Accuracy of a smartphone application for triage of skin lesions based on machine learning algorithms

A. Udrea, G.D. Mitra, D. Costea, E.C. Noels, M. Wakkee, D.M. Siegel, T.M. de Carvalho, T.E.C. Nijsten

JEADV; accepted for publication.

ABSTRACT

Background Machine learning algorithms achieve expert level accuracy in skin lesion classification based on clinical images. However, it is not yet shown whether these algorithms could have high accuracy when embedded in a smartphone application, where image quality is lower and there is high variability in image taking scenarios by users. In the past, these applications were criticized due to lack of accuracy.

Objective In this study we evaluate the accuracy of the newest version of a smartphone application (SA) for risk assessment of skin lesions.

Methods This SA uses a machine learning algorithm to compute a risk rating. The algorithm is trained on 131,873 images taken by 31,449 users in multiple countries between January 2016 and August 2018 and rated for risk by dermatologists. To evaluate the sensitivity of the algorithm we use 285 histopathologically validated skin cancer cases (including 138 malignant melanomas), from two previously published clinical studies (195 cases) and from the SA user database (90 cases). We calculate the specificity on a separate set from the SA user database containing 6000 clinically validated benign cases.

Results The algorithm scored a 95.1% (95% CI, 91.9% - 97.3%) sensitivity in detecting (pre) malignant conditions (93% for malignant melanoma and 97% for keratinocyte carcinomas and precursors). This level of sensitivity was achieved with a 78.3% (95% CI, 77.2%-79.3%) specificity.

Conclusions This SA provides a high sensitivity to detect skin cancer, however there is still room for improvement in terms of specificity. Future studies are needed to assess the impact of this SA on health systems and its users.

INTRODUCTION

Access to medical care is being transformed with the advent of mobile applications which can monitor, diagnose and provide health information. In the field of dermatology there is considerable interest in developing algorithms for image analysis of skin lesions to evaluate the risk of malignant melanoma (MM) and keratinocyte carcinomas (KC).[1-4] Early detection of skin cancer is important since the prognosis is highly dependent on whether the cancer is detected at an early stage, when it is easier to treat and its prognosis is more favorable.[5, 6]

Smartphones equipped with applications that analyse skin lesions can become a powerful tool for early detection of skin cancer. Besides early diagnosis, they could help to lower the burden of skin cancer by reducing the need of physician consultations and the number of referrals to secondary care by general practitioners (GP). However, usage of smartphone applications (SA) for self-assessment of skin cancer (including the one here presented) has been criticized due to lack of evidence on its diagnostic accuracy.[7-10] They may have a relatively low sensitivity, which instead of facilitating early detection of skin cancer could cause a sense of false reassurance.

The mHealth application evaluated here is called SkinVision (developed by SkinVision B.V., Netherlands). It is a SA that allows laypersons to self-assess skin lesions for the risk of skin cancer. With this application, a user may take a picture of a lesion with his smartphone and have it analysed by a proprietary algorithm. This will return a risk rating for the lesion, with recommendation to see a doctor for high risk cases. This SA had its accuracy tested in two clinical studies. At first, it was only able to detect MM and its rule-based fractal algorithm scored a 73% sensitivity and 83% specificity.[11] In a later study, this algorithm was re-calibrated to be able to process both pigmented and non-pigmented lesions and detect, aside from MM, KC and some premalignant conditions.[12] The recalibrated version had a sensitivity of 80% and a 78% specificity.

The SA maintains a database with validated cases, where histopathological reports are shared voluntarily by some of its users and clinical validation is done by at least one dermatologist on all images that are being processed by the algorithm. This is used for recalibration of the algorithm. Since the previous clinical studies[11, 12], the available data for training the algorithm has substantially increased. In this paper, we present the new version of the risk assessment algorithm which uses a machine learning approach, and evaluate its sensitivity and specificity to detect skin cancer.

MATERIALS AND METHODS

Materials and data acquisition

The data used to train and test the algorithms for skin cancer risk assessment is obtained retrospectively from two previously published clinical studies[11, 12] and the SA user database. The Munich and the Eindhoven studies had approval from the local ethics committee, with respective numbers, No. 529-12 and No. 2014-41.[11, 12] All users gave their consent electronically when registering for the SA. The Munich University Hospital and Eindhoven hospital datasets contain cases seen in consecutive patients in the periods December 2012 to December 2013 and December 2014 to April 2016, respectively. The SA user database contains lesions imaged by users from several countries between January 2016 and October 2018. In Table 1, we show how we split the data to train and test the algorithm.

