Dermatology

Validation of a Market-Approved Artificial Intelligence Mobile Health App for Skin Cancer Screening: A Prospective Multicenter Diagnostic Accuracy Study

Tobias Sangers[a]   Suzan Reeder[b]   Sophie van der Vet[a]   Sharan Jhingoer[a]
Antien Mooyaart[c]   Daniel M. Siegel[d]   Tamar Nijsten[a]   Marlies Wakkee[a]

[a]Department of Dermatology, Erasmus MC Cancer Institute, University Medical Center Rotterdam, Rotterdam, the Netherlands; [b]Department of Dermatology, Albert Schweitzer Hospital, Dordrecht, the Netherlands; [c]Department of Pathology, Erasmus MC Cancer Institute, University Medical Center Rotterdam, Rotterdam, the Netherlands; [d]Department of Dermatology, SUNY Downstate Health Sciences University and the VA New York Harbor Health Care System, Brooklyn, NY, USA

Abstract
Background: Mobile health (mHealth) consumer applications (apps) have been integrated with deep learning for skin cancer risk assessments. However, prospective validation of these apps is lacking. ***Objectives:*** To identify the diagnostic accuracy of an app integrated with a convolutional neural network for the detection of premalignant and malignant skin lesions. ***Methods:*** We performed a prospective multicenter diagnostic accuracy study of a CE-marked mHealth app from January 1 until August 31, 2020, among adult patients with at least one suspicious skin lesion. Skin lesions were assessed by the app on an iOS or Android device after clinical diagnosis and before obtaining histopathology. The app outcome was compared to the histopathological diagnosis, or if not available, the clinical diagnosis by a dermatologist. The primary outcome was the sensitivity and specificity of the app to detect premalignant and malignant skin lesions. Subgroup analyses were conducted for different smartphone types, the lesion's origin, indication for dermatological consultation, and lesion location. ***Results:*** In total, 785 lesions, including 418 suspicious and 367 benign control lesions, among 372 patients (50.8% women) with a median age of 71 years were included. The app performed at an overall 86.9% (95% CI 82.3–90.7) sensitivity and 70.4% (95% CI 66.2–74.3) specificity. The sensitivity was significantly higher on the iOS device compared to the Android device (91.0 vs. 83.0%; $p = 0.02$). Specificity calculated on benign control lesions was significantly higher than suspicious skin lesions (80.1 vs. 45.5%; $p < 0.001$). Sensitivity was higher in skin fold areas compared to smooth skin areas (92.9 vs. 84.2%; $p = 0.01$), while the specificity was higher for lesions in smooth skin areas (72.0 vs. 56.6%; $p = 0.02$). ***Conclusion:*** The diagnostic accuracy of the mHealth app is far from perfect, but is potentially promising to empower patients to self-assess skin lesions before consulting a health care professional. An additional prospective validation study, particularly for suspicious pigmented skin lesions, is warranted. Furthermore, studies investigating mHealth implementation in the lay population are needed to demonstrate the impact on health care systems.

© 2022 The Author(s)
Published by S. Karger AG, Basel

Correspondence to:
Tobias Sangers, t.sangers@erasmusmc.nl
Marlies Wakkee, m.wakkee@erasmusmc.nl

Introduction

Health care systems are challenged by high volumes of skin cancers, requiring optimal use of health care resources [1–3]. In the clinical arena, it is well documented that experience increases diagnostic accuracy [4]. In line with this observation, in countries with a closed health system that position the general practitioner (GP) acts as a gatekeeper to specialized health care, skin cancer detection accuracy is thought to be suboptimal [5]. Teledermatology might close this gap to some extent but requires the involvement of GPs and dermatologists [6]. Additionally, as has been learned in the COVID-19 pandemic, many patients with concerns about melanoma cannot capture a quality image for review by the dermatologist due to inadequate use of technology in their possession [7].

In 2017, deep learning algorithms achieved skin cancer detection accuracy comparable to dermatologists [8]. Two years later, these algorithms were already integrated in a mobile health (mHealth) application for consumer smartphone devices [9]. Using the smartphone camera, an app can provide an instant risk assessment of skin lesions. In light of the potential benefit, health insurers across Europe, Australia, and New Zealand have already introduced a form of reimbursement of at least one mHealth app [10–13]. However, a recent systematic review warned about the limited evidence currently available regarding the accuracy of the available apps [14]. Until now, no prospective validation studies of deep learning in mHealth apps for skin cancer detection have been performed. In this study, we aim to validate an mHealth app currently approved for consumers in Europe, Australia, and New Zealand, which uses a deep-learning convolutional neural network (CNN) for skin premalignancy and malignancy detection in the setting of a dermatology department.

