

Artificial Intelligence in Dermatology: Assessing Predictability in Clinical Diagnosis



Madina Mohamed Hubail^{1,*}, Ahmed Abdel Khabir¹, Doaa Shokry Al Emam² and Sara Hamdy Fouad¹

¹Dermatology, Andrology & STDs Department, Faculty of Medicine, Mansoura University, Mansoura, Egypt

²Public Health and Preventive Medicine, Faculty of Medicine, Mansoura University, Mansoura, Egypt

Abstract:

Introduction: The use of Artificial Intelligence (AI) for image-based diagnosis in dermatology is increasing rapidly. The clinical accuracy of AI in diagnosing different skin conditions remains under evaluation. This study aimed to evaluate the diagnostic performance of an AI application in comparison to confirmed clinical diagnoses by dermatologists.

Method: A cross-sectional study was carried out on 400 patients with different skin conditions, including acne, alopecia, eczema, pigmentary disorders, psoriasis, immunological disorders, tumors, infections, and infestations. The study analyzed AI-based predictions using the Tibot AI application, comparing them against dermatologists' diagnoses.

Results: The AI application demonstrated high diagnostic accuracy for certain dermatological conditions such as adnexal disorders (AUC 0.93-0.98), pigmentary disorders (AUC 0.88-0.94), and cutaneous tumors (AUC 0.87-0.95). Sensitivity for adnexal disorders was 88.9% (top one) and 94.4% (top three), and for Pigmentary disorders, it was 75.8% and 87.9% for top one and top-three predictions, respectively. However, AI performance was lower for immunological disorders (31.3% sensitivity) and cutaneous infestations (22.2%). Overall accuracy improved across all conditions when considering the top-three predictions.

Discussion: Tibot AI-application demonstrated high diagnostic accuracy for conditions with distinct morphological features such as adnexal, pigmentary disorders, and cutaneous tumors. It showed lower sensitivity for immunological disorders and infestations, indicating the need for further AI training with more diverse datasets.

Conclusion: AI-based diagnostic accuracy improved significantly when considering the top-three diagnoses, indicating its value as a differential diagnostic tool. It showed promising accuracy in adnexal, pigmentary disorders, and cutaneous tumors. However, it is less robust for immunological skin diseases and infections, highlighting the need for further refinement.

Keywords: Artificial Intelligence, Deep learning, Machine learning, Dermatology, Adnexal disorders, Pigmentary disorders.

© 2025 The Author(s). Published by Bentham Open.

This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Public License (CC-BY 4.0), a copy of which is available at: <https://creativecommons.org/licenses/by/4.0/legalcode>. This license permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

*Address correspondence to this author at the Dermatology, Andrology & STDs, Mansoura University, Mansoura, Egypt; Tel: +973 33005606; E-mail: madina.hubail@gmail.com

Cite as: Hubail M, Abdel Khabir A, Al Emam D, Fouad S. Artificial Intelligence in Dermatology: Assessing Predictability in Clinical Diagnosis. Open Dermatol J, 2025; 19: e18743722410124.
<http://dx.doi.org/10.2174/0118743722410124250901081508>



Received: May 09, 2025
Revised: May 31, 2025
Accepted: July 25, 2025
Published: September 02, 2025



Send Orders for Reprints to
reprints@benthamscience.net

1. INTRODUCTION

The application of AI in medical image assessment has significantly enhanced the diagnostic accuracy, reduced physicians' workload, and minimized diagnostic errors, thereby improving disease prediction and detection [1]. AI systems function intelligently and autonomously, enabling them to anticipate and address challenges as they arise [2]. Their strength lies in their ability to analyze vast multidimensional datasets, extracting patterns that can support precise clinical decision-making. Additionally, AI models are dynamic, capable of adapting to new inputs and continuously refining their performance [3].

Machine learning (ML), a crucial AI subfield, has transformed healthcare by facilitating disease diagnosis, drug discovery, and risk assessment. Recent advances in big data and electronic medical records have further strengthened ML applications [4]. Neural networks and fuzzy logic algorithms play a pivotal role in automating predictive analysis and diagnostic processes [5].

In dermatology, AI is particularly valuable due to the field's reliance on morphological and visual pattern recognition. AI can leverage extensive databases of clinical, dermatoscopic, and histopathological images to improve diagnostic accuracy [6, 7]. AI-based tools have demonstrated efficacy in the early detection of skin malignancies, inflammatory dermatoses, pigmentary disorders, and hair abnormalities, thus augmenting dermatologists' diagnostic capabilities [8].

One such AI tool, Tibot, provides dermatological assessments by analyzing uploaded images and relevant clinical data. However, studies evaluating its diagnostic accuracy compared to dermatologists remain essential [9]. The demand for automated AI-driven diagnosis is increasing due to variability in dermatological presentations, unequal access to specialists, and the necessity for timely and precise diagnosis [10].