Table 1. Data used to train and test the algorithm for disease classification.

Dataset	Sample	Period Image Acquisition	Exclusion criteria	Studies
University Hospital Munich	Sensitivity Test Set I: 40 MM cases from 195 lesions in 195 consecutive patients.	December 2012 - December 2013	-	Maier et al 2014
Catharina Hospital Eindhoven	Sensitivity Test Set II: 155 skin cancer cases from 341 lesions in 256 consecutive patients (147 KC and precursors, 8 MM)	December 2014 - April 2016	-	Thissen et al 2017
Smartphone Application User Database [a]	Training Set: 131,873 images, half of the images used were previously classified by a dermatologist as high risk, the other half as low risk. Sensitivity Test Set III: 90 histopathologically validated cases of MM. Specificity Test Set: 6,000 cases which were clinically validated by dermatologists as benign.	January 2016 - October 2018	Skin on the Fitzpatrick scale V-VI; if lesion is located on mucosal surfaces, close to a visible scar or located under the nails; if lesion and surrounding skin contain foreign matter (e.g. tattoos) or are not intact (e.g. contains an ulcer).	-

[a] All cases clinically validated as low risk were randomly selected from the user database either for training or testing. Since there are a lot less cases rated as high risk or with histopathological report, we used all of them for either training or testing.

Training dataset

The training dataset consists of 131,873 images acquired with different iPhone and Android devices by 31,449 SA users. All images were given a risk rating by a dermatologist (Supplemental Material Appendix 1). Half of the training set consisted of images rated as high risk and half rated as low risk. Cases clinically validated as low risk were randomly

selected from the user database, while all cases rated as high risk were used (since there are a lot less cases rated as high risk).

Test dataset for sensitivity

The test dataset for sensitivity contains 285 cases (195 from previous clinical studies and 90 from the SA user database) which were histopathologically validated for skin cancer. The participant flowcharts of each dataset are given in Figures 1-3.

We used 40 MM cases from the Munich study (Figure 1) and 155 cases of different types of skin cancer or pre-cancer from the Eindhoven study (Figure 2).[11, 12] Previously, some cases were excluded due to bad imaging[11] or were split in training and validation set.[12] For this algorithm version, almost all malignant cases are used to validate the sensitivity, with the exception of 3 basal cell carcinoma (BCC) cases due to poor image quality. In the real user setup this would mean that the user would be asked to take a new image of the lesion.

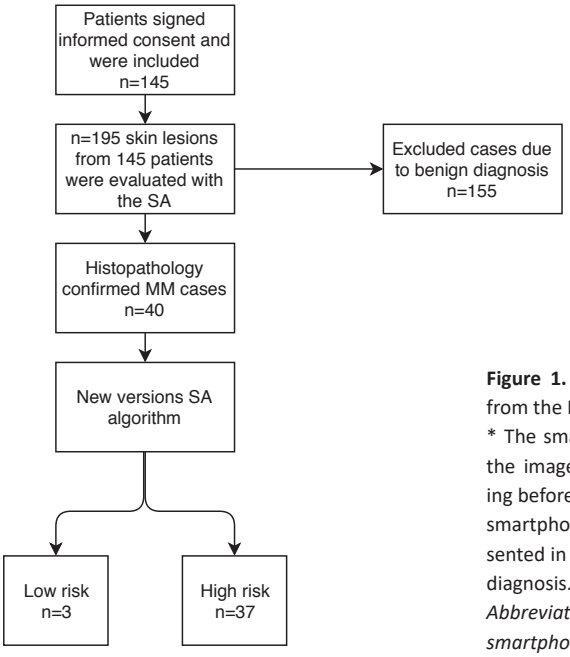


Figure 1. Participant flowchart of malignant cases from the Munich Study.*

* The smartphone application version used to take the images was developed during pre-clinical testing before the Munich study. The new version of the smartphone application algorithm is the version presented in this study. All cases excluded had a benign diagnosis.

Abbreviations: MM, malignant melanoma; SA, smartphone application.

The 90 MM cases from the user database (Figure 3) correspond to assessments made by users for which the user shared the final histopathology diagnosis with SA between October 2016 - October 2018. All user assessments were reviewed by an affiliated dermatologists. Users with high risk rating by the quality control of the dermatologist were messaged by the Customer Service team. However, the SA studied here cannot oblige the users to share their final diagnosis, this is a voluntary decision of the user. In total, at

the time of writing the study (October 2018) 338 users shared their final histopathology diagnosis with the SA, of which 178 were MMs. Out of these, we randomly selected 90 (51%) to be included in the test set (Figure 3).

