Dermatology Image Quality Assessment (DIQA): Artificial intelligence to ensure the clinical utility of images for remote consultations and clinical trials

To the Editor: Dermatological imaging is extensively used both in clinical practice and clinical trials, especially for diagnosis and severity assessment. The problem is that images require a minimum visual quality in order to be analyzed, be it by a doctor or an algorithm. Images that lack visual quality can derail the clinical process, disrupt clinical trials, and pose a risk to patient safety.[1,2] Clinical images are widely used in dermatology to capture the state of conditions in a non-invasive way. However, these images are subject to huge variability in lightning, distance to the lesion, focus, and other factors, which impact the semantic content and therefore the clinical usefulness of an image. To solve this problem, we developed an artificial intelligence–based tool that gauges the dermatological image quality and ensures the quality and clinical utility of images during remote consultations and clinical trials.

We gathered a dataset comprising 934 dermatology images (clinical and dermoscopic) taken with smartphones, digital cameras, and dermatoscopes. The images were evaluated by 40 non-expert observers, according to International Telecommunication Union's recommendations (ITU-T P.910), in addition to following an evaluation protocol that considered factors such as lightning, focus, or distance to the lesion of interest. Based on these criteria, the observers rated every image a final quality score from 1 (worst) to 10 (best). In the end, each image had a mean opinion score that reflected the overall opinion of the 40 observers.

We also used other image quality assessment datasets to make sure our models were presented with as many different types of distortions (either real or artificial) as possible, as some may not be observed in our dataset. These datasets included mean opinion scores from their own observer groups that were transformed from their original scale into our 1 to 10 scale.

We split each dataset into a training and validation set and trained a convolutional neural network to predict the quality score of images (Fig 1).[3] The model was tested on the dermatological image validation set by comparing the output quality score to the original mean opinion score using the mean absolute error (MAE). We also evaluated performance (Table I) in terms of linear correlation and Spearman's rank correlation.

The first experiments used transfer learning: a model pre-trained on the general-domain datasets was fine-tuned on the dermatological dataset. The second experiment used all datasets at once during training.

From the first set of experiments, the model that was pretrained on real distorted images and fine-tuned in dermatological images yielded the best correlation. However, the model from the second experiment (trained with all the datasets) produced an even lower mean absolute error and comparable correlation.

In future work, we will expand our clinical image dataset with artificial distortions to consider other artifacts such as motion blur and compression, as in the general-domain datasets. We will also conduct experiments with dermatologists instead of non-expert observers to assess not only visual quality but also clinical meaning.

Dermatology image quality assessment shows promise as a quality-check tool that improves remote consultation and clinical trials, especially those in which patients with skin pathologies report their condition remotely. Furthermore, artificial intelligence can be implemented into computer-aided diagnosis systems to ensure that only high-quality images are uploaded.

Ignacio Hernández Montilla, MSc,[a] *Taig Mac Carthy, MSc,*[b] *Andy Aguilar, MBA,*[b] *and Alfonso Medela, MSc*[a]

From the Department of Medical Data Science, Legit.Health, Bilbao, Spain[a]*; Department of Clinical Endpoint Innovation, Legit.Health, Bilbao, Spain.*[b]

Funding sources: None.

IRB approval status: Not applicable.

Patient consent: Not applicable.

Key words: artificial intelligence; image quality assessment; perceptual image quality; remote care; teledermatology.

Reprints not available from the authors.

Correspondence to: Ignacio Hernández Montilla, MSc, Plaza de Ingenieros Torres Quevedo 1, Legit.Health, 48013, Bilbao, Spain

E-mail: ignaciohernandez@legit.health

Conflicts of interest

The authors state no conflict of interest.

