

Evaluation of the Diagnostic Accuracy of an Online Artificial Intelligence Application for Skin Disease Diagnosis

Oscar ZAAR^{1,2}, Alexander LARSON^{1,2}, Sam POLESIE^{1,2}, Karim SALEH³, Mikael TARSTEDT⁴, Antonio OLIVES⁵, Andrea SUAREZ⁶, Martin GILLSTEDT^{1,2} and Noora NEITTAANMÄKI^{1,2,7}

¹Department of Dermatology and Venereology, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, ²Region Västra Götaland, Sahlgrenska University Hospital, Department of Dermatology and Venereology, Gothenburg, ³Division of Dermatology and Venereology, Department of Clinical Sciences, Lund University, Skåne University Hospital, Lund, ⁴Department of Dermatology, Karlskoga Hospital, Karlskoga, Sweden, ⁵Department of Dermatology, Juaneda General Hospital, Menorca, Spain, ⁶Teladoc Health, Inc., New York, USA, and ⁷Departments of Pathology and Dermatology, Institutes of Biomedicine and Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

Artificial intelligence (AI) algorithms for automated classification of skin diseases are available to the consumer market. Studies of their diagnostic accuracy are rare. We assessed the diagnostic accuracy of an open-access AI application (Skin Image Search™) for recognition of skin diseases. Clinical images including tumours, infective and inflammatory skin diseases were collected at the Department of Dermatology at the Sahlgrenska University Hospital and uploaded for classification by the online application. The AI algorithm classified the images giving 5 differential diagnoses, which were then compared to the diagnoses made clinically by the dermatologists and/or histologically. We included 521 images portraying 26 diagnoses. The diagnostic accuracy was 56.4% for the top 5 suggested diagnoses and 22.8% when only considering the most probable diagnosis. The level of diagnostic accuracy varied considerably for diagnostic groups. The online application demonstrated low diagnostic accuracy compared to a dermatologist evaluation and needs further development.

Key words: artificial intelligence; online diagnostics; dermatology; skin disease.

Accepted Aug 21, 2020; Epub ahead of print Aug 27, 2020

Acta Derm Venereol 2020; 100: adv00260.

Corr: Oscar Zaar, MD, PhD, Department of Dermatology and Venereology, Sahlgrenska University Hospital, Gröna stråket 16, SE-413 45 Gothenburg, Sweden. E-mail: oscar.zaar@vregion.se

Skin diseases are common and one of the leading causes of disability worldwide (1). Population-based studies show similar overall prevalence of skin conditions of approximately 22–35% and as much as 12–15% of visits to primary care in industrialized countries relate to skin problems (2–4). However, access to dermatologists is limited. Artificial intelligence (AI) is an emerging technology with the promise of assisting clinicians in making correct healthcare decisions faster and more reliably (5–9). AI based on convolutional neural networks (CNNs) have shown results of equal or even superior diagnostic accuracy when compared to dermatologists (8, 10–13).

SIGNIFICANCE

In research settings, artificial intelligence algorithms for automated classification of skin diseases have shown promising results with a diagnostic accuracy that have performed equal to or even outperformed dermatologists. Similar online artificial intelligence applications are available to the consumer market, readily accessible to anyone with a smartphone. Nevertheless, external and independent validation investigations of the diagnostic accuracy of these applications are lacking. The studied artificial intelligence application achieved an unsatisfactory overall diagnostic accuracy. The level of diagnostic accuracy varied greatly for diagnostic groups as well as for individual diagnoses. Online automated classification systems should be further developed to ensure appropriate accuracy.

With the increasing development of healthcare applications for smartphones – over 165,000 in 2017 and an increasing number of dermatology-associated smartphone applications (14, 15) – there is a risk that consumers rely on some of these to get answers to their dermatological concerns (16). A recent review of the medical smartphone applications for the assessment of melanoma showed a pooled specificity of 84% and sensitivity of 73% (14). However, various studies evaluating smartphone applications in the use of skin cancer detection have shown poor performances (17–23). While smartphone applications can potentially be informative for the general public, the results of using such systems may generate concern or provide a false sense of reassurance, perhaps leading to missed or delayed diagnosis. Thus, reliable technology and correct information is fundamental. Studies on smartphone applications with a more generalized dermatological assessment covering both skin cancer and inflammatory dermatoses including genital diseases are lacking.