2. METHODS

A cross-sectional study was conducted at the Dermatology, Andrology & STDs Department, Mansoura University Hospitals, Egypt, from November 2023 to November 2024 to evaluate the predictability of artificial intelligence in dermatological diagnosis. The ethical approval was provided by the Institutional Review Board of the medical research ethics committee at Mansoura University, Egypt with acceptance code MS:23.12.2668. All patients provided informed consent to participate.

The study enrolled 400 patients with acne, rosacea, alopecia, tumors, eczema, immunological skin disorders, pigmentary disorders, psoriasis, and infections (bacterial, fungal, viral, and infestations). Exclusion criteria included prior dermatological treatment or refusal to consent to image use.

The data collection involved clinical evaluation, dermoscopy (using a DermLite DL5), and laboratory or histopathological assessments when needed. Lesion

images were captured using an iPhone 15 Pro Max under controlled lighting. Two dermatologists reviewed diagnoses before images were uploaded to Tibot AI, which generated three top probabilistic diagnoses per case. A summary of the methodology is shown in Fig S1.

The diagnostic performance of Tibot AI-application was statistically assessed using sensitivity, specificity, predictive values, Receiver Operating Characteristic (ROC), and Area Under the Curve (AUC). Tibot diagnoses were compared to dermatologist-confirmed diagnoses to assess the accuracy of AI.

3. RESULTS

This study included 400 patients attending the Dermatology, Andrology & STDs Department at Mansoura University Hospitals. The median age was 33 years (range: 0.5 - 79 years), with 56.3% males and 43.8% females. The majority of participants (42.5%) were between 20 and 40 years old, followed by 26.8% in the 40-60 age group.

The most prevalent category of skin conditions was cutaneous infections and infestations (26.0%), followed by inflammatory disorders (25.8%), adnexal disorders (16.8%), cutaneous tumors (15.3%), pigmentary disorders (8.3%), and immunological skin disorders (8.0%). Among cutaneous infections, bacterial infections accounted for 5.5% viral infections 9.3% fungal infections 9.0%, and 2.3% skin infestations

The diagnostic performance of the AI application in identifying various skin conditions is outlined as follows. For bacterial infections, sensitivity is 36.4% in the top one prediction, rising to 72.7% in the top three, with a specificity of 99.5% and an overall accuracy of 96%. In viral infections, sensitivity increases from 67.6% in the top one to 86.5% in the top three, while specificity is 95.3%, and total accuracy reaches 92.8%. Fungal infections exhibit a sensitivity of 58.3% in the top one prediction, improving to 69.4% in the top three, with a specificity of 93.7% and an overall accuracy of 90.5% Table 1.

For skin infestations, sensitivity remains low at 22.2% for both top one and top three predictions; however, specificity is high at 99%, with an overall accuracy of 97.3%. In the diagnosis of benign tumors, sensitivity increases from 79.1% in the top one to 95.3% in the top three, with a specificity of 95.5% and an overall accuracy of 93.8%. Malignant (suspicious) tumors demonstrate a sensitivity of 77.8% in the top one, improving to 88.9% in the top three, with a specificity of 99.2% and an overall accuracy of 98.3%.

For pigmentary disorders, sensitivity is 75.8% in the top one, increasing to 87.9% in the top three, with a specificity of 99.2% and a total accuracy of 97.3%. Immunological skin disorders have a sensitivity of 31.3% in the top one prediction, rising to 56.3% in the top three, with a specificity of 97% and an overall accuracy of 91.8%. Sensitivity for psoriasis improves from 69% in the top one to 84.5% in the top three, with a specificity of 97.1% and an accuracy of 93.3%.

Table 1. The sensitivity, specificity, positive predictive value, negative predictive value, and total accuracy of artificial intelligence software for the diagnosis of various dermatoses in the study group (N=400).

Skin conditions	Sensitivity		Specificity	PPV	NPV	Accuracy
	Top one	Top three				
I. Cutaneous infections and infestations	36.4%	72.7%	99.5%	80%	96.4%	96%
■ Bacterial infection	67.6%	86.5%	95.3%	59.5%	96.6%	92.8%
■ Viral infection	58.3%	69.4%	93.7%	47.7%	95.8%	90.5%
■ Fungal infection	22.2%	22.2%	99%	33.3%	98.2%	97.3%
■ Skin infestations (parasitic and insect bites)						
II. Cutaneous tumors	79.1%	95.3%	95.5%	68%	97.4%	93.8%
■ Benign tumor	77.8%	88.9%	99.2%	82.4%	99%	98.3%
■ Malignant (suspicious) tumor						
III. Pigmentary disorders	75.8%	87.9%	99.2%	89.3%	97.8%	97.3%
IV. Immunological skin disorders	31.3%	56.3%	97%	47.6%	94.2%	91.8%
V. Inflammatory disorders	69%	84.5%	97.1%	80%	94.9%	93.3%
■ Psoriasis	62.2%	86.7%	91%	46.7%	95%	87.8%
■ Eczema						
VI. Adnexal disorders	88.9%	94.4%	97.5%	78%	98.9%	96.8%
■ Acne and rosacea	90.3%	96.8%	99.2%	90.3%	99.2%	98.5%
■ Alopecia						

PPV: positive predictive value. NPV: negative predictive value.