Material and Methods

Study Design and Participants

We performed a prospective cross-sectional multicenter diagnostic accuracy study at the dermatology outpatient clinics of the Erasmus MC Cancer Institute and Albert Schweitzer Hospital in the Netherlands from January 1 until August 31, 2020. The study was designed and results were reported following the updated Standards for Reporting Diagnostic Accuracy Studies [15]. The need for ethical approval to conduct this study was waived by the Medical Research Ethics Committee (MREC) of the Erasmus MC University Medical Center, registered under MEC-2019-041, after screening the study design. All patients gave written informed consent.

All patients aged ≥18 years with an appointment at the dermatology outpatient clinics with one or more suspicious skin lesions were eligible to participate in our study. Suspicious skin lesions were defined as skin lesions for which a patient was either referred to the dermatology outpatient departments by a GP, or which were considered premalignant or malignant during follow-up visits of known patients at the dermatology departments by a dermatologist at the outpatient clinics. Exclusion criteria were patients who were unable to provide consent or skin lesions with a prior biopsy, obscuring by foreign matter (e.g., tattoos), or lack of a clear clinical or histopathological diagnosis.

Procedures

Eligible patients were approached by dermatologists during consultations at the outpatient clinics and subsequently included by one of the researchers (T.S., S.V.). After obtaining the clinical diagnosis for one or more suspicious skin lesions, risk assessments were acquired by one of the researchers (T.S., S.V.) who photographed skin lesions using the mHealth app (SkinVision, Amsterdam, the Netherlands). In addition to the suspicious lesion, at least one clinically diagnosed benign skin lesion, which the patient usually indicated as another potential lesion that they would like to be assessed, was included. No specific preference was given towards lesions in terms of location, size, or lesion type.

For this study, we used two different smartphone devices with a 12-megapixel camera running either an Android 10 (Galaxy S9, Samsung, Seoul) or iOS 13 (iPhone Xr, Apple Inc., Cupertino) operating system, which the researcher randomly chose. The app automatically checked acquisition conditions such as lighting and photo quality before a photo was accepted for assessment. Photos that could not be successfully taken within 1 min were recorded as a failure for the device type. In this case, the alternative smartphone device was used. If the app on both devices was unable to take a photo of a skin lesion successfully, this was recorded as a failed inclusion.

Algorithm Assessment

The app performed a risk assessment of the photographed skin lesions using a CNN (version RD-174). Skin lesions were classified as low or high risk of skin cancer within 30 s by the CNN, in which premalignant and malignant skin lesions were trained as high-risk lesions (suppl. Table 1; for all online suppl. material, see www.karger.com/doi/10.1159/000520474) [9]. Users of the mHealth app are advised to "monitor" a skin lesion when a lesion is categorized as low risk of skin malignancy or premalignancy by the algorithm. For lesions that are considered high risk by the algorithm, users are advised to visit a doctor.

Histopathology and Follow-Up

Histopathology of suspicious skin lesions was obtained through a biopsy or excision based on clinical indication. The decision to obtain histology was made by the dermatologist before risk assessment with the app to ensure that the CNN outcome did not affect routine clinical care. Histopathology was not obtained from skin lesions that were clinically benign and for lesions that, according to the Dutch guideline, do not require histological confirmation such as actinic keratosis and superficial basal cell carcinoma. If a biopsy was followed by an excision, the diagnosis of the latter was used as a gold standard. To gauge the likelihood of false-negative outcomes among skin lesions of which no histopathology was obtained, all patients' medical files were followed up to 3 months.