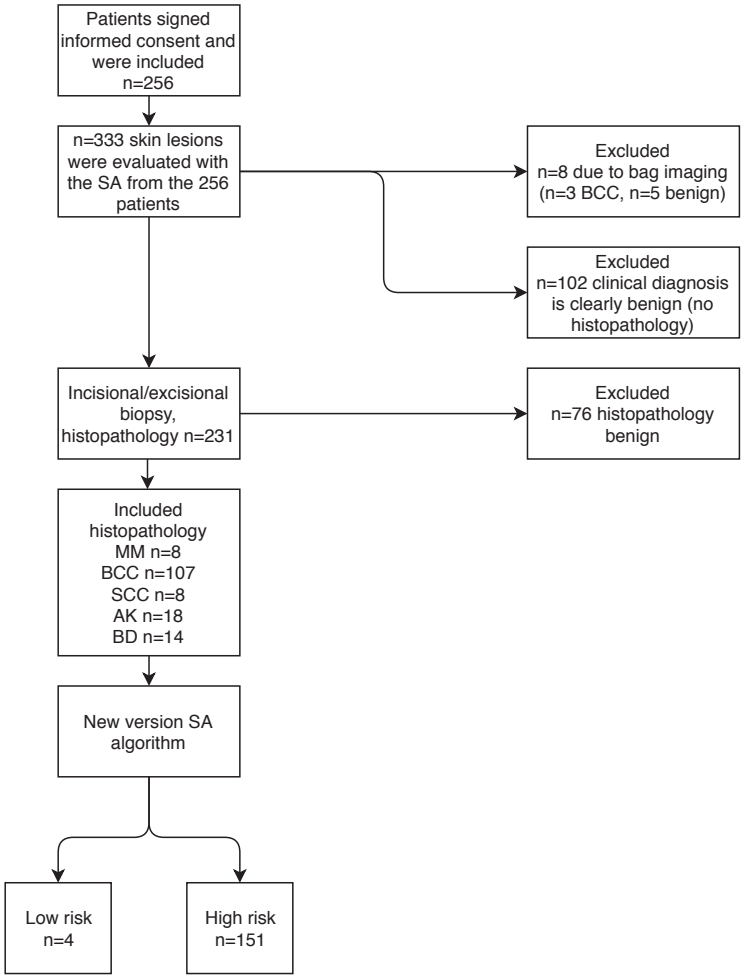


Figure 2. Participant flowchart of malignant cases from the Eindhoven Study.*

* The smartphone application version used to take the images was developed during Munich Study.[11] The new version of the smartphone application algorithm is the version presented in this study. In total, 155 cases were included to test the sensitivity and 186 cases were excluded. The 4 cases denoted as low risk by the smartphone application algorithm had BCC as final histopathology diagnosis.

Abbreviations: AK, actinic keratosis; BCC, basal cell carcinoma; BD, Bowen's disease; SCC, squamous cell carcinoma; MM, malignant melanoma; SA, smartphone application.

