Lindsey Furton, Kaitlin Hartlage, and Laurent Billot, PhD for their contributions to the statistical analyses.

#### ORCID iDs

Erik Jaklitsch  <https://orcid.org/0000-0002-9311-1660>

Elizabeth V. Seiverling  <https://orcid.org/0009-0007-9115-6148>

Daniel M. Siegel  <https://orcid.org/0000-0002-3918-9483>

#### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was sponsored by DermaSensor, Inc.

#### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

#### Supplemental Material

Supplemental material for this article is available online.

#### References

1. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *CA Cancer J Clin.* 2023;73(1):17-48. doi:10.3322/CAAC.21763
2. American Cancer Society. Basal and Squamous Cell Skin Cancer Statistics | American Cancer Society. Published October 31, 2023. <https://www.cancer.org/cancer/types/basal-and-squamous-cell-skin-cancer/about/key-statistics.html>
3. Rogers HW, Weinstock MA, Feldman SR, Coldiron BM. Incidence estimate of nonmelanoma skin cancer (Keratinocyte Carcinomas) in the US population, 2012. *JAMA Dermatol.* 2015;151(10):1081-1086. doi:10.1001/JAMADERMATOL.2015.1187
4. Karia PS, Han J, Schmults CD. Cutaneous squamous cell carcinoma: estimated incidence of disease, nodal metastasis, and deaths from disease in the United States, 2012. *J Am Acad Dermatol.* 2013;68(6):957-966. doi:10.1016/J.JAAD.2012.11.037
5. Nehal KS, Bichakjian CK. Update on Keratinocyte Carcinomas. *N Engl J Med.* 2018;379(4):363-374. doi:10.1056/NEJMRA1708701/SUPPL\_FILE/NEJMRA1708701\_DISCLOSURES.PDF
6. Tsang MW, Resneck JS. 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):54-58. doi:10.1016/J.JAAD.2006.04.001
7. Hester T, Thomas R, Cederna J, et al. Increasing access to specialized dermatology care: a retrospective study investigating clinical operation and impact of a university-affiliated free clinic. *Dermatol Ther.* 2021;11(1):105-115. doi:10.1007/S13555-020-00462-Z
8. Coups EJ, Geller AC, Weinstock MA, Heckman CJ, Manne SL. Prevalence and correlates of skin cancer screening among middle-aged and older white adults in the United

- States. *Am J Med.* 2010;123(5):439-445. doi:10.1016/J.AMJMED.2009.10.014
9. Goldsmith SM. A unifying approach to the clinical diagnosis of melanoma including “D” for “Dark” in the ABCDE criteria. *Dermatol Pract Concept.* 2014;4(4):75. doi:10.5826/DPC.0404A16
  10. Malvey J, Hauschild A, Curiel-Lewandrowski C, et al. Clinical performance of the Nevisense system in cutaneous melanoma detection: an international, multicentre, prospective and blinded clinical trial on efficacy and safety. *Br J Dermatol.* 2014;171(5):1099-1107. doi:10.1111/BJD.13121
  11. Monheit G, Cognetta AB, Ferris L, et al. The performance of MelaFind: a prospective multicenter study. *Arch Dermatol.* 2011;147(2):188-194. doi:10.1001/ARCHDERMATOL.2010.302
  12. Emery JD, Hunter J, Hall PN, Watson AJ, Moncrieff M, Walter FM. Accuracy of SIAscopy for pigmented skin lesions encountered in primary care: development and validation of a new diagnostic algorithm. *BMC Dermatol.* 2010;10:9. doi:10.1186/1471-5945-10-9
  13. Ulrich M, Von Braunmuehl T, Kurzen H, et al. The sensitivity and specificity of optical coherence tomography for the assisted diagnosis of nonpigmented basal cell carcinoma: an observational study. *Br J Dermatol.* 2015;173(2):428-435. doi:10.1111/BJD.13853
  14. Higgins HW, Lee KC, Leffell DJ. Point of care cutaneous imaging technology in melanoma screening and mole mapping. *F1000Prime Rep.* 2014;6:34. doi:10.12703/P6-34
  15. Pyun SH, Min W, Goo B, et al. Real-time, in vivo skin cancer triage by laser-induced plasma spectroscopy combined with a deep learning-based diagnostic algorithm. *J Am Acad Dermatol.* 2023;89(1):99-105. doi:10.1016/J.JAAD.2022.06.1166
  16. Sun MD, Kentley J, Mehta P, Dusza S, Halpern AC, Rotemberg V. Accuracy of commercially available smartphone applications for the detection of melanoma. *Br J Dermatol.* 2022;186(4):744. doi:10.1111/BJD.20903
  17. Bigio IJ, Boyer J, Mourant JR. Theoretical and experimental investigations of elastic scattering spectroscopy as a potential diagnostic for tissue pathologies. Paper presented at: Advances in Optical Imaging and Photon Migration (1994), paper OPTTM265; 21–March 23, 1994; Florida; OPTTM.265. doi:10.1364/AOIPM.1994.OPTTM.265
  18. Manolagos D, Patrick G, Geisse JK, et al. Use of an elastic-scattering spectroscopy and artificial intelligence device in the assessment of lesions suggestive of skin cancer: a comparative effectiveness study. *JAAD Int.* 2023;14:52-58. doi:10.1016/J.JDIN.2023.08.019
  19. Hartman RI, Trepanowski N, Chang MS, et al. Multicenter prospective blinded melanoma detection study with a handheld elastic scattering spectroscopy device. *JAAD Int.* 2024;15:24-31. doi:10.1016/j.jdin.2023.10.011
  20. Jaklitsch E, Thames T, de Campos Silva T, Coll P, Oliviero M, Ferris LK. Clinical utility of an AI-powered, handheld elastic scattering spectroscopy device on the diagnosis and management of skin cancer by primary care physicians. *J Prim Care Community Health.* 2023;14:21501319231205979. doi:10.1177/21501319231205979
  21. Rodriguez-Diaz E, Manolagos D, Christman H, et al. Optical spectroscopy as a method for skin cancer risk assessment. *Photochem Photobiol.* 2019;95(6):1441-1445. doi:10.1111/PHP.13140
  22. Tepedino M, Baltazar D, Hanna K, Bridges A, Billot L, Zeitouni NC. Elastic scattering spectroscopy on patient-selected lesions concerning for skin cancer. *J Am Board Fam Med.* 2024;37(3):427-435. doi:10.3122/JABFM.2023.230256R2
  23. Merry S, Croghan I, McCormick B, Chatha K, Leffell D. Clinical performance of novel elastic scattering spectroscopy (ESS) in detection of skin cancer: a blinded, prospective, multicenter clinical trial [initial results]. *Cutis.* 2022;110(6 Suppl):31.
  24. Tepedino M, Baltazar D, Hucks C, Chatha K, Zeitouni NC. Use of elastic scattering spectroscopy on patient selected lesions that are concerning for skin cancer. *Cutis.* 2022;110(6 Suppl):35-36.
  25. Salmon P, Bonning M. Use of elastic-scattering spectroscopy and machine learning when assessing skin lesions suggestive of skin cancer. *J Dermatol Phys Assist.* 2021;15(4):64-65.
  26. Hartman R, Tepedino K, Fung MA, et al. Validation of a handheld elastic-scattering spectroscopy device on lesions suggestive of melanoma. *J Dermatol Physician Assist.* 2022;16(4):51.
  27. Manolagos D, Rabinovitz H, Geisse J, Bonning M, Rodriguez-Diaz E, Bigio I CA. Clinical validation of a handheld elastic scattering spectroscopy artificial intelligence device. Paper presented at: Podium Presentation, American Academy of Dermatology Innovation Academy; July 20-24, 2022; Vancouver, CA.
  28. Obuchowski NA, Rockette HE. Hypothesis testing of diagnostic accuracy for multiple readers and multiple tests an anova approach with dependent observations. *Commun Stat Simul Comput.* 1995;24(2):285-308. doi:10.1080/03610919508813243
  29. Esserman LJ, Thompson IM, Reid B, et al. Addressing overdiagnosis and overtreatment in cancer: a prescription for change. *Lancet Oncol.* 2014;15(6):e234-e242. doi:10.1016/S1470-2045(13)70598-9

## Hartman et al. (2024)

## ORIGINAL ARTICLE

## Multicenter prospective blinded melanoma detection study with a handheld elastic scattering spectroscopy device



Rebecca I. Hartman, MD, MPH,<sup>a,b,c</sup> Nicole Trepanowski, MD,<sup>a,d</sup> Michael S. Chang, MD,<sup>a,b</sup> Kelly Tepedino, MD,<sup>e</sup> Christopher Gianacas, MBIostat,<sup>f,g</sup> Jennifer M. McNiff, MD,<sup>h</sup> Maxwell Fung, MD,<sup>i</sup> Naiara Fraga Braghiroli, MD, PhD,<sup>j</sup> and Jane M. Grant-Kels, MD<sup>k,l</sup>

**Background:** The elastic scattering spectroscopy (ESS) device (DermaSensor Inc., Miami, FL) is a noninvasive, painless, adjunctive tool for skin cancer detection.

**Objectives:** To investigate the performance of the ESS device in the detection of melanoma.

**Methods:** A prospective, investigator-blinded, multicenter study was conducted at 8 United States (US) and 2 Australian sites. All eligible skin lesions were clinically concerning for melanoma, examined with the ESS device, subsequently biopsied according to dermatologists' standard of care, and evaluated with histopathology. A total of 311 participants with 440 lesions were enrolled, including 44 melanomas (63.6% in situ and 36.4% invasive) and 44 severely dysplastic nevi.

**Results:** The observed sensitivity of the ESS device for melanoma detection was 95.5% (95% CI, 84.5% to 98.8%, 42 of 44 melanomas), and the observed specificity was 32.5% (95% CI, 27.2% to 38.3%). The positive and negative predictive values were 16.0% and 98.1%, respectively.

**Limitations:** The device was tested in a high-risk population with lesions selected for biopsy based on clinical and dermoscopic assessments of board-certified dermatologists. Most enrolled lesions were pigmented.

**Conclusion:** The ESS device's high sensitivity and NPV for the detection of melanoma suggest the device may be a useful adjunctive, point-of-care tool for melanoma detection. (JAAD Int 2024;15:24-31.)

**Key words:** AI; artificial intelligence; automated; biopsy; DermaSensor; DERM-ASSESS III; detection; elastic scattering device; elastic scattering spectroscopy; ESS; handheld; melanoma; non-invasive; NPV; pigmented lesion; PPV; sensitivity; skin cancer; specificity; spectroscopic; spectroscopy; technology.

From the Department of Dermatology, Brigham and Women's Hospital, Boston, Massachusetts<sup>a</sup>; Harvard Medical School, Boston, Massachusetts<sup>b</sup>; Department of Dermatology, VA Integrated Service Network (VISN-1), Jamaica Plain, Massachusetts<sup>c</sup>; Boston University School of Medicine, Boston, Massachusetts<sup>d</sup>; North Florida Dermatology, Lake City, Florida<sup>e</sup>; The George Institute for Global Health, UNSW Sydney, Sydney, Australia<sup>f</sup>; School of Population Health, UNSW Sydney, Sydney, Australia<sup>g</sup>; Departments of Dermatology and Pathology, Yale School of Medicine, New Haven, Connecticut<sup>h</sup>; University of California Davis School of Medicine, Sacramento, California<sup>i</sup>; Dermatology Department, Miami Cancer Institute, Miami, Florida<sup>j</sup>; Department of Dermatology, University of Connecticut School of Medicine, Farmington, Connecticut<sup>k</sup>; and Department of Dermatology, University of Florida College of Medicine, Gainesville, Florida.<sup>l</sup>

Drs Hartman and Trepanowski are cofirst authors.

Funding sources: This study was funded by DermaSensor, Inc. Dr Hartman is funded by the Melanoma Research Foundation and the VA Integrated Service Network 1 (VISN-1).

Patient consent: Obtained.

IRB approval status: All sites received IRB approval under a central IRB, WCG-Copernicus Group IRB (Protocol #CSP-20-0002) in the United States and the Bellberry Human Research Ethics Committee in Australia (Protocol #2020-08-804). Approved by the Brigham and Women's Hospital/Dana Farber Cancer Institute IRB (Protocol #20-365).

Accepted for publication October 26, 2023.

Correspondence to: Rebecca I. Hartman, MD, MPH, Department of Dermatology, Brigham and Women's Hospital, 221 Longwood Ave, Boston, MA 02115. E-mail: [rhartman@bwh.harvard.edu](mailto:rhartman@bwh.harvard.edu). 2666-3287

Published by Elsevier Inc. on behalf of the American Academy of Dermatology, Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.1016/j.jdin.2023.10.011>

## INTRODUCTION

Melanoma incidence has increased faster than nearly all other cancers.<sup>1</sup> Although representing less than 5% of cutaneous malignancies, melanoma accounts for most skin cancer deaths.<sup>1</sup> Novel therapies approved since 2011 have improved survival for late-stage melanomas but are costly and have significant side effects.<sup>2,3</sup>

Dermatologists utilize tools to evaluate lesions for melanoma including dermoscopy, confocal microscopy, longitudinal photography and digital dermoscopy, gene expression testing, and artificial intelligence (AI).<sup>4</sup> Use of these technologies varies by training and access. The number of skin lesions needed to biopsy (NNB) to diagnose 1 melanoma varies from 14.8 for clinicians, 7.5 for dermatologists, and 13.2 for US-based dermatologic practitioners (including dermatologists and advanced practice professionals).<sup>5</sup> Biopsied benign lesions increase morbidity and healthcare costs.

A handheld device (DermaSensor, DermaSensor Inc) using elastic scattering spectroscopy (ESS) and AI has been developed as an adjunctive tool for skin cancer detection.<sup>6,7</sup> ESS utilizes the spectral recording of photons scattered back from refractive-index gradients to gain information about cellular and sub-cellular structures.<sup>8</sup> Different tissue types and histopathologic changes, such as malignancy, exhibit specific optical signatures.<sup>8</sup> In the primary care setting, the ESS device has a sensitivity of 90.0% to 95.5% and a specificity of 20.7% to 60.7% for detecting melanoma and keratinocyte carcinomas.<sup>6,7</sup> Our objective was to investigate the real-world performance of the ESS device for melanoma detection using skin lesions identified as potential melanomas by board-certified dermatologists.

## METHODS

### Ethical conduct

The study was approved by a central institutional review board (IRB) and local IRBs or ethics committees and was conducted in accordance with the revised Declaration of Helsinki, Good Clinical Practice guidelines, national and data protection laws, and applicable regulatory requirements. The study sponsor developed the trial protocol, provided

ESS devices and funding to carry out the study, and engaged a contract research organization (CRO) to oversee study operations and a biostatistics firm to conduct data analyses. As the study was noninterventional, clinical trial registration was not required.

## CAPSULE SUMMARY

- Novel technologies for melanoma detection may improve diagnostic performance, increase access, reduce resource utilization, and decrease healthcare costs.
- In this study, the high sensitivity and negative predictive value of a noninvasive and easy-to-use elastic scattering spectroscopy device suggest that this point-of-care tool may be a useful adjunct for melanoma detection.

### Study design and data acquisition

The DERMaSensor Use in the ASSESSment of Skin Lesions Suggestive of Melanoma III (DERM-ASSESS III) was a blinded, prospective study performed at 8 sites in the US and 2 sites in Australia from December 2020 through October 2021. One site was an academic institution (Brigham and Women's Hospital/Dana Farber Cancer Institute); the remaining 9 sites

were high melanoma volume private practice dermatology clinics. All 12 enrolling dermatologist-investigators were board-certified. Potential participants were screened according to inclusion and exclusion criteria (Supplementary Table I, available via Mendeley at <https://data.mendeley.com/datasets/fcgghmx8f4/1>). All participants provided written informed consent, and participation did not impact treatment. Electronic case records recorded participant demographics and melanoma risk factors. All enrolled lesions were clinically suspicious for melanoma. Dermatologist-investigators performed dermoscopy assessments per their standard of care and were asked to clinically predict histopathologic diagnoses as benign, melanoma, or severely dysplastic nevus (SDN).

Dermatologist-investigators and the sponsor were blinded to device output. A CRO blinded case records, ESS device measurements, and histopathologic reports from the sponsor.

### Histopathologic evaluation as the reference standard

Biopsies were performed using dermatologist-investigators' standard of care and were assessed with dermatopathology consensus review via a primary and secondary consensus process. Each lesion was independently assessed by a local board-certified dermatopathologist and a subsequent dermatopathology over-read by a board-certified dermatopathologist was performed (primary consensus). For lesions diagnosed as benign during primary consensus, no further review was required. For lesions diagnosed as melanoma,

**Abbreviations used:**

AI:	artificial intelligence
AUROC:	area under the receiver operating characteristic curve
BCC:	basal cell carcinoma
CI:	confidence interval
CRO:	contract research organization
DERM-ASSESS III:	DERMaSensor Use in the ASSESSment of Skin Lesions Suggestive of Melanoma III
ESS:	elastic scattering spectroscopy
NNB:	number needed to biopsy
NPV:	negative predictive value
PPV:	positive predictive value
SCC:	squamous cell carcinoma
SD:	standard deviation
SDN:	severely dysplastic nevus

SDN, moderately DN, or other potentially high-risk or inconclusive diagnoses during primary consensus, further blinded review was conducted by 2 or 3 additional expert dermatopathologists (secondary consensus). The histopathologic reference standard refers to the diagnosis after final consensus was achieved. Only lesions with histopathologic reference standard diagnoses were included in the analysis population.

**ESS Device**

The ESS device is a noninvasive, point-and-click spectrometer system weighing 1.9 kg in total, with the wireless handheld piece weighing 0.3 kg (Supplementary Fig 1, available via Mendeley at <https://data.mendeley.com/datasets/fcgghmx8f4/1>).<sup>9</sup> The unit assesses skin lesions  $\geq 2.5$  mm in diameter in less than 30 seconds. The device emits light pulses of hundreds of distinct wavelengths and receives and analyzes spectral data from cellular and sub-cellular structures. The unit captures 5 recordings to generate a spectral reading which is analyzed by a locked algorithm in the device's software. The volume of tissue that the device evaluates with each spectral recording was estimated computationally with Monte Carlo simulations to approximately 0.7 mm (length)  $\times$  0.4 mm (width)  $\times$  0.5 mm (depth).<sup>10</sup> The device classifies lesions as "Investigate Further" or "Monitor," with the "Investigate Further" results accompanied by a spectral score of 1 to 10, with 10 being most similar to previously validated malignant lesions.<sup>7</sup> Lesions classified as "Monitor" do not receive a spectral score. Lesions classified as "Investigate Further" by the ESS device are interpreted as high risk, that is high risk of malignancy,

whereas "Monitor" results are interpreted as low risk, that is low risk of malignancy.<sup>8</sup>

**Outcome measures**

Primary endpoints were device performance for distinguishing histopathologically benign lesions from those histopathologically diagnosed as melanoma (excluding other malignant diagnoses, eg, keratinocyte carcinomas and SDN) and the combined outcome of melanoma and SDN.

Statistical analyses included concordance analyses (sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV)) using dermatopathology with clustered confidence intervals (CIs) based on Wilson score and subgroup analyses based on participant and lesion clinical characteristics.<sup>11</sup> Area under the receiver operating characteristic curve (AUROC) statistics for device and dermatologist-investigators were used for comparative effectiveness analyses. PPV analyses were conducted to assess device performance across various spectral score groupings, and Wald test *P*-values were presented. NNB for melanoma detection was calculated using the inverse of the PPV.<sup>12</sup> *P* < .05 and nonoverlapping CIs were considered statistically significant. Analyses were performed using SAS 9.4 (SAS Institute), R version 4.2 (R Core Team), and STATA 17 (StataCorp).

**RESULTS**

Of 332 recruited subjects, the final study population contained 311 subjects with 440 lesions suspicious for melanoma (Supplementary Fig 2, available via Mendeley at <https://data.mendeley.com/datasets/fcgghmx8f4/1>). The average participant was 62.0  $\pm$  15.4 years-old, and participants included 53.7% males and 46.3% females (Table I). Most participants identified as non-Hispanic (96.1%) and white (97.7%). Participants had an average of 1.4  $\pm$  0.8 lesions enrolled.

Lesions were located on the trunk (61.6%), head and neck (15.5%), upper extremity (13.0%), and lower extremity (10.0%) (Table II). The average lesion size was 5.8 mm  $\times$  4.7 mm. Lesions were predominantly flat (83.6%), smooth (91.1%), and pigmented (96.8%). Histopathologic diagnoses included benign (74.1%), SDN (10.0%), melanoma (10.0%), or malignant other (5.9%, basal cell carcinoma, squamous cell carcinoma, and atypical fibroxanthoma) (Table III). Overall study cancer prevalence was 15.9%, and NNB for melanoma detection by dermatologist-investigators was 10.

The dichotomous outcome of the ESS device was compared with the histopathologic reference standard. Sensitivity of the ESS device for detecting

**Table I.** Participant characteristics

Characteristics	Participants with eligible lesions <i>n</i> = 311, <i>n</i> (%)
Age - years, Mean (SD)	62.0 (15.4)
Sex	
Male	167 (53.7)
Female	144 (46.3)
Ethnicity	
Hispanic or Latino	8 (2.6)
Not Hispanic or Latino	299 (96.1)
Unknown	4 (1.3)
Race	
White	304 (97.7)
Other	3 (1.0)
Native Hawaiian or other Pacific Islander	2 (0.6)
White, Native Hawaiian or other Pacific Islander	1 (0.3)
Asian	1 (0.3)
Fitzpatrick skin type*	
I - Always burns, never tans	32 (10.3)
II - Always burns, tans minimally	165 (53.1)
III - Sometimes mild burn, tans uniformly	66 (21.2)
IV - Burns minimally, always tans well	20 (6.4)
V - Very rarely burns, tans very easily	20 (6.4)
VI - Never burns, never tans	8 (2.6)
Melanoma risk factors†	
Ultraviolet light exposure (natural or tanning bed)	143 (46.0)
Personal history of skin cancer	119 (38.3)
Fair skin, freckling, light hair	107 (34.4)
Family history of skin cancer	70 (22.5)
Many moles and/or dysplastic nevi	58 (18.6)
New or changing lesion(s)	41 (13.2)
Weakened immune system	8 (2.6)
No risk factors reported	57 (18.3)
Person who discovered the lesion(s) of concern	
Health care provider	224 (72.0)
Patient	80 (25.7)
Family member/partner	7 (2.3)

SD, Standard deviation.