Sangers/Reeder/van der Vet/Jhingoer/
Mooyaart/Siegel/Nijsten/Wakkee
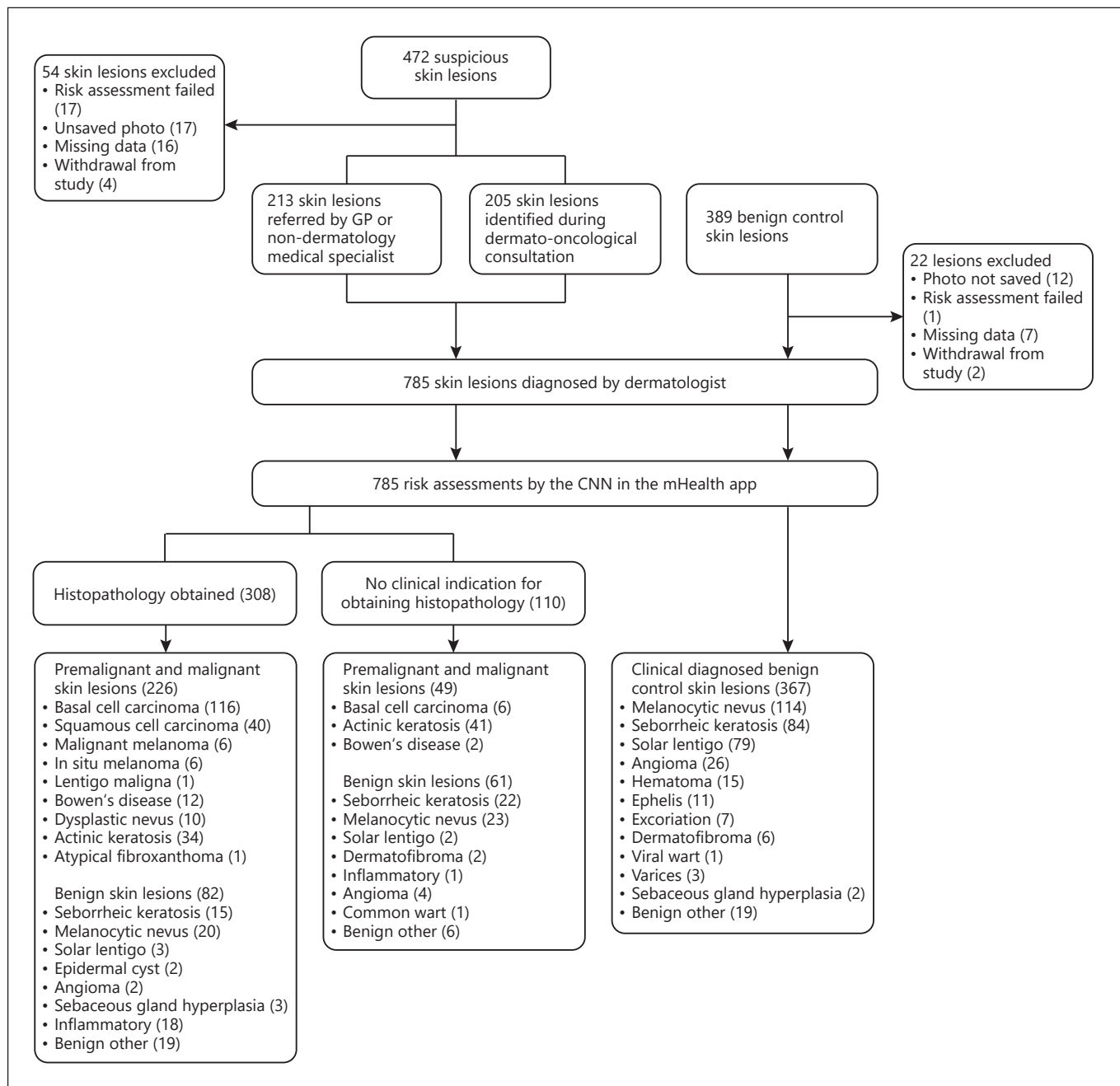
Fig. 1. Flow chart of included skin lesions. CNN, convolutional neural network.

Outcomes

The primary outcome of this study was the sensitivity and specificity of the app to detect premalignant and malignant skin lesions. The CNN outcome was compared to the histopathological diagnosis or, if there was no clinical indication for obtaining histopathology (e.g., clearly benign lesions and actinic keratoses), the clinical diagnosis from the treating dermatologist. Secondary outcomes were the positive and negative predictive values, positive likelihood ratio, negative likelihood ratio, and the overall accuracy of the app.

Exploratory subgroup analyses were performed for the different smartphone devices, the origin of the lesion (melanocytic vs. non-melanocytic) or indication for dermatological consultation (suspicious lesion identified by nondermatologist vs. lesions identified during follow-up visits of known patients at the dermatology departments). As factors surrounding a skin lesion on a photo can influence a CNN outcome, a subgroup analysis was performed between lesions located on, or close to, skin fold lesion locations (e.g., nasolabial fold) versus smooth skin lesion areas [16].