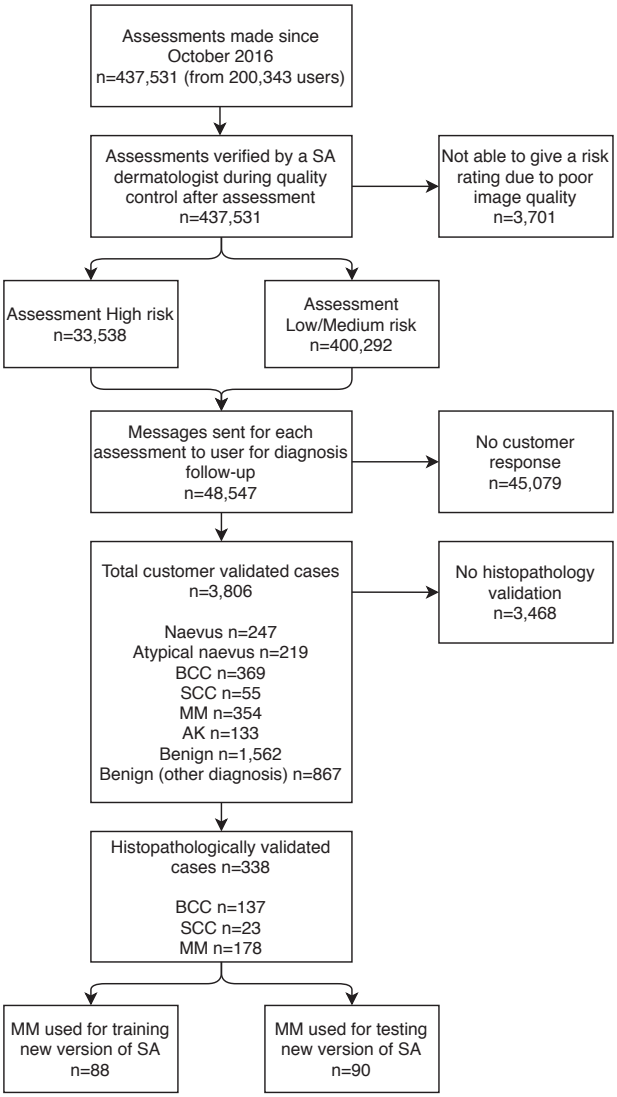


Figure 3. Participant flowchart of malignant cases from the smartphone application user database.*

* Images were taken with the smartphone application algorithm version available at the time of assessment. We included cases for which users contacted the customer service up to October 2018. All users with cases deemed high risk were messaged, a small proportion of low risk cases (n=15,009) were contacted as well, namely if at quality control the dermatologists disagreed with the algorithm result.

Abbreviations: AK, actinic keratosis; BCC, basal cell carcinoma; BD, Bowen's disease; SCC, squamous cell carcinoma; MM, malignant melanoma; SA, smartphone application.

Test dataset for specificity

The test dataset for specificity is based on 6,000 cases from the user database. The participant flowchart for this dataset is given in Figure 4. The 6,000 benign cases included were randomly selected from more than 65,000 consecutive cases in the period June, July and August 2018 and were not included in the training of the algorithm. All cases were assessed, remotely, by an affiliated dermatologist and the cases were included if they were classified as "low risk" by the dermatologist (Supplemental Material Appendix 1).

Algorithm to analyse the skin lesion image

There are several steps needed to analyse the image of the lesion provided by the user in order to calculate a risk degree, namely, lesion segmentation, noise filtering, feature extraction and final risk classification based on the lesion's features.

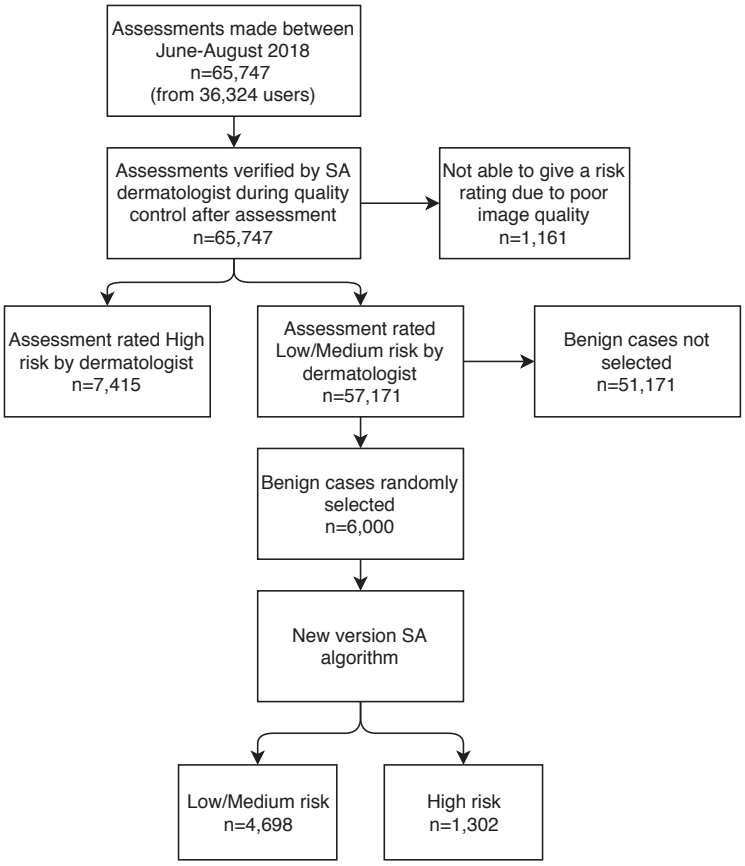


Figure 4. Participant flowchart of benign cases from the smartphone application user database used to calculate specificity.*

* Images were taken with the smartphone application algorithm version available at the time of assessment.
Abbreviations: SA, smartphone application.