The Skin Image Search™ (SIS) application (iDoc24, CA, USA, available from <https://www.firstderm.com/ai-dermatology/>) is an AI algorithm evolved by training a CNN with a dataset of approximately 58,000 smartphone images obtained from the consumers. The application includes 33 common skin diseases (**Table I**). The pro-

Table I. The included diagnoses. Diagnostic groups and the 33 diagnoses in the Skin Image Search™ application with number of clinical cases in each category

Diagnostic groups and 33 diagnoses in the AI application (ICD code)	Clinical images n (%)
Tumours	176 (33.8)
Basal cell carcinoma (C44.9E)	41 (7.9)
Malignant melanoma (C43.9, C43.9C, D03.9)	31 (6.0)
Nevus, NOS (D22.9X, including D22.9C)	24 (4.6)
Actinic keratosis (L57.0)	24 (4.6)
Seborrhoeic keratosis (L82.9X)	22 (4.2)
Atypical nevus (D22.9F)	16 (3.1)
Lentigo solaris (L81.4)	11 (2.1)
Dermatofibroma (D23.9)	7 (1.3)
Dermal nevus (D22.9C)	0 (0.0)
Dermatitis	48 (9.2)
Dermatitis, NOS (L30.9, including L20.0, L23.9 and L28.0)	45 (8.6)
Seborrhoeic dermatitis (L21.9)	3 (0.6)
Atopic dermatitis (L20.0)	0 (0.0)
Contact dermatitis (L23.9)	0 (0.0)
Neurodermatitis (L28.0)	0 (0.0)
Genital diseases	32 (6.1)
Balanitis (N48.1, N48.1D)	23 (4.4)
Condyloma (A63.0)	9 (1.7)
Herpes simplex (B00.9)	0 (0.0)
Genital herpes (A60.0)	0 (0.0)
Sebaceous glands (N48.89)	0 (0.0)
Follicular diseases	33 (6.3)
Folliculitis (L73.9)	28 (5.4)
Epidermal cyst (L72.0)	3 (0.6)
Furuncle (L02.2)	2 (0.4)
Facial diseases	73 (14.0)
Acne vulgaris (L70.9, L70.0)	54 (10.4)
Rosacea (L71.9)	19 (3.6)
Viral diseases	10 (1.9)
Verruca vulgaris (B07.9)	6 (1.2)
Molluscum contagiosum (B08.1)	4 (0.8)
Fungal diseases	19 (3.6)
Pityriasis versicolor (B36.0)	12 (2.3)
Tinea corporis (B35.4)	7 (1.3)
Psoriasis (L40.9, L40.0)	71 (13.6)
Lichen ruber planus (L43.0)	18 (3.5)
Urticaria (L50.9)	15 (2.9)
Insect bite(s) (T14.0D)	15 (2.9)
Post-infl. hyperpigmentation (L81.0)	11 (2.1)

gram is free, available online, and can be accessed with a web browser (24). The user uploads an image showing a skin condition, and the online service searches its dataset for matching images and provides a list with the top 5 most likely differential diagnoses, in falling order from 1–5, hereby referred to as the “top 5”. According to the information provided by the company, the AI algorithm has shown results of 40% diagnostic accuracy in ranking the correct diagnosis as the top most likely diagnosis. The accuracy in the top 5, i.e. the presence of the true diagnosis among the 5 proposed differential diagnoses, is claimed to be 80% (25). Nevertheless, independent and systematic validation of the application has not been performed.

The aim of the study was to analyse the diagnostic accuracy of the SIS application.

MATERIALS AND METHODS

This study was designed as an observational study carried out at the Department of Dermatology and Venereology at Sahlgrenska University Hospital in Gothenburg from April 2018 to May 2019.

The images included were obtained from patients giving their consent to have their skin conditions photographed. Clinical photos were taken by professional health care workers using smartphone cameras (iPhone 8 plus, Apple Inc., CA, USA). Dermoscopy images were not included.