Statistical Analyses

Descriptive statistics were used to compare the distribution of baseline characteristics. Sensitivity, specificity, positive and negative predictive values, positive likelihood ratio, negative likelihood ratio, and overall accuracy were calculated with a 95% Clopper-Pearson confidence interval. Sensitivity and specificity were compared between subgroups using a two-proportion Z test. A χ^2 test was performed to compare lesion characteristics between successful and failed risk assessments. All analyses were performed using IBM SPSS Statistics for Windows, version 25.0 (IBM Corp., Armonk, NY, USA).

Sample Size

The app's sensitivity was estimated between 85 and 95% with a specificity between 70 and 80% [17, 18]. An absolute precision level (margin of error) of 5.0% was considered satisfactory to report the app's sensitivity and specificity. For a 95% level of confidence, at least 196 premalignant or malignant skin lesions needed to be included in this study to report the lowest expected boundary of 85% sensitivity. To report the lowest expected boundary of 70% specificity, at least 323 benign skin lesions were required in the study (suppl. Fig. 1).

Results

Study Population

Of the initial 861 lesions among 392 patients evaluated with the app, 785 lesions (91.2%, 785/861) among 372 patients were included in the complete case analysis (Fig. 1). In 48 cases (5.6%, 48/861), risk assessments failed during the first attempt, of which 60% (29/48) failed with the iOS device. Risk assessments failed significantly more often on skin lesions located in the head and neck area, and in skin folds ($p < 0.001$). After switching devices, 30 assessments were successful during the second attempt. Skin lesions that failed on both devices were in the majority diagnosed as basal cell carcinoma (72.2%, 13/18) and were frequently located on the nose (50.0%, 9/18), ear (16.7%, 3/18), or other parts of the head (22.2%, 4/18). Additionally, 58 (6.7%, 58/861) lesions had to be excluded mainly because of unsaved risk assessments ($n = 29$), missing data with respect to incomplete registration ($n = 23$), and withdrawal from the study ($n = 6$).

The final data set of 372 patients, as described in Table 1, comprised 418 suspicious and 367 benign control skin lesions. The median age of the included patients was 71 years (IQR 58–78), with equal gender distribution and over 80% (308/372) having a Fitzpatrick skin type 1-2. The majority of benign control lesions were located on the extremities (53.4%, 196/367) and rarely in the head and neck area (16.9%, 62/367), whereas almost half of all suspicious lesions were located in the head and neck area (48.8%, 204/367).

**Table 1.** Characteristics of the study population and the assessed lesions

Characteristics	Number	Percent
Patients	372	
Median age (IQR), years	71 (58–78)	
Sex		
Male	183	49.2
Female	189	50.2
Fitzpatrick skin type		
1	42	11.3
2	266	71.5
3	39	10.5
4	4	1.1
Missing	21	5.6
Hospital		
EMC	108	29.0
ASZ	264	71.0
Assessed lesions	785	
Suspicious lesions	418	53.2
Location		
Head and neck	204	48.8
Back	41	9.8
Thorax and abdomen	67	16.0
Extremities	106	25.4
Benign control lesions	367	46.8
Location		
Head and neck	62	16.9
Back	41	11.2
Thorax and abdomen	68	18.5
Extremities	196	53.4

ASZ, Albert Schweitzer Hospital; EMC, Erasmus MC, University Medical Center; IQR, interquartile range.

After clinical evaluation of all suspicious and benign control lesions by a dermatologist, histopathology was obtained from 308 (39.2%, 308/785) skin lesions (Fig. 1). Of the biopsied lesions, 226 (73.4%, 226/308) were considered premalignant or malignant. The remaining 477 (60.8%, 477/785) skin lesions were diagnosed based on clinical inspection only as benign ($n = 428$), premalignant ($n = 43$), and malignant ($n = 6$). In total, 275 lesions turned out to be premalignant or malignant, representing 35.0% (275/785) of all included skin lesions. Three months of follow-up of patient records for the benign skin lesion showed no dermatology consultations suggesting malignant transformation.