For eczema, sensitivity increases from 62.2% in the top one to 86.7% in the top three, with a specificity of 91% and an overall accuracy of 87.8%. Both acne/rosacea and alopecia exhibit high sensitivity, starting at 88.9% in the top one and improving to 94.4% in the top three, with specificity values of 97.5% and overall accuracies of 96.8% for both conditions (Tables 1 and 2).

Area Under the Curve (AUC) was the highest for the Adnexal disorders (0.930.98), followed by benign cutaneous tumors (AUC: 0.87-0.95) and pigmentary disorders (AUC: 0.88-0.94), indicating high diagnostic ability. Values were the lowest for immunological disorders (AUC = 0.64-0.77) (Table 2 and Fig. 1).

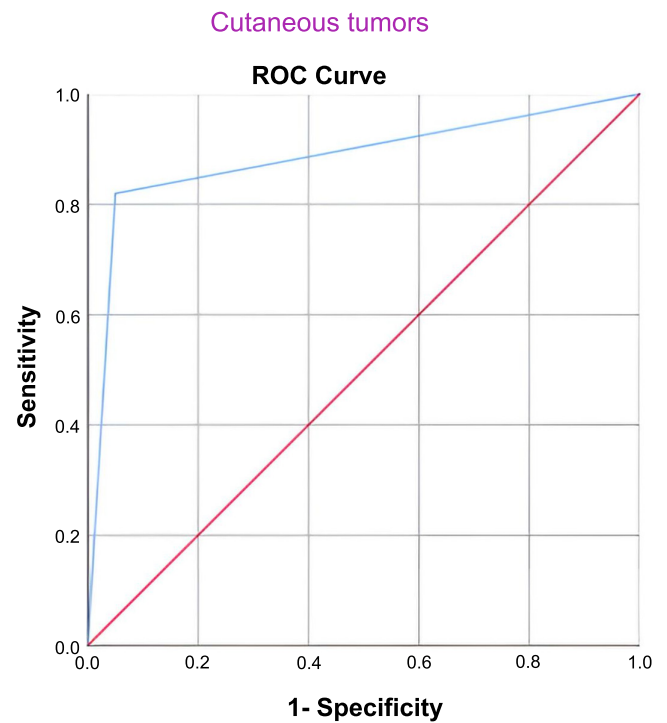
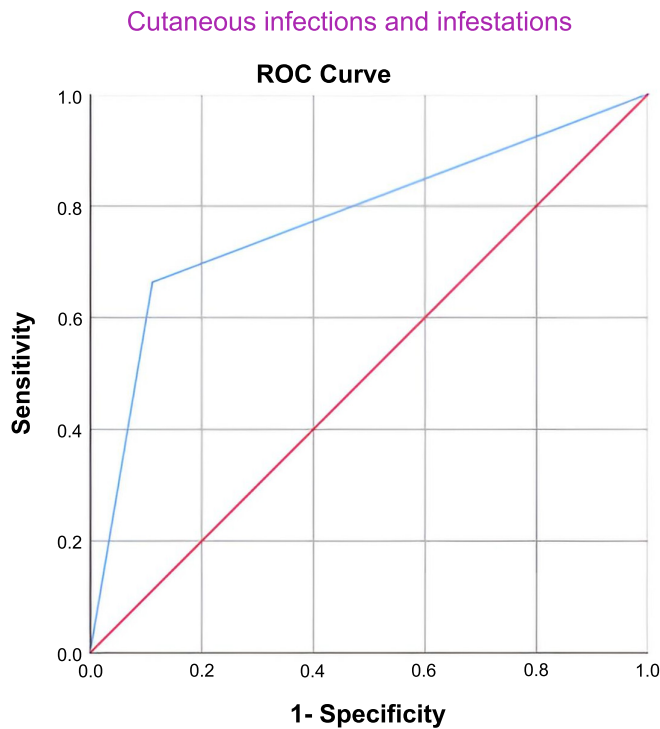


Fig. 1 contd....