The data collection was done by consecutive sampling, a method in which images were collected until the desired sample size of at least 20 images per diagnosis of the 33 different diagnoses available in the SIS application and/or a total of 500 images were acquired. The diagnosis had to be clinically confirmed by a dermatologist and/or histopathological examination. All images with an ambiguous diagnosis were excluded. All images were anonymous since they contained no personal data and the patient could not be recognized by the image. When necessary this was achieved by cropping the images in Microsoft Paint (Microsoft corporation, WA, USA). Images of poor quality or images portraying different kinds of solitary lesions (e.g. nevus near a seborrhoeic keratosis) were excluded or if possible cropped in a way that made the confirmed diagnosis become the main visible lesion in the image. Non-assessable images due to low image quality (e.g. bad lighting or blurry focus) or unfit image composition (e.g. camera not aimed at the lesion, image taken from an inadequate distance or angle) were excluded. Images showing a ruler or pen marking (e.g. circles or arrows) were excluded, as well as an applied bandage or sticking plaster partially or completely covering the lesion in question. Images showing lesions undergoing some kind of surgical intervention (e.g. shaving, excision or biopsy), or bleeding following such treatment, were also excluded.

The following additional information was collected from the patient charts: patient ID, age, gender, skin type, image type (overview or macro), body site and final diagnosis (defined as the clinical diagnosis or, when available, the histopathological diagnosis).

Collected images were uploaded for automated classification by the SIS application using a demonstration version that used the same algorithm as the online version. The demonstration version was provided by the developer and was more flexible compared to the regular version available online, allowing the classification process to be made by uploading a single image, instead of two images used in the online version. However, according to the developer, the online version analyses only one selected image and thus there was assumed to be no difference in the accuracy between the demonstration version and the online version.

The top 5 differential diagnoses provided by the SIS application were imported into Microsoft Excel (Microsoft corporation, WA, USA). Matching the image's confirmed diagnosis with the top 5, each classification was given a score from 1–5, depending on which position the confirmed diagnosis had. If the confirmed diagnosis was absent from the top 5, the classification was given a score of 0 and marked as incorrectly diagnosed. If multiple differential diagnoses were correct for the same image, the best score was chosen. The overall diagnostic accuracy was analysed as well as the diagnostic accuracy for separate diagnoses and diagnostic groups with a common aetiology (e.g. tumours, viral diseases and fungal diseases) or body site (e.g. genital diseases and facial dermatoses).

All data were analysed using R version 3.0.3 (The R Foundation for Statistical Computing, Vienna, Austria). Fisher's exact test was applied in the analysis of contingency tables. Ethical approval was granted by the ethical review board of Linköping (approval number: 2019-02186).

RESULTS

A total of 891 clinical images were collected. Of these, 241 images were excluded according to the exclusion cri-

teria above and 129 images were dismissed as duplicates, resulting in a final number of 521 images obtained from 215 patients. Often several images were included from the same patient. The median age for the population (131 males [60.9%] and 84 females [39.1%]) was 53.4 years (range 1–94 years) with the three most common body sites being trunk, face/neck, and lower extremities, accounting for 23.0%, 21.5%, and 16.9%, respectively. The skin phototypes were the following: type I 16.7%, II 59.5%, III 17.2%, IV 4.2%, V 0.9%, VI 1.4%. The patients' diagnoses were confirmed by histopathological examination in 52.6% (113/215) of the cases and in the remaining 47.4% (102/215) cases the diagnoses were made clinically by dermatologists.

In the combined results of all 521 images uploaded for classification, the AI algorithm was able to classify the correct diagnosis among the top 5 in 56.4% (294/521) of the images and failed to give the correct diagnosis in 43.6% (227/521) of the images. The SIS application gave the true diagnosis ranked as the most likely in 22.8% (119/521) of the images (**Fig. 1**).

When focusing only on diagnostic groups containing several diagnoses, the SIS application had most success in accurately classifying (i.e. the correct diagnosis present in the top 5) dermatitis (91.7%, 44/48) and genital diseases (87.5%, 28/32), followed by facial diseases (68.9%, 51/73) and tumours (67.0%, 118/176). The least successful in diagnostic accuracy were viral diseases (40.0%, 4/10), fungal diseases (42.1%, 8/19) and follicular diseases (48.5%, 17/33) (**Fig. 2**).

For individual diagnosis, it was easiest for the system to detect dermatitis, tumours (e.g. nevus, atypical nevus, malignant melanoma and seborrhoeic keratosis), common genital diseases (e.g. balanitis and condyloma) and the facial disease acne, all having a diagnostic accuracy of $\geq 75\%$ in the top 5, (**Fig. 3**). Acne had the highest frequency of number 1 scores (66.7%, 36/54) followed by balanitis (65.2%, 15/23) and seborrhoeic keratosis (59.1%, 13/22). Inferior results were shown in diagnostic accuracy for psoriasis (11.3%, 8/71), lichen planus (11.1%, 2/18) and post-inflammatory hyperpigmentation (27.3%, 3/11). Interestingly, melanoma also had a relatively low frequency of number 1 scores (9.7%, 3/31).