\*Determined by clinical judgment or assessment.

†Percentages do not sum to 100% because participants may have had more than 1 melanoma risk factor.

melanoma was 95.5% (95% CI, 84.5% to 98.8%, 42 of 44 melanomas) (Table IV). Specificity of the ESS device was 32.5% (95% CI, 27.2% to 38.3%), NPV was 98.1% (95% CI, 91.8% to 99.6%), and PPV was 16.0% (95% CI, 11.6% to 21.7%). The 2 device false negative melanomas were a 10 mm × 6 mm flat, smooth, and pigmented lesion on the trunk that was histopathologically diagnosed as a “Level 2 superficial

**Table II.** Lesion characteristics

Characteristics	<i>n</i> = 440, <i>n</i> (%)
Anatomic location	
Trunk	271 (61.6)
Head and neck	68 (15.5)
Upper extremity	57 (13.0)
Lower extremity	44 (10.0)
Length (mm), mean (SD)	5.8 (2.7)
Width (mm), mean (SD)	4.7 (2.0)
Flat or elevated	
Flat	368 (83.6)
Elevated	72 (16.4)
Texture	
Smooth	401 (91.1)
Rough	39 (8.9)
Pigmentation	
Pigmented	426 (96.8)
Non-pigmented	14 (3.2)

*mm*, Millimeter; *SD*, standard deviation.

spreading melanoma with 0.3 mm thickness” (Supplementary Fig 3, available via Mendeley at <https://data.mendeley.com/datasets/fcggmx8f4/1>) and a 9 mm × 4.6 mm flat, smooth, and pigmented lesion on the lower extremity that was histopathologically diagnosed as a “Level 1 superficial spreading melanoma in situ” (Supplementary Fig 4, available via Mendeley at <https://data.mendeley.com/datasets/fcggmx8f4/1>).

Comparatively, device performance for detecting melanoma and SDN yielded a sensitivity of 90.9% (95% CI, 83.1% to 95.3%), specificity of 32.5% (95% CI, 27.2% to 38.3%), NPV of 93.0% (95% CI, 85.7% to 96.7%), and PPV of 26.7% (95% CI, 21.5% to 32.6%) (Supplementary Table II, available via Mendeley at <https://data.mendeley.com/datasets/fcggmx8f4/1>). The device yielded similar sensitivities with overlapping CIs for lesions that were clinically suspicious for melanoma but were histopathologically diagnosed as other malignant diagnoses (Supplementary Table III and IV, available via Mendeley at <https://data.mendeley.com/datasets/fcggmx8f4/1>).

Specificity of the device was higher for lesions <6 mm (39.3%, 95% CI, 32.2% to 46.8%) than for lesions ≥6 mm (19.6%, 95% CI, 13.2% to 28.2%) (Supplementary Figs 5 and 6, available via Mendeley at <https://data.mendeley.com/datasets/fcggmx8f4/1>). Sensitivity of the device by lesion size could not be compared due to study sample size.

The PPV for melanoma detection for lesions with spectral scores of 6-10 was 31.9% (95% CI, 20.2% to 46.4%), compared to 12.6% (95% CI, 8.4% to 18.4%) for lesions with spectral scores of 1-5 ( $P < .001$ , Table V). The NNB to detect melanoma was 3.1 and

**Table III.** Clinical and histopathologic diagnoses of lesions

Characteristics	<i>n</i> = 440, <i>n</i> (%)
Dermatologist-investigator prediction	
Benign	224 (50.9)
Severely dysplastic nevus	128 (29.1)
Melanoma	88 (20.0)
Dermatologist-investigator level of confidence	
High	319 (72.5)
Low	121 (27.5)
Histopathologic diagnosis	
Benign, <i>n</i> = 326	
Benign nevus	108 (24.5)
Mildly dysplastic nevus	88 (20.0)
Seborrheic keratosis	57 (13.0)
Solar lentigo	22 (5.0)
Actinic keratosis	12 (2.7)
Benign - discordant*	10 (2.3)
Moderately dysplastic nevus	8 (1.8)
Lichenoid keratosis	6 (1.4)
Benign other <sup>†</sup>	5 (1.1)
Simple lentigo	4 (0.9)
Blue nevus	3 (0.7)
Verruca	1 (0.2)
Epidermal cyst	1 (0.2)
Dermatofibroma	1 (0.2)
Severely dysplastic nevus, <i>n</i> = 44	
Atypical melanocytic hyperplasia	30 (68.2)
Atypical melanocytic proliferation	14 (31.8)
Malignant, <i>n</i> = 70	
Melanoma	
Thickness, <i>n</i> = 44	
In situ	28 (63.6)
≤1.0 mm	12 (27.3)
>1.0-2.0 mm	1 (2.3)
>2.0-4.0 mm	2 (4.5)
>4.0 mm	1 (2.3)
Basal cell carcinoma	13 (3.0)
Squamous cell carcinoma	12 (2.7)
Malignant other <sup>‡</sup>	1 (0.2)

*mm*, Millimeter.

\*Benign-discordant lesions were consistently considered benign/low risk histopathologic diagnoses during review by the expert dermatopathologists but received different benign diagnoses during primary review and secondary review processes.

<sup>†</sup>Benign other diagnoses include poroma, post-inflammatory pigmentary alternation, inflamed verrucoid keratosis, ochronosis, and foreign body reaction.

<sup>‡</sup>Other malignant diagnosis was atypical fibroxanthoma (malignant fibrous histiocytoma).

7.9 for spectral scores of 6-10 and 1-5, respectively. Spectral scores of 8-10 yielded a PPV of 47.4% (95% CI, 24.9% to 69.8%) ( $P < .001$ , scores 8-10 vs scores 4-7), scores of 4-7 yielded a PPV of 20.5% (95% CI, 12.7% to 31.5%) ( $P = .003$ , scores 4-7 vs scores 1-3), and scores of 1-3 yielded a PPV of

**Table IV.** Concordance between ESS device\* and histopathologic reference standard<sup>†</sup> for detection of melanoma

Histopathologic diagnosis <sup>†</sup>	ESS Device*	
	High risk	low risk
Melanoma, <i>n</i> = 44, <i>n</i> (%)	42 (95.5)	2 (4.5)
Benign, <i>n</i> = 326, <i>n</i> (%)	220 (67.5)	106 (32.5)
Specificity (95% CI <sup>‡</sup> )	0.325 (0.272-0.383)	
Sensitivity (95% CI <sup>‡</sup> )	0.955 (0.845-0.988)	
NPV (95% CI <sup>‡</sup> )	0.981 (0.918-0.996)	
PPV (95% CI <sup>‡</sup> )	0.160 (0.116-0.217)	

CI, Confidence interval; ESS, elastic scattering spectroscopy; NPV, negative predictive value; PPV, positive predictive value.

\*Lesions that were classified by the ESS device as "Investigate Further" were interpreted as high risk, whereas lesions classified as "Monitor" were interpreted as low risk.

<sup>†</sup>Severely dysplastic nevus (*n* = 44) and other malignant diagnoses [(basal cell carcinoma (*n* = 13), squamous cell carcinoma (*n* = 12), and atypical fibroxanthoma (*n* = 1)] are excluded from this table.

<sup>‡</sup>95% CIs were calculated accounting for the within-subject correlation using the Wilson method.

10.3% (95% CI, 6.3% to 16.5%). For spectral scores of 8-10, 4-7, and 1-3, the NNB to detect melanoma was 2.1, 4.9, and 9.7, respectively. Findings were similar for the combined outcome of melanoma and SDN (Supplementary Table V, available via Mendeley at <https://data.mendeley.com/datasets/fcggmx8f4/1>). Of the SDN classified by the device as high risk (86.4%), 94.7% had a spectral score of 1-5.

AUROC comparisons for clinical predictions by the device versus dermatologist-investigators yielded similar results for the detection of melanoma (AUROC 0.758 vs 0.747;  $P = .829$ ) and for melanoma and SDN (AUROC 0.652 vs 0.633;  $P = .700$ , Supplementary Table VI, available via Mendeley at <https://data.mendeley.com/datasets/fcggmx8f4/1>).

## DISCUSSION

The ESS device was highly sensitive with a high NPV for detecting melanoma and the combined outcome of melanoma and SDN. In real-world use, a negative ESS device result may be useful in reducing healthcare resource utilization and morbidity. The previously reported sensitivity for melanoma detection by referral centers/experts with clinical examination (69%) and dermoscopy (87%) is similar to the sensitivity of the ESS device for melanoma detection in this study (95.5%), although the values are not directly comparable due to differences in study designs and populations.<sup>13</sup> The specificity of the ESS device for melanoma detection in this study was lower (32.5%) than the reported

**Table V.** Spectral score<sup>a</sup> breakdown for detection of melanoma in lesions classified by the ESS device as high risk

Spectral score <sup>a</sup>	Benign <sup>†</sup> , n = 220, n (%)	Melanoma <sup>†</sup> , n = 42, n (%)	PPV (95% CI) <sup>‡</sup>	NNB	P-value <sup>§</sup>
Group 1					
1-5	188 (85.5)	27 (64.3)	0.126 (0.084-0.184)	7.9	
6-10	32 (14.5)	15 (35.7)	0.319 (0.202-0.464)	3.1	<.001
Group 2					
1-3	148 (67.3)	17 (40.5)	0.103 (0.063-0.165)	9.7	
4-7	62 (28.2)	16 (38.1)	0.205 (0.127-0.315)	4.9	.003 <sup>  </sup>
8-10	10 (4.5)	9 (21.4)	0.474 (0.249-0.698) <sup>¶</sup>	2.1	<.001 <sup>#</sup>

CI, Confidence interval; NNB, number needed to biopsy; PPV, positive predictive value.

<sup>a</sup>Lesions that were classified by the ESS device as "Investigate Further" were interpreted as high risk, whereas lesions classified as "Monitor" were interpreted as low risk. "Investigate Further" results are accompanied by a spectral score of 1 to 10, with 10 being most similar to previously validated malignant lesions.

<sup>†</sup>Histopathologic diagnoses.

<sup>‡</sup>The PPV calculation only considers lesions classified by the device as high risk. 95% CIs were calculated accounting for the within-subject correlation using the Wilson method.

<sup>§</sup>The Wald method was used to compare PPVs.

<sup>||</sup>The P-value compares PPVs for lesions with spectral scores of 1-3 vs 4-7.

<sup>¶</sup>Where the sample size was too small to utilize the Wilson method for calculating CIs, the Wald method was used.

<sup>#</sup>The P-value compares PPVs for lesions with spectral scores of 4-7 vs 8-10.

specificity of clinical examination (88%) and dermoscopy (91%) by referral centers/experts,<sup>13</sup> but all lesions in this study were preselected for biopsy by dermatologist-investigators, and thus, these values are not directly comparable as well. The AUROC is an additional measure of diagnostic accuracy that was similar in this study for the device and dermatologist-investigators in the detection of melanoma (AUROC 0.758 vs 0.747;  $P = .829$ ).<sup>14</sup>

Our findings suggest that the ESS device is likely best used when there is clinical suspicion for melanoma, with a negative ESS result interpreted as highly unlikely to be melanoma or SDN. ESS device use may allow for prioritization of individuals with high-risk skin lesions as classified by the device. The device may also yield benefit as an adjunct in teledermatology settings as dermatologists have been found to have a sensitivity of 67.7% and specificity of 38.6% for melanoma detection with dermoscopic and clinical images alone.<sup>15</sup> The ESS device is handheld, portable, requires minimal training, and does not require additional materials/reagents.

The ESS device may be particularly useful in areas that are cosmetically-sensitive or at risk for poor wound healing and in patients with numerous atypical-appearing lesions.<sup>16</sup> Although the ESS device had a PPV of 16.0% and specificity of 32.5%, the device correctly classified 32.5% of benign lesions that were biopsied by dermatologist-investigators. The PPV of 16.0% equates to a NNB of 6.3, compared to a NNB of 10 for the dermatologist-investigators. The device yielded a negative result of "Monitor" for

2 histopathologic melanomas (4.5%), and thus, biopsy is warranted when there is sufficient clinical and/or dermoscopic suspicion for melanoma. Additionally, the ESS device does not account for patient history and similarity to other lesions, which are critical in clinical decision-making.<sup>17</sup> Device performance differed by lesion size, with a lower specificity for detecting melanoma in lesions  $\geq 6$  mm than for lesions  $< 6$  mm. This finding could be due to higher variability between the 5 device recordings during examination of larger lesions.

Although the PPV for melanoma detection with a spectral score of 6-10 was 2.5 times higher than those with a spectral score of 1-5, there were histopathologic melanomas in the 1-5 spectral score group. A spectral score of 8-10 exhibited a PPV approaching 50%; such lesions should be triaged quickly to biopsy. Further research is needed to evaluate characteristics of melanomas that received a lower spectral score to improve device performance.

Previously FDA-approved devices for melanoma detection include Nevisense (SciBase AB) and MelaFind (MELA Sciences Inc). The Nevisense system uses similar spectroscopy technology, electrical impedance spectroscopy, to detect melanoma and has a sensitivity of 96.6% (one-sided 95% lower confidence bound of 94.2%) and specificity of 34.4% (95% CI, 32.0% to 36.9%).<sup>17</sup> MelaFind's AI multispectral camera system for melanoma detection has a sensitivity of 98.4% (one-sided 95% lower confidence bound of 95.6%) and specificity of 9.5% (95% CI, 6.1% to 12.9%).<sup>18</sup> Compared to these devices, the ESS device is simple to use, requires

minimal training, and provides an immediate result. Additionally, unlike Nevisense and MelaFind, the ESS device was developed for detection of all common skin cancer types, not melanoma specifically.<sup>17,18</sup>

Other AI technologies for melanoma detection currently lack FDA approval. Moleanalyzer Pro (Fotofinder Systems) uses a convolutional neural network to classify skin lesions using dermoscopic images.<sup>19</sup> Moleanalyzer Pro has been shown to have a comparable sensitivity (81.6%, 95% CI, 66.6% to 90.8%) and a higher specificity (88.9%, 95% CI, 83.7% to 92.7%) for melanocytic lesions than dermatologists, including lesions in special sites (ie, acral, nails).<sup>19</sup> However, the time and effort required to take and upload dermoscopic images may limit use of this technology. Publicly available smartphone and web-based dermatology apps using AI diagnostics have an overall sensitivity of 28% (95% CI, 17% to 39%) and a specificity of 81% (95% CI, 71% to 91%).<sup>20</sup> As these apps are publicly available, use may cause false reassurance or alternatively, increased health-care utilization.<sup>20</sup>

Limitations include study design, as this study was conducted by dermatologists in high-volume melanoma dermatology clinics with a high cancer prevalence. Device performance in amelanotic melanomas is unknown as the majority of lesions included were pigmented. Lesions in nonaccessible sites (eg, under nails), adjacent to or on scars/areas of past surgical intervention, acral surfaces, mucosal surfaces, or near the eye were not included; device performance in these settings is unknown. All lesions in this study were clinically suspicious for melanoma per board-certified dermatologists and selected for biopsy; therefore, the benign lesions in this study were likely not representative of such lesions in the general population. Due to study design, we lack information about dermatologist-investigators' false negative rate and biopsy sensitivity as well as device performance in lesions not selected for biopsy by dermatologist-investigators. This may limit the generalizability of the sensitivity and specificity in this study with respect to the general population, who may have lower skin cancer baseline risk and may present with lower risk skin lesions.

## CONCLUSION

The handheld, noninvasive ESS device reported herein exhibited high sensitivity and negative predictive value for melanoma detection and had AUROC comparable to that of dermatologist-investigators. Coupled with skin examination findings, this device may aid clinicians in melanoma detection. The device may be particularly useful as a triage tool for high-risk lesions, for lesions in

cosmetically sensitive locations or areas with poor wound healing, in patients with numerous atypical-appearing lesions, and as an adjunctive tool for teledermatology.

## Conflicts of interest

Drs Tepedino, McNiff, Fung, and Hartman were all provided funding for their participation in the study. Mr Gianacas is a paid consultant for DermaSensor, Inc. Dr Grant-Kels is an uncompensated member of the Advisory Board for DermaSensor, Inc.

## REFERENCES

1. Matthews NH, Li WQ, Qureshi AA, Weinstock MA, Cho E. Epidemiology of melanoma. In: Ward WH, Farma JM, eds. *Cutaneous Melanoma: Etiology and Therapy*. Codon Publications; 2017:3-22. <https://doi.org/10.15586/codon.cutaneousmelanoma.2017.ch1>
2. Arnold M, Singh D, Laversanne M, et al. Global burden of cutaneous melanoma in 2020 and projections to 2040. *JAMA Dermatol*. 2022;58(5):495-503. <https://doi.org/10.1001/jama Dermatol.2022.0160>
3. Seiger K, Schmults CD, Silk AW, Ruiz ES. Cost and utilization of immunotherapy and targeted therapy for melanoma: cross-sectional analysis in the Medicare population, 2013 and 2015. *J Am Acad Dermatol*. 2020;82(3):761-764. <https://doi.org/10.1016/j.jaad.2019.10.023>
4. Leachman SA, Cassidy PB, Chen SC, et al. Methods of melanoma detection. *Cancer Treat Res*. 2016;167:51-105. [https://doi.org/10.1007/978-3-319-22539-5\\_3](https://doi.org/10.1007/978-3-319-22539-5_3)
5. Nelson KC, Swetter SM, Saboda K, Chen SC, Curiel-Lewandrowski C. Evaluation of the number-needed-to-biopsy metric for the diagnosis of cutaneous melanoma: a systematic review and meta-analysis. *JAMA Dermatol*. 2019;155(10):1167. <https://doi.org/10.1001/jamadermatol.2019.1514>
6. Merry S, Chatha K, Croghan I, Nguyen V, McCormick B, Leffell D. Clinical performance of novel elastic scattering spectroscopy (ESS) in detection of skin cancer: a blinded, prospective, multi-center clinical trial. *J Clin Aesthetic Dermatol*. 2023;16(4 Suppl 1):S5-S31.
7. Tepedino M, Baltazar D, Hucks C, Chatha K, Zeitouni N. Use of elastic-scattering spectroscopy on patient-selected lesions that are concerning for skin cancer. *Cutis*. 2022;110(6S):35-36.
8. Rodriguez-Diaz E, Manolakos D, Christman H, et al. Optical spectroscopy as a method for skin cancer risk assessment. *Photochem Photobiol*. 2019;95(6):1441-1445. <https://doi.org/10.1111/php.13140>
9. Device. DermaSensor. Accessed May 24, 2023. <https://www.dermasensor.com.au/device/>
10. Boyer J, Mourant J, Bigio I. Theoretical and experimental investigations of elastic scattering spectroscopy as a potential diagnostic for tissue pathologies. *Osti.gov*. Accessed May 24, 2023. <https://www.osti.gov/biblio/10138239>
11. Saha KK, Miller D, Wang S. A comparison of some approximate confidence intervals for a single proportion for clustered binary outcome data. *Int J Biostat*. 2016;12(2):20150024. <https://doi.org/10.1515/ijb-2015-0024>
12. Marchetti MA, Yu A, Nanda J, et al. Number needed to biopsy ratio and diagnostic accuracy for melanoma detection. *J Am Acad Dermatol*. 2020;83(3):780-787. <https://doi.org/10.1016/j.jaad.2020.04.109>
13. Vestergaard ME, Macaskill P, Holt PE, Menzies SW. Dermoscopy compared with naked eye examination for the

- diagnosis of primary melanoma: a meta-analysis of studies performed in a clinical setting. *Br J Dermatol.* 2008;159(3):669-676. <https://doi.org/10.1111/j.1365-2133.2008.08713.x>
14. Simundić AM. Measures of diagnostic accuracy: basic Definitions. *EJIFCC.* 2009;19(4):203-211.
  15. Winkelmann RR, Yoo J, Tucker N, White R, Rigel DS. Assessment of a diagnostic predictive probability model provided by a multispectral digital skin lesion analysis device for melanoma and other high-risk pigmented lesions and its impact on biopsy decisions. *J Clin Aesthetic Dermatol.* 2014;7(12):16-18.
  16. Shah A, Hynstrom J, Florell SR, Grossman D. Use of the pigmented lesion Assay to rapidly screen a patient with numerous clinically atypical pigmented lesions. *JAAD Case Rep.* 2019;5(12):1048-1050. <https://doi.org/10.1016/j.jdc.2019.10.004>
  17. Malvey J, Hauschild A, Curiel-Lewandrowski C, et al. Clinical performance of the Nevisense system in cutaneous melanoma detection: an international, multicentre, prospective and blinded clinical trial on efficacy and safety. *Br J Dermatol.* 2014;171(5):1099-1107. <https://doi.org/10.1111/bjd.13121>
  18. Monheit G, Cognetta AB, Ferris L, et al. The performance of MelaFind: a prospective multicenter study. *Arch Dermatol.* 2011;147(2):188. <https://doi.org/10.1001/archdermatol.2010.302>
  19. Winkler JK, Blum A, Kommos K, et al. Assessment of diagnostic performance of dermatologists cooperating with a convolutional neural network in a prospective clinical study: human with Machine. *JAMA Dermatol.* 2023;159(6):621. <https://doi.org/10.1001/jamadermatol.2023.0905>
  20. Sun MD, Kentley J, Mehta P, Dusza S, Halpern AC, Rotemberg V. Accuracy of commercially available smartphone applications for the detection of melanoma. *Br J Dermatol.* 2022;186(4):744-746. <https://doi.org/10.1111/bjd.20903>

## Jaklitsch et al. (2025)

## RESEARCH LETTER

**Prospective evaluation of an AI-enabled elastic scattering spectroscopy device for triage of patient-identified skin lesions in dermatology clinics**



*To the Editor:* Early detection of skin cancer holds promise for reducing mortality and surgical morbidity.<sup>1</sup> However, wait time for suspicious lesions can be long, and although many patient-identified lesions are benign, initial skin cancer presentations are often first noticed by patients.<sup>2,3</sup> A highly sensitive Elastic Scattering Spectroscopy (ESS) device integrated with artificial intelligence (AI) utilizing spectral scans was cleared for adjunct usage in primary care and is the only FDA-cleared tool for evaluating all 3 common types of skin cancer (melanoma, basal cell carcinoma [BCC], and squamous cell carcinoma [SCC]).