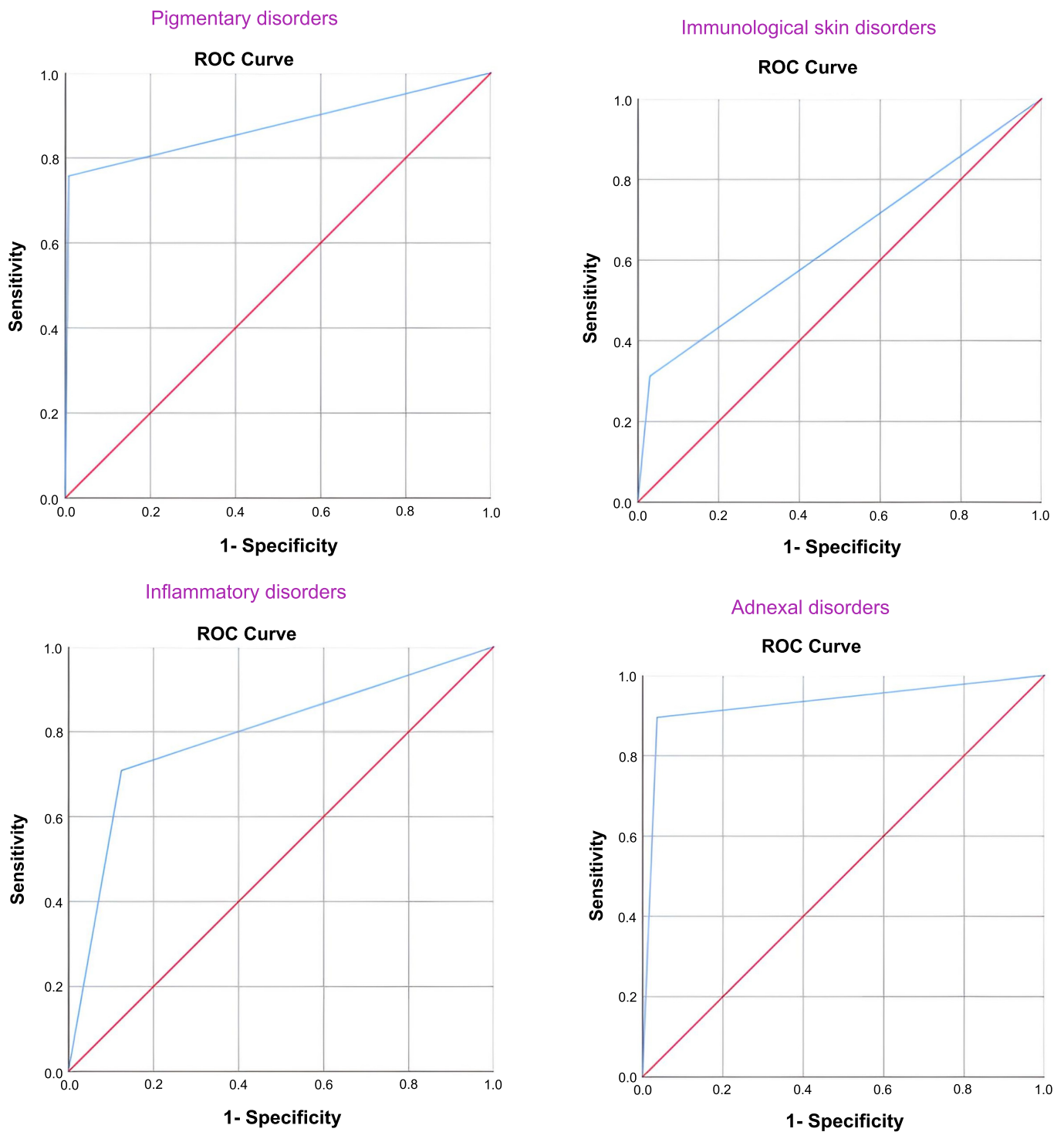


Fig. (1). ROC curves for artificial intelligence software for the diagnosis of the main categories of skin conditions in the study group.

The application of Artificial Intelligence demonstrated high sensitivity and positive predictive value (PPV) (Fig. 2) for adnexal disorders, particularly acne and rosacea (sensitivity 88.9%, PPV 78%) and alopecia (sensitivity and

PPV both 90.3%), indicating strong diagnostic performance. Similarly, AI showed good sensitivity for psoriasis (69%) and suspicious tumors (77.8%), with high PPVs of 80% and 82.4%, respectively.

Table 2. Receiver operating characteristic (ROC) for artificial intelligence software for the diagnosis of main categories of skin conditions in the study group (N=400).

Skin conditions	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV	Accuracy
I. Cutaneous infections and infestations	0.78 (0.72 - 0.83)	66.3%	88.9%	67.6%	88.3%	76.3%
II. Cutaneous tumors	0.89 (0.83 - 0.94)	82%	95%	74.6%	96.7%	93.8%
III. Pigmentary disorders	0.88 (0.79 - 0.96)	75.8%	99.2%	89.3%	97.8%	97.3%
IV. Immunological skin disorders	0.64 (0.53 - 0.76)	31.3%	97%	47.6%	94.2%	91.8%
V. Inflammatory disorders	0.79 (0.74 - 0.85)	70.9%	87.5%	66.4%	89.7%	83.3%
VI. Adnexal disorders	0.93 (0.89 - 0.97)	89.6%	96.4%	83.3%	97.9%	95.3%

PPV: positive predictive value. NPV: negative predictive value.