DISCUSSION

A total of 521 images were collected and classified by the online open-access AI application. The results show an overall diagnostic accuracy of 56.4% in the top 5, and 22.8% accuracy for the most probable diagnosis (i.e. score=1). This would probably be interpreted as a poor diagnostic performance if translated into use in clinical practice as a diagnostic aid.

Several studies have shown poor agreement between dermatologist and online smartphone application with low sensitivity and specificity (18, 21–23). Our results broaden the research field by assessing not only skin cancer detection with smartphone applications but also including inflammatory and genital diseases.

Computer vision has successfully been used in other medical fields including radiology (26). In a recent survey 77.3% of dermatologists agreed that AI will improve dermatology (27). However, there are still challenges in the development of AI in dermatology. In contrast to radiology, where all the images are taken with professional equipment, photographs used in dermatological diagnostics vary in quality and can be taken by both health care professionals and patients themselves. The variation in the quality of the images can cause problems when training and testing the AI algorithms.

The majority of patients with skin conditions are handled and treated by primary care physicians. Since access to specialized care is limited, only the minority of patients visit a dermatologist. The images collected in this study were all gathered from the Department of Dermatology and Venereology at Sahlgrenska University Hospital, and it is questionable whether these images are representative of the images taken by the general public which were used to develop the SIS application. One could expect that the clinical images used in this study portrayed more advanced or rare variations of skin diseases. Such imparity could explain why some diagnoses had very low diagnostic accuracy. An example could be psoriasis, for which the AI app achieved very low accuracy with only 11.3% (8/71) of the images in the top 5. Further, it has been observed in earlier studies that diagnostic accuracy tends to drop when tested on an independent dataset (28).

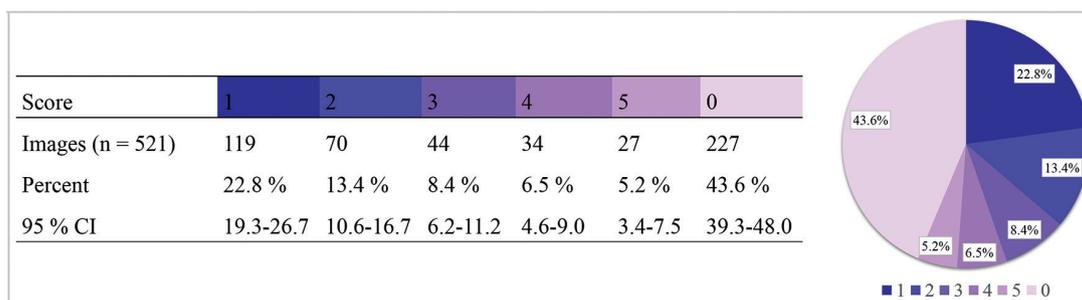
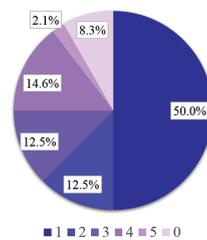


Fig. 1. The overall diagnostic accuracy of the Skin Image Search™ application. The scores 0–5 are colour-coded by different shades of blue with darker shades portraying more accurate scores. The artificial intelligence application was able to place the correct diagnosis among the top 5 in 56.4% (294/521) of the images and failed to give the correct diagnosis in 43.6% (227/521) of the images. CI: confidence interval.

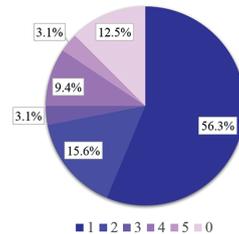
Dermatitis

Score	1	2	3	4	5	0
Images (n = 48)	24	6	6	7	1	4
Percent	50.0 %	12.5 %	12.5 %	14.6 %	2.1 %	8.3 %
95 % CI	35.2-64.8	4.7-25.2	4.7-25.2	6.1-27.8	0.1-11.1	2.3-20.0



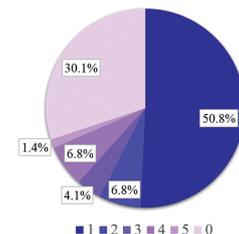
Genital diseases

Score	1	2	3	4	5	0
Images (n = 32)	18	5	1	3	1	4
Percent	56.3 %	15.6 %	3.1 %	9.4 %	3.1 %	12.5 %
95 % CI	37.7-73.6	5.3-32.8	0.1-16.2	2.0-25.0	0.1-16.2	3.5-29.0