Algorithm Accuracy

As presented in Figure 2, the app identified 239 of the 275 premalignant and malignant lesions as a high-risk
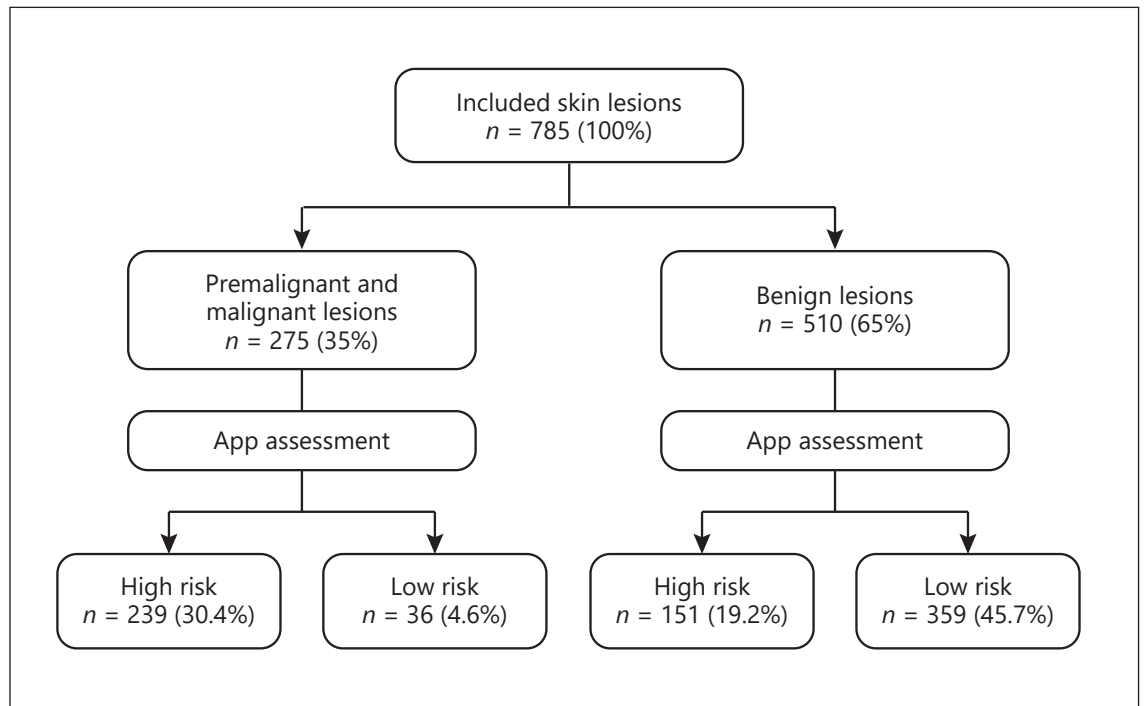
Fig. 2. Flow chart of the app risk assessment outcome for the included premalignant, malignant, and benign skin lesions.

Table 2. Overall sensitivity and specificity of the app in detecting skin premalignancy and malignancy, including subgroup analyses between the iOS and Android devices, melanocytic versus nonmelanocytic skin lesions, skin fold lesion areas versus smooth skin lesion areas, suspicious skin lesions versus benign control lesions, skin lesions identified after GP and nondermatology referrals versus follow-up consultations of known patients at the outpatient clinics

Assessment type	N (%)	Sensitivity (95% CI), %	p value	Specificity (95% CI), %	p value
Overall app accuracy	785 (100)	86.9 (82.3–90.7)		70.4 (66.2–74.3)	
Android device	425 (54.1)	83.0 (75.7–88.8)*	0.02	71.5 (65.9–76.7)	0.27
iOS device	360 (45.9)	91.0 (84.9–95.3)*		69.0 (62.6–75.0)	
Melanocytic skin lesions	179 (22.8)	81.8 (59.7–94.8)	0.26	73.3 (65.6–80.0)	0.17
Nonmelanocytic skin lesions	606 (77.2)	87.4 (82.6–91.2)		69.1 (64.0–73.9)	
Skin fold lesion areas	138 (17.6)	92.9 (85.3–97.4)*	0.01	56.6 (42.3–70.2)*	0.02
Smooth skin lesion areas	647 (82.4)	84.2 (78.2–89.1)*		72.0 (67.6–76.1)*	
Suspicious skin lesions	418 (53.3)	86.9 (82.3–90.7)		45.5 (37.1–54.0)***	<0.001
Benign control lesions	367 (46.7)	–		80.1 (75.7–84.1)***	
GP and nondermatology referrals	213 (27.1)	89.6 (83.4–94.1)	0.09	39.1 (27.6–51.6)	0.07
Follow-up consultations	205 (26.1)	84.0 (76.6–89.8)		51.4 (39.4–63.2)	

* p < 0.05, *** p < 0.001.