This investigator-initiated, matched comparison study aims to evaluate device performance on patient-identified lesions in a dermatology setting to assess its utility for screening lesions appropriate for referral across clinical contexts. We prospectively evaluated 150 patients presenting with self-identified lesions of concern to 3 outpatient dermatology offices. Diagnostic accuracy of device readout (classifying lesions as either “Monitor,” alone, or “Investigate Further,” alongside a score of 1 to 10, indicating spectral similarity to malignant tissue) was compared to blinded dermatologist management decisions and histopathology.

The study included 150 lesions in 72 males and 78 females, with a mean age of 59.6 years. Malignant lesions were identified in 14.7% of cases (22/150), with melanoma, BCC, and SCC accounting for 3, 9, and 10, respectively (Table I). The ESS device identified 138 lesions as “Investigate Further” and 12 as “Monitor.” Sensitivity for detecting malignant lesions was 100%, negative predictive value (NPV) was 100%, specificity was 9.4%, and positive predictive value (PPV) was 15.9%. For lesions with a spectral score of 7 to 10, sensitivity was 86.4%, NPV 96.6%, with specificity increasing to 67.2% and PPV to 31.1%. Compared to dermatologists’ decisions, sensitivity was 95.5%, NPV 75%, specificity 10.7%, and PPV 45.7% (Table II). Overall AUC of the device compared to malignant lesions was 0.79 (Supplementary Fig 1, available via Mendeley at <https://data.mendeley.com/datasets/vpxw84996p/1>), identical to a validation study in

**Table I.** Patient demographics, research for concern, lesion location, classification, and dermatologist management decision

Patient demographics	n (%)
Sex (male)	72 (48.0)
Sex (female)	78 (52.0)
Age (mean ± SD; range)	59.6 ± 17.5; 22-92
Reason for concern	n (%)
Worry for skin cancer	120 (80.0)
Causes discomfort	15 (10.0)
Aesthetic concern	8 (5.3)
Other	7 (4.7)
Lesion location	n (%)
Trunk	50 (33.3)
Head	42 (28.0)
Upper extremity	19 (12.7)
Lower extremity	23 (15.3)
Other	16 (10.7)
Lesion classification	n (%)
Melanoma	3 (2.0)
Basal cell carcinoma	9 (6.0)
Squamous cell carcinoma	10 (6.7)
Actinic keratosis	15 (10.0)
Seborrheic keratosis	57 (38.0)
Dermatofibroma	5 (3.3)
Nevus	34 (22.6)
Other	17 (11.3)
Dermatologist management	n (%)
Biopsy	66 (44.0)
Treat	31 (20.7)
Monitor	13 (8.7)
Reassurance	40 (26.7)

primary care.<sup>4</sup> Frequency counts of benign and malignant lesions compared to corresponding spectral score are also visualized (Supplementary Fig 2, available via Mendeley at <https://data.mendeley.com/datasets/vpxw84996p/1>).

These results demonstrate that the ESS device maintains robust sensitivity and NPV for detecting malignant lesions, even in a dermatology setting with patient-identified lesions, and support its role as an adjunctive rule-out tool in skin cancer triage. Despite a relatively low specificity (9.4%) and PPV (15.9%), most false positive results (59.5%) were actively managed by dermatologists through biopsy or treatment. Furthermore, given dermatologist assessment was used as the gold standard, all unbiopsied lesions had a biased specificity of 100%. Spectral score was a reliable positive predictor of malignancy, with these results (using a threshold of spectral score 7 to 10 for

© 2025 The Author(s). Published by Elsevier Inc. on behalf of the American Academy of Dermatology, Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

**Table II.** Sensitivity, specificity, NPV, and PPV as compared to histopathology and dermatologist decision to biopsy

	Compared to pathology (benign vs malignant, 0 score negative)	Compared to pathology (1-2 score negative)	Compared to pathology (1-6 score negative)	Compared to derm decision to biopsy
Overall sensitivity	100% (22/22)	100% (22/22)	86.4% (19/22)	95.5% (63/66)
Melanoma sensitivity	100% (3/3)	100% (3/3)	66.7% (2/3)	
BCC sensitivity	100% (9/9)	100% (9/9)	88.9% (8/9)	
SCC sensitivity	100% (10/10)	100% (10/10)	90% (9/10)	
Overall specificity	9.4% (12/128)	21.9% (28/128)	67.2% (86/128)	10.7% (9/84)
BMN specificity	8.8% (3/34)	26.5% (9/34)	79.4% (27/34)	
SK specificity	14% (8/57)	28.1% (16/57)	68.4% (39/57)	
Benign other specificity (including AK)	2.7% (1/37)	8.1% (3/37)	54.1% (20/37)	
AK specificity	0% (0/15)	0% (0/15)	26.7% (4/15)	
Benign other specificity (excluding AK)	4.5% (1/22)	13.6% (3/22)	72.7% (16/22)	
Overall NPV	100% (12/12)	100% (28/28)	96.6% (86/89)	75.0% (9/12)
Overall PPV	15.9% (22/138)	18% (22/122)	31.1% (19/61)	45.7% (63/138)

tissue of the highest similarity to malignant tissue) nearly identical to a prior study.<sup>5</sup>

Though limited by sample size and a single dermatologist reference, this study is unique in including only patient-selected lesions, with outcomes compared to dermatologists' management decisions. This study suggests that the ESS device, as a complementary tool, may enhance triage and decision-making for patient-identified lesions to be employed by non-specialists in any clinical setting. Still, physicians using the device should utilize appropriate clinical context to minimize unnecessary interventions.

Erik Jaklitsch, MD,<sup>a</sup> Selina Chang, BS,<sup>a</sup> Sabrina Bruno, BS,<sup>a</sup> Nicholas D'Angelo, BS,<sup>a</sup> Joe K. Tung, MD,<sup>a,b</sup> and Laura K. Ferris, MD, PhD<sup>b,c</sup>

From the University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania<sup>a</sup>; Department of Dermatology, University of Pittsburgh, Pittsburgh, Pennsylvania<sup>b</sup>; and Department of Dermatology, University of North Carolina, Chapel Hill, North Carolina.<sup>c</sup>

Funding sources: None.

Patient consent: Written informed consent was obtained from all patients.

IRB approval status: Reviewed and approved by University of Pittsburgh IRB, STUDY23030050.

Key words: artificial intelligence; basal cell carcinoma; diagnostic tools; elastic scattering

spectroscopy; machine learning; melanoma; skin cancer; squamous cell carcinoma; technology.

Correspondence to: Laura K. Ferris, MD, PhD, Department of Dermatology, University of North Carolina, 410 Market St Suite 400, Chapel Hill, NC 27516

E-mail: [ferrislk@email.unc.edu](mailto:ferrislk@email.unc.edu)

#### Conflicts of interest

None disclosed.

#### REFERENCES

- Henrikson NB, Ivlev I, Blasi PR, et al. Skin cancer screening: updated evidence report and systematic review for the US preventive services task force. *JAMA*. 2023;329(15):1296-1307. <https://doi.org/10.1001/JAMA.2023.3262>
- Jiang A, Jefferson IS, Robinson SK, et al. Skin cancer discovery during total body skin examinations. *Int J Womens Dermatol*. 2021;7(4):411. <https://doi.org/10.1016/J.IJWD.2021.05.005>
- Weinstock MA, Risica PM, Martin RA, et al. Melanoma early detection with thorough skin self-examination: the "Check It Out" randomized trial. *Am J Prev Med*. 2007;32(6):517. <https://doi.org/10.1016/J.AMEPRE.2007.02.024>
- Merry SP, Chatha K, Croghan I, Nguyen VL, McCormick B, Leffell D. Clinical performance of novel elastic scattering spectroscopy (ESS) in detection of skin cancer: a blinded, prospective, multi-center clinical trial. *J Clin Aesthet Dermatol*. 2023;16(4 Suppl):s16.
- Tepedino M, Baltazar D, Hanna K, Bridges A, Billot L, Zeitouni NC. Elastic scattering spectroscopy on patient-selected lesions concerning for skin cancer. 2024;37:427-435. <https://doi.org/10.3122/jabfm.2023.230256R2>

<https://doi.org/10.1016/j.jdin.2025.07.007>

## Lee and Tung (2024)

## NOTES &amp; COMMENTS

**Comment on “Use of an elastic-scattering spectroscopy and artificial intelligence device in the assessment of lesions suggestive of skin cancer: A comparative effectiveness study”**



*To the Editor:* We read with great interest the article by Manolakos et al,<sup>1</sup> which assesses the performance of an elastic-scattering spectroscopy (ESS) device to evaluate lesions suggestive of skin cancer. The authors used ESS to categorize clinically suspicious skin lesions into “monitor” (negative result) or “investigate further” (positive result) and compared the results with gross and histopathologic diagnoses by dermatologists and dermatopathologists. The device achieved high sensitivity in predicting all lesion types; however, the specificity was low across all major outcomes. The article did not provide information about the training data set used by the ESS algorithm, but the discrepancy between sensitivity and specificity suggests there might be some class imbalance in the underlying data set, likely skewed toward benign lesions. Imbalance data sets are not uncommon in health care associated data,<sup>2</sup> but it could limit the generalizability of a machine learning model if class balance is significantly different in the test data set. Because the study participants in this article have a higher pretest probability, the differences are not amenable to posterior probability adjustment. This is confirmed by the fact that the model specificity is higher in unbiopsied benign lesions when compared with that of biopsied lesions. Although high sensitivity, even at the cost of lower specificity, is generally acceptable for screening tools, it is instrumental to apply this device algorithm on a patient population similar to the one it was originally trained on to ensure the suggested goal of prioritizing referrals and avoiding unnecessary procedures. It is possible to leverage the “spectral similarity score” to fine-tune cut-offs for positive predictions based on patients’ pretest probabilities.

In addition, representation of different skin types also needs to be considered when applying the device algorithm in the primary care setting. Of the authors’ 383 patients across testing and cross-validation groups, only one participant was non-White. It is well-known that artificial intelligence

(AI) can perpetuate and exacerbate racial and ethnic disparities.<sup>3</sup> Health care professionals report lower clinical confidence in managing skin conditions in dark skin tones.<sup>4</sup> Any device utilizing an AI algorithm should be developed and cross-validated with this in mind to avoid exacerbating known disparities. Therefore, it is important to include more skin of color patients in research studies given that ESS patterns may vary across different skin tones.

Altogether, we emphasize the importance of applying AI algorithms to populations similar to the training data set and class balance. We should also make a concerted effort to address the underrepresentation of darker skin types and non-White populations in AI research to avoid inadvertently propagating racial bias in health care.<sup>5</sup>

*Tai-Lin (Irene) Lee, MD,<sup>a</sup> and Joe K. Tung, MD, MBA<sup>b,c</sup>*

*From the Department of Radiology and Imaging Science, Emory University School of Medicine, Atlanta, Georgia<sup>a</sup>; Department of Dermatology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania<sup>b</sup>; and Department of Dermatology, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania.<sup>c</sup>*

*Funding sources: None.*

*IRB approval status: Not applicable.*

*Key words: artificial intelligence; class imbalance; class balance; elastic-scattering spectroscopy; machine learning.*

*Correspondence to: Joe K. Tung, MD, MBA, Department of Dermatology, University of Pittsburgh Medical Center, 3708 Fifth Ave, Pittsburgh, PA 15213*

*E-mail: tungjk@upmc.edu*

**Conflicts of interest**

None disclosed.

**REFERENCES**

1. Manolakos D, Patrick G, Geisse JK, et al. Use of an elastic-scattering spectroscopy and artificial intelligence device in the assessment of lesions suggestive of skin cancer: a comparative effectiveness study. *JAAD Int.* 2024;14:52-58. <https://doi.org/10.1016/j.jdin.2023.08.019>
2. Zhao Y, Wong ZSY, Tsui KL. A framework of rebalancing imbalanced healthcare data for rare events’ classification: a case of look-alike sound-alike mix-up incident detection. *J Healthc Eng.* 2018;2018:6275435. <https://doi.org/10.1155/2018/6275435>

© 2024 by the American Academy of Dermatology, Inc. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

JAAD INT  
VOLUME 17

Notes & Comments **123**

3. Tipton K, Leas BF, Flores E, et al. Impact of healthcare algorithms on racial and ethnic disparities in health and healthcare. Agency for Healthcare Research and Quality (AHRQ); 2023. <https://doi.org/10.23970/AHRQEPCCER268>
  4. Harrison J. A scoping review exploring the confidence of healthcare professionals in assessing all skin tones. *Br Paramed J*. 2023; 8(2):18-28. <https://doi.org/10.29045/14784726.2023.9.8.2.18>
  5. Fahmy LM, Karantza IM, Schreidah CM, Lapolla BA, Geskin LJ. Skin of color representation in dermoscopy studies distinguishing benign from malignant skin lesions: a scoping review. *JAAD Int*. 2023;13:179-180. <https://doi.org/10.1016/j.jdin.2023.08.018>
- <https://doi.org/10.1016/j.jdin.2024.06.005>

## Merry et al. (2025)

Original Research

## Primary Care Physician Use of Elastic Scattering Spectroscopy on Skin Lesions Suggestive of Skin Cancer

Journal of Primary Care & Community Health  
 Volume 16: 1–11  
 © The Author(s) 2025  
 Article reuse guidelines:  
[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)  
 DOI: 10.1177/21501319251344423  
[journals.sagepub.com/home/jpc](https://journals.sagepub.com/home/jpc)  


Stephen P. Merry<sup>1</sup>, Ivana T. Croghan<sup>2</sup>, Kimberly A. Dukes<sup>3</sup>,  
 Brian C. McCormick<sup>4</sup>, Gerard T. Considine<sup>5</sup>, Michelle J. Duvall<sup>1</sup>,  
 Curtis T. Thompson<sup>6</sup>, and David J. Leffell<sup>7</sup>

### Abstract

**Objectives:** To evaluate the performance of noninvasive, elastic scattering spectroscopy, algorithm-powered device (DermaSensor) to detect melanoma and basal and squamous cell cancers in the primary care setting.

**Patients & Methods:** DERM-SUCCESS, a blinded, prospective, multicenter pivotal study, enrolled adult patients between August 17, 2020, and December 9, 2021, with lesions that their primary care physicians (PCPs) suspected of skin cancer at clinics in the US (n = 18) and Australia (n = 4). These lesions were assessed by PCPs and scanned with the DermaSensor device. Biopsy specimens were collected, and histopathologic analysis was performed by dermatopathologists. The diagnostic performance of the device, dermatopathologist discordance, and subgroup analyses of clinical interest were calculated.

**Results:** Of the 1579 skin lesions enrolled, dermatopathologic analysis identified 224 (14.2%) cancers. Device sensitivity was 95.5% (95% CI, 91.7%–97.6%) overall and 96.3% (92.9%–98.4%) for patients in the FDA-approved age group 40 years and older (90.2% for melanoma, 97.8% for basal cell carcinoma, and 97.7% for squamous cell carcinoma). Device specificity was 20.7%. The negative predictive value was 96.6%, and the positive predictive value was 16.6% (NNB 6). The device misclassified as “monitor” rather than “investigate further” 4 keratinocyte carcinomas and 4 melanomas in patients aged 40 years or older (n = 8, 0.5% of lesions, 3.7% of cancers biopsied).

**Conclusions:** The DermaSensor device is an easy-to-use, point-of-care, hand-held skin cancer adjunctive diagnostic device with high sensitivity and NPV to help inform PCP decision-making about skin lesions suspicious for cancer that need further evaluation and those that may be monitored.

### Keywords

basal cell carcinoma, DermaSensor, elastic scattering spectroscopy, melanoma, primary care, skin cancer, squamous cell carcinoma

Received: March 12, 2025; revised: April 28, 2025; accepted: April 29, 2025

### Introduction

The incidence of basal cell carcinoma (BCC), squamous cell carcinoma (SCC) (collectively, keratinocyte carcinomas [KC]), and malignant melanoma is increasing worldwide. Skin cancer will be diagnosed in 1 of 5 Americans during their lifetime, making it the most common type of cancer in the US.<sup>1</sup> In 2022, an estimated 5.4 million cases of KC were diagnosed in 3.3 million people, and an estimated 2500 to 15 000 patients died of SCC.<sup>1–4</sup> An estimated 99 780 new cases of invasive melanoma and 97 920 cases of

melanoma in situ were diagnosed in 2022, and 7650 people died of melanoma.<sup>1</sup> Most skin cancer is curable if diagnosed early, including KC, melanoma in situ, and nonulcerated malignant melanoma with a Breslow depth less than 1 mm.<sup>1</sup>

Skin biopsy and subsequent histopathologic analysis is the standard for skin cancer diagnoses.<sup>5</sup> The number needed to biopsy (NNB) ratio of total biopsies to identify 1 histopathologic-confirmed skin cancer is commonly used to compare the diagnostic accuracy of providers who perform skin biopsy and is the inverse of positive predictive value (PPV).<sup>6</sup> Training and experience in dermatologic care



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

correlates inversely with NNB and directly with diagnostic specificity and PPV.<sup>6-8</sup> One study at a large academic medical center found an NNB to find 1 melanoma of 14.33 for dermatologists, 27.80 for PCPs, and 53.56 for other primary care providers.<sup>8</sup> And, PCP sensitivity for correctly managing (i.e., refer or biopsy) skin cancer by naked eye examination is well below that of dermatologists, ranging from 54% to 88%.<sup>9,10</sup>

With only 1 dermatologist for every 30 000 people in the US, early skin cancer detection cannot depend on dermatologists. Improved PCP performance is necessary to improve detection and avoid unnecessary biopsies and referrals of benign lesions for biopsy.<sup>11</sup> Dermoscopy training greatly improves PCP sensitivity and specificity,<sup>10,12-17</sup> however, the requirements for additional training and ongoing practice to maintain dermoscopy competency has limited its use in the US to only 8% of PCPs.<sup>18,19</sup> Despite availability of a range of education and algorithm interventions to improve PCP diagnostic ability, outcomes are inconsistent, difficult to replicate, and challenging to maintain.<sup>20</sup>

The Nevisense system (Scibase) is an electrical impedance spectroscopy device approved by the US Food and Drug Administration (FDA) for melanoma detection (sensitivity, 96.6%; specificity, 34.4%). Its use is restricted to dermatologists and only lesions concerning for melanoma.<sup>21</sup> Because seborrheic keratoses generate high electrical impedance scores, they must be excluded via prescreening.

Artificial intelligence-assisted smartphone applications that assess clinical and dermoscopic images of skin lesions are becoming increasingly available though none are authorized for use by the FDA.<sup>22,23</sup> Although such image-based software may eventually achieve dermatologist-level sensitivity and specificity,<sup>23-26</sup> current performance (early 2025) is below standard of care,<sup>27,28</sup> increasing use of health care resources without added benefits.<sup>29,30</sup> These applications are unable to evaluate cellular characteristics like ESS.

An affordable, automated point-of-care device is needed to improve PCP skin cancer detection and reduce unnecessary referrals for biopsy. Elastic scattering spectroscopy (ESS) is a noninvasive method for obtaining a spectral recording of photons scattered by chromophores in tissues,

facilitating detection of cellular elements indicative of malignant neoplasms in skin<sup>31</sup> and other tissues.<sup>32-36</sup>

## Methods

### Trial Design

The DermaSensor Study of Primary Care Physician Use of Elastic Scattering Spectroscopy (DERM-SUCCESS) was a prospective, blinded validation trial designed in collaboration with the FDA for clearance of a novel ESS, algorithm-powered device (DermaSensor device, manufactured by DermaSensor, Inc.) for evaluating lesions suggestive of skin cancer. The primary aim of the DERM-SUCCESS trial was to measure the performance of the device for lesions suggestive of melanoma, BCC, and/or SCC. The secondary aim was to compare device sensitivity to a performance goal of 90%, which is based on the published dermatologist sensitivity range of 81% to 96%.<sup>37-40</sup>

This study was reviewed and approved by the 3 Institutional Review Boards, each overseeing separate sites. Written informed consent was obtained for all participants before enrollment. Study participants received a small stipend.

### Setting

Participants were enrolled in a convenience sample of suburban and rural family medicine and internal medicine practices, 18 in the US and 4 in Australia, of experienced PCPs who conduct a high volume of skin exams and perform their own skin biopsies from August 17, 2020, through December 9, 2021. Nine of the 22 study sites were in towns with populations under 15 000 people.

### Participants

All patients seeing study PCPs during study enrollment with a suspicious lesion who met eligibility criteria, and were willing to participate were enrolled (Figure 1). Eligibility criteria included patients aged 22 years or older with a skin lesion(s) suggestive of melanoma, BCC, and/or SCC (based on visual inspection by PCP) requiring biopsy.