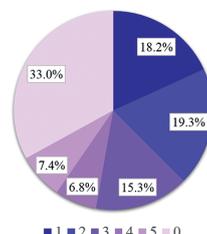
Facial diseases

Score	1	2	3	4	5	0
Images (n = 73)	37	5	3	5	1	22
Percent	50.8 %	6.8 %	4.1 %	6.8 %	1.4 %	30.1 %
95 % CI	38.7-62.6	2.7-15.3	0.9-11.5	2.3-15.3	0.0-7.4	19.9-42.0



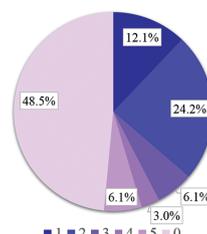
Tumors

Score	1	2	3	4	5	0
Images (n = 176)	32	34	27	12	13	58
Percent	18.2 %	19.3 %	15.3 %	6.3 %	7.4 %	33.0 %
95 % CI	12.8-24.7	13.8-25.9	10.4-21.5	3.6-11.6	4.0-12.3	26.1-40.4



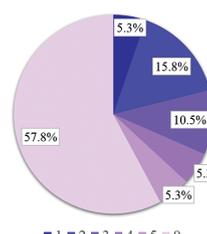
Follicular diseases

Score	1	2	3	4	5	0
Images (n = 33)	4	8	2	1	2	16
Percent	12.8 %	24.2 %	6.1 %	3.0 %	6.1 %	48.5 %
95 % CI	3.4-28.2	11.1-42.3	0.7-20.2	0.1-15.8	0.7-20.2	30.0-66.5



Fungal diseases

Score	1	2	3	4	5	0
Images (n = 19)	1	3	2	1	1	11
Percent	5.3 %	15.3 %	10.5 %	5.3 %	5.3 %	57.8 %
95 % CI	0.1-26.0	3.4-39.6	1.3-33.1	0.1-26.0	0.1-26.0	33.5-79.7



Viral diseases

Score	1	2	3	4	5	0
Images (n = 10)	0	3	0	1	0	6
Percent	0.0 %	30.0 %	0.0 %	10.0 %	0.0 %	60.0 %
95 % CI	0.0-30.8	6.7-65.2	0.0-30.8	0.3-44.5	0.0-30.8	26.2-87.8



Fig. 2. Diagnostic accuracy for the diagnostic groups.