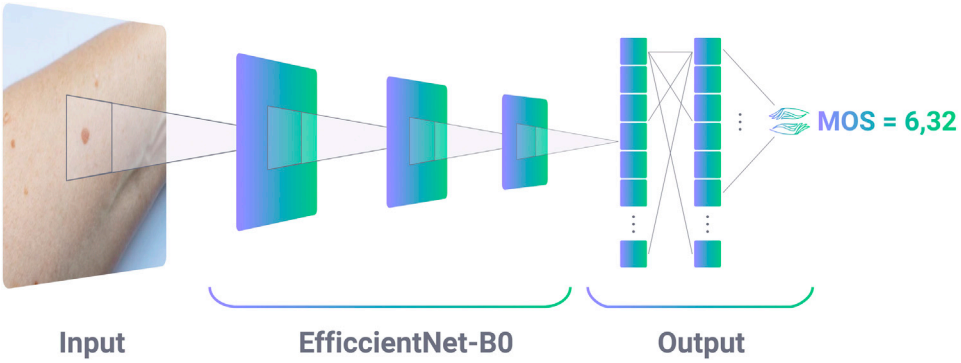

Fig 1. Dermatology image quality assessment. A deep learning model (EfficientNet-B0) processes an image and outputs the predicted perceived quality score, also called mean opinion score. *MOS*, Mean opinion score.

Table I. Results on the validation set of the dermatological image dataset. The top 3 rows correspond to the first experiment and the bottom row to the second experiment

Training data	Distortion type	Data used for fine-tuning	LCC	SROCC	MAE
General domain images	Synthetic	Dermatological images	0.657	0.686	0.887
General domain images	Real	Dermatological images	0.806	0.802	0.746
General domain images	Synthetic + real	Dermatological images	0.750	0.747	0.761
General domain + dermatological images	Synthetic + real	No fine-tuning	0.737	0.734	0.472

LCC, Linear correlation; *MAE*, mean absolute error; *SROCC*, Spearman's rank correlation.

REFERENCES

1. Pala P, Bergler-Czop BS, Gwiżdż JM. Teledermatology: idea, benefits and risks of modern age - a systematic review based on melanoma. *Postepy Dermatol Alergol*. 2020;37(2):159-167. https://doi.org/10.5114/ada.2020.94834
2. Maier K, Zaniolo L, Marques O. Image quality issues in teledermatology: a comparative analysis of artificial intelligence solutions. *J Am Acad Dermatol*. 2021;87(1):240-242.
3. Talebi H, Milanfar P. NIMA: neural image assessment. *IEEE Trans Image*. 2018;27(8):3998-4011.

https://doi.org/10.1016/j.jaad.2022.11.002

Variation in Medicare Part D topical steroid prescription costs

To the Editor: The increasing cost of topical steroids has placed a financial burden on both patients and the health care system.[1] The factors contributing to cost variability are poorly understood. Herein, we aim to evaluate variations in out-of-pocket generic topical steroid costs for patients with Medicare Part D.

Medicare.gov was used to collect prescription topical steroid costs for Medicare Prescription Drug plan (PDP) enrollees. In January 2021, we selected a range of topical steroids, geographic locations, pharmacy types, and PDP suppliers (Supplementary Table I, available via Mendeley at https://doi.org/10.17632/ktxpnvrwc7.1) to collect drug prices reflecting current costs for topical steroids dependent on the date Medicare.gov was accessed. Linear regressions were used to evaluate the associations between drug costs and variables. All analyses were conducted in R Version 4.0.

When analyzed by potency groups, the costs of high- and low-potency topical steroids were \$34.50 ± 2.30 and \$43.30 ± 2.40 higher, respectively, than mid-potency steroids per 15 g of medication ($P < .001$; Table I). Drug costs varied significantly by PDP supplier, with some patients being charged up to \$38.50 more per prescription compared to the cheapest recorded prescription ($P < .001$, Table I). Triamcinolone in the mid-potency group was found to be uniformly cheapest across Medicare PDPs, while high- and low-potency topical steroids did not demonstrate a consistently cheaper option (Table II). Drug potency accounted for 18.9% of drug price variability, whereas an additional 9.2% was attributable to the type of PDP. Multivariable analysis of all other variables including geographic region, urban versus rural location, and pharmacy type did not significantly impact topical steroid costs (Table I). Overall, only 28.2% of the variation among drug prices was accounted for by the variables investigated in this study.

The results from our study demonstrate substantial variability in topical steroid drug price as a function of drug potency and PDP type. A Medicare patient may pay up to 327 times more for a 15-gram tube of a low-