lesion, resulting in an overall sensitivity of 86.9% (95% CI 82.3–90.7). The 36 (13.1%, 36/275) lesions that were incorrectly classified as low risk included two in situ melanomas, one invasive melanoma, two cutaneous squamous cell carcinomas, and 13 basal cell carcinomas (suppl. Table 2a, b). The overall specificity of the app was 70.4% (95% CI 66.2–74.3) (Table 2), leading to an overall accuracy of 76.2% (95% CI 73.0–79.1). The calculated

positive and negative likelihood ratios were 2.9 (2.6–3.4) and 0.2 (0.1–0.3), respectively. Of all 390 high-risk assessments, 239 were confirmed by dermatological and/or histopathological evaluation resulting in a positive predictive value of 61.3% (95% CI 57.9–64.6), and the negative predictive value was 90.9% (95% CI 88.0–93.2). Lesions that were rather frequently falsely positive identified by the app were melanocytic nevi (42, 27.8%), seborrheic keratoses (26, 17.2%), and angiomas (16, 10.6%). A detailed list of all false-positive and false-negative cases is presented in supplementary Table 2a–d. The algorithm outcome calculated solely on histopathology-validated skin lesions revealed an 89.8% (95% CI 85.1–93.4) sensitivity and a 32.9% (95% CI 22.9–44.2) specificity.

Subgroup Analyses

In the exploratory subgroup analysis of the different devices, we noticed that the app performed at a significantly higher sensitivity on the iOS device (91.0%, 95% CI 84.9–95.3) compared to 83.0% (95% CI 75.7–88.8) on the Android device ($p = 0.02$). There was no significant difference in specificity (71.5 vs. 69.0%; $p = 0.27$) between devices. Subgroup analysis of the app performance between melanocytic and nonmelanocytic skin lesions revealed no difference in sensitivity (81.8 vs. 87.4%; $p = 0.26$) and specificity (73.3 vs. 69.1%; $p = 0.17$). Although the number of lesions in skin fold areas (17.6%, 138/785) was relatively low, the sensitivity of the app was higher in the skin fold areas compared to smooth skin areas (92.9 vs. 84.2%; $p = 0.01$), while the specificity was higher for lesions on smooth skin areas (72.0 vs. 56.6%; $p = 0.02$). The sensitivity (89.6 vs. 84.0%; $p = 0.09$) and specificity (39.1 vs. 51.4%; $p = 0.07$) did not differ between suspicious lesions identified by nondermatologists versus lesions identified during follow-up of known patients at the dermatology department. The specificity calculated solely on benign control lesions was 80.1% (95% CI 75.7–84.1), which was significantly higher compared to the 45.5% (95% CI 37.1–54.0) specificity of suspicious skin lesions ($p < 0.001$). Cross-tabulations of these results are added to the supplementary material (suppl. Table 3a–l).

Discussion

The results of this first prospective diagnostic accuracy study of a market-approved mHealth app with a deep learning algorithm show an overall 86.9% sensitivity and 70.4% specificity in detecting skin premalignancy and malignancy in dermatology departments. Compared to a recent systematic review, our findings show that inclusion of the deep learning algorithm results in a sensitivity at the highest end of the previous reported confidence interval [14].

mHealth apps related to skin cancer are often available for both iOS and Android operating systems [19]. The observed difference in sensitivity emphasizes the need for validation on both platforms. In addition, the specificity of the app calculated on histopathology-verified lesions was lower compared to the overall specificity. This drop in specificity is expected because the lesions that a dermatologist selects for histological evaluation are the most difficult lesions to categorize correctly. Besides the lesion type, the inclusion of lesions located on or near skin folds resulted in a higher sensitivity but lower specificity. Although the exact reason for this could not be identified due to the *black box* aspect of the deep learning algorithm, we hypothesize that this effect was attributed to skin fold lesion areas that contain more lines and irregularities compared to smooth skin areas, which may be more difficult for the app to distinguish from benign lesions.

The diagnostic accuracy of the current study was lower than a recent retrospective validation study of the app (95% sensitivity and 78% specificity) [18]. This difference can be attributed in part to the larger clinical variety of included skin lesions and eligibility of the skin lesions (in the retrospective study, benign lesions were only included after considered benign in a teledermatology assessment). In the current study, specificity was calculated based on both benign control *and* suspicious lesions. This addition of suspicious skin lesions resulted in a lower specificity because they were also more difficult to distinguish from malignant skin lesions for the app. When restricting the analysis to benign control lesions, the specificity was 80.1% which is in accordance with the retrospective validation study [18]. However, including the suspicious lesions in the analysis better reflects the real-world situation in which the app is used.