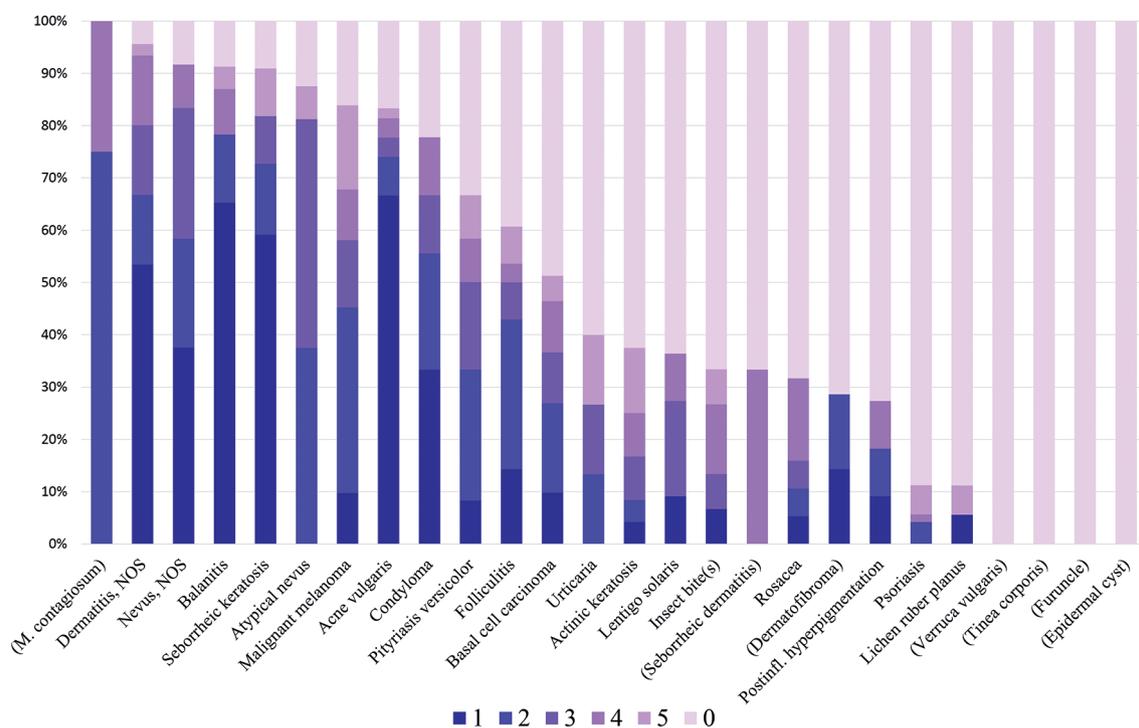


Fig. 3. Diagnostic accuracy for individual diagnoses. The bars are encoded with different colours depending on the given scores (0–5). Diagnoses with <10 images uploaded for classification are presented in parentheses.

Clinical photos, including images taken by smartphones, can vary significantly in terms of image factors such as distance to object, angle, quality and light, making the classification process considerably more challenging (16, 21, 26). Esteva et al. (10) overcame this challenge to some extent by demonstrating a new approach in CNN for melanoma classification based on clinical images without restrictions to man-made criteria (e.g. asymmetry, border, colour and size) that matched the diagnostic performance of dermatologists. Without the restriction of specialized equipment, such systems could potentially provide widespread automated diagnostic aid for all medical practitioners and would also be available for anyone with access to a smartphone (14). However, further development is needed before these applications can be used by consumers.

Strengths in our study comprised a considerably large number of clinical images, compared to similar studies on smartphone applications (17–21). The selection process of images was carried out with consecutive sampling and predefined inclusion and exclusion criteria. Histopathological verification of the diagnoses was available in more than half of the patients. In the rest of the cases the diagnoses were made by the specialists working in a university hospital. However, a limitation was that we did not use a standardized consensus opinion including several dermatologists for the diagnoses that were not histologically verified. Unfortunately, not all 33 diagnoses in the AI application were represented and 8 diagnoses were represented by fewer than 10 images,

making it difficult to draw any conclusions for these individual diagnoses. The images were not evenly distributed among the individual diagnoses, which could create a bias in the overall accuracy approaching those holding more images. As many as 241 images were excluded, which may illustrate the presence of a large percentage of clinical images of poor quality and inadequate image composition. A limitation was that all the images were collected in the same department.

Future studies are warranted in order to more accurately evaluate the studied application, foremost having enough data to represent all diagnoses. It is important to make independent studies of diagnostic accuracy for online classification systems within dermatology and venereology since they are available for anyone in the general population and could have an impact on the users' health. Further, because the development of such systems will be continuous, studies have to be renewed and updated to keep up with the current quality of such services.

In conclusion, the AI application in this study achieved an unsatisfactory level of overall diagnostic accuracy. The level of diagnostic accuracy varied greatly for diagnostic groups as well as for individual diagnoses. Our study examined one of the online automated classification systems available to consumers, and users of this application are advised to approach this technology with curiosity about the developing field and with an understanding that this system does not replace the need for clinical examination by clinicians.

Conflict of interest: OZ, KS, AO, AS and MAT previously worked as consultants for iDoc24, the company developing Skin Image Search™.