		Actual diagnosis												Total	PPV
		Acne and Rosacea	Alopecia	Bacterial infection	Benign tumor	Eczema	Fungal infection	Immunological skin disorder	Infestation	Pigmentation	Psoriasis	Suspicious tumor	Viral infection	Total	PPV
Predicted diagnosis	Acne and rosacea	32	0	0	2	2	0	2	0	2	0	0	1	41	78%
	Alopecia	0	28	0	0	0	2	0	1	0	0	0	0	31	90.3%
	Bacterial infection	0	0	8	1	1	0	0	0	0	0	0	0	10	80%
	Benign tumor	0	1	2	34	0	2	1	1	3	2	1	3	50	68%
	Eczema	2	0	5	0	28	3	11	5	0	2	1	3	60	46.7%
	Fungal infection	1	0	1	0	8	21	2	0	1	6	0	4	44	47.7%
	Immunological skin disorder	1	2	1	0	0	2	10	0	1	3	0	1	21	47.6%
	Infestation	0	0	0	0	1	0	1	2	0	2	0	0	6	33.3%
	Pigmentation	0	0	0	2	0	1	0	0	25	0	0	0	28	89.3%
	Psoriasis	0	0	0	0	3	2	4	0	0	40	1	0	50	80%
	Suspicious tumor	0	0	0	1	0	0	0	0	0	2	14	0	17	82.4%
	Viral infection	0	0	5	3	2	3	1	0	1	1	1	25	42	59.5%
Total		36	31	22	43	45	36	32	9	33	58	18	37	400	
Sensitivity		88.9%	90.3%	36.4%	79.1%	62.2%	58.3%	31.3%	22.2%	75.8%	69%	77.8%	67.6%		

Fig. (2). Confusion matrix of actual diagnosis versus artificial intelligence–predicted top diagnosis, along with sensitivity and positive predictive value for individual skin conditions. Dark blue cells represent true positives.

Conversely, bacterial infections had a low sensitivity of 36.4% but a high PPV of 80%, suggesting that while AI correctly identifies positive cases, it frequently misses true infections. Eczema (sensitivity 62.2%, PPV 46.7%) and fungal infections (sensitivity 58.3%, PPV 47.7%) had relatively low PPVs, indicating a higher number of false-positive results. Immunological skin disorders had the lowest sensitivities (31.3%) and a PPV of 47.6%, reflecting difficulties in AI-based identification.

Skin infestations had the weakest performance, with a sensitivity of 22.2% and a PPV of 33.3%, highlighting the need for algorithm improvements in detecting these conditions. Pigmentary disorders were well identified (sensitivity 75.8%, PPV 89.3%), while viral infections showed moderate performance (sensitivity 67.6%, PPV 59.5%).

4. DISCUSSION

Artificial intelligence is revolutionizing dermatology, with the Tibot app enhancing diagnostic accuracy, especially for conditions with distinct morphological features. Powered by convolutional neural networks (CNNs), Tibot analyzes skin lesion images, providing predictions that align closely with clinical evaluations, making it a valuable tool in dermatological practice [11].

In this study, the median age was 33 years (0.5-79), with 42.5% aged 20-40 years and 26.8% aged 40-60 years, showing a slight male predominance (56.3% males, 43.8% females).

Infections are common at 20-40 years, while scabies affects all ages [12]. Benign tumors peak at 40-60 years, melanomas increase with age, especially in men [13, 14].

Melasma is more frequent in women aged 20-40 years, while vitiligo can occur at any age [15]. Lupus and pemphigus are the most common in women 20-60 years [16]. Psoriasis, eczema, acne, and rosacea peak in early adulthood, while alopecia and male-pattern baldness increase with age [17-19]. These trends align with our study. Overall, our study group's age and male predominance align with conditions common in adults aged 20-60 years.

In this current study, cutaneous infections and infestations constitute the largest group (26.0%), with viral (9.3%) and fungal infections (9.0%) being the most common, underscoring the need for public health initiatives targeting these issues. Inflammatory disorders, including psoriasis (14.5%) and eczema (11.3%), account for 25.8%, highlighting the burden of chronic conditions that require long-term management. Adnexal disorders, such as acne and rosacea (9.0%) and alopecia (7.8%), collectively represent 16.8%, reflecting the impact of hormonal and environmental factors.

Cutaneous tumors are also significant in the present study, with benign tumors (10.8%) being more prevalent than suspicious malignant ones (4.5%), emphasizing the importance of early detection and screening. Additionally, pigmentary disorders (8.3%) and immunological skin conditions (8.0%) highlight diverse dermatological challenges, suggesting a need for specialized care across all categories.

The Tibot AI app demonstrated strong diagnostic accuracy, particularly for adnexal disorders like acne/rosacea and alopecia (AUC: 0.93-0.98). It also performed excellently in identifying cutaneous tumors (AUC: 0.87-0.95) and showed high diagnostic reliability for pigmentary disorders (AUC: 0.88-0.94) and inflammatory conditions such as psoriasis and eczema (AUC: 0.77-0.91) (Table 2, Fig. 1).

However, the model performed less reliably for skin infestations (AUC = 0.61) and immunological disorders (AUC = 0.64-0.77), possibly due to overlapping clinical features or limited representation in the training dataset.

Notably, the improved accuracy across all categories with top-three predictions highlights the app's potential in differential diagnosis. For example, bacterial infections showed a significant increase in AUC from 0.68 (95% CI: 0.54-0.82) (top one) to 0.86 (95% CI: 0.75-0.97) (top three), demonstrating its ability to capture a broader diagnostic spectrum with added flexibility.