Lesion segmentation

The goal of this step is to separate the lesion from the surrounding skin (Supplemental Material Appendix 2 and Figures S1 and S2). This is a difficult task due to the fact that both pigmented and skin coloured lesions can be imaged with the SA. For a precise segmentation we use a Conditional Adversarial Network Algorithm, which is a refined version of the procedure presented by Udrea and Mitra.[13] The network was trained on 5,000 images, with equal class size for pigmented lesions and non-pigmented lesions. For tests we have used 583 images (segmented by a dermatologist). We quantify the segmentation precision considering the sensitivity in detecting pixels that represent the lesion and specificity in correctly identifying pixels representing skin (Supplemental Material Appendix 2 and Table S1).

Noise removal

Once the separation between the lesion (Figure 5 a) and b)) and the skin is done (Figure 5 c) and d)), the area surrounding the lesion is processed and all hair and small lesions (e.g. freckles) are filtered out/eliminated by applying an inpainting procedure on the grey scaled image (Figure 5 e) and f)).[14]

Lesion features extraction

Several geometric, texture and colour descriptors are calculated for each lesion (Supplemental Material Appendix 3). They are the features used in the classification algorithm to compute the risk class associated with each lesion. In total, for risk classification we use 24 colour, texture and shape features.

Support Vector Machine (SVM) classifier

In order to assign a specific lesion to a risk class (low or high), we use a SVM classifier with Radial Basis Functions kernel. For each image we apply the lesion segmentation, noise removal and feature extraction steps in order to compute the 24 features associated with the lesion that are used for classification. The parameters of the SVM classifier are chosen using Particle Swarm Optimization [15]. We find the combination of parameters which maximizes the sensitivity given the constraint that the specificity cannot be lower than 80%.

Statistical evaluation

The sensitivity and specificity, with a 95% Clopper-Pearson confidence interval are computed using the online statistical software MedCalc (https://www.medcalc.org).

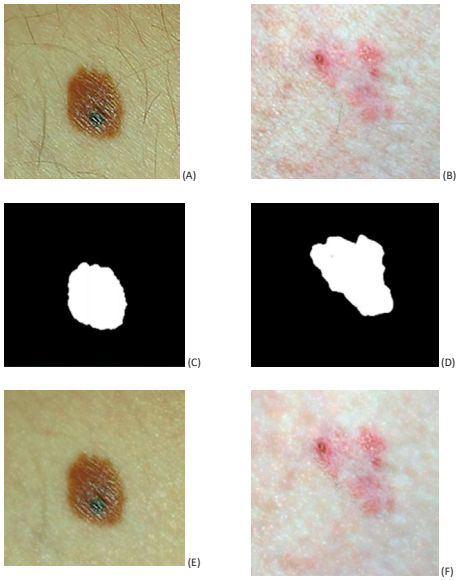


Figure 5. Example of a noise removal procedure for two lesions.*
*The original skin lesion is shown in 5 a) and b), the segmented black and white image in c) and d) and the noise-free image in e) and f).

RESULTS

Image Characteristics

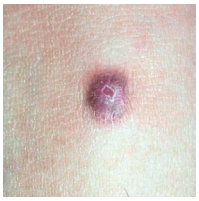
All images from the clinical studies were taken with an iPhone. In the user database most of the images (87%) used for training were taken with iPhones, as the application is more popular for this group of users (Supplemental Material Appendix 4, Table S2).

Demographic Characteristics

All images were taken by users 18 years or older. The images in the user database come from users from several countries, but mainly United Kingdom, the Netherlands, Australia and New Zealand.

Risk Classification

We tested the algorithm for risk classification for sensitivity on 285 images which are linked to a malignant histopathological result (Table 2). In figure 6, we present two MM cases from the user database: one correctly identified and one false negative. Overall, the algorithm has a sensitivity of 95% (95% CI, 91.9% - 97.3%) to detect skin cancer. In particular, the sensitivity to detect melanoma is 92.8% (95% CI, 87.8% - 96.5%) and the sensitivity in detecting KC and their precursors is 97.3% (95% CI, 93.2% - 99.3%). The sensitivity to detect melanoma in the user database is similar to that of the clinical studies. The results of the accuracy for lesion segmentation are presented in Supplemental Material Appendix 2.



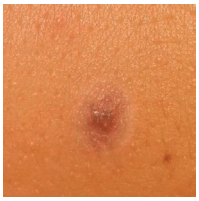


Figure 6. Example of a correctly and incorrectly classified MMs.*
*The skin lesion shown in 6 a) is correctly classified as MM, the skin lesion in 6 b) is incorrectly classified as MM.
Abbreviations: MM, malignant melanoma.