<sup>1</sup>Department of Family Medicine, Mayo Clinic, Rochester, MN, USA

<sup>2</sup>Department of Internal Medicine, Division of General Internal Medicine, Mayo Clinic, Rochester, MN, USA

<sup>3</sup>Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA

<sup>4</sup>Hampton Family Practice, Hampton, VA, USA

<sup>5</sup>Clare Medical Centre, Clare, SA, Australia

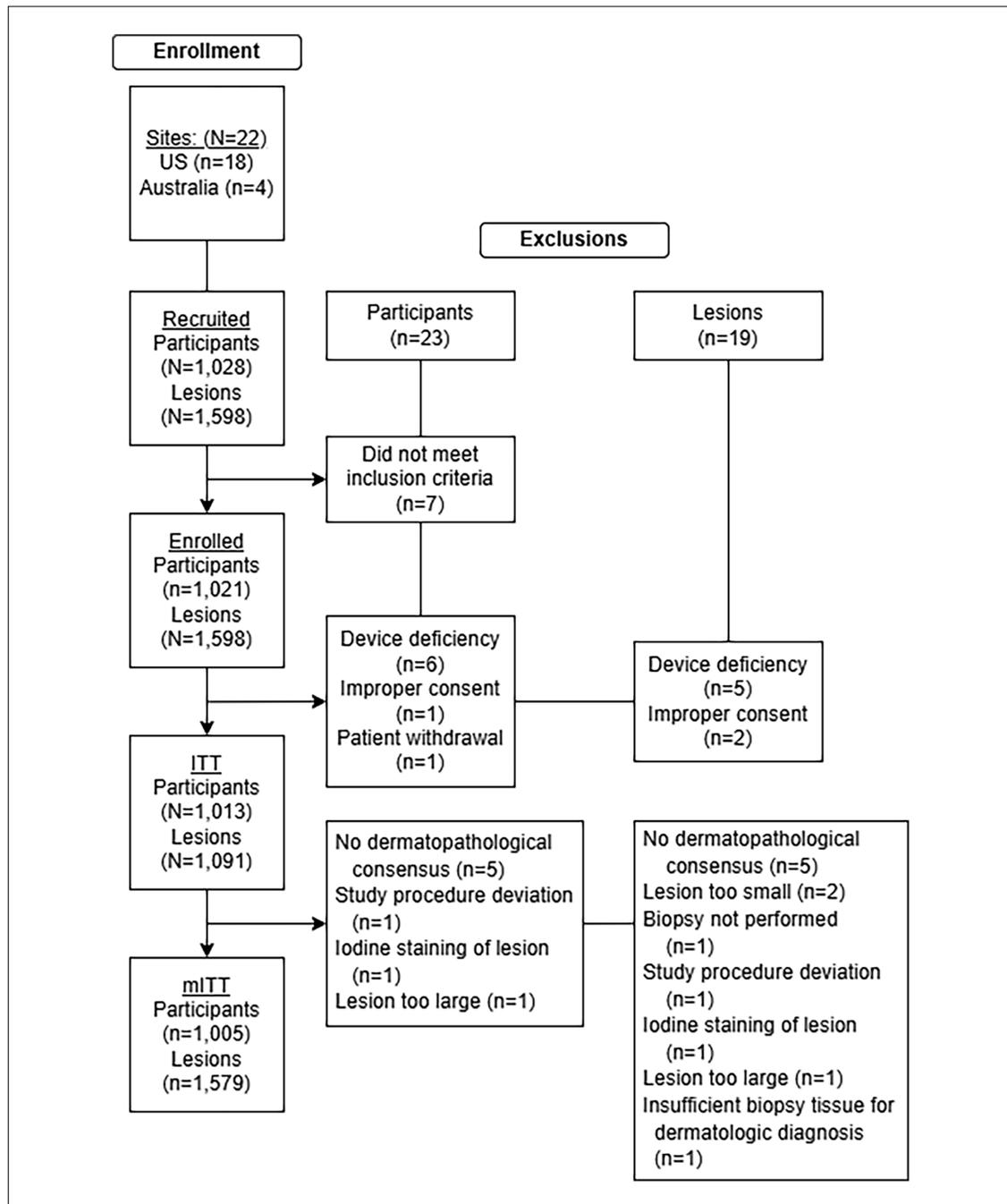
<sup>6</sup>Department of Dermatology and Department of Pathology, Oregon Health & Science University, Portland, Oregon, USA; CTA Pathology: Curtis Thompson, MD & Associates, Beaverton, OR, USA

<sup>7</sup>Department of Dermatology, Section of Dermatologic Surgery and Cutaneous Oncology, Yale School of Medicine, New Haven, CT, USA

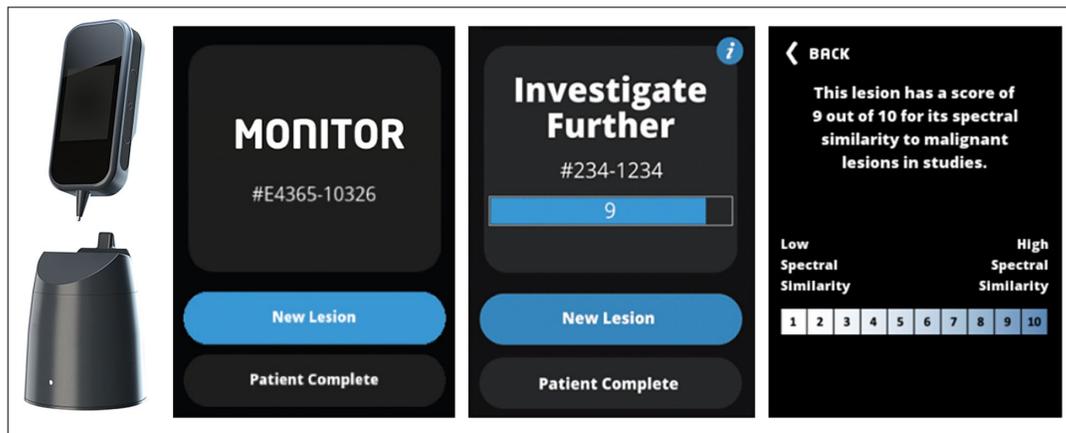
#### Corresponding Author:

Stephen P. Merry, Department of Family Medicine, Mayo Clinic, 200 First St SW, Rochester, MN 55905, USA.

Email: merry.stephen@mayo.edu



**Figure 1.** Flow diagram of included and excluded study data. The ITT population comprised all patients and lesions that were included in the safety analysis, and the mITT population comprised all patients and lesions that were included in the effectiveness analysis.  
Abbreviations: ITT, intent-to-treat; mITT, modified intent-to-treat.



**Figure 2.** Elastic scattering spectroscopy device for analysis of suspected skin cancer lesions. An example patient number (#234-1234), with a negative monitor or positive investigate further result, is shown. To investigate further results, a representative spectral score and corresponding definition are shown. (© 2023 DermaSensor, Inc; used with permission).

Excluded lesions included those less than 2.5 mm or greater than 15 mm in diameter (smaller lesions were excluded due to risk of the 2.5 mm probe tip placement on perilesional skin; larger lesions were excluded since the 5 recordings sample a limited volume of the lesional tissue); located on mucosal or acral skin, within 1 cm of the eye, at a previous biopsy or surgical site, or in acute sunburned skin; completely covered with thick crust, erosion, or ulceration (without a nonulcerated area  $\geq 2.5$  mm) or containing foreign matter. Dermoscopy was not used for lesion assessment. After target numbers of KCs were reached (see protocol below), participating PCPs were instructed to continue to enroll patients with pigmented lesions until the number of melanomas required by the FDA for this trial were enrolled. Participant sex, age, ethnicity, and self-reported race were recorded, and Fitzpatrick skin type was assessed by the patient and recorded by PCPs.

### Device

The battery-operated, handheld DermaSensor device (Figure 2) emits pulses of light (200 microseconds/pulse) that span wavelengths from near-ultraviolet, through visible, to near-infrared. Two optical fibers are enclosed in a biocompatible 2.5-mm sterilizable fiberoptic tip that is placed in contact with the lesion surface. The volume of tissue evaluated with the device for each spectral recording was estimated computationally with Monte Carlo simulations to be approximately  $0.7 \times 0.4 \times 0.5$  mm (length  $\times$  width  $\times$  depth). Five spectral scans at various points on the lesion can be performed in about 10 s total time per lesion to record light reflectance of tissue structure and architecture (e.g., nuclear and chromatin characteristics) with the algorithm, then comparing those spectral readings to lesions in the training data set.

The device AI developed algorithm was trained on more than 10 000 spectral scans and 2000 lesions. In accordance with requirements by the FDA for a pivotal study, the algorithm was developed before study enrollment. The algorithm was upgraded and then used for analysis of the blinded dataset based on data from 3 other clinical trials before study completion. The device output to which investigators and dermatopathologists were blinded is either *monitor* or *investigate further*, providing a spectral score of 1 through 10 for the latter (see Figure 2). A higher spectral score indicates greater spectral similarity with that of malignant lesions.

### Study Sequence

The PCPs electronically recorded a lesion(s) description. The lesion(s) was then scanned with the ESS device. Both PCPs and patients were blinded to the results. A biopsy specimen was collected according to the PCP standard of care and provided to an independent dermatopathologic core laboratory for histopathologic analysis. As with the PCPs, the dermatopathologists were blinded to device output.

All specimens were analyzed by 2 dermatopathologists who had to independently reach diagnostic consensus using MPATH-Dx criteria for melanocytic lesions<sup>41</sup>. Discordance (benign or malignant) between the first 2 dermatopathologists was adjudicated by a third dermatopathologist and by 2 additional dermatopathologists for lesions diagnosed by at least 2 of the initial dermatopathologists as highly atypical, with differing diagnoses (malignant, benign, or atypical) by each of the initial 3 dermatopathologists, and/or diagnosed as malignant by 1 dermatopathologist but not as malignant by the adjudicator. If the 2 additional dermatopathologists disagreed, the lesion was excluded from analysis.

Patients' participation in the study was complete after their enrolled lesion(s) was biopsied. All biopsy results were communicated to patients by their PCP and standard of care followed.

### Data Analysis

Device sensitivity was tested with an  $\alpha$  level of .025 and compared by using a method-of-moments approach for clustered matched-pair data<sup>41</sup> to account for possible within-patient correlation for patients with multiple lesions. The device's ability to detect malignant lesions better than random chance (sensitivity + specificity > 1) was tested with generalized estimating equations logistic regression with a noninferiority margin of 10% by using a 1-sided exact binomial test with an  $\alpha$  level of .025, with each patient treated as a random effect and assuming an exchangeable within-patient correlation structure to account for multiple within-patient lesions. The secondary hypothesis was tested for performance in comparison with a 90% threshold. Level of agreement between dermatopathologists was assessed with the Cohen  $\kappa$  statistic. For all other statistical inferences, 2-sided Wilson or exact confidence intervals were determined as appropriate, and unless specified,  $P < .05$  was considered statistically significant. All statistical analyses were performed with SAS software, v9.4 (SAS Institute Inc.).

### Results

A total of 1028 patients (1598 lesions) consented to participate in the study (Figure 1), of whom 7 did not meet the inclusion criteria and 16 (19 lesions) were excluded because their lesions did not meet eligibility criteria, device deficiencies, or a dermatopathologic consensus was not reached. The resulting 1005 patients (1579 lesions, 1451 in the US and 128 in Australia) were included in our analyses.

Slightly more than half of participants were women (51.5%), and the majority (97.1%) were White (Table 1). Mean (SD) participant age was 58.5 (15.1) years, and 138 patients (13.7%) were younger than 40 years. Most participants (72.5%) had Fitzpatrick skin types I through III, with most (35.0%) having type III skin. Most patients had 1 (65.4%) or 2 (20.6%) eligible lesions analyzed in the study. Of 1579 lesions, 998 (63.2%) had darker pigmentation, and 581 (36.8%) had lighter pigmentation. Nearly two-thirds (60.6%) of lesions were elevated or papular, and 67.8% were discovered by the patient. Mean (SD) lesion size was 6.7 (2.9) mm in length and 5.2 (2.2) mm in width. All biopsy specimens collected were clinically suspicious for a malignant neoplasm. Histopathologic analysis by the dermatopathologists indicated that 1355 lesions were benign. Of 224 malignant lesions, 176 were KCs (90 BCC and 86 SCC) and 48 were melanocytic high-risk lesions (19 severely atypical nevi/atypical

**Table 1.** Patient (n = 1005) and Lesion (n = 1579) Characteristics.

Characteristic	Value <sup>a</sup>
<b>Patient characteristics</b>	
Sex	
Men	487 (48.5)
Women	518 (51.5)
Age, y	58.5 (15.1)
Race	
Asian	9 (0.9)
Black or African American	7 (0.7)
Native Hawaiian or Other Pacific Islander	3 (0.3)
White	976 (97.1)
Other/multiracial	10 (1.0)
Fitzpatrick skin type	
I (always burns, never tans)	99 (9.9)
II (always burns, tans minimally)	278 (27.7)
III (sometimes mild burn, tans uniformly)	352 (35.0)
IV (burns minimally, always tans well)	148 (14.7)
V (very rarely burns, tans very easily)	110 (10.9)
VI (never burns)	18 (1.8)
Risk factors	
Weakened immune system	32 (3.2)
Family history of skin cancer	332 (33.0)
Fair skin, freckling, light hair	362 (36.0)
Many moles and/or dysplastic nevi	331 (32.9)
Number of eligible lesions	
1	657 (65.4)
2	207 (20.6)
3	78 (7.8)
4	41 (4.1)
5	22 (2.2)
<b>Lesion characteristics (n=1,579)</b>	
Anatomic location	
Head	247 (15.6)
Arm	299 (18.9)
Leg	207 (13.1)
Trunk	826 (52.3)
Length, mm	6.7 (2.9)
Width, mm	5.2 (2.2)
Elevation	
Flat	622 (39.4)
Elevated	957 (60.6)
Surface texture	
Smooth	864 (54.7)
Rough	715 (45.3)
Pigmentation	
Lighter	581 (36.8)
Darker	998 (63.2)
Discovery	
Patient	1,070 (67.8)
Family member/partner	79 (5.0)
Health care clinician	430 (27.2)

<sup>a</sup>Patient age, lesion length, and lesion width summarized as mean (SD), and all other variables summarized as No. (%) of patients or lesions.

**Table 2.** Dermatopathologic Risk Classifications and Diagnoses (n = 1579).

Dermatopathologic classification or diagnosis	No. (%)
Malignant lesions	224 (14.2)
Melanoma <sup>a</sup>	48 (3.0)
Basal cell carcinoma	90 (5.7)
Squamous cell carcinoma	86 (5.4)
Benign lesions	1,355 (85.8)
Benign melanocytic nevus	500 (31.7)
Seborrheic keratosis	490 (31.0)
Actinic keratosis	71 (4.5)
Lentigo	65 (4.1)
Other	229 (14.5)

<sup>a</sup>Includes 29 melanomas (invasive and in situ) and 19 highly atypical lesions (atypical junctional melanocytic proliferations and severely atypical nevi).

junctional melanocytic proliferations and 29 invasive and in situ melanomas) (Table 2). Of the 48 melanocytic lesions, 7 were on patients younger than 40 years. No KCs were identified for patients younger than 40 years. Severely atypical nevi, atypical junctional melanocytic, and melanoma in situ lesions were grouped in the analysis as high-risk melanocytic lesions with melanomas by FDA requirement to ensure referral to dermatology due to the close histopathological relationship and lack of reproducibility of diagnostic differentiation between them on histopathology.<sup>41,42</sup>

The level of agreement between the 2 core laboratory dermatopathologists was 81.3% (Cohen  $\kappa$ , -0.1; 95% CI, -0.2 to 0.0); they agreed that 39 of the 48 melanomas were melanoma (or its near-equivalent, e.g., severe atypia). One diagnosed 44 of 48 (91.7%) melanomas; the other diagnosed 43 of 48 (89.6%) melanomas. For patients 40 years and older, the sensitivity of the 2 dermatopathologists was 90.2% (37 of 41 melanomas). Thus, the device sensitivity of 90.2% for melanoma was identical to that of both dermatopathologists (i.e., all 3 correctly diagnosed 37 of 41 melanomas). Five lesions excluded from primary analysis because of a lack of consensus had various diagnoses on the spectrum from compound melanocytic nevi to severely atypical nevi. Exclusion of these lesions was documented before study database lock and unblinding of the study results.

Overall sensitivity of the ESS device for detecting skin cancer was 95.5% (95% CI, 91.7%-97.6%) (Table 3). For the FDA-approved target group of patients aged 40 years and older, the device sensitivity for skin cancer was 96.3% (95% CI, 92.9%-98.4%), including 90.2% (95% CI, 76.9%-97.3%) for melanoma, 97.8% (92.2%-99.7%) for BCC, and 97.7% (91.9%-99.7%) for SCC. For all ages,

**Table 3.** Sensitivity of Device Using Dermatopathologic Diagnoses as Standard

Diagnosis	Sensitivity, % (95% CI) <sup>a</sup>
All skin cancers	95.5 (91.7-97.6)
All skin cancers (patient level) <sup>b</sup>	97.0 (93.6-98.9)
All skin cancers (patients $\geq$ 40 years)	96.3
Melanoma <sup>c</sup>	87.5 (76.4-93.8)
Melanoma (patient level) <sup>b</sup>	93.6 (82.5-98.7)
Melanoma (patients $\geq$ 40 years)	90.2 (76.9-97.3)
Melanoma (patients $\geq$ 40 years, patient level) <sup>b</sup>	95.1 (83.5-99.4)
BCC <sup>d</sup>	97.8 (91.3-99.5)
BCC (patient level) <sup>b</sup>	97.4 (90.9-99.7)
SCC <sup>d</sup>	97.7 (91.1-99.4)
SCC (patient level) <sup>b</sup>	98.7 (93.1-100.0)

Abbreviations: BCC, basal cell carcinoma; SCC, squamous cell carcinoma.

<sup>a</sup>Values for 95% CI calculated accounting for within-subject correlations by using the Wilson method described by Saha et al.<sup>43</sup>

<sup>b</sup>Patients with at least 1 positive device result were coded as device positive at the patient level.

<sup>c</sup>For the device's indicated use for patients aged 40 years and older, the device sensitivity was 90.2%.

<sup>d</sup>For BCC and SCC, no malignant lesions were recorded for patients younger than 40 years; therefore, overall and patient-level sensitivity were unchanged for the indicated use population.

device sensitivity was 87.5% (95% CI, 76.4%-93.8%) for melanoma. Patient-level sensitivity, defined as the proportion of participants with skin cancer for whom the device correctly detected at least 1 malignant lesion was 97.4% (95% CI, 94.2%-99.2%) for all skin cancers and 95.1% (95% CI, 83.5%-99.4%) for melanoma for patients aged 40 years and older (see Table 3). For patients aged 40 years and older, patient-level sensitivity was 97.4% (95% CI, 94.2%-99.2%) for all skin cancers and 95.1% (95% CI, 83.5%-99.4%) for melanoma for patients aged 40 years and older. When compared with the performance goal of 90%, device sensitivity was both noninferior ( $P < .001$ ) and superior ( $P = .002$ ). The device also had significantly higher odds (odds ratio, 4.93; 95% CI, 2.84-8.56;  $P < .001$ ) of detecting malignant lesions than random chance.

*Monitor* (i.e., negative) device results did not occur for any melanomas that exceeded 0.7 mm in depth (i.e., stage IA, T1a). The device misclassified 1 severely atypical nevus, 1 melanoma in situ, and 4 melanomas with 0.7 mm or less thickness although higher spectral scores provided by the device directly correlated with increased malignancy nearly reaching an *investigate further* (i.e., positive) result (See Table 4). Overall specificity of the ESS device was 20.7% (95% CI, 18.5%-23.1%) for correctly classifying biopsied lesions as benign, and 20.3% (95% CI, 18.0%-22.7%) for patients aged 40 and above.

**Table 4.** Clinical Characteristics and Histopathologic Findings for Missed Melanomas.

Clinical characteristics	Device result	Algorithm output <sup>a</sup>	Histopathologic findings	Body location
Flat, smooth, dark, 4 × 4 mm	Monitor	0.305	Lentiginous compound melanocytic nevus with highly atypical characteristics	Upper back
Elevated, smooth, dark, 5 × 4 mm	Monitor	0.418	Melanoma in situ	Upper back
Flat, smooth, dark, 10 × 6 mm	Monitor	0.392	0.4-mm thickness melanoma	Forearm
Flat, smooth, dark, 4 × 4 mm	Monitor	0.412	0.4-mm thickness melanoma	Shoulder
Flat, smooth, dark, 10 × 8 mm	Monitor	0.442	0.6-mm thickness melanoma	Face, cheek
Elevated, rough, dark, 7 × 7 mm	Monitor	0.449	0.7-mm thickness melanoma	Upper leg

<sup>a</sup>The device algorithm internally generates a numeric value based on a spectral score between 0 and 1, which is then converted to the device output of either a negative result of *monitor* (i.e., 0) or positive result of *investigate further* with a corresponding confidence score (1-10). The algorithm cutpoint for a *investigate further* result is a numeric value of 0.450; thus, 4 melanomas were within 0.032 of being a true positive result, and the highly atypical lesion was the only lesion missed by a large margin.

Although the dataset included fewer patients with Fitzpatrick skin type IV through VI (27.5%), no meaningful differences in sensitivity and specificity were evident between patients with Fitzpatrick skin type I through III (sensitivity, 96.5%; 95% CI, 92.6%-98.7%; specificity, 18.7%; 95% CI, 16.2%-21.5%) and IV through VI (sensitivity, 92.2%; 95% CI, 81.1%-97.8%; specificity, 25.1%; 95% CI, 20.9%-29.7%).

The AUROC curve for the device was calculated as 0.7796 which would be classified as *very good* for diagnostic devices.<sup>44</sup>

The overall NPV of the device was 96.6% (95% CI, 93.5%-98.2%), which indicates that a lesion with a device result of *monitor* had a 3.4% chance of being malignant. Overall PPV of the device for skin cancer detection was 16.6% (95% CI, 14.2%-19.3%), which indicates that a lesion with a device result of *investigate further* had a 16.6% chance of being malignant (i.e., NNB of 6.0). For *investigate further* results, device PPV increased with increasing spectral score. Of further results, 37% of lesions had a spectral score of 1-3, which correlated with a 5.9% probability of malignancy, increasing to a PPV of 49.6% PPV with spectral scores of 9 to 10 (See Figure 3).