REFERENCES

1. Karimkhani C, Dellavalle RP, Coffeng LE, Flohr C, Hay RJ, Langan SM, et al. Global skin disease morbidity and mortality: an update from the Global Burden of Disease Study 2013. *JAMA Dermatol* 2017; 153: 406–412.
2. Lindberg M, Isacson D, Bingeors K. Self-reported skin diseases, quality of life and medication use. A nationwide pharmaco-epidemiological survey in Sweden. *Acta Derm Venereol* 2014; 94: 188–191.
3. Rea JN, Newhouse ML, Halil T. Skin disease in Lambeth. A community study of prevalence and use of medical care. *Br J Prev Soc Med* 1976; 30: 107–114.
4. Augustin M, Herberger K, Hintzen S, Heigel H, Franzke N, Schafer I. Prevalence of skin lesions and need for treatment in a cohort of 90 880 workers. *Br J Dermatol* 2011; 165: 865–873.
5. Janda M, Soyer HP. Can clinical decision making be enhanced by artificial intelligence? *Br J Dermatol* 2019; 180: 247–248.
6. Quer G, Muse ED, Nikzad N, Topol EJ, Steinhubl SR. Augmenting diagnostic vision with AI. *Lancet* 2017; 390: 221.
7. Lim BCW, Flaherty G. Artificial intelligence in dermatology: are we there yet? *Br J Dermatol* 2019; 181: 190–191.
8. Miller DD, Brown EW. Artificial Intelligence in Medical Practice: The Question to the Answer? *Am J Med* 2018; 131: 129–133.
9. Suzuki K: Overview of deep learning in medical imaging. *Radiol Phys Technol* 2017; 10: 257–273.
10. Esteve 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: 115–118.
11. Han SS, Park GH, Lim W, Kim MS, Na JI, Park I, et al. Deep neural networks show an equivalent and often superior performance to dermatologists in onychomycosis diagnosis: Automatic construction of onychomycosis datasets by region-based convolutional deep neural network. *PLoS One* 2018; 13: e0191493.
12. Marchetti MA, Codella NCF, Dusza SW, Gutman DA, Helba B, Kalloo A, et al. Results of the 2016 International Skin Imaging Collaboration International Symposium on Biomedical Imaging challenge: Comparison of the accuracy of computer algorithms to dermatologists for the diagnosis of melanoma from dermoscopic images. *J Am Acad Dermatol* 2018; 78: 270–277.
13. Fujisawa Y, Otomo Y, Ogata Y, Nakamura Y, Fujita R, Ishitsuka Y, et al. Deep-learning-based, computer-aided classifier developed with a small dataset of clinical images surpasses board-certified dermatologists in skin tumour diagnosis. *Br J Dermatol* 2019; 180: 373–381.
14. Buechi R, Faes L, Bachmann LM, Thiel MA, Bodmer NS, Schmid MK, et al. Evidence assessing the diagnostic performance of medical smartphone apps: a systematic review and exploratory meta-analysis. *BMJ Open* 2017; 7: e018280.
15. Steeb T, Wessely A, French LE, Heppt MV, Berking C. Skin cancer smartphone applications for German-speaking patients: review and content analysis using the mobile app rating scale. *Acta Derm Venereol* 2019; 99: 1043–1044.
16. Ransohoff JD, Li S, Sarin KY. Assessment of accuracy of patient-initiated differential diagnosis generation by google reverse image searching. *JAMA Dermatol* 2016; 152: 1164–1166.
17. Nabil R, Bergman W, Kukutsch NA. Conflicting results between the analysis of skin lesions using a mobile-phone application and a dermatologist's clinical diagnosis: a pilot study. *Br J Dermatol* 2017; 177: 583–584.
18. 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: e175–e182.
19. 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: 422–426.
20. Dorairaj JJ, Healy GM, McInerney A, Hussey AJ. Validation of a Melanoma Risk Assessment Smartphone Application. *Dermatol Surg* 2017; 43: 299–302.
21. Chung Y, van der Sande AAJ, de Roos KP, Bekken MW, de Haas ERM, Kelleners-Smeets NWJ, et al. Poor agreement between the automated risk assessment of a smartphone application for skin cancer detection and the rating by dermatologists. *J Eur Acad Dermatol Venereol* 2019; 34: 274–278.
22. Chuchu N, Takwoingi Y, Dinnes J, Matin RN, Bassett O, Moreau JF, et al. Smartphone applications for triaging adults with skin lesions that are suspicious for melanoma. *Cochrane Database Syst Rev* 2018; 12: CD013192.
23. 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.
24. Internet, available from: <https://www.firstderm.com/ai-dermatology/>.
25. Internet, available from: <https://www.firstderm.com/best-practice-artificial-intelligence-algorithm-skin/>.
26. Bi WL, Hosny A, Schabath MB, Giger ML, Birkbak NJ, Mehrtash A, et al. Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA Cancer J Clin* 2019; 69: 127–157.
27. Polesie S, Gillstedt M, Kittler H, Lallas A, Tschandl P, Zalaudek I, et al. Attitudes towards artificial intelligence within dermatology: an international online survey. *Br J Dermatol* 2020; 183: 159–161.
28. Mollersén K, Zortea M, Schopf TR, Kirchesch H, Godtliëben F. Comparison of computer systems and ranking criteria for automatic melanoma detection in dermoscopic images. *PLoS One* 2017; 12: e0190112.