Current state of machine learning for non-melanoma skin cancer

Ajay Sharma MD,[1] Samantha Shwe BS, [1] Natasha Atanaskova Mesinkovska MD, PhD[1]

[1]University of California Irvine School of Medicine, Department of Dermatology, Irvine, CA, USA

Background

- Artificial intelligence (AI) has and will transform the practice of dermatology, providing an additional clinical tool for diagnostic, prognostic, and therapeutic predictions.
- Within the realm of AI lies machine learning (ML) and its subset deep learning (DL), techniques which employ complex algorithms on large data sets to generate models capable of these predictions [1].
- ML and DL have been increasingly utilized for skin cancer screening, primarily of melanomas, but also of non-melanoma skin cancers (NMSC).

Objective

- This study presents the first quantitative review of the success of these techniques in NMSC screening.

Methods

- Adhering to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines, a search of the online databases MEDLINE, arXiv, and PubMed was performed.
- All studies involving ML techniques and NMSC screening were captured in the initial search, comprising 537 articles.
- 85 full-text articles were read, resulting in 77 articles that met final inclusion criteria.
- Of these, 52 manuscripts were included.

Table 1. Mean sensitivity, specificity, and area under the curve.

	N	Mean (%)	Standard Deviation (%)	Range (%)
Sensitivity	52	89.15	7.66	70.00-100.00
Specificity	44	81.13	21.90	12.00-100.00

Table 2. Mean sensitivity and specificity for skin cancer type, algorithm type, diagnostic standard, dataset source, and dataset size.

	N	Sensitivity (mean)	Specificity (mean)
Skin Cancer Type			
Basal Cell Carcinoma	44	89.19%	84.66%
Both	26	88.97%	75.44%
Squamous Cell Carcinoma	5	83.10%	97.60%
Algorithm Type			
Neural Network	25	87.57%	78.76%
Support Vector Machine	10	89.30%	87.84%
Classification	7	91.16%	81.50%
Segmentation	5	87.02%	88.83%
Other	5	96.08%	69.50%
Diagnostic Standard			
Histopathology	41	88.65%	78.75%
Dermoscopy	28	91.22%	85.43%
Dataset Source			
Private	24	89.52%	74.85%
Public	22	88.02%	85.20%
Dataset Size			
0-100	7	89.67%	92.38%
101-500	11	88.13%	67.78%
501-1,000	8	91.78%	79.63%
1,000-10,000	15	90.51%	85.41%
>10,000	9	86.28%	80.80%

Results

- Mean sensitivity and specificity of included manuscripts were calculated (Table 1).
- Basal cell carcinoma had an 89% sensitivity, 85% specificity; squamous cell carcinoma had an 83% sensitivity, 98% specificity for diagnosis by machine learning (Table 2).

Discussion

- Reviews have reported the sensitivity and specificity of basal cell carcinoma (BCC) diagnosis as 89% and 86% and squamous cell carcinoma (SCC) diagnosis as 79 and 87% [2, 3].
- Compared to the results of our study, there were no significant differences between the sensitivity and specificity of ML techniques and dermoscopy (p-values > 0.05).
- Thus, there is insufficient evidence to conclude that an ML algorithm is superior at NMSC screening than a trained dermatologist utilizing dermoscopy for either BCC or SCC.
- With AI technology easily scalable to non-dermatologists or less experienced physicians, the performance of computer-aided diagnosis of NMSC remains promising.

References

1. Murphree DH, Puri P, Shamim H,, et al. Journal Pre-proof Deep Learning for Dermatologists: Part I Fundamental Concepts. J Am Acad Dermatol 2020;doi: 10.1016/j.jaad.2020.05.056.

2. Reiter O, Mimouni I, Gdalevich M, et al. The diagnostic accuracy of dermoscopy for basal cell carcinoma: A systematic review and meta-analysis. J Am Acad Dermatol 2019;80:1380–8.

3. Lallas A, Argenziano G, Zendri E, et al. Expert Review of Anticancer Therapy Update on non-melanoma skin cancer and the value of dermoscopy in its diagnosis and treatment monitoring 2014;doi: 10.1586/era.13.38.