For specificity, we tested the algorithm on 6,000 images clinically validated as benign (without histopathology report). The specificity of the algorithm is 78.3% (95% CI, 77.24%-79.34%).

Table 2. Risk classification algorithm results for different types of benign and skin cancer lesions. [a]

Skin cancer type	**Low risk**	**High risk**	**Sensitivity (95% CI)** [b]
MM [c]	10	128	92.8% (87.8%-96.5%)
BCC	4	103	96.3% (90.7%-99.0%)
SCC	0	8	100% (63.1%-100%)
AK	0	18	100% (81.5%-100%)
BD	0	14	100% (76.8%-100%)
(Pre)malignant cases	14	271	95.1% (91.9%-97.3%)
No skin cancer	**Low risk**	**High risk**	**Specificity (95% CI)** [b]
Clinically Validated Benign cases [d]	4,698	1,302	78.3% (77.2%-9.3%)

[a] MM denotes Malignant Melanoma, BCC denotes basal cell carcinoma, SCC denotes squamous cell carcinoma, AK denotes Actinic Keratosis and BD denotes Bowen's disease.
[b] Sensitivity equals the number of (pre)malignant lesions correctly classified as high risk (True Positives) divided by the total number of (pre)malignant lesions. Specificity equals the number of benign lesions correctly classified as low risk (True Negatives) divided by the total number of benign lesions
[c] Sensitivity for the MM cases from the smartphone app user database is 83/90 = 92.22%, while the sensitivity for MM from clinical studies is 45/48 = 93.75%
[d] Clinically validated by two dermatologists, no histopathological report available.
Abbreviations: AK, actinic keratosis; BCC, basal cell carcinoma; BD, Bowen's disease; SCC, squamous cell carcinoma; MM, malignant melanoma.

DISCUSSION

The SA has a sensitivity to detect various forms of skin cancer including MM, BCC and squamous cell carcinoma (SCC) of 95%, given a 78% specificity. This is a substantial improvement over previous versions of this SA, which had a sensitivity of 80% and a specificity of 78%.[12] The sensitivity to detect MM in the user database is similar to the clinical studies, which indicates that in a context where pictures are taken by lay users instead of the researcher or clinician, the accuracy to detect skin cancer is not necessarily lower.

There were several improvements made in this algorithm in order to obtain a better sensitivity. For the algorithm used in the Eindhoven study[12], only 233 skin lesions were used for training which is likely not sufficient to reflect the full spectrum of skin lesion diversity. In this version, the SA database of more than 130,000 images was used for training. The previous version of the lesion segmentation method was based on Otsu thresholding, while skin lesion classification was rule-based.[11] The current method uses a Conditional Adversarial Network Algorithm (lesion segmentation) combined with a Support Vector Machines (SVM) classifier.[13, 15]

Previous studies on the usage of SAs to detect skin cancer, which included previous versions of this SA and its competitors, criticized their low capability to detect skin cancer.[7-10] However, this is a rapidly evolving field, suggesting frequent re-evaluations are crucial.[1] In this study, we show that the new algorithm for risk assessment has a high accuracy for self-assessment of skin lesions for skin cancer. This result is comparable to some of the recently published, best skin cancer disease classification algorithms.[2-4]

Compared to these studies[2-4], it is more challenging to obtain high accuracy in our setup, as images taken with a smartphone have a lot more variability than clinical or dermoscopic images. Image acquisition within this SA is partially controlled, since the camera module only accepts a picture if it satisfies some minimal quality requirements, namely, the lesion should be present and contained in the image, there should be no shadows or hair covering the lesion, and the picture should be focused and centred. However, the images may be of lower quality, and there is variation associated with angle, distance, luminosity and the characteristics of the smartphone. For these reasons, a large dataset is necessary to cover as many image taking scenarios as possible.

In practice, we believe the level of required diagnostic accuracy for the SA should be higher than that of a regular GP[16-18], and as close as possible to that of a dermatologist. Two recent studies compared the sensitivity of dermatologists and a machine learning algorithm to detect skin cancer, against histopathology. In the first study, 21 dermatologists scored, on average, a sensitivity and specificity higher than 90% and 70%, respectively.[3] In the second study, the average sensitivity of 58 dermatologists is 87%, for a 71% specificity.[4] Though there are some differences in the setup, this is comparable to the level shown by the risk assessment algorithm evaluated here.