There is no consensus on the diagnostic accuracy required for implementing an algorithm in skin cancer diagnosis but depends on where the test is positioned in the patient journey (i.e., laypeople, GPs or dermatologists). The sensitivity in this validation study was comparable to rates reported about untrained laypeople, but showed a much higher specificity (i.e., reducing unnecessary GP consultations) [20]. The sensitivity of GPs to visually diagnose skin cancer ranges from 25 to 91%, with a specificity varying between 55 and 92%, suggesting the app is at

least as good as, and possibly exceeds the accuracy of, most GPs. A recent Cochrane review reported a 94.9% (95% CI 90.1–97.4) sensitivity and 84.3% (95% CI 48.5–96.8) specificity of teledermatology to identify cutaneous malignancies, suggesting that the mHealth app needs improvement to match the accuracy of teledermatology [6].

Although the mHealth app is expected to further improve in time due to technological developments and an improved algorithm, there are also other options that could be explored to improve the app's accuracy. One option would be to explore the optimal integration of a deep learning algorithm with a targeted teledermatology function to achieve maximum accuracy with the least human effort. Another option would be to explore the addition of dermoscopy pictures which is known to improve the diagnostic accuracy of dermatologists, but this would require users to buy an additional accessory for their mobile phone, providing a potential barrier for accessible use [21, 22].

This study has several limitations. First, the results reported in this study are based on an artificial setting in which photos were taken by trained researchers and not by patients or their partners. However, we expect a limited impact of the assessor on this external validation because the integrated quality check only accepts high-quality images. Second, photos were taken at outpatient departments and not in the intended setting at home, in which the app would typically be used. Third, lesions that were not successfully assessed by the app were excluded from the calculation of the primary outcome. Fourth, in contrast to the distribution of benign and malignant lesions in our study, we expect the number of benign lesions to far exceed the number of malignant lesions typically assessed by people in the general population. Given the relatively low specificity calculated on benign control lesions (80.1%), the use of the app with the current algorithm in the general population without a targeted approach could lead to high volumes of benign skin lesions being categorized as high risk, and unnecessary consultations [23]. Fifth, over 80% of our study population had Fitzpatrick skin type I or II, limiting the generalizability of the algorithm accuracy found in this study for darker skin types. Sixth, we used two smartphone devices with a relatively high-resolution camera. Potential variations in accuracy on iOS and Android devices with lower resolution cameras should be explored in future studies. Finally, the number of melanomas included in this study was relatively low ($n = 12$), of which two in situ melanomas and one invasive melanoma were falsely categorized as low risk by the app. This relatively high proportion empha-sizes the need to perform an adequately powered additional validation study for suspicious pigmented skin lesions.

The app's diagnostic accuracy is far from perfect, but it is potentially promising to empower patients to self-assess skin lesions before consulting a health care professional. An additional prospective validation study, particularly for suspicious pigmented skin lesions, is warranted. In addition, studies investigating the health care impact of mHealth implementation in the lay population are needed. mHealth apps based on deep learning algorithms could be part of the solution in managing the skin cancer epidemic. Furthermore, mHealth with deep learning could serve as a valuable triage tool in the ongoing pandemic era where local conditions may limit in-person evaluation by dermatologists.

Key Message

Prospective validation of a smartphone app for skin cancer screening, available to download by the general population, reveals an 86.9% sensitivity and 70.4% specificity in detecting premalignant and malignant skin lesions.

Statement of Ethics

The need for ethical approval was waived by the medical ethical committee of the Erasmus MC University Medical Center after screening of the study design (MEC-2019-041). Written informed consent was obtained from all participants.

Conflict of Interest Statement

Funding Sources

Author Contributions

T.S. and M.W. verified the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. T.S., S.R., S.V., T.N., M.W. conceived and designed the study. T.S., S.V., J.S., and M.W. analyzed the data. S.R., A.M., D.S., T.N., and M.W. provided strategic guidance and oversight. T.S., T.N., and M.W. drafted the manuscript with input from all authors. The final version of the paper has been approved by all authors.

Data Availability Statement

De-identified data used in this study are currently not publicly available. Researchers interested in data access should contact M.W. (m.wakkee@erasmusmc.nl). Data requests will need to undergo ethical and legal approval by the relevant institutions.

References

1 Karimkhani C, Green AC, Nijsten T, Weinstock MA, Dellavalle RP, Naghavi M, et al. The global burden of melanoma: results from the Global Burden of Disease Study 2015. Br J Dermatol. 2017;177(1):134–40.

2 Koelink CJ, Kollen BJ, Groenhof F, van der Meer K, van der Heide WK. Skin lesions suspected of malignancy: an increasing burden on general practice. BMC Fam Pract. 2014;15(1):29.