## Discussion

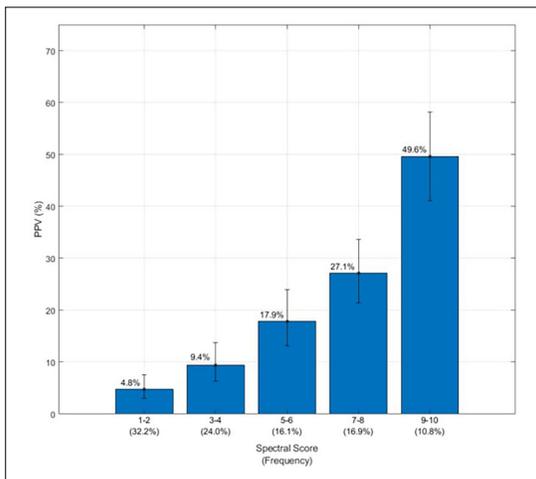
The DermaSensor device's 96.3% sensitivity for detecting skin cancer and 96.6% NPV classifies it a very good diagnostic device for risk stratifying suspicious skin lesions. The device was approved by the FDA in January 2024 based on this data for skin cancer risk assessment for patients 40 years and older, making it the first such device available for use by PCPs.<sup>23</sup>

Since spectral scores are directly related to increasing malignancy risk, the device may find good utility for prioritizing timing of further care (e.g., urgency of referral to dermatology). The NNB for all skin cancers in this data set skewed towards a high percentage of melanocytic lesions

was 7.1 for PCPs and 6.0 for the device. If PCPs had been unblinded and device output used to inform biopsy decisions, they would have decreased their biopsy frequency in the approved 40 years and older population by 18.4% (1288 vs 1579 biopsies) and would have monitored rather than investigated further 3.7% of cancers (4 KCs and 4 melanomas or 0.5% of lesions). For the 37% of *investigate further* lesions with a low spectral score of 1 to 3 which represented a 5.9% likelihood of malignancy, PCPs may have further decreased unnecessary care by monitoring lesions that were also of low clinical concern (e.g., lesions not suggestive of melanoma and/or patients with limited risk factors).

The device correctly classified 90.2% of melanomas as *investigate further* for patients 40 years and older (87.5% of melanomas overall)—an accuracy very similar to the individual dermatopathologists when compared to dermatopathology consensus. Discordance for melanoma diagnoses between dermatopathologists is common as observed in a trial of 187 US dermatopathologists randomly assigned to interpret the same lesions months apart where intraobserver and interobserver concordance was under 50% for diagnoses ranging from moderately dysplastic nevi to early-invasive melanoma, though higher for lesions at the ends of the spectrum—benign nevi and invasive melanoma.<sup>42</sup>

Patient-level device sensitivity was 95.1% for the device to read *investigate further* for at least 1 lesion in any given patient with melanoma. Two recently published studies of the same device and algorithm in dermatology practices reported similar high device sensitivity for both KC and melanoma and a NNB congruent with that of dermatologists.<sup>45,46</sup> Another recently published small study using the same device and algorithm on lesions of concern to patients in 1 primary care practice found the device sensitivity to be 90%, but a specificity for correctly classifying benign lesions of concern to patients of 60.7% overall and 76.9% for pigmented benign lesions.<sup>47</sup> This study will need to be repeated in additional practices; but the device may be able to decrease unnecessary referrals for biopsy and prioritize referrals.



**Figure 3.** Positive Predictive Value (PPV) According to Spectral Score for Investigate Further Results. Device PPV ranged from 6.4% (95% CI, 3.7%-10.9%) for the lowest spectral score of 1 to 61.3% (95% CI, 48.3%-72.9%) for the highest score of 10. The device PPV was 7.4% (95% CI, 5.5%-9.9%; NNB, 13.5) for low (1-5) spectral scores and 33.2% (95% CI, 28.8%-37.9%; NNB, 3.0) for high (6-10) scores ( $P < .001$ ). When spectral scores were further categorized as low (1-3), medium (4-7), or high (8-10), the PPV was 5.9% (95% CI, 4.1%-8.5%; NNB, 16.9) for low scores, 18.4% (95% CI, 14.8%-22.7%; NNB, 5.4) for medium scores, and 39.6% (95% CI, 33.4%-46.2%; NNB, 2.5) for high scores ( $P < .001$ ).

By analyzing ESS data from suspicious lesions, the device provides immediate risk information to the PCP to inform lesion management. This may improve referrals by PCPs and enhance referral adherence by patients. Pigmented lesions of concern for melanoma with a *monitor* result should still be monitored for changes, as should all concerning skin lesions. Future studies will help clarify re-examination frequency and lesion selection criteria for device use and device impact on unnecessary referral rates from PCP practices.

### Limitations

Our study was limited by the exclusion criteria (e.g., fully crusted lesions which may have excluded some squamous cell cancers and very small and large lesions) and included only lesions identified as concerning and warranting biopsy by PCPs who frequently perform biopsies for skin cancer and who were blinded to device results. Enrollment of only lesions suspicious for melanoma after achieving the target sample size for BCC and SCC resulted in a high proportion of melanocytic lesions

(about 2/3) making overall comparisons of NNB to an average PCP practice inaccurate. Other limitations of our study included a patient population that was only 12.7% Fitzpatrick V-VI.

### Conclusions

The DermaSensor device is an easy-to-use, point-of-care, hand-held, AI-developed, algorithm-powered skin cancer adjunctive diagnostic device with high sensitivity and NPV for use in the primary care setting. PCPs can use the device to prioritize skin lesions that merit further evaluation and those that may be monitored.

### Authors' Note

#### Presentations

Subsets of the data were presented as posters at (1) the 2nd Annual Innovations in Dermatology Fall Conference; November 3-5, 2022; Las Vegas, Nevada; (2) Maui Derm Hawaii 2023; January 23-27, 2023; Maui, Hawaii; presented as (3) a podium presentation at the 11th Annual American Dermoscopy Meeting; July 13-15, 2023; Stowe, Vermont; (4) as a podium presentation at the American Academy of Dermatology Innovation Academy Meeting; August 10-13, 2023; Tampa, Florida; and (5) the Skin of Color Update conference; October 6-8, 2023; New York, New York. Another subset of the data was presented as a podium presentation at the Society of Teachers of Family Medicine (STFM) Annual Spring Conference; April 29-May 3, 2023; Tampa, Florida.

Portions of this manuscript have been published in abstract form: *Cutis*. 2022;110(6 Suppl):31; *J Clin Aesthet Dermatol*. 2023;16(4 Suppl 1):S16; and *J Am Acad Dermatol*. 2023;89(3 Suppl):AB233.

### Acknowledgments

We commend Joseph Massaro, PhD (deceased), for writing the initial study statistical plan and for his contributions to the study design. We acknowledge Irving Bigio, PhD, Eladio Rodriguez-Diaz, PhD, and Preston Hoang, MS, for their contributions developing the device technology. We thank Na Wang, PhD, Margaret Shea, Kanisha Mittal, Lindsey Furton, Kaitlin Hartlage, Cason Hucks, and Matthew Bullard for their contributions to the statistical analyses. We also thank Phillip Leboit, MD, Maxwell Fung, MD, and Jennifer McNiff, MD, for serving as expert dermatopathology panelists for this study. All persons acknowledged above or their institutions were compensated for their contributions to the study. We especially thank all study participants and investigators; without their participation, this study would not have been possible.

Cody Simmons, MS (cofounder and chief executive officer of DermaSensor, Inc), and Kiran Chatha, MD, MPH (DermaSensor, Inc employee), reviewed the manuscript and provided data verification and validation support. Nisha Badders, PhD, ELS, Mayo Clinic, provided editorial suggestions on an earlier draft of the manuscript.

**ORCID iDs**

Stephen P. Merry  <https://orcid.org/0000-0002-3816-8154>  
 Ivana T. Croghan  <https://orcid.org/0000-0003-3464-3525>  
 Kimberly A. Dukes  <https://orcid.org/0000-0002-5135-8809>  
 Brian C. McCormick  <https://orcid.org/0009-0008-6009-4471>  
 Gerard T. Considine  <https://orcid.org/0000-0002-6313-0721>  
 Curtis T. Thompson  <https://orcid.org/0000-0003-0788-346X>  
 David J. Leffell  <https://orcid.org/0009-0008-8766-3793>

**Ethics and Considerations**

This study was reviewed and approved by the Western Institutional Review Board-Copernicus Group Review Board (#20182730), the Australian Bellberry Human Research Ethics Committee (#2018-08-620), and the Mayo Clinic Institutional Review Board (#20-011463).

**Consent to Participate**

Before enrollment, written informed consent was obtained for all participants. Study participants received a small stipend.

**Author Contributions**

Concept and design: Leffell  
 Acquisition of data: Merry, Croghan, McCormick, Considine, Thompson  
 Analysis and interpretation of data: Dukes, Merry, Croghan, Leffell  
 Drafting of the manuscript: Merry, Croghan  
 Critical revision of manuscript: Duvall, McCormick, Considine, Thompson, Leffell, Dukes  
 Statistical analysis: Dukes  
 Obtained funding: The institutions of Dukes, Merry, Croghan, Duvall, McCormick, Considine, and Thompson received funding for conducting the clinical trial  
 Administrative, technical, and material support: Thompson  
 Study supervision: Leffell  
 No use of artificial intelligence-assisted technologies were used in the production of this manuscript.

**Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was funded by DermaSensor, Inc. The sponsor designed the study and hired a contract research organization to oversee the study conduct and data collection and management. Data analysis was completed by an external biostatistics group at Boston University. The sponsor did not prepare the manuscript but did review it.

**Declaration of Conflicting Interests**

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Drs Merry and Croghan conducted the study at Mayo Clinic in Rochester, Minnesota, with the support of the sponsor DermaSensor, Inc. Dr Dukes performed statistical analysis under contract from the sponsor at Boston University School of Public

Health. Drs McCormick and Considine were site principal investigators at their respective family medicine private practices in Virginia and Australia. Dr Thompson performed dermatopathologic analyses under contract. None have received personal remuneration from the company. Dr Leffell is a member of the DermaSensor Scientific Advisory Board and is compensated by the company.

**Abbreviations**

AI=artificial intelligence  
 AUROC=area under the receiver operating characteristic  
 APP=advanced practice provider  
 BCC=basal cell carcinoma  
 DERM-SUCCESS=**DERM**aSensor **S**tudy of Primary **C**are Physician use of **E**lastic **S**cattering **S**pectroscopy on skin lesions suggestive of skin cancer  
 ESS=elastic scattering spectroscopy  
 FDA=US Food and Drug Administration  
 KC=keratinocyte carcinoma  
 NNB=number needed to biopsy  
 NPV=negative predictive value  
 PCP=primary care physician  
 PPV=positive predictive value  
 SCC=squamous cell carcinoma

**Trial Registry Information**

Trial registration: NCT06690086

**Data Availability Statement**

All data supporting the study findings are contained in this manuscript. Kimberly A. Dukes, PhD, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Stephen P. Merry, MD, MPH, takes responsibility for the integrity of the work as a whole, from inception to published article.

**References**

1. American Cancer Society. Cancer Facts & Figures 2022. American Cancer Society. Accessed October 11, 2022. <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2022/2022-cancer-facts-and-figures.pdf>
2. Rogers HW, Weinstock MA, Feldman SR, Coldiron BM. Incidence Estimate of Nonmelanoma Skin Cancer (Keratinocyte Carcinomas) in the U.S. Population, 2012. *JAMA Dermatol.* 2015;151(10):1081-1086. (In eng). doi:10.1001/jamadermatol.2015.1187
3. Muzic JG, Schmitt AR, Wright AC, et al. Incidence and trends of basal cell carcinoma and cutaneous squamous cell carcinoma: a population-based study in Olmsted County, Minnesota, 2000 to 2010. *Mayo Clin Proc.* 2017;92(6):890-898. (In eng). doi:10.1016/j.mayocp.2017.02.015
4. Mansouri B, Housewright CD. The treatment of actinic keratoses-the rule rather than the exception. *JAMA Dermatol.* 2017;153(11):1200. (In eng). doi:10.1001/jamadermatol.2017.3395

5. Greenwood JD, Merry SP, Boswell CL. Skin biopsy techniques. *Prim Care*. 2022;49(1):1-22. doi:10.1016/j.pop.2021.10.001
6. Marchetti MA, Yu A, Nanda J, et al. Number needed to biopsy ratio and diagnostic accuracy for melanoma detection. *J Am Acad Dermatol*. 2020;83(3):780-787. doi:10.1016/j.jaad.2020.04.109
7. Nelson KC, Swetter SM, Saboda K, Chen SC, Curiel-Lewandrowski C. Evaluation of the number-needed-to-biopsy metric for the diagnosis of cutaneous melanoma: a systematic review and meta-analysis. *JAMA Dermatol*. 2019;155(10):1167-1174. (In eng). doi:10.1001/jamadermatol.2019.1514
8. Privalle A, Havighurst T, Kim K, Bennett DD, Xu YG. Number of skin biopsies needed per malignancy: comparing the use of skin biopsies among dermatologists and nondermatologist clinicians. *J Am Acad Dermatol*. 2020;82(1):110-116. doi:10.1016/j.jaad.2019.08.012
9. Chen SC, Bravata DM, Weil E, Olkin I. A comparison of dermatologists' and primary care physicians' accuracy in diagnosing melanoma: a systematic review. *Arch Dermatol*. 2001;137(12):1627-134. DOI: 10.1001/archderm.137.12.1627.
10. Argenziano G, Puig S, Zalaudek I, et al. Dermoscopy improves accuracy of primary care physicians to triage lesions suggestive of skin cancer. *J Clin Oncol*. 2006;24(12):1877-1882. doi:10.1200/JCO.2005.05.0864
11. Shahwan KT, Kimball AB. Should we leave the skin biopsies to the dermatologists? *JAMA Dermatol*. 2016;152(4):371-372. doi:10.1001/jamadermatol.2015.5051
12. Rogers T, Marino ML, Dusza SW, et al. A clinical aid for detecting skin cancer: the triage amalgamated dermoscopic algorithm (TADA). *J Am Board Fam Med*. 2016;29(6):694-701. doi:10.3122/jabfm.2016.06.160079
13. Dinnes J, Deeks JJ, Chuchu N, et al. Dermoscopy, with and without visual inspection, for diagnosing melanoma in adults. *Cochrane Database Syst Rev*. 2018;12(12):Cd011902. (In eng). doi:10.1002/14651858.CD011902.pub2
14. Seiverling EV, Ahms HT, Greene A, et al. Teaching benign skin lesions as a strategy to improve the triage amalgamated dermoscopic algorithm (TADA). *J Am Board Fam Med*. 2019;32(1):96-102. doi:10.3122/jabfm.2019.01.180049
15. Sawyers EA, Wigle DT, Marghoob AA, Blum A. Dermoscopy training effect on diagnostic accuracy of skin lesions in canadian family medicine physicians using the triage amalgamated dermoscopic algorithm. *Dermatol Pract Concept*. 2020;10(2):e2020035. doi:10.5826/dpc.1002a35
16. Westerhoff K, McCarthy WH, Menzies SW. Increase in the sensitivity for melanoma diagnosis by primary care physicians using skin surface microscopy. *Br J Dermatol*. 2000;143(5):1016-1020. doi:10.1046/j.1365-2133.2000.03836.x
17. Menzies SW, Emery J, Staples M, et al. Impact of dermoscopy and short-term sequential digital dermoscopy imaging for the management of pigmented lesions in primary care: a sequential intervention trial. *Br J Dermatol*. 2009;161(6):1270-1277. doi:10.1111/j.1365-2133.2009.09374.x
18. Morris JB, Alfonso SV, Hernandez N, Fernandez MI. Examining the factors associated with past and present dermoscopy use among family physicians. *Dermatol Pract Concept*. 2017;7(4):63-70. doi:10.5826/dpc.0704a13
19. Fee JA, McGrady FP, Rosendahl C, Hart ND. Dermoscopy use in primary care: a scoping review. *Dermatol Pract Concept*. 2019;9(2):98-104. (In eng). doi:10.5826/dpc.0902a04
20. Brown AE, Najmi M, Duke T, Grabell DA, Koshelev MV, Nelson KC. Skin cancer education interventions for primary care providers: a scoping review. *J Gen Intern Med*. 2022;37(9):2267-2279. doi:10.1007/s11606-022-07501-9
21. Malvey J, Hauschild A, Curiel-Lewandrowski C, et al. Clinical performance of the Nevisense system in cutaneous melanoma detection: an international, multicentre, prospective and blinded clinical trial on efficacy and safety. *Br J Dermatol*. 2014;171(5):1099-1107. (In eng). doi:10.1111/bjd.13121
22. Hogarty DT, Su JC, Phan K, et al. Artificial intelligence in dermatology-where we are and the way to the future: a review. *Am J Clin Dermatol*. 2020;21(1):41-47. (In eng). doi:10.1007/s40257-019-00462-6
23. Beltrami EJ, Brown AC, Salmon PJM, Leffell DJ, Ko JM, Grant-Kels JM. Artificial intelligence in the detection of skin cancer. *J Am Acad Dermatol*. 2022;87(6):1336-1342. (In eng). doi:10.1016/j.jaad.2022.08.028
24. Sangers TE, Kittler H, Blum A, et al. Position statement of the EADV Artificial Intelligence (AI) Task Force on AI-assisted smartphone apps and web-based services for skin disease. *J Eur Acad Dermatol Venereol*. 2024;38(1):22-30. (In eng). doi:10.1111/jdv.19521
25. Stafford H, Buell J, Chiang E, et al. Non-melanoma skin cancer detection in the age of advanced technology: a review. *Cancers (Basel)*. 2023;15(12):3094. (In eng). doi:10.3390/cancers15123094
26. Valova I, Dinh P, Gueorguieva N. Mobile application for skin cancer classification using deep learning. *J Mach Intell Data Sci*. 2023;4:15-26. doi:10.11159/jmids.2023.003
27. Brenes H, Silva F, Aladab A, et al. The role of smartphone-based applications and artificial intelligence aiding in screening, faster diagnosis and prevention in skin cancer: a systematic mini-review. *Principles Pract Clin Res*. 2023;9(3):12-17. doi:10.21801/ppcrj.2023.93.4
28. Sun MD, Kentley J, Mehta P, Dusza S, Halpern AC, Rotemberg V. Accuracy of commercially available smartphone applications for the detection of melanoma. *Br J Dermatol*. 2022;186(4):744-746. (In eng). doi:10.1111/bjd.20903
29. Jahn AS, Navarini AA, Cerminara SE, et al. Over-detection of melanoma-suspect lesions by a CE-certified smartphone app: performance in comparison to dermatologists, 2D and 3D convolutional neural networks in a prospective data set of 1204 pigmented skin lesions involving patients' perception. *Cancers (Basel)*. 2022;14(15):3829. (In eng). doi:10.3390/cancers14153829
30. Smak Gregoor AM, Sangers TE, Bakker LJ, et al. An artificial intelligence based app for skin cancer detection evaluated in a population based setting. *NPJ Digit Med*. 2023;6(1):90. (In eng). doi:10.1038/s41746-023-00831-w
31. Rodriguez-Diaz E, Manolagos D, Christman H, et al. Optical spectroscopy as a method for skin cancer risk assessment. *Photochem Photobiol*. 2019;95(6):1441-1445. (In eng). doi:10.1111/php.13140

32. Bigio IJ, Bown SG, Briggs G, et al. Diagnosis of breast cancer using elastic-scattering spectroscopy: preliminary clinical results. *J Biomed Opt.* 2000;5(2):221-228. (In eng). doi:10.1117/1.429990
33. Keshtgar MR, Chicken DW, Austwick MR, et al. Optical scanning for rapid intraoperative diagnosis of sentinel node metastases in breast cancer. *Br J Surg.* 2010;97(8):1232-1239. (In eng). doi:10.1002/bjs.7095
34. Lovat LB, Johnson K, Mackenzie GD, et al. Elastic scattering spectroscopy accurately detects high grade dysplasia and cancer in Barrett's oesophagus. *Gut.* 2006;55(8):1078-1083. (In eng). doi:10.1136/gut.2005.081497
35. Mourant JR, Bigio IJ, Boyer J, Conn RL, Johnson T, Shimada T. Spectroscopic diagnosis of bladder cancer with elastic light scattering. *Lasers Surg Med.* 1995;17(4):350-357. (In eng). doi:10.1002/lsm.1900170403
36. Suh H, A' Amar O, Rodriguez-Diaz E, Lee S, Bigio I, Rosen JE. Elastic light-scattering spectroscopy for discrimination of benign from malignant disease in thyroid nodules. *Ann Surg Oncol.* 2011;18(5):1300-1305. (In eng). doi:10.1245/s10434-010-1452-y
37. Stanganelli I, Serafini M, Bucchi L. A cancer-registry-assisted evaluation of the accuracy of digital epiluminescence microscopy associated with clinical examination of pigmented skin lesions. *Dermatology.* 2000;200(1):11-16. doi:10.1159/000018308
38. Carli P, Nardini P, Crocetti E, De Giorgi V, Giannotti B. Frequency and characteristics of melanomas missed at a pigmented lesion clinic: a registry-based study. *Melanoma Res.* 2004;14(5):403-407. doi:10.1097/00008390-200410000-00011
39. Soyer HP, Argenziano G, Zalaudek I, et al. Three-point checklist of dermoscopy. A new screening method for early detection of melanoma. *Dermatology.* 2004;208(1):27-31. doi:10.1159/000075042
40. Dinnes J, Deeks JJ, Grainge MJ, et al. Visual inspection for diagnosing cutaneous melanoma in adults. *Cochrane Database Syst Rev.* 2018;12(12):CD013194. doi:10.1002/14651858.CD013194
41. Carney PA, Reisch LM, Piepkorn MW, et al. Achieving consensus for the histopathologic diagnosis of melanocytic lesions: use of the modified Delphi method. *J Cutaneous Pathol* 2016;43(10):830-837.
42. Elmore JG, Barnhill RL, Elder DE, et al. Pathologists' diagnosis of invasive melanoma and melanocytic proliferations: observer accuracy and reproducibility study. *BMJ* 2017;357:j2813. doi:10.1136/bmj.j2813
43. Saha KK, Miller D, Wang S. A comparison of some approximate confidence intervals for a single proportion for clustered binary outcome data. *Int J Biostat.* 2016;12(2):1-18. (In eng). doi:10.1515/ijb-2015-0024
44. Simundic AM. Measures of diagnostic accuracy: basic definitions. *EJIFCC* 2009;19(4):203-211. <https://www.ncbi.nlm.nih.gov/pubmed/27683318>
45. Hartman RI, Trepanowski N, Chang MS, et al. Multicenter prospective blinded melanoma detection study with a handheld elastic scattering spectroscopy device. *JAAD Int.* 2024;15:24-31. (In eng). doi:10.1016/j.jdin.2023.10.011
46. Manolagos D, Patrick G, Geisse JK, et al. Use of an elastic-scattering spectroscopy and artificial intelligence device in the assessment of lesions suggestive of skin cancer: a comparative effectiveness study. *JAAD Int.* 2024;14:52-58. (In eng). doi:10.1016/j.jdin.2023.08.019
47. Tepedino M, Baltazar D, Hanna K, Bridges A, Billot L, Zeitouni NC. Elastic scattering spectroscopy on patient-selected lesions concerning for skin cancer. *J Am Board Fam Med.* 2024;37(3):427-435.