Overall, the Tibot AI app demonstrated high diagnostic accuracy for distinct skin conditions, such as adnexal disorders and cutaneous tumors (AUC > 0.90 for top-one and top-three predictions), while immunological disorders and skin infestations had comparatively lower accuracy.

In this study, the AI software demonstrated strong diagnostic performance, particularly for adnexal disorders (acne, rosacea, alopecia) and cutaneous tumors. Sensitivities for top-one and top-three predictions were 88.9% and 90.3% for adnexal disorders and 79.1% and 77.8% for tumors, with high specificities (94.4%-96.8% for

adnexal disorders and 95.3%-88.9% for tumors). Pigmentary disorders also showed high accuracy (97.8%-97.3%). However, lower sensitivity and positive predictive values for immunological disorders and cutaneous infections suggest areas for improvement. Overall, the software demonstrated high accuracy, reinforcing its potential as a reliable diagnostic tool for dermatology.

The confusion matrix revealed valuable insights into the AI model's performance across different skin conditions. Adnexal disorders had the highest sensitivity (89.6%) and a strong positive predictive value (PPV) (83.3%), highlighting the AI's robust capability in diagnosing acne and alopecia. Similarly, cutaneous tumors showed good diagnostic performance, with a sensitivity of 82% and a PPV of 74.6%, demonstrating effectiveness in detecting both benign and malignant growths.

Pigmentary disorders had a high PPV (89.3%), indicating accurate identification when predicted by the AI, though sensitivity was lower at 75.8%. Inflammatory disorders and cutaneous infections showed moderate sensitivity (70.9% and 66.3%, respectively) with PPVs of 66.4% and 67.6%, suggesting room for improvement. Immunological skin disorders had the lowest sensitivity (31.3%) and PPV (47.6%), highlighting a key limitation. Overall, while the AI demonstrates strong diagnostic potential for many skin conditions, further refinement and additional data are needed to enhance its accuracy, particularly for less prevalent or complex cases.

A study by Marri *et al.* assessing Tibot AI's diagnostic accuracy reported near-perfect performance for adnexal disorders, with top-three prediction accuracies of 98.6% (acne/rosacea) and 100% (alopecia), and top-one accuracies of 91.7% and 97.7%, respectively [20]. Suspicious tumors also showed strong accuracy (100% for top-three and 81.8% for top-one predictions). Eczema and pigmentary disorders had high top-three accuracies (100% and 98.8%), though their top-one predictions were lower (75% and 88.5%). Bacterial and fungal infections had top-three accuracies of 83.3% and 96.5%, but lower top-one values (50% and 82.9%). Viral infections and immunological disorders showed moderate performance, with top-three accuracies of 94.5% and 95%, and top-one values of 63% and 75%. Psoriasis and skin infestations had the lowest top-one accuracies (70.2% and 68.7%), though their top-three predictions were higher (91.4% and 93.7%) [20].

Also, Marri *et al.* reported that Tibot AI demonstrated varied diagnostic performance across different skin conditions. The software exhibited strong sensitivity for eczema (91.66%) and fungal infections (96.85%), indicating a high detection ability for these conditions. Alopecia achieved perfect sensitivity (100%) and positive predictive value (PPV) (100%), reflecting exceptional diagnostic accuracy. However, viral infections had the lowest sensitivity, at 26.66%, highlighting potential limitations in accurately detecting these conditions. Overall, Tibot showed high sensitivity and PPV for many common skin disorders, though performance varied for

more complex or less distinct conditions such as immunological disorders and viral infections [20].

The results of Marri *et al.* are broadly consistent with our findings regarding the conditions with the highest and lowest prediction accuracies. However, their study reported higher accuracy values overall [20]. These discrepancies may be attributed to differences in study design, sample size, and inherent biases. Our study's reliance on confirmed clinical diagnoses by three different dermatologists minimizes the likelihood of misclassification, thereby enhancing its real-world applicability. Additionally, the larger sample size in the study by Marri *et al.* (on 600 clinical images vs. 400 patients in our study) may have contributed to their observed higher accuracy values.

In a study by Patil *et al.*, the AI software Tibot demonstrated varied diagnostic performance across different skin conditions. It showed particularly strong sensitivity for eczema (91.66%) and fungal infections (96.85%), indicating a high ability to detect these conditions. Alopecia achieved perfect sensitivity (100%) and a positive predictive value (PPV) of 100%, suggesting exceptional accuracy in diagnosing this condition. In contrast, viral infections had the lowest sensitivity at 26.66%, reflecting limitations in detecting these cases accurately. The PPV was generally high for conditions such as alopecia (100%), eczema (94.3%), and infestations (94.44%), indicating that when Tibot predicted a positive result, it was more likely to be correct. However, immunological disorders showed a lower PPV of 42.10%, highlighting challenges in accurately diagnosing these conditions [9].

Overall, Tibot demonstrated high sensitivity and PPV for many common and distinct skin conditions, though there was variability in performance for more complex or less distinct diagnoses, such as immunological disorders and viral infections [9]. Notably, the study of Patil *et al.* was sponsored and supported by Polyfins Technology Inc., the company developing Tibot, introducing a potential conflict of interest bias that may have influenced the reported outcomes [9].