3 Noels E, Hollestein L, Luijkx K, Louwman M, de Uyl-de Groot C, van den Bos R, et al. Increasing costs of skin cancer due to increasing incidence and introduction of pharmaceuticals, 2007–2017. Acta Derm Venereol. 2020;100(10):adv00147.

4 Kittler H, Pehamberger H, Wolff K, Binder M. Diagnostic accuracy of dermoscopy. Lancet Oncol. 2002;3(3):159–65.

5 Ahmadi K, Prickaerts E, Smeets JGE, Joosten VHMJ, Kelleners-Smeets NWJ, Dinant GJ. Current approach of skin lesions suspected of malignancy in general practice in the Netherlands: a quantitative overview. J Eur Acad Dermatol Venereol. 2018;32(2):236–41.

6 Chuchu N, Dinnes J, Takwoingi Y, Matin RN, Bayliss SE, Davenport C, et al. Teledermatology for diagnosing skin cancer in adults. Cochrane Database Syst Rev. 2018;12(12):CD013193.

7 Farshchian M, Potts G, Kimyai-Asadi A, Mehregan D, Daveluy S. Outpatient teledermatology implementation during the COVID-19 pandemic: challenges and lessons learned. J Drugs Dermatol. 2020;19(6):683.

8 Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115–8.

9 de Carvalho TM, Noels E, Wakkee M, Udrea A, Nijsten T. Development of smartphone apps for skin cancer risk assessment: progress and promise. JMIR Dermatol. 2019;2(1):e13376.

10 SkinVision reimbursement CZ [cited 2021 May 7]. Available from: https://www.cz.nl/vergoedingen/skinvision.

11 Medicash Launches Skin Cancer Detection App [cited 2021 May 7]. Available from: https://www.medicash.org/article/news/medicash-skin-cancer-app-skinvision.

12 CBHS partners with SkinVision to provide digital skin checks at no cost for members [cited 2021 May 7]. Available from: https://www.cbhs.com.au/health-well-being-blog/blog-article/2020/06/12/cbhs-partners-with-skinvision-to-provide-digital-skin-checks-at-no-cost-for-members.

13 SkinVision reimbursement Accuro NZ [2021 May 7]. Available from: https://www.accuro.co.nz/about/skinvision/.

14 Freeman K, Dinnes J, Chuchu N, Takwoingi Y, Bayliss SE, Matin RN, et al. Algorithm based smartphone apps to assess risk of skin cancer in adults: systematic review of diagnostic accuracy studies. BMJ. 2020;368:m127.

15 Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ. 2015;351:h5527.

16 Winkler JK, Fink C, Toberer F, Enk A, Deinlein T, Hofmann-Wellenhof R, et al. Association between surgical skin markings in dermoscopic images and diagnostic performance of a deep learning convolutional neural network for melanoma recognition. JAMA Dermatol. 2019;155(10):1135–41.

17 Thissen M, Udrea A, Hacking M, von Braunmuehl T, Ruzicka T. mHealth app for risk assessment of pigmented and nonpigmented skin lesions – a study on sensitivity and specificity in detecting malignancy. Telemed J E Health. 2017;23(12):948–54.

18 Udrea A, Mitra GD, Costea D, Noels EC, Wakkee M, Siegel DM, et al. Accuracy of a smartphone application for triage of skin lesions based on machine learning algorithms. J Eur Acad Dermatol Venereol. 2020;34(3):648–55.

19 Kong FW, Horsham C, Ngoo A, Soyer HP, Janda M. Review of smartphone mobile applications for skin cancer detection: what are the changes in availability, functionality, and costs to users over time?. Int J Dermatol. 2021 Mar;60(3):289–308.

20 Girardi S, Gaudy C, Gouvernet J, Teston J, Richard MA, Grob JJ. Superiority of a cognitive education with photographs over ABCD criteria in the education of the general population to the early detection of melanoma: a randomized study. Int J Cancer. 2006;118(9):2276–80.

21 Dinnes J, Deeks JJ, Chuchu N, Matin RN, Wong KY, Aldridge RB, et al. Visual inspection and dermoscopy, alone or in combination, for diagnosing keratinocyte skin cancers in adults. Cochrane Database Syst Rev. 2018;12(12):CD011901.

22 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):669–76.

23 Sangers TE, Nijsten T, Wakkee M. Mobile health skin cancer risk assessment campaign using artificial intelligence on a population-wide scale: a retrospective cohort analysis. J Eur Acad Dermatol Venereol. 2021 Nov;35(11):e772–4.