## Motavaf et al. (2025)

ARTICLE IN PRESS

HEALTH POLICY &amp; PRACTICE

## Current landscape of AI Regulation in Dermatology

Mateen Motavaf, BS,<sup>a</sup> Danyal Raza, MS,<sup>a</sup> William Frye, BS,<sup>a</sup> and Melodi Javid Whitley, MD, PhD<sup>a,b</sup>

**Key words:** 510(k); AI; Artificial Intelligence; Deep Learning; DermaSensor; FDA; Machine Learning; MelaFind; ML; Regulation.

The rise of artificial intelligence (AI) in medicine during the last decade has presented transformative enhancements to patient care and clinical workflows. Dermatology, a highly visual field relying on skin morphology, texture, and coloration to diagnose diseases, holds promise for implementing AI diagnostics.<sup>1</sup>

Despite the rapid pace of innovation, regulatory oversight mechanisms remain outdated. Many products reach patients without regulatory approval, threatening patient safety. Addressing gaps in existing post- and premarket regulatory frameworks is necessary for adoption.<sup>2</sup>

The Food and Drug Administration (FDA) currently classifies AI-augmented products under 3 primary pathways: the 510(k) premarket notification, the de novo classification, and Premarket Approval (PMA) (Fig 1). Given the dynamic nature and volume of AI products, the same level of scrutiny cannot be broadly applied, and a risk-based framework may allow for efficient resource allocation.

Direct-to-consumer AI dermatology applications additionally present safety concerns. A cross-sectional study involving 41 dermatology-focused mobile applications found that only 15% had associated supporting peer-reviewed publications,<sup>3</sup> among which only 1 was subject to a multicenter diagnostic clinical trial. These products may worsen patient outcomes by providing false reassurance and delaying timely consultation with board-certified dermatologists.

One product, DermaSensor, obtained de novo authorization following the DERM-SUCCESS clinical

### Abbreviations used:

AI: artificial intelligence  
FDA: Food and Drug Administration  
PMA: Premarket Approval

trial with >1000 participants, demonstrating 96% sensitivity in skin cancer detection,<sup>4</sup> exemplifying the necessary rigor when regulating AI products. Rather than trying to fit products into specific FDA pathways, real-world validation must be prioritized as it determines clinical utility and success. Pre-launch testing standards should be representative of intended clinical use and include peer-reviewed publications and clinical trials.

Unlike traditional diagnostic algorithms and devices, AI/ML models are not static and are constantly evolving through retraining. However, FDA regulations ignore post-deployment model drift, leaving tools vulnerable to degradation in clinical settings, particularly in cases requiring differentiation between rare melanoma variants or inflammatory diseases. Dynamic oversight is critical even in cases where a model is static, which can reveal unknown biases and catch deterioration that can occur when the model is used in changing clinical contexts, a process known as concept drift.<sup>5</sup> Without careful oversight, AI tools risk reinforcing existing disparities and deteriorate patient care.<sup>5</sup>

Moving forward, regulators must adopt a dynamic oversight framework by mandating the use of “sandboxes” (constructed clinical environments) in the

From the Duke University School of Medicine, Durham, North Carolina<sup>a</sup>; and Department of Dermatology, Duke University School of Medicine, Durham, North Carolina.<sup>b</sup>

Funding sources: None.

Patient consent: Not applicable.

IRB approval status: Not applicable.

Accepted for publication November 12, 2025.

Correspondence to: Melodi Javid Whitley, MD, PhD, Department of Dermatology, Duke University School of Medicine, 40 Medicine Circle, Durham, NC, 27710. E-mail: [melodi.whitley@duke.edu](mailto:melodi.whitley@duke.edu).

Published online December 10, 2025

J Am Acad Dermatol 2025;■■■:■■■.

0190-9622/\$36.00

© 2025 Published by Elsevier Inc. on behalf of the American Academy of Dermatology, Inc.

<https://doi.org/10.1016/j.jaad.2025.11.035>

1

### FDA Regulatory Pathways for Class II and III Medical Devices

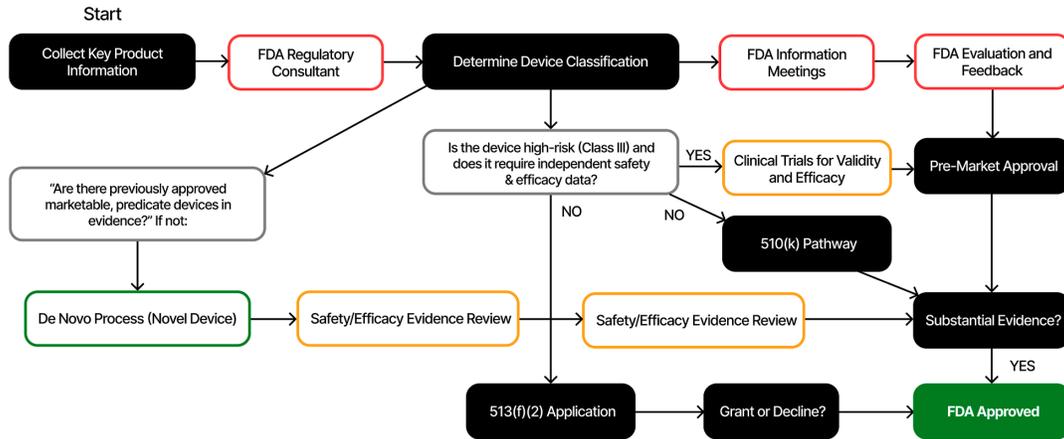


Fig 1. Overview of FDA regulatory pathways for class II and III medical devices in 2025.

premarket stage to ensure dynamic performance. Regulators must also enforce prospective, dermatologist-supervised trials that evaluate performance across features such as classification stratified by skin type, sensitivity (eg, melanoma vs SCC), and minimizing unwarranted biopsies. Additionally, detailed performance across Fitzpatrick skin types, Lesion type (eg, pigmented vs non-pigmented), clinical settings (primary care vs dermatology clinic vs telederm), and diagnostic tiering must be evaluated. Differentiating between benign edge cases (eg, seborrheic keratoses, dysplastic nevi) must also be considered. Trials also must address proprietary training datasets, which are a barrier to thorough independent assessment of bias and representativeness.

While artificial intelligence holds great potential for transforming dermatological care, regulatory frameworks must evolve to implement safeguards for equity, privacy, and accountability, as many eHealth apps fall outside of the Health Insurance

Portability and Accountability Act’s protections and the legal question of liability remains unresolved.

**Conflicts of interest**

None disclosed.

**REFERENCES**

1. Esteva A, Kuprel B, Novoa R, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542:115-118. <https://doi.org/10.1038/nature21056>
2. Warraich HJ, Tazbaz T, Califf RM. FDA perspective on the regulation of artificial intelligence in health care and biomedicine. *JAMA*. 2025;333(3):241-247. <https://doi.org/10.1001/jama.2024.21451>
3. Wongvibulsin S, Yan MJ, Pahalyants V, Murphy W, Daneshjou R, Rotemberg V. Current state of dermatology mobile applications with artificial intelligence features. *JAMA Dermatol*. 2024;160(6):646-650. <https://doi.org/10.1001/jamadermatol.2024.0468>
4. Venkatesh KP, Kadakia KT, Gilbert S. Learnings from the first AI-enabled skin cancer device for primary care authorized by FDA. *NPJ Digit Med*. 2024;7:156. <https://doi.org/10.1038/s41746-024-01161-1>
5. Cukras AR. MelaFind is approved by the FDA—Where does it fit in dermatology? *J Drugs Dermatol*. 2012;11(4):420-422.

## Seiverling et al. (2025)

## ORIGINAL RESEARCH

## Enhancing Diagnostic Precision in Primary Care: A Multireader Multicase (MRMC) Study of an AI-Powered Handheld Elastic Scattering Spectroscopy Device for Informed Referral Decisions in Melanoma Evaluation

by ELIZABETH V. SEIVERLING, MD; ASGHAR SHAH, BA; MARTIN A. WEINSTOCK, MD, PhD; JANE M. GRANT-KELS, MD; NATHAN P. FALK, MD, MBA; and DANIEL M. SIEGEL, MD

*Dr. Seiverling is with the Department of Dermatology at the University of Pittsburgh in Pittsburgh, Pennsylvania. Mr. Shah is with the Warren Alpert Medical School of Brown University in Providence, Rhode Island. Dr. Weinstock is with Brown University and the Providence VA Medical Center in Providence, Rhode Island. Dr. Grant-Kels is with the Department of Dermatology at the University of Connecticut School of Medicine in Farmington, Connecticut and the Department of Dermatology at the University of Florida College of Medicine in Gainesville, Florida. Dr. Falk is with the Department of Family Medicine and Rural Health at Florida State University College of Medicine in Tallahassee, Florida. Dr. Siegel is with SUNY Downstate Health Sciences University in Brooklyn, New York.*

*J Clin Aesthet Dermatol.* 2025;18(10):59–65.

**BACKGROUND:** Skin cancer is the most common cancer faced by adults in the United States. Melanoma, while a less common subtype of skin cancer compared to basal cell and squamous cell carcinomas, is associated with greater rates of metastases, mortality, and morbidity, and its rate of incidence is projected to increase. Primary care physicians (PCPs) can play an important role in skin cancer detection and in the decision to refer a patient to a dermatologist. Technologies such as the elastic scattering spectroscopy (ESS) device (DermaSensor, Inc.), a handheld, noninvasive assistive tool, may help in the evaluation of a skin growth and improve appropriate referral decision making. **METHODS:** A total of 50 malignant and 50 benign lesions were assessed by each of the 118 physicians (board-certified internal and family medicine physicians), yielding 5,900 malignant and benign lesion assessments without the device and 5,900 with the ESS device. Physicians were also surveyed regarding their confidence in their management decision. **RESULTS:** The study met the primary endpoint; the area under the receiver operating characteristic (AUROC) of the PCPs aided with the device was 0.671 (95% confidence interval [CI]: 0.611-0.732) compared with the AUROC unaided by the device of 0.630 (95% CI: 0.582-0.678), a significant increase ( $p=0.036$ ). When asked whether the device would provide value to their decision making, 91.5% of respondents either agreed or strongly agreed. **CONCLUSION:** The ESS device improved PCP accuracy in managing lesions suggestive of melanoma and increased their sensitivity for all skin cancers and melanoma. Participating internal medicine and family medicine physicians reported increased confidence in their assessments with the device. The ESS device can improve PCP decision making when managing lesions suggestive of melanoma. **KEYWORDS:** Primary care, elastic scattering spectroscopy, melanoma, skin cancer detection

Skin cancer is the most common type of cancer in the United States (US), and, compared to basal cell and squamous cell cancers, melanoma has the highest mortality rate.<sup>1</sup> Melanoma is projected to become the second most common cancer in the US by 2040 with 219,000 cases.<sup>2</sup> Early detection reduces the likelihood of disease-specific morbidity and mortality.<sup>3</sup> However, there are drastic shortages of dermatologists,<sup>4</sup> and patients often seek care for dermatologic conditions in the primary care setting.

While the American Academy of Family Physicians includes some skin cancer detection training as an important part of all family medicine resident training, access to training can be limited and many primary care physicians (PCPs) feel ill-equipped to evaluate skin growths.<sup>5</sup>

Recent advances have been made in artificial intelligence (AI) technology to improve melanoma detection and have the potential to aid PCPs in the evaluation of skin growths.<sup>6-8</sup> Efforts to implement AI

algorithms for melanoma detection have also been applied in primary cancer screening.<sup>9</sup> One such AI-based tool that has been investigated and cleared by the US Food and Drug Administration (FDA) is the elastic scattering spectroscopy (ESS) device (DermaSensor, Inc.), a handheld, noninvasive, and painless assistive tool for skin cancer detection. In a prospective multicenter study, the ESS device exhibited a sensitivity of 97.04% in detecting skin cancer lesions, including a sensitivity of 96.67% for melanoma, with sensitivity and accuracy (ie, area under the curve) found to be comparable to dermatologists' in-person performance, which holds promise for improving PCPs' ability to assess suspicious lesions and appropriately refer suspicious lesions to dermatologists.<sup>10</sup> Results of the DermaSensor Use in the Assessment of Skin Lesions Suggestive of Melanoma III (DERM-ASSESS III) study, a prospective blinded melanoma validation study, demonstrated similar findings for device sensitivity and accuracy compared to dermatologists, reporting that the ESS device had a

**FUNDING:** This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**DISCLOSURES:** Dr. Seiverling is a consultant for Melatech and was compensated by DermaSensor, Inc. for clinical research activities for this study. Drs. Grant-Kels and Siegel are advisors to DermaSensor. The remaining authors have no conflicts of interest to disclose.

**CORRESPONDENCE:** Asghar Shah, BA; Email: asghar\_shah@brown.edu

## ORIGINAL RESEARCH

melanoma sensitivity of 95.5% and a negative predictive value (NPV) of 98.1%, highlighting its potential benefit as a nonspecialist adjunctive tool for melanoma detection at the point of care.<sup>11</sup> Finally, results from studies involving PCPs found the device sensitivity to be 95.5% and NPV to be 96.6% across all skin cancers,<sup>12</sup> with a companion utility study showing that PCPs' use of the ESS device significantly improved diagnostic sensitivity (10.6%,  $p=0.0085$ ), management sensitivity (9.4%,  $p=0.003$ ), and physician confidence, indicating its potential to enhance PCP skin cancer diagnosis and confidence in management.<sup>13</sup> Moreover, results from a study of the ESS device suggest that device use may improve PCP sensitivity for skin cancer from 83.0% to 95.5% for high-risk lesions, and could rule out 20.7% of suspicious lesions from further evaluation.<sup>14</sup>

Given the importance of PCP decision making regarding lesions suspicious of skin cancer, particularly melanoma, it is postulated that use of the ESS device may assist PCPs in referral decision making and enhance provider confidence in managing these cases. Here, we describe the findings of a multireader multicase (MRMC) study to assess the referral performance of PCPs when evaluating lesions suggestive of melanoma with and without the aid of the ESS device. Since previous studies were not adequately powered or designed to demonstrate the device impact on melanoma management, this study was accomplished with intentional inclusion of more melanoma cases.

## METHODS

**ESS device.** This handheld, noninvasive tool uses ESS and machine learning (ML) to aid in evaluation of skin lesions. The device's algorithm has been trained on more than 10,000 recordings from more than 2,000 skin lesions to distinguish malignant from benign skin lesions, including histologically confirmed melanoma and keratinocyte carcinoma, as well as unbiopsied benign lesions, diagnosed by dermatologists. None of the ESS spectral recordings used in the training process were employed in the lesion testing set used in this study.

For each lesion, the device classifies the observed spectral pattern as either a positive result of having malignant characteristics ("investigate further") or a negative result of having benign characteristics ("monitor"). For lesions classified as "investigate further," a

spectral similarity score from 1 to 10 is reported, with higher values corresponding to the amount of spectral similarity a lesion has to malignant lesions in prior studies used to develop the algorithm.<sup>10-12</sup>

**Study design.** This melanoma-focused reader study was a web-based, MRMC investigation using clinical information, digital images, and ESS device result data. PCPs were asked to perform 200 reads for 100 skin lesion cases, each presented without and then with ESS device output.

The study examined two aspects of PCP decisions concerning skin evaluation using the device: (1) the impact of the device result on the PCPs' management decision on whether the lesion should be referred for further evaluation by a dermatologist, and (2) the device's impact on the PCP's diagnostic assessment of whether the lesion was malignant or benign. A study platform with traceability for each physician's responses was used to conduct this reader study. The study sponsor was blinded to the physician lesion case responses during study enrollment.

For every lesion case evaluation (both unaided and aided), physicians completed a questionnaire about their diagnosis of the lesion, their recommended management decision, and their confidence level in that decision. Physicians were instructed to maintain consistent evaluation criteria, based on their own clinical judgement, throughout the study, including for the 1 to 10 confidence assessment. A series of 10 questions were included at the end of the study to assess physician perceptions of the device and its benefits. The questions were asked as a select multiple response or with a 5-point Likert scale assessment ranging from "strongly agree" to "strongly disagree."

**Study physicians.** PCP reader eligibility was limited to currently practicing, board-certified PCPs (family medicine or internal medicine); they could not have participated in any of the clinical studies pertaining to this device from 2020 to 2022. Physicians with board certification in surgery or dermatology were excluded. A total of 118 PCP readers, none of whom participated in prior studies pertaining to the device, completed the study and were eligible for the effectiveness analysis. The study began after IRB approval, and proper informed consent was obtained from patients relating to the cases acquired as part of the previous clinical study.

**Lesion image selection.** This study used a randomly selected subset of lesion images and

accompanying clinical information (ie, lesion cases) that were acquired during a previous clinical study (DERM-ASSESS III clinical study).<sup>11</sup> Photos were acquired using a standardized approach via an iPad and clip-on handscope.

The photography procedure and training provided to study sites were intended to produce high-quality lesion images that would be usable for reader studies. Study sites in the DERM-ASSESS III clinical study were representative of the demographics of the patient population in the US for the device's intended use. The images and cases collected and used were therefore representative of those patients with evaluable lesions suggestive of melanoma.

The lesion images and the accompanying clinical information that met the quality standards of the DERM-ASSESS III study protocol (ie, in the study effectiveness population) were reviewed by a panel of three independent physicians (one dermatologist and two PCPs) to confirm quality eligibility for the reader study by confirming that the lesion photos allowed for an assessment clinically comparable to a real-world clinical care assessment. In total, 100 high-resolution digital clinical lesion cases for 50 malignant lesions including melanoma and non-melanoma skin cancers [NMSC; ie, basal cell carcinoma [BCC] and squamous cell carcinoma [SCC]], and 50 benign lesions, all biopsy-proven, including nevi, seborrheic keratosis (SK), and other benign lesions that had passed evaluation by the physician panel were randomly selected. Given the limited number of BCC and SCC lesion cases enrolled for lesions suggestive of melanoma, an overrepresentation of BCCs and SCCs were included to limit physician bias toward melanoma diagnoses. To minimize bias, every physician reader had a unique, randomized order in which cases were reviewed. Case randomization was performed using an algorithm within the survey platform.

**Study outcomes.** The device was expected to significantly increase physician sensitivity and area under the receiver operating characteristic (AUROC) curve based on the results of prior reader studies.<sup>13,15</sup> However, this study was conducted to provide supplemental evidence for regulatory purposes; thus, only noninferiority hypothesis tests were used as endpoints, but both noninferiority and superiority were tested for. The primary endpoint was to determine whether the AUROC of PCPs aided by knowledge of the ESS output was noninferior to the PCPs' unaided

ORIGINAL RESEARCH

AUROC, using biopsy-verified skin lesions as the reference standard.

The secondary endpoints were: (1) noninferiority of device-aided PCP sensitivity compared to unaided sensitivity for referral decisions for all malignancies included in the study; (2) noninferiority of device-aided PCP sensitivity for histopathology-confirmed melanoma only compared to unaided sensitivity for referral decisions; and (3) noninferiority of device-aided PCP specificity compared to unaided specificity for referral decisions.

All included lesions were biopsied and dermatopathology was used as the reference standard for sensitivity and specificity calculations. Further analyses included aided and unaided AUROC for melanoma, BCC, SCC, and NMSC (BCC and SCC collectively) and aided and unaided referral specificity for benign melanocytic nevi, seborrheic keratoses, and other benign lesions. In addition, physician confidence in their referral decisions was compared with and without the device output availability.

RESULTS

**Study physicians.** Overall, 118 participants included board-certified internal and family medicine physicians (49.2% and 50.8%, respectively) with a broad distribution of years in practice (range: 1-21+ years) (Table 1). There were 72.0% male and 28.0% female physicians. Years in practice varied with most physicians reporting 6 to 10 years and 21+ years in practice (27.1% and 31.4%, respectively), followed by 1 to 5 years (20.3%), 11 to 15 years (16.9%), and 16 to 20 years in practice (4.2%). Type of practice also varied, with most physicians reporting being employed in a multi-specialty group private practice (25.4%) or a hospital-owned practice (20.3%) and most commonly practicing in an urban area (66.9%). Physicians recruited from each state was reflective of US physician population numbers in each state, ranging from 0.8–12.7% of physicians recruited, with most physicians recruited from California (12.7%). The frequency they reported evaluating patients with concerns regarding their skin lesions was most often reported as "always" (57.6%) followed by "usually" (27.1%). With regards to referring patients to dermatology, 57.6% of physicians reported "sometimes," followed by "most of the time" (39.8%) and "rarely" (2.5%). In accordance with the device indication for use and the study eligibility requirements, no physicians reported

**TABLE 1. Physician characteristics**

CHARACTERISTICS	ALL PHYSICIANS (n=118)
<b>Sex, n (%)</b>	
Male	85 (72.0)
Female	33 (28.0)
<b>Area of board certification and active practice, n (%)</b>	
Internal medicine	58 (49.2)
Family medicine	60 (50.8)
<b>Year practicing medicine, n (%)</b>	
1-5 years	24 (20.3)
6-10 years	32 (27.1)
11-15 years	20 (16.9)
16-20 years	5 (4.2)
21+ years	37 (31.4)
<b>Geographic region of United States, n (%)</b>	
Midwest	24 (20.3)
Northeast	33 (28.0)
South	37 (31.4)
West	24 (20.3)
<b>Type of area, n (%)</b>	
Urban area (population > 50,000)	79 (66.9)
Urban cluster (population between 2,500 and 50,000)	30 (25.4)
Rural (population <2,500)	9 (7.6)

"never" or "always" referring patients for skin lesions. Self-rated competence in skin lesion assessment was most reported as "advanced" (49.2%) followed by "intermediate" (45.8%). Only 5.1% of physicians self-reported being an "expert" in skin lesion assessment. Nearly one-third of physicians reported completing an elective rotation in dermatology (31.4%).