For this study we used three skin lesion datasets to test the sensitivity of the SA, two based on clinical studies and one on the user database. The advantage of having data from clinical studies is that participants' lesions are reviewed by a dermatologist and the lesions are clearly linked with the histopathology report. Furthermore, including cases from the user database allows us to compare the performance of the SA in a clinical and real-life setting. The limitation of the clinical studies is that they include mainly high risk skin lesions, therefore they may be inadequate to evaluate the specificity of the SA. The dataset used to test the algorithm for specificity is based on cases from the user database.

The dermatologist assessment was made based on the image taken by the user, without any further information, and the cases were clinically validated as benign without a histopathological report to confirm there is actually no cancer. Therefore, it is possible that the more prominent benign cases were rendered of low risk.

The main risk associated with the usage of this SA by lay users is that a MM or KC is incorrectly classified as low risk (i.e. false negative) and its diagnosis and treatment is delayed. The specificity of the SA is 78%. Even though, the SA does not recommend an immediate visit to the doctor for all high risk cases, usage of this SA may cause detrimental psychological effects on the user, by becoming anxious about a benign lesion or cause unnecessary doctor visits. For a user checking the skin lesions with this SA it may prove difficult to evaluate all relevant skin lesions, as lay users may not know which lesions could be dangerous[19] and/or which the user cannot see (namely, for users taking an image without a partner and/or if the lesions are located in places which are hard to reach). On the other hand, this may also occur in current clinical practice as previous studies suggested GPs usually do not perform a full body skin examination.[20-22]

These risks are mitigated by the SA in several ways. First, every pair of image and corresponding algorithm-based risk rating gets a quality control by a dermatologist. Second, for lesions classified as high risk or for cases upgraded/downgraded by a SA dermatologist, a user will get a message within 48 hours from the Costumer Service team, concerning the degree of urgency of the case (Supplementary Material Appendix 1). Finally, it is also recommended for every user of this application to check their skin regularly.

Beyond the application efficacy, there are still open questions about how mHealth applications for skin lesion assessment should be employed in the health system. This SA is not a tool to diagnose skin cancer, and it doesn't replace the intervention of a medical doctor. The application may have a higher impact in health systems where access to specialized care is difficult due to long waiting times, cost of a medical consultation or geographical distance. For instance, in countries such as the USA, where there may be long waits to see a dermatologist[23, 24], this SA may allow patients to "jump the queue" when the SA shows a high risk rating for a skin lesion, resulting in early diagnosis and management. This SA could also be useful, in countries like the Netherlands, where the GP is the gatekeeper to secondary care, since, 69% of GP visits related to suspicious skin lesions and about 40% of referrals from GP to dermatologist result in a benign diagnosis.[16, 25]

While this SA has the potential, when used by lay persons, to reduce healthcare expenditures and improve the use of resources there are still several steps needed in order to assess the health impact of the SA, before it can be widely implemented. In this study we cannot calculate the true PPV or NPV of the SA. For this, we would need complete follow-up for all SA users, while not all users agree to share their final diagnosis with the SA or decide to visit a doctor for further follow-up. In order to show the health impact of the SA, a possible avenue is to establish partnerships with large insurance companies and carry

out a study (possibly using a quasi-experimental design approach[26]) to test within the insured population whether the SA can reduce the number of doctor consultations and/or whether it could result in a lower average cost of treatment due to earlier detection.

In conclusion, in this study we tested the accuracy of the newest version of the SA for skin lesion risk assessment. The new risk assessment algorithm has a high sensitivity (95%) to detect skin cancer and it may therefore be a valuable tool for early detection of premalignant and malignant skin lesions. The accuracy of the application, namely its specificity, may be further improved over time as more data is available to train the risk classification algorithm.