A recent study evaluating the diagnostic sensitivity of a machine learning model in dermatology on 100 cases found that its top-1 accuracy (39%) was lower than that of general practitioners (64%) and dermatologists (72%). The model performed best in detecting benign lesions (with 96% sensitivity in Top-5), malignant tumors (83.5%), and infectious diseases (75%). Unlike our study, this research focused on comparing the accuracy of AI to that of human diagnosis [21].

A recent study used a machine learning (ML) model to analyze images from 100 clinical cases in order to evaluate its sensitivity in diagnosing various dermatological conditions. The performance of the ML model was compared to that of general practitioners and dermatologists in terms of diagnostic accuracy and found that the top-1 accuracy of the ML model (39%) was lower than that of GPs (64%) and dermatologists (72%).

Diagnostic sensitivity of benign lesions was the highest at 96% as the Top-5 predictable diagnoses, followed by 83.5% sensitivity in detecting malignant skin tumors. For infectious diseases, the model's sensitivity in the Top-5 was 75% [21]. Unlike our study, this research focused on comparing the accuracy of AI to that of human diagnosis.

Convolutional neural networks-based AI models show strong potential in dermatological diagnostics. Wu *et al.* reported high accuracy for eczema (92.57%) and psoriasis (89.46%), outperforming Tibot, which had lower top-1 sensitivities (62.2% and 69%, respectively) [22]. Fujisawa *et al.*, achieved 76.5% accuracy in classifying benign and malignant tumors, while Tibot excelled in malignant tumor prediction (with 98.3% accuracy, and 77.8% sensitivity) [23].

5. LIMITATIONS

- The study was conducted in a single tertiary care center, limiting the demographic and geographic diversity of the sample.
- Certain underrepresented conditions, particularly immunological and infectious dermatoses, showed reduced diagnostic sensitivity, which might be due to insufficient training datasets.
- The cross-sectional design did not allow assessment of how AI performs over time or adapts to clinical changes during disease progression or treatment.

SUGGESTIONS FOR FUTURE RESEARCH

- Multicenter studies should be conducted, involving ethnically and geographically diverse populations to enhance generalizability.
- Longitudinal designs need to be implemented to evaluate the consistency of AI diagnostics over time and across treatment stages.
- Clinical metadata and AI training must be integrated to enhance the diagnostic accuracy for immunological disorders and skin infestation cases.
- More pediatric and elderly populations should be recruited to test the AI's performance across a wider age spectrum.

CONCLUSION

The Tibot AI application demonstrated high accuracy in diagnosing adnexal disorders and cutaneous tumors and pigmentary disorders, but performed less reliably in case of immunological disorders and skin infestations. Its sensitivity improved with top-three predictions, highlighting its potential as a diagnostic aid rather than a standalone tool. Enhancing AI training with diverse datasets could improve performance, especially for conditions with overlapping features. Integrating AI with clinical expertise is essential for optimizing dermatological diagnostics and patient care.

AUTHORS' CONTRIBUTIONS

The authors confirm their contribution to the paper as follows: M.H.: Writing - reviewing and editing; A.A.K.:

Validation; D.A.E.: Analysis and interpretation of results; S.F.: Review, methodology;. All authors reviewed the results and approved the final version of the manuscript.

LIST OF ABBREVIATIONS

AI	=	Artificial Intelligence
ML	=	Machine learning
AUC	=	Area Under the Curve
ROC	=	Receiver Operating Characteristic

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The ethical approval was provided by the Institutional Review Board of the medical research ethics committee at Mansoura University, Egypt acceptance code no. MS:23.12.2668.

HUMAN AND ANIMAL RIGHTS

No animals were used in this research. All procedures performed in studies involving human participants were in accordance with the ethical standards of institutional and/or research committee and with the 1975 Declaration of Helsinki, as revised in 2013.

CONSENT FOR PUBLICATION

All patients provided informed consent to participate.

STANDARDS OF REPORTING

STROBE guidelines were followed.

AVAILABILITY OF DATA AND MATERIAL

The data supporting the findings of the article is available within the article.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none.

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publisher's website along with the published article.