**Subject and lesion demographics.** The population of included images were from patients (n=100) that were 52.0% male and 48.0% female with a mean age of 62.3 (standard deviation [SD]: 15.06). Over half of patients were aged between 61-80 (52.0%). A total of 88 (88.0%) patients were of lighter skin types (ie, Fitzpatrick skin types I-III), with 59.0% designated as Fitzpatrick skin type II and 20.0% Fitzpatrick skin type III. For the subjects' most often reported risk factors, 14.0% had a new or changing lesion, 57.0% had ultraviolet light exposure, 36.0% had lighter skin tone, freckling, and light hair, and 20.0% had many moles and/or dysplastic nevi.

The reader study lesions (n=100) were mostly located on the trunk (61%), followed by head (18%), leg (11%), and arm (10%). Half of the 100 lesion cases were malignant (n=50). Among all lesions, 79% were flat (21% elevated), and 88%

**TABLE 1, CONT. Physician characteristics**

CHARACTERISTICS	ALL PHYSICIANS (n=118)
<b>Type of practice, n (%)</b>	
Independent primary care physician (PCP) - solo private practice	16 (13.6)
Owner in group private practice (primary care only)	8 (6.8)
Owner in group private practice (multi-specialty)	9 (7.6)
Employed in group private practice (primary care only)	18 (15.3)
Employed in group private practice (multi-specialty)	30 (25.4)
Hospital-owned practice	24 (20.3)
Federally qualified health center	3 (2.5)
Academic center	8 (6.8)
Locum tenens	1 (0.8)
Other, occupational health practice	1 (0.8)
<b>Self-reported frequency of evaluating patients with concerns regarding their skin lesions, n (%)</b>	
Always	68 (57.6)
Usually	32 (27.1)
Sometimes	18 (15.3)
Rarely	0 (0.0)
Never	0 (0.0)
<b>Referral of patients with skin lesions to dermatologist, n (%)</b>	
Always	0 (0.0)
Most of the time	47 (39.8)
Sometimes	68 (57.6)
Rarely	3 (2.5)
Never	0 (0.0)
<b>Self-rated skin lesion assessment competence, n (%)</b>	
Expert	6 (5.1)
Advanced	58 (49.2)
Intermediate	54 (45.8)
Beginner	0 (0.0)
No competence	0 (0.0)
<b>Completed elective rotation in dermatology, n (%)</b>	
Yes	37 (31.4)
No	81 (68.6)

were smooth (12% rough). Among the malignant lesions, 68% were melanoma, with an equal proportion of BCC and SCC at 16% each. Among the histologically proven benign lesions, 62% were benign melanocytic nevi (BMN), and 20% were seborrheic keratosis (SK). The average lesion width was 4.67 mm (SD: 1.875) (Table 2 and Table 3).

ORIGINAL RESEARCH

**TABLE 2. Reader study patient characteristics**

CHARACTERISTICS	READER STUDY PARTICIPANTS (n=100)
<b>Sex, n (%)</b>	
Male	52 (52.0)
Female	48 (48.0)
<b>Age</b>	
Mean (standard deviation)	62.3 (15.06)
Median	65
Minimum, maximum	22, 86
<b>Age (years), n (%)</b>	
<20	0 (0.0)
21-30	2 (2.0)
31-40	8 (8.0)
41-50	15 (15.0)
51-60	14 (14.0)
61-70	25 (25.0)
71-80	27 (27.0)
81-90	9 (9.0)
90+	0 (0.0)
<b>Race, n (%)</b>	
White	99 (99.0)
Non-white	1 (1.0)
<b>Fitzpatrick skin type, n (%)</b>	
I. Always burns, never tans	9 (9.0)
II. Always burns, tans minimally	59 (59.0)
III. Sometimes mild burn, tans uniformly	20 (20.0)
IV. Burns minimally, always tans well	7 (7.0)
V. Very rarely burns, tans very easily	4 (4.0)
VI. Never burns	1 (1.0)
<b>Risk factors, n (%)</b>	
New or changing lesion(s)	14 (14.0)
Ultraviolet light exposure (natural or tanning bed)	57 (57.0)
Lighter skin tone, freckling, light hair	36 (36.0)
Family history of skin cancer	22 (22.0)
Many moles and/or dysplastic nevi	20 (20.0)
Personal history of skin cancer	35 (35.0)
Weakened immune system	3 (3.0)

**Primary outcome.** A total of 50 malignant and 50 benign lesions were assessed by each of the 118 physicians, yielding 5,900 lesion assessments without the device and 5,900 with the ESS device result. The study met the primary endpoint; the AUROC for referral decisions

**TABLE 3. Reader study lesion characteristics**

CHARACTERISTICS	READER STUDY LESIONS (n=100)
<b>Anatomic location, n (%)</b>	
Head	18 (18.0)
Arm	10 (10.0)
Leg	11 (11.0)
Trunk	61 (61.0)
<b>Flat or elevated, n (%)</b>	
Flat	79 (79.0)
Elevated	21 (21.0)
<b>Smooth or rough, n (%)</b>	
Smooth	88 (88.0)
Rough	12 (12.0)
<b>Overall pathology, n (%)</b>	
Malignant	50 (50.0)
Benign	50 (50.0)
<b>Malignant lesion pathology (n = 50), n (%)</b>	
Basal cell carcinoma (BCC)	8 (16.0)
Melanoma	34 (68.0)
Squamous cell carcinoma (SCC)	8 (16.0)
<b>Benign lesion pathology (n = 50), n (%)</b>	
Benign melanocytic nevi (BMN)	31 (62.0)
Seborrheic keratosis (SK)	10 (20.0)
Other benign	9 (18.0)
<b>Length (mm)</b>	
Mean (standard deviation)	6.13 (2.893)
Median	5.50
Minimum, maximum	2.5, 15.0
<b>Width (mm)</b>	
Mean (standard deviation)	4.67 (1.875)
Median	4.00
Minimum, maximum	2.5, 11.0

regarding all lesions of the PCPs aided with the device was 0.671 (95% confidence interval [CI]: 0.611-0.732) compared to the AUROC unaided by the device of 0.630 (95% CI: 0.582-0.678). The aided AUROC for referral decisions regarding all lesions was noninferior to the unaided AUROC ( $p < 0.001$ ). Furthermore, a statistical test for superiority found that the aided AUROC was significantly higher than the unaided AUROC (absolute difference of 0.041;  $p = 0.036$ ). See Figure 1 for physician sensitivity and AUROC distributions and Figure 2 for primary and second outcomes.

**Secondary outcomes.** Aided sensitivity of the PCPs was noninferior to unaided sensitivity across all skin cancer types, and superior to unaided assessments for melanoma detection. Specificity

did not decrease significantly, as demonstrated through the noninferiority test (Table 4).

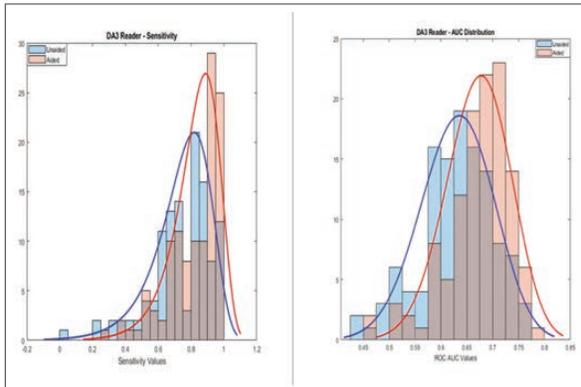
**Physician confidence assessment.** From unaided to aided assessments, the mean scores changed from  $6.4 \pm 1.59$  (range: 1-10) to  $7.0 \pm 1.74$  (range: 1-10). There was a significant difference in the mean score (0.67; 95% CI: 0.52-0.81;  $p < 0.001$ ) between aided and unaided equal. With the aid of the device result, confidence scores decreased in 16.0% ( $n = 1890$ ) of readings, while 33.1% ( $n = 3903$ ) of scores stayed the same and 50.9% ( $n = 6007$ ,  $p < 0.001$ ) of scores increased for all readings.

Sensitivity increased with increasing confidence levels. When evaluating confidence groupings into low (1-3), mid (4-7), and high (8-10) spectral scores, sensitivity increased with availability of a device result. For low device scores (1-3), PCPs' unaided sensitivity of 77.3% increased to 84.2% for aided assessments, while specificity decreased from an unaided value of 33.0% to 25.2% when PCPs were aided with the device. For malignant lesions in which the PCPs' had low confidence in their unaided assessments (242 lesions), the unaided sensitivity was 77.3% and aided sensitivity was 88.0%, a difference of 10.7%. For low confidence benign lesions (276), the unaided specificity was 33.0% and aided specificity was 31.9%, a difference of -1.1%.

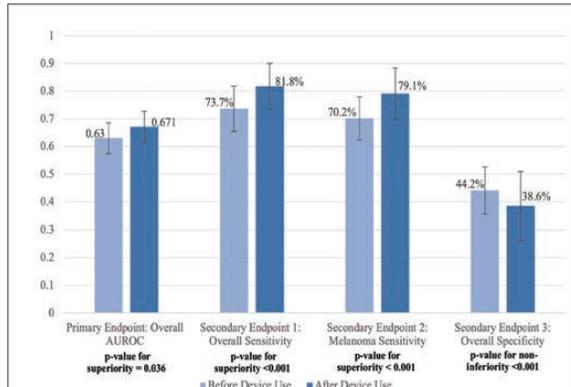
**Diagnostic sensitivity and specificity.** The diagnostic sensitivity for PCPs in detecting malignancies with the device result was 62.1% (95% CI: 54.0-70.2%), which was higher than sensitivity without the device result of 59.8% (95% CI: 52.7-66.9%). Physician diagnostic specificity for benign lesions with the device output available at 64.2% (95% CI: 57.1-71.3%) was also higher than without the device output available at 62.4% (95% CI: 56.3-68.6%). While use of the device to inform diagnostic assessments is not the device's indication for use, these results suggest that even if the device were used in this manner for lesions suggestive of melanoma, neither sensitivity nor specificity would decrease and that one or both may increase.

**Subgroup analyses.** While the present study was not powered to assess lesion subtypes and there was no formal hypothesis testing within subgroups, numerical increases in AUROC were seen and AUROC was noninferior for each cancer subtype. For example, when assessing sensitivity by cancer types (melanoma, BCC, SCC, NMSC), the study results demonstrated superior sensitivity for melanoma when aided by the device and

ORIGINAL RESEARCH



**FIGURE 1.** Reader sensitivity and area under the receiver operating characteristic (AUROC) distribution.  
DA3: DERM-ASSESS III



**FIGURE 2.** Results for primary endpoint (overall area under the receiver operating characteristic [AUROC]) and secondary endpoints 1-3 (overall sensitivity, melanoma sensitivity, and overall specificity)

**TABLE 4.** Secondary outcomes of average reader sensitivity and specificity

TEST	SENSITIVITY <sub>AIDED</sub>		SENSITIVITY <sub>UNAIDED</sub>		AIDED-UNAIDED DIFFERENCE				
	ESTIMATE	95% CI <sup>4</sup>	ESTIMATE	95% CI <sup>4</sup>	ESTIMATE	90% CI <sup>4</sup>	p-VALUE <sup>5</sup>	95% CI <sup>4</sup>	SUPER p-VALUE <sup>6</sup>
Reader sensitivity for referral by assessment type <sup>1</sup>	81.8%	76.0-87.6%	73.7%	67.7-79.6%	8.1%	4.2-12.0%	<0.001	3.5-12.8%	<0.001
Reader sensitivity for referral for melanoma by assessment type <sup>2</sup>	79.1%	72.4-85.7%	70.2%	62.9-77.6%	8.8%	4.2-13.4%	<0.001	3.4-14.3%	<0.001
	SENSITIVITY <sub>AIDED</sub>		SENSITIVITY <sub>UNAIDED</sub>		AIDED-UNAIDED DIFFERENCE				
	ESTIMATE	95% CI <sup>4</sup>	ESTIMATE	95% CI <sup>4</sup>	ESTIMATE	90% CI <sup>4</sup>	p-VALUE <sup>7</sup>	95% CI <sup>4</sup>	-
Reader specificity for referral by assessment type <sup>3</sup>	38.6%	30.9-46.2%	44.2%	38.1-50.1%	-5.6%	-10.7 to -0.6%	<0.001	-11.6 to 0.4%	-

<sup>1</sup>Analysis includes 50 malignant lesions and 118 readers.

<sup>2</sup>Analysis includes 34 malignant lesions and 118 readers.

<sup>3</sup>Analysis includes 50 benign lesions and 118 readers.

<sup>4</sup>Average reader sensitivity and standard errors obtained from iMRMC package from U-statistics approach.

<sup>5</sup>p-value for noninferiority hypothesis  $H_0: Se_{Aided} - Se_{Unaided} \leq 0.05$ .

<sup>6</sup>p-value for superiority hypothesis  $H_0: Se_{Aided} - Se_{Unaided} \leq 0$ .

<sup>7</sup>p-value for noninferiority hypothesis  $H_0: Sp_{Aided} - Sp_{Unaided} \leq 0.2$  (20%).

a meaningful sensitivity increase for NMSC. Average physician AUROC for melanoma unaided by the device (0.605, 95% CI: 0.548-0.661) was numerically higher when aided by device use (0.637, 95% CI: 0.570-0.703). Moreover, PCPs' sensitivity was observed to be higher for each malignant lesion subgroup when comparing their unaided sensitivity to aided sensitivity. Overall PCP sensitivity with device output for melanoma was 79.1% (95% CI: 72.4-85.7%), and for NMSC was 87.6% (95% CI: 77.2-97.9%). This was an increase from PCP sensitivity without device output of 70.2% (95% CI: 62.9- 77.6%) for

melanoma and 80.9% (95% CI: 73.7-88.1%) for NMSC.

Specificity was observed to be lower across benign lesion subgroups when comparing PCP-aided specificity to unaided specificity. Specificity for benign melanocytic nevi was 42.5% (95% CI: 33.0-52.0%) with the device compared to 47.2% (95% CI: 40.5-53.8%) without the device. For SKs, specificity was 32.7% (95% CI: 17.5-47.9%) with device output and 35.6% (95% CI: 26.5-44.6%) without device output. For other benign lesions, specificity was 31.5% (95% CI: 18.0-44.9%) with device output and 43.4% (95% CI: 29.8-57.0%)

without device output.

**Physician perceptions of device.** When asked about whether the device use would provide value to their decision making, 91.5% of respondents either agreed or strongly agreed, with 7.6% reporting neutral and 0.8% reporting they disagree or strongly disagree. When asked about the inclusion of the spectral score with the "investigate further" result, 4.2% disagreed there was a benefit, 14.4% were neutral, while 81.4% either agreed (47.5%) or strongly agreed (33.9%) that there was a benefit. In addition, 93.2% of respondents also agreed or strongly agreed that

ORIGINAL RESEARCH

**TABLE 5.** Overall shift in referral decisions from unaided-to-aided assessment

LESION TYPE	TOTAL ASSESSMENTS <sup>1</sup>	UNAIDED ASSESSMENT CORRECT			UNAIDED ASSESSMENT INCORRECT		
		Unaided, correct <sup>2</sup> n (%)	Aided DermaSensor Results for Malignant or Benign (% Based on Unaided Correct)		Unaided, correct <sup>2</sup> n (%)	Aided DermaSensor Results Correct for Malignant or Benign (% Based on Unaided Correct)	
			Aided, correct <sup>3</sup> n (%)	Aided, incorrect <sup>3</sup> n (%)		Aided, correct <sup>4</sup> n (%)	Aided, incorrect <sup>4</sup> n (%)
Malignant	5900	4346 (73.7)	4020 (92.5)	326 (7.5)	1554 (26.3)	805 (51.8)	749 (48.2)
Melanoma	4012	2818 (70.2)	2590 (91.9)	228 (8.1)	1194 (29.8)	582 (48.7)	612 (51.3)
NMSC (BCC or SCC)	1888	1528 (80.9)	1430 (93.6)	98 (6.4)	360 (19.1)	223 (61.9)	137 (38.1)
Basal cell carcinoma (BCC)	944	735 (77.9)	700 (95.2)	35 (4.8)	209 (22.1)	140 (67.0)	69 (33.0)
Squamous cell carcinoma (SCC)	944	793 (84.0)	730 (92.1)	63 (7.9)	151 (16.0)	83 (55.0)	68 (45.0)
Benign	5900	2606 (44.2)	1671 (64.1)	935 (35.9)	3294 (55.8)	604 (18.3)	2690 (81.7)

<sup>1</sup>Total assessments from 118 primary care physician readers.

<sup>2</sup>Percentage of unaided assessments correct or incorrect out of total unaided assessments.

<sup>3</sup>Percentage of lesions with correct or incorrect aided assessments out of number of lesions with correct unaided assessments.

<sup>4</sup>Percentage of lesions with correct or incorrect aided assessments out of number of lesions with incorrect unaided assessments.

they would benefit from having NPV and positive predictive value (PPV) available in instructional materials associated with the device.

When asked about their own standard of care performance in assessing skin lesions, most respondents rated their unaided sensitivity between 61 to 80% (52.5%) and their specificity between 51 to 70% (39.8%). Respondents reported that the availability of the ESS device output would alter their behavior encouraging them to perform better quality skin lesion evaluations. Also, regarding expected benefits of the device in real-world care with their patients, 81% of respondents agreed that the device would provide an immediate, objective result to help with management of suspicious skin lesions, 75% agreed that they would detect more skin cancer, and 71% agreed they would have greater confidence in their clinical assessments and management decisions. Only 1% of respondents said they would not expect any benefit with the device.

**DISCUSSION**

In this MRMCC study, we found both a noninferior and superior increase in overall management sensitivity and in AUROC with physician use of the adjunctive device. Aided sensitivity of melanoma identification was also

noninferior and superior to unaided sensitivity. While the specificity decreased by 5.6 to 38.6% aided, this was noninferior to the specificity endpoint. Diagnostic sensitivity also increased from 59.8 to 62.1% with the availability of device output, and specificity increased from 62.4 to 64.2%.

The only two AI tools to have been approved prior to this ESS device were MelaFind and Nevisense, and both were restricted for use to only dermatologists.<sup>16</sup> Despite FDA approval in 2011, MelaFind was discontinued for sale and clinical use in 2017. The limitations of this device included unnecessary biopsies owing to low specificity of 10%, high device cost and workflow burden, and device use restricted to only pigmented lesions. Resembling prior studies of the ESS device, device use improves PCPs' evaluation and management decisions regarding all skin cancers. However, this is the first study demonstrating the potential of the device to significantly improve melanoma detection.

**Clinical implications.** PCP use of the ESS device may improve referral decisions. For example, correctly referred aided melanomas increased by 48.7% (n=582) for melanomas that would have been incorrectly managed if unaided. For each skin cancer subgroup, this trend was

consistent and resulted in an overall increase in correctly aided referrals for melanoma as well as BCCs and SCCs (Table 5).

There has been a rising incidence of skin cancer in the US amid the changing climate. A 2°C increase in temperature due to global warming has been postulated to increase skin cancer incidence by 10% annually; this temperature rise is expected by 2050. Therefore, the ESS device may help meet the need for PCPs to diagnose these skin cancers and refer patients appropriately.<sup>17,18</sup>

Noninvasive, point-of-care tools for PCPs have the potential to aid in rapid skin cancer detection and earlier access to dermatologists while also reducing unnecessary referrals of benign lesions. When the tool is used as an adjunct to careful consideration of the clinical appearance and available clinical information (patient skin phototype, age, location on the body, risk factors, changes of the lesion reported by the patient), there is greater potential for more accurate triage of potential skin cancers.

This study was strengthened by the multireader design, which allowed for multiple physicians to assess lesions, each with a different randomized order of lesion cases, potentially reducing individual bias. The study was also strengthened by its inclusion of a representative mix of common skin cancers, with two-thirds of cancerous

## ORIGINAL RESEARCH

lesions being melanoma. Limitations of our study impacted both sensitivity and specificity calculations. Limitations include the limited sample size in certain subgroups including those with Fitzpatrick skin type VI and I, limited sample size of patients ages 21-30, and limited sample of flat lesions. Additionally, this study was conducted in a theoretical setting, not direct patient care.

## CONCLUSION

Skin cancer is the most common type of cancer in the US, with projected increases in diagnoses anticipated. While less common than BCCs and SCCs, melanoma spreads more aggressively and is associated with decreased survival rates. As the primary entry point into the healthcare system, PCPs play a pivotal role in connecting patients with specialized dermatological care when warranted.

We present the findings of a reader study assessing the referral performance of PCPs when evaluating lesions suggestive of melanoma with and without the ESS result. The device significantly improved physicians' accuracy in managing lesions suggestive of melanoma, with significant increases in sensitivity for all skin cancers and melanoma, while not significantly raising the rate of unnecessary referrals. Moreover, the physicians reported increased confidence in their assessments with the availability of the device result. As such, the device can improve PCP decision making and confidence when managing lesions suggestive of melanoma.