REFERENCES

1. Ngoo A, Finnane A, McMeniman E, Soyer HP, Janda M. Fighting Melanoma with Smartphones: A Snapshot of Where We are a Decade after App Stores Opened Their Doors. Int J Med Inform. 2018;118:99-112.
2. Codella NCF, Gutman D, Emre Celebi M, Helba B, Marchetti MA, Dusza SW, et al. Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC)2017.
3. 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.
4. Haenssle HA, Fink C, Schneiderbauer R, Toberer F, Buhl T, Blum A, et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Ann Oncol. 2018;29(8):1836-42.
5. SEER. Cancer Stat Facts: Melanoma of the Skin. Available from: https://seer.cancer.gov/statfacts/html/melan.html.
6. Statistics OfN. Cancer survival in England: national estimates for patients followed up to 2017. 2019. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancersurvivalinengland/nationalestimatesforpatientsfollowedupto2017#how-does-stage-at-diagnosis-affect-net-survival.
7. Kassianos AP, Emery JD, Murchie P, Walter FM. Smartphone applications for melanoma detection by community, patient and generalist clinician users: a review. Br J Dermatol. 2015;172(6):1507-18.
8. Ngoo A, Finnane A, McMeniman E, Tan JM, Janda M, Soyer HP. Efficacy of smartphone applications in high-risk pigmented lesions. Australas J Dermatol. 2018;59(3):e175-e82.
9. Rat C, Hild S, Rault Sérandour J, Gaultier A, Quereux G, Dreno B, et al. Use of Smartphones for Early Detection of Melanoma: Systematic Review. J Med Internet Res. 2018;20(4):e135.
10. Wolf JA, Moreau JF, Akilov O, Patton T, English JC, 3rd, Ho J, et al. Diagnostic inaccuracy of smartphone applications for melanoma detection. JAMA Dermatol. 2013;149(4):422-6.
11. Maier T, Kulichova D, Schotten K, Astrid R, Ruzicka T, Berking C, et al. Accuracy of a smartphone application using fractal image analysis of pigmented moles compared to clinical diagnosis and histological result. J Eur Acad Dermatol Venereol. 2015;29(4):663-7.
12. 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.
13. Udrea A, Mitra GD, editors. Generative Adversarial Neural Networks for Pigmented and Non-Pigmented Skin Lesions Detection in Clinical Images. 2017 21st International Conference on Control Systems and Computer Science (CSCS); 2017 29-31 May 2017.
14. Telea A. An Image Inpainting Technique Based on the Fast Marching Method. Journal of Graphics Tools. 2004;9(1):23-34.
15. Li J, Li B, editors. Parameters Selection for Support Vector Machine Based on Particle Swarm Optimization2014; Cham: Springer International Publishing.
16. Ahmadi K, Prickaerts E, Smeets JGE, Joosten V, 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.
17. Goulding JM, Levine S, Blizard RA, Deroide F, Swale VJ. Dermatological surgery: a comparison of activity and outcomes in primary and secondary care. Br J Dermatol. 2009;161(1):110-4.

18. Koelink CJ, Vermeulen KM, Kollen BJ, de Bock GH, Dekker JH, Jonkman MF, et al. Diagnostic accuracy and cost-effectiveness of dermoscopy in primary care: a cluster randomized clinical trial. J Eur Acad Dermatol Venereol. 2014;28(11):1442-9.
19. Aldridge RB, Zanotto M, Ballerini L, Fisher RB, Rees JL. Novice identification of melanoma: not quite as straightforward as the ABCDs. Acta Derm Venereol. 2011;91(2):125-30.
20. Johnson MM, Leachman SA, Aspinwall LG, Cranmer LD, Curiel-Lewandrowski C, Sondak VK, et al. Skin cancer screening: recommendations for data-driven screening guidelines and a review of the US Preventive Services Task Force controversy. Melanoma Manag. 2017;4(1):13-37.
21. Noels EC, Hollestein LM, van Egmond S, Lugtenberg M, van Nistelrooij LPJ, Bindels PJE, et al. Healthcare utilization and management of actinic keratosis in primary and secondary care: a complementary database analysis. Br J Dermatol. 2019.
22. Noels EC, Lugtenberg M, van Egmond S, Droger SM, Buis PAJ, Nijsten T, et al. Insight into the management of actinic keratosis: a qualitative interview study among general practitioners and dermatologists. Br J Dermatol. 2019.
23. Jayakumar KL, Samimi SS, Vittorio CC, Micheletti RG, Steele KT, Ogunleye TA, et al. Expediting patient appointments with dermatology rapid access clinics. Dermatol Online J. 2018;24(6).
24. Kimball AB, Resneck JS, Jr. The US dermatology workforce: a specialty remains in shortage. J Am Acad Dermatol. 2008;59(5):741-5.
25. van Rijsingen MC, Hanssen SC, Groenewoud JM, van der Wilt GJ, Gerritsen MJ. Referrals by general practitioners for suspicious skin lesions: the urgency of training. Acta Derm Venereol. 2014;94(2):138-41.
26. Geldsetzer P, Fawzi W. Quasi-experimental study designs series-paper 2: complementary approaches to advancing global health knowledge. J Clin Epidemiol. 2017;89:12-6.