REFERENCES

- [1] Ghaffar Nia N, Kaplanoglu E, Nasab A. Evaluation of artificial intelligence techniques in disease diagnosis and prediction. *Discov Artif Intell* 2023; 3(1): 5. <http://dx.doi.org/10.1007/s44163-023-00049-5> PMID: 40478140
- [2] Shukla Shubhendu S, Vijay J. Applicability of artificial intelligence in different fields of life. *Int J Sci Eng Res* 2013; 1(1): 28-35.
- [3] Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med* 2019; 17(1): 195. <http://dx.doi.org/10.1186/s12916-019-1426-2> PMID: 31665002
- [4] Sahu M, Gupta R, Ambasta RK, Kumar P. Artificial intelligence and machine learning in precision medicine: A paradigm shift in big data analysis. *Prog Mol Biol Transl Sci* 2022; 190(1): 57-100. <http://dx.doi.org/10.1016/bs.pmbts.2022.03.002> PMID: 36008002
- [5] Yin J, Ngiam KY, Teo HH. Role of artificial intelligence applications in real-life clinical practice: Systematic review. *J Med Internet Res* 2021; 23(4): e25759. <http://dx.doi.org/10.2196/25759> PMID: 33885365
- [6] De A, Sarda A, Gupta S, Das S. Use of artificial intelligence in dermatology. *Indian J Dermatol* 2020; 65(5): 352-7. http://dx.doi.org/10.4103/ijid.IJD_418_20 PMID: 33165383
- [7] Jartarkar SR, Patil A, Wollina U, *et al.* New diagnostic and imaging technologies in dermatology. *J Cosmet Dermatol* 2021; 20(12): 3782-7. <http://dx.doi.org/10.1111/jocd.14499> PMID: 34652880
- [8] Li Z, Koban KC, Schenck TL, Giunta RE, Li Q, Sun Y. Artificial intelligence in dermatology image analysis: Current developments and future trends. *J Clin Med* 2022; 11(22): 6826. <http://dx.doi.org/10.3390/jcm11226826> PMID: 36431301
- [9] Patil A, Patil S, Rao DN, Basar F, Bate S. Assessment of tibot® artificial intelligence application in prediction of diagnosis in dermatological conditions: Results of a single centre study. *Indian Dermatol Online J* 2020; 11(6): 910-4. http://dx.doi.org/10.4103/idoj.IDOJ_61_20 PMID: 33344338
- [10] Bajwa MN, Muta K, Malik MI, *et al.* Computer-aided diagnosis of skin diseases using deep neural networks. *Appl Sci* 2020; 10(7): 2488. <http://dx.doi.org/10.3390/app10072488>
- [11] Sengupta D. Artificial intelligence in diagnostic dermatology: Challenges and the way forward. *Indian Dermatol Online J* 2023; 14(6): 782-7. http://dx.doi.org/10.4103/idoj.idoj_462_23 PMID: 38099026
- [12] Bezie Z, Deboch B, Ayele D, *et al.* Common skin diseases. *Int Dev* 2005; 13: 200-4.
- [13] Nikolaou V, Stratigos AJ. Emerging trends in the epidemiology of melanoma. *Br J Dermatol* 2014; 170(1): 11-9. <http://dx.doi.org/10.1111/bjd.12492> PMID: 23815297
- [14] Khan M, Martin-Clavijo A. Benign and malignant conditions of the skin. *Scott-Brown's Otorhinolaryngology and Head and Neck Surgery*. CRC Press 2018; pp. 1321-36.
- [15] Majid I, Aleem S. Melasma: Update on epidemiology, clinical presentation, assessment, and scoring. *J Skin Stem Cell* 2021; 8(4): 10-20. <http://dx.doi.org/10.5812/jssc.120283>
- [16] Egami S, Yamagami J, Amagai M. Autoimmune bullous skin diseases, pemphigus and pemphigoid. *J Allergy Clin Immunol* 2020; 145(4): 1031-47. <http://dx.doi.org/10.1016/j.jaci.2020.02.013> PMID: 32272980
- [17] Silverwood RJ, Forbes HJ, Abuabara K, *et al.* Severe and predominantly active atopic eczema in adulthood and long term risk of cardiovascular disease: Population based cohort study. *BMJ* 2018; 361: k1786. <http://dx.doi.org/10.1136/bmj.k1786> PMID: 29792314
- [18] Anastassakis K, Anastassakis K. Hormonal and genetic etiology of male androgenetic alopecia. *Androgenetic Alopecia From A to Z*. Springer 2022; pp. 135-80. http://dx.doi.org/10.1007/978-3-030-76111-0_11
- [19] Plewig G. Acne and rosacea. *Braun-Falco's Dermatology*. Springer 2009; pp. 993-1017. http://dx.doi.org/10.1007/978-3-540-29316-3_67
- [20] Marri SS, Inamadar AC, Janagond AB, Albadri W. Analyzing the predictability of an artificial intelligence app (Tibot) in the diagnosis of dermatological conditions: A cross-sectional study. *JMIR Dermatol* 2023; 6(1): e45529. <http://dx.doi.org/10.2196/45529> PMID: 37632978
- [21] Escalé-Besa A, Vidal-Alaball J, Miró Catalina Q, Gracia VH, Marin-Gomez FX, Fuster-Casanovas A. The use of artificial intelligence for skin disease diagnosis in primary care settings: A systematic review. *Healthcare* 2024; 12(12): 1192. <http://dx.doi.org/10.3390/healthcare12121192> PMID: 38921305

[22] Wu H, Yin H, Chen H, *et al.* A deep learning, image based approach for automated diagnosis for inflammatory skin diseases. *Ann Transl Med* 2020; 8(9): 581.
<http://dx.doi.org/10.21037/atm.2020.04.39> PMID: 32566608

[23] Fujisawa Y, Otomo Y, Ogata 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(2): 373-81.
<http://dx.doi.org/10.1111/bjd.16924> PMID: 29953582