By enhancing PCP management accuracy and confidence, this FDA-approved device has the potential to optimize resource utilization by minimizing unnecessary referrals and biopsies, saving healthcare dollars, and reducing patient morbidity and concern, while helping to prioritize access for patients with high-risk lesions. Given the rising rates of skin cancer in the US and limited access to dermatologic care, this device has the potential to make a significant impact in skin cancer detection in the primary care setting.

## REFERENCES

- Aggarwal P, Knabel P, Fleischer AB. United States burden of melanoma and non-melanoma skin cancer from 1990 to 2019. *J Am Acad Dermatol*. 2021;85(2):388-95.
- Rahib L, Wehner MR, Matrisian LM, Nead KT. Estimated projection of US cancer incidence and death to 2040. *JAMA Netw Open*. 2021;4(4):e214708.
- Schuchter L, Schultz DJ, Synnestvedt M, et al. A prognostic model for predicting 10-year survival in patients with primary melanoma. *Ann Intern Med*. 1996;125(5):369-75.
- Balboul S, Gronbeck C, Feng H. Dermatology workforce projections in the United States, 2021 to 2036. *Arch Dermatol Res*. 2024;316(5):192.
- Najmi M, Brown AE, Harrington SR. A systematic review and synthesis of qualitative and quantitative studies evaluating provider, patient, and health care system-related barriers to diagnostic skin cancer examinations. *Arch Dermatol Res*. 2022;314(4):329-40.
- Gareau DS, Correa da Rosa J, Yagerman S, et al. Digital imaging biomarkers feed machine learning for melanoma screening. *Exp Dermatol*. 2017;26(7):615-8.
- Gareau DS, Browning J, Correa Da Rosa J, et al. Deep learning-level melanoma detection by interpretable machine learning and imaging biomarker cues. *J Biomed Opt*. 2020;25(11):112906.
- Marchetti MA, Cowen EA, Kurtansky NR, et al. Prospective validation of dermoscopy-based open-source artificial intelligence for melanoma diagnosis (PROVE-AI study). *NPJ Digit Med*. 2023;6(1):1-11.
- Giavina-Bianchi M, Sousa RM de, Paciello VZ de A, et al. Implementation of artificial intelligence algorithms for melanoma screening in a primary care setting. *PLoS One*. 2021;16(9):e0257006.
- Manolakos D, Patrick G, Geisse JK, et al. Use of an elastic-scattering spectroscopy and artificial intelligence device in the assessment of lesions suggestive of skin cancer: a comparative effectiveness study. *JAAD Int*. 2024;14:52-8.
- Hartman RI, Trepanowski N, Chang MS, et al. Multicenter prospective blinded melanoma detection study with a handheld elastic scattering spectroscopy device. *JAAD Int*. 2024;15:24-31.
- Merry SP, Croghan IT, Dukes KA, et al. Primary Care Physician Use of Elastic Scattering Spectroscopy on Skin Lesions Suggestive of Skin Cancer. *Journal of Primary Care & Community Health*. 2025;16.
- Jaklitsch E, Thames T, de Campos Silva T, Coll P, Oliviero M, Ferris LK. Clinical utility of an AI-powered, handheld elastic scattering spectroscopy device on the diagnosis and management of skin cancer by primary care physicians. *J Prim Care Community Health*. 2023;14:21501319231205979.
- Merry SP, Chatha K, Croghan I, Nguyen V, McCormick B, Leffell D. Clinical performance of novel elastic scattering spectroscopy (ESS) in detection of skin cancer: a blinded, prospective, multi-center clinical trial. *J Clin Aesthet Dermatol*. 2023;16(4 Suppl 1):S5-31.
- Ferris LK, Jaklitsch E, Seiverling EV, et al. DERM-SUCCESS FDA pivotal study: a multireader multicase evaluation of primary care physicians' skin cancer detection using AI-enabled elastic scattering spectroscopy. *J Prim Care Community Health*. 2025;16:21501319251342106.
- Beltrami EJ, Brown AC, Salmon PJM, Leffell DJ, Ko JM, Grant-Kels JM. Artificial intelligence in the detection of skin cancer. *J Am Acad Dermatol*. 2022;87(6):1336-42.
- Kaffenberger BH, Shetlar D, Norton SA, Rosenbach M. The effect of climate change on skin disease in North America. *J Am Acad Dermatol*. 2017;76(1):140-7.
- Lin MJ, Torbeck RL, Dubin DP, et al. Climate change and skin cancer. *J Eur Acad of Dermatol Venereol*. 2019;33(9):e324-e325. **JCAD**

## Venkatesh et al. (2024)

npj | digital medicine

Published in partnership with Seoul National University Bundang Hospital

News &amp; Views

<https://doi.org/10.1038/s41746-024-01161-1>

# Learnings from the first AI-enabled skin cancer device for primary care authorized by FDA

Kaushik P. Venkatesh, Kushal T. Kadakia &amp; Stephen Gilbert

Check for updates

The U.S. Food and Drug Administration's (FDA) recent authorization of DermaSensor, an AI-enabled device for skin cancer detection in primary care, marks a pivotal moment in digital health innovation. Clinically, the authorization of the first AI-enabled device for use by non-specialists for detecting skin cancer reinforces the feasibility of digital health technologies to bridge gaps in access and expertise in medical practice. The authorization also establishes a new regulatory precedent for FDA authorization of medical devices incorporating AI and machine learning (ML) technologies within dermatology. Together, this article uses the DermaSensor authorization to examine the clinical evidence and regulatory implications of emerging AI-enabled technologies in dermatology.

primary care office. This article uses the DermaSensor authorization to examine the clinical evidence and regulatory implications of emerging AI-enabled technologies in dermatology.

## Regulatory landscape and new horizons

FDA's framework for medical device regulation was established when most products were hardware-based technologies. However, many devices today incorporate some elements of software technology, including AI and machine learning (ML), with 171 AI/ML devices authorized by FDA as of October 2023. Some products, such as DermaSensor, are physical devices that rely on AI/ML software for functionality, and are termed "Software in a Medical Device" (SiMD)<sup>4,5</sup>. Other products, such as Digital Diagnostic's autonomous diabetic retinopathy software<sup>6</sup>, are primarily software-based products, and are termed "Software as a Medical Device" (SaMD)<sup>4</sup>.

In a 2021 Action Plan, FDA clarified its intent to continue regulating such devices using its existing review pathways, while also developing new regulatory processes for issues unique to devices incorporating software (e.g., change control planning, mitigation of algorithmic bias)<sup>7</sup>. These review pathways consist of premarket clearance (510(k)), De Novo classification, and premarket approval (PMA). Review pathways are selected based on the potential risk posed by a device, which in turn determines requirements for pre- and post-market evidence generation.

The first dermatology AI device to gain FDA authorization was Melafind (multispectral spectroscopy) approved in 2017 via the PMA pathway (Table 1)<sup>8</sup>. This pathway is reserved for high-risk devices and requires manufacturers submit clinical evidence of safety and effectiveness (e.g., evidence generated from a clinical trial). Melafind's indication for use was limited to dermatologists and melanoma detection (Table 2). However, the device was discontinued due to low specificity (10%) leading to unnecessary biopsies, narrow use cases, poor integrability into workflow, and limited coverage<sup>9</sup>. The next dermatology AI device, Nevisense (electrical impedance spectroscopy), was also authorized under the PMA pathway in 2017 for use by dermatologists alone for melanoma detection (sensitivity 96%, specificity 34%), and remains on the market as of 2024<sup>10</sup>.

In contrast to Melafind and Nevisense, DermaSensor (elastic scattering spectroscopy) was reviewed under FDA's De Novo pathway. This pathway is intended for novel devices of low-to-moderate risk for which the manufacturer or FDA lack similar, authorized devices to compare to. Consequently, in addition to enabling a new product (DermaSensor) to progress to market, De Novo authorization also establishes a new product classification code which future dermatology devices can use as regulatory precedent to gain FDA authorization via the 510(k) pathway—a unique regulatory development compared to the PMA approach of AI-enabled dermatology devices to date (Table 2)<sup>11</sup>. While 510(k) is the predominant route to market entry for all medical devices—including AI/ML-enabled

On January 17, 2024, the U.S. Food and Drug Administration (FDA) authorized DermaSensor, the first artificial intelligence (AI)-enabled medical device intended for use in the primary care setting for skin cancer detection<sup>1</sup>. The device, which was already authorized for use in the European Union and Australia, establishes a new regulatory precedent in the U.S. for medical device innovation as the first AI-enabled dermatologic device indicated for use by non-specialist. However, the approval also raises questions about the need for continuous evidence generation and modern approaches to regulation.

Clinical diagnosis in dermatology is highly reliant on visual assessment. Differentiating between "normal" and "abnormal" skin lesions—which in turn determines the need for further diagnostic studies and intervention—depends on a dermatologist's pattern recognition skills, which compound over the course of a clinical career. Performing nuanced visual analysis requires years of training, posing both a challenge for non-specialists<sup>2</sup> and an opportunity for innovation through AI/ML pattern recognition. Skin cancer—including squamous cell carcinoma, basal cell carcinoma, and melanoma—is the most common malignancy worldwide<sup>3</sup> and is primarily screened through gross assessment and dermatoscopy prior to diagnosis by biopsy or excision. Primary care physicians often screen patients for concerning lesions prior to referral to dermatology<sup>2,3</sup>. With this context, DermaSensor and similar PCP-targeted AI-enabled devices provide critical opportunities to impact the broader healthcare system by bringing specialist care into the

**Table 1 | Comparison of pathways to FDA authorization for medical devices, including AI-enabled dermatology devices**

Category	PMA	De Novo	510(k)
AI-enabled dermatology device	NeviSense & MelaFind	DermaSensor	None
Risk classification	High risk	Low-moderate risk	Low-moderate risk
Qualifications for pathway	Novel, life-sustaining without predicate	Novel device without predicate	Substantially equivalent predicate device
Required data to submit to FDA	Both non-clinical and clinical evidence of safety and effectiveness	Non-clinical and sometimes clinical data demonstrating the device's safety and effectiveness	Demonstration that the new device is similar to a previously-authorized device; primarily relies on non-clinical evidence of device performance, but may in some cases include clinical data of safety and effectiveness
Review time and cost	Slower (<180 days) and more expensive due to requirement for clinical data	Generally faster (<150 days)	Fastest review (<90 days)
Post-market	Often required to complete a post-approval study, with requirements for regular data reports for FDA	May be subject to requirements for postmarket evidence generation	Rarely subjected to formal postmarket surveillance

**Table 2 | Comparison of pivotal study sizes and primary endpoints among FDA-authorized AI-enabled medical devices for skin cancer detection**

Device	Imaging technology	Indication	Patient <i>n</i> (Lesions <i>n</i> )	Sensitivity	Specificity	Lesions of Fitzpatrick V/VI (%)
MelaFind <sup>21</sup>	Multispectral spectroscopy	Disease: Melanoma User: Dermatologist	1383 (1631)	98.3%	9.9%	9 (0.6%)
NeviSense <sup>22</sup>	Electrical impedance spectroscopy	Disease: Melanoma User: Dermatologist	1951 (2416)	97.0%	31.3%	29 (1.6%)
DermaSensor <sup>13</sup>	Elastic scattering spectroscopy	Disease: Melanoma, squamous cell carcinoma, and basal cell carcinoma in patients >40 years User: Primary care physician	1005 (1579)	96.6%	21.0%	128 (12.7%) <sup>a</sup>

<sup>a</sup>Publicly-available FDA documents for DermaSensor report data by patients rather than lesions.

devices—the pathway has [raised concerns](#) about the adequacy of its evidentiary requirements for assessing the safety of new digital health tools<sup>12</sup>.

### Examining the evidence for the DermaSensor authorization

FDA's authorization of DermaSensor broke new clinical ground, as the device is indicated for use by non-dermatologist physicians to evaluate skin lesions raising concern for melanoma, basal cell carcinoma, and squamous cell carcinoma in patients 40 years or older. Evidence supporting the authorization was generated from [three studies](#): a pivotal trial, a supplemental validation study for melanoma, and a clinical utility study<sup>13</sup>.

The pivotal study, titled DERM-SUCCESS, was conducted on 1579 lesions of 1005 patients from 22 primary care centers; the supplemental validation study, DERM-ASSESS, included 311 patients with 440 lesions evaluated by dermatologists from 10 sites. The device's output was used in addition to traditional clinical history and examination to decide whether suspicious lesions should be monitored versus further evaluated (dermatologist referral and/or biopsy). DERM-SUCCESS reported 95.5% device sensitivity compared with 83% for primary care physicians (PCPs), an NPV (negative result by device confirmed by biopsy) of 96.6%, and non-inferiority to dermatologist sensitivity of 90%. However, specificity was low at 20.7%.

Notably, [FDA's authorization](#) specifies that DermaSensor is approved for use by “physicians who are not dermatologists,” with the manufacturer identifying PCPs as a target population<sup>14</sup>. The clinical utility study of 108 PCPs and >10,000 lesions showed that PCP device use increased management sensitivity (91.4% vs. 82.0%) and diagnostic sensitivity (81.7% vs.

71.1%) and decreased false negative referrals by half (8.6% from 18%). However, there was a statistically significant decrease in specificity (44.2% to 32.4%) for referrals.

Taken together, the results from DermaSensor's premarket clinical studies demonstrate the device's capability for sensitive detection and application in the primary care setting. However, DermaSensor also, illustrates that issues with specificity (which affected MelaFind and Nevisense) remain a challenge for AI-enabled medical devices in dermatology, carrying the risk of potentially unnecessary referrals that may result in biopsies and further costs.

### Clinical implications for access and quality

DermaSensor provides new horizons for dermatologic care by extending the diagnostic capacity of PCPs for skin cancer; a key differentiator from previous FDA-authorized AI-enabled medical devices in dermatology. Using DermaSensor could strengthen the diagnostic abilities of PCPs who normally refer such cases, and could therefore meaningfully address access limitations in dermatology, which affect more than a third of patients<sup>15</sup> with an [average wait time of 35 days](#)<sup>16</sup>.

There are analogous situations of the introduction of novel AI-based diagnostic devices into primary care from other medical disciplines. For example, FDA has authorized AI-enabled devices for [autonomous diabetic retinopathy screenings](#) for use in primary care settings<sup>17</sup>. This technology-enabled shift toward greater “specialist” diagnosis by PCPs raises important questions about the expanding scope of generalist clinical practice. For

DermaSensor, an important question is whether the device will improve PCP diagnostic performance enough to warrant deferral of dermatologist evaluation. Additionally, realizing the aspiration of improved diagnostic access is also dependent on the uptake of new technologies among PCPs serving the populations with the greatest needs. For example, rural patients face the [longest wait-times and greatest shortages of dermatologists](#)<sup>18</sup>; however, only 5% of PCPs in DERM-SUCCESS were considered rural. Device deployment among providers in medically underserved areas warrants further postmarket research.

DermaSensor's authorization also comes with notable regulatory conditions. FDA has imposed requirements for post-market performance testing in underrepresented populations, as notably 97.1% of patients in the pivotal DERM-SUCCESS trial were of White race and only 13% were from the most pigmented skin types (Fitzpatrick V/VI). While improving clinical trial diversity is a priority across all FDA-regulated products, the implications are particularly salient in dermatology, where the lack of darker skin tones in images used to train AI models have resulted in [poorer diagnostic performance](#) on darker skin phenotypes<sup>19</sup>. The lack of representation risks re-entrenching diagnostic biases for patients who are already systematically and historically [marginalized](#) in dermatology care<sup>20</sup>. In addition to investigating equitable performance across patient populations, FDA should also monitor DermaSensor's real-world performance with particular attention to diagnostic specificity and the risk of triggering unnecessary care cascades through false positives; an issue that contributed to the discontinuation of MelaFind.

**The path forward.** DermaSensor's authorization is the latest in the ongoing wave of AI-enabled medical devices to gain regulatory approval in the U.S. and beyond. Understanding the opportunities and limitations of these technologies is critical to their safe and effective deployment in real-world clinical practice. The approach taken by FDA is notable for its equity-focused lens to ensure alignment between new product development and population health needs via post-market evidence generation. This approach should be proactively implemented in the regulatory approach to AI-enabled devices more broadly, and ideally at the pre-market stage. Taken together, while the short-term impact of DermaSensor's authorization is the addition of a new AI-enabled specialty tool in the primary care toolkit, the device's long-term legacy may be a milestone for the regulation of AI-enabled medical devices.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Kaushik P. Venkatesh**<sup>1</sup>, **Kushal T. Kadakia**<sup>1</sup> & **Stephen Gilbert**<sup>2</sup> ✉

<sup>1</sup>Harvard Medical School, Boston, MA, USA. <sup>2</sup>Else Kröner Fresenius Center for Digital Health, TUD Dresden University of Technology, Dresden, Germany. ✉e-mail: [stephen.gilbert@ukdd.de](mailto:stephen.gilbert@ukdd.de)

Received: 27 March 2024; Accepted: 10 June 2024;

Published online: 15 June 2024

## References

- Office of the Commissioner. FDA Roundup: January 16, FDA. <https://www.fda.gov/news-events/press-announcements/fda-roundup-january-16-2024> (2024).
- Chen, S. C. et al. Diagnosing and managing cutaneous pigmented lesions: primary care physicians versus dermatologists. *J. Gen. Intern. Med.* **21**, 678–682 (2006).
- Guy, G. P. et al. Vital signs: melanoma incidence and mortality trends and projections—United States, 1982–2030. *Morb. Mortal. Wkly. Rep.* **64**, 591–596 (2015).
- Health C for D and R. Software as a Medical Device (SaMD). FDA. <https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd> (2024).

- Health C for D and R. Artificial Intelligence and Machine Learning (AI/ML)-enabled medical devices. FDA. <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-ai-ml-enabled-medical-devices> (2023).
- Abramoff, M. D. et al. Autonomous artificial intelligence increases real-world specialist clinic productivity in a cluster-randomized trial. *npj Digital Med.* **6**, 1–8 (2023).
- Health C for D and R. Artificial intelligence and machine learning in software as a medical device. FDA. <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device> (2024).
- Singer, N. Dissent over a device to help find melanoma. *The New York Times*. <https://www.nytimes.com/2013/07/21/business/dissent-over-a-device-to-help-find-melanoma.html> (2013).
- Cukras, A. R. On the comparison of diagnosis and management of melanoma between dermatologists and MelaFind. *JAMA Dermatol.* **149**, 622–623 (2013).
- Ollmar, S. & Grant, S. Nevusense: improving the accuracy of diagnosing melanoma. *Melanoma Manag.* **3**, 93–96 (2016).
- US Food and Drug Administration. Device Classification Under Section 513(f)(2)(De Novo). <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/denovo.cfm?id=DEN230008> (2024).
- Muehlethaler, U. J., Bluethgen, C. & Vokinger, K. N. FDA-cleared artificial intelligence and machine learning-based medical devices and their 510(k) predicate networks. *Lancet Digital Health* **5**, e618–e626 (2023).
- Clinical Studies. DermaSensor. <https://www.dermasensor.com/clinical-studies/> (2024).
- US FDA. FDA authorization of DermaSensor documentation. [https://www.accessdata.fda.gov/cdrh\\_docs/pdf23/DEN230008.pdf](https://www.accessdata.fda.gov/cdrh_docs/pdf23/DEN230008.pdf) (2024).
- Venkatesh KP, Brito G, Nelson K, Friedman A. Barriers to care-seeking and treatment adherence among dermatology patients: a cross-sectional national survey study. *J. Drugs Dermatol.* <https://jddonline.com/articles/barriers-to-care-seeking-and-treatment-adherence-among-dermatology-patients-a-cross-sectional-national-survey-study-S1545961622P0677X/> (2024).
- American Academy of Family Physicians. Appointment wait times increase across specialties, but not in family medicine. Brand. <https://www.aafp.org/pubs/afm/blogs/inpractice/entry/wait-times.html> (2024).
- Office of the Commissioner. FDA permits marketing of artificial intelligence-based device to detect certain diabetes-related eye problems. FDA. <https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-artificial-intelligence-based-device-detect-certain-diabetes-related-eye> (2020).
- Brumbaugh, B., Goldman, N., Nambudiri, V. & LaChance, A. H. The Resident Physician Shortage Reduction Act: an opportunity to address the rural dermatology workforce deficit. *J. Am. Acad. Dermatol.* **87**, 1461–1464 (2022).
- Daneshjoui, R. et al. Disparities in dermatology AI performance on a diverse, curated clinical image set. *Sci. Adv.* **8**, eabq6147 (2022).
- Narla, S., Heath, C. R., Alexis, A. & Silverberg, J. I. Racial disparities in dermatology. *Arch. Dermatol. Res.* **315**, 1215–1223 (2023).
- US Food and Drug Administration. Premarket Approval (PMA). <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=p090012> (2024).
- US Food and Drug Administration. Premarket Approval (PMA). <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P150046> (2024).

## Author contributions

K.P.V. and K.T.K. developed the concept of the manuscript and drafted the manuscript. S.G. provided critical oversight and revisions. All authors had final approval of the completed manuscript and take accountability for all aspects of the work in ensuring accuracy and integrity of the work.

## Competing interests

K.P.V. declares no nonfinancial interests and no competing financial interests. K.T.K. reports fees from the Common Health Coalition (through ChangeLab Solutions) and the Journal of the American College of Cardiology, all unrelated to this manuscript. S.G. declares a nonfinancial interest as an Advisory Group member of the EY-coordinated “Study on Regulatory Governance and Innovation in the Field of Medical Devices” conducted on behalf of the DG SANTE of the European Commission. S.G. declares the following competing financial interests: he has or has had consulting relationships with Una Health GmbH, Lindus Health Ltd., Flo Ltd., Thymia Ltd., FORUM Institut für Management GmbH, High-Tech Gründerfonds Management GmbH, and Ada Health GmbH and holds share options in Ada Health GmbH. S.G. is a News and Views Editor for npj Digital Medicine. S.G. played no role in the internal review or decision to publish this News and Views article.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41746-024-01161-1>.

**Correspondence** and requests for materials should be addressed to Stephen Gilbert.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

---

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024