Real-world use of a deep convolutional neural network to assist in the diagnosis of pyoderma gangrenosum

Emma L. Hodson, MS,[a] Iman Salem, MD, MS,[a] Mattias Birkner, PhD,[b] Aravindhan Sriharan, MD,[c] Alicia T. Dagrosa, MD, MBA,[a] Matthew J. Davis, MD,[a] and Carsten R. Hamann, MD, PhD[a,c,d,e]

Key words: artificial intelligence; pyoderma gangrenosum; wound care.

INTRODUCTION

Early diagnosis of pyoderma gangrenosum (PG) can be challenging. A delayed or inaccurate diagnosis can lead to significant morbidity. PG can be misdiagnosed as cellulitis, ecthyma, or other infectious disease processes, leading to unnecessary debridement or other surgeries that can significantly worsen PG. There has been an abundance of research using artificial intelligence-powered machine learning and deep learning to create models to assist in dermatologic diagnosis. However, these tools often remain embedded in research institutions with limited use in real-world settings.[1] A machine learning algorithm to aid in distinguishing PG from venous ulcers based on clinical images alone was recently developed, ultimately reaching a sensitivity of 97%.[2]

Abbreviation used:

PG: Pyoderma gangrenosum

CASE REPORT

A 70-year-old woman presented to our clinic with a 2-month history of a painful violaceous ulcer on her right leg (Fig 1). She presented a history of CREST syndrome, atrial fibrillation, and esophageal dyskinesia. She reported no inciting trauma and relatively gradual spread of the wound over several weeks. She reported to an outside hospital, was diagnosed with a necrotizing infection and underwent surgical debridement, complicated by postoperative infection that led to sepsis and an extended inpatient hospital stay. On physical examination, there were 2 discrete necrotic appearing plaques with notable thick black eschars and the absence of undermined borders, although a violaceous color at the edge of a portion of the wound was present. Initial differentials included PG, calciphylaxis, vasculitis, or an arterial ulcer.

Two 4-mm punch biopsies were performed, both at the plaque's edge. They reported findings compatible with reactive changes adjacent to an ulcer but were not specific as to the cause. The subcutaneous fat was not available for evaluation of calciphylaxis. There was no evidence of vasculitis. Extensive laboratory work-up was significant for elevated antinuclear antibody (1:640) and anticentromere antibodies. A prednisone taper starting at 60 mg daily was initiated with a presumptive diagnosis of PG, despite some atypical findings. Over the next 3 months, she experienced modest improvement on low dose of prednisone (10 mg), low dose of cyclosporine (2 mg/kg), and other adjunctive therapies, such as topical metronidazole cream, collagenase ointment, and gentamycin cream. Given this improvement, a diagnosis of PG was strongly favored. Infliximab was added to the above regimen but was discontinued owing to side effects, specifically morbilliform rash and headaches.

To add further support for this diagnosis, photographs of the patient's initial ulcer were analyzed by

From the Department of Dermatology, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire[a]; Institute of Medical Physics, Paracelsus Medical University Nuremberg, City Hospital of Nuremberg, Nürnberg, Germany[b]; Department of Pathology and Laboratory Medicine, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire[c]; HonorHealth Dermatology Residency, Scottsdale AZ[d]; and Contact Dermatitis Institute, Phoenix, Arizona.[e]

Funding Sources: None.

IRB approval: Not applicable.

Correspondence to: Carsten R. Hamann, MD, PhD, Contact Dermatitis Institute, 3400 E McDowell Rd, Phoenix, AZ 85008. E-mail: hamann511@gmail.com.

JAAD Case Reports 2023;38:8-10.

2352-5126

https://doi.org/10.1016/j.jdcr.2023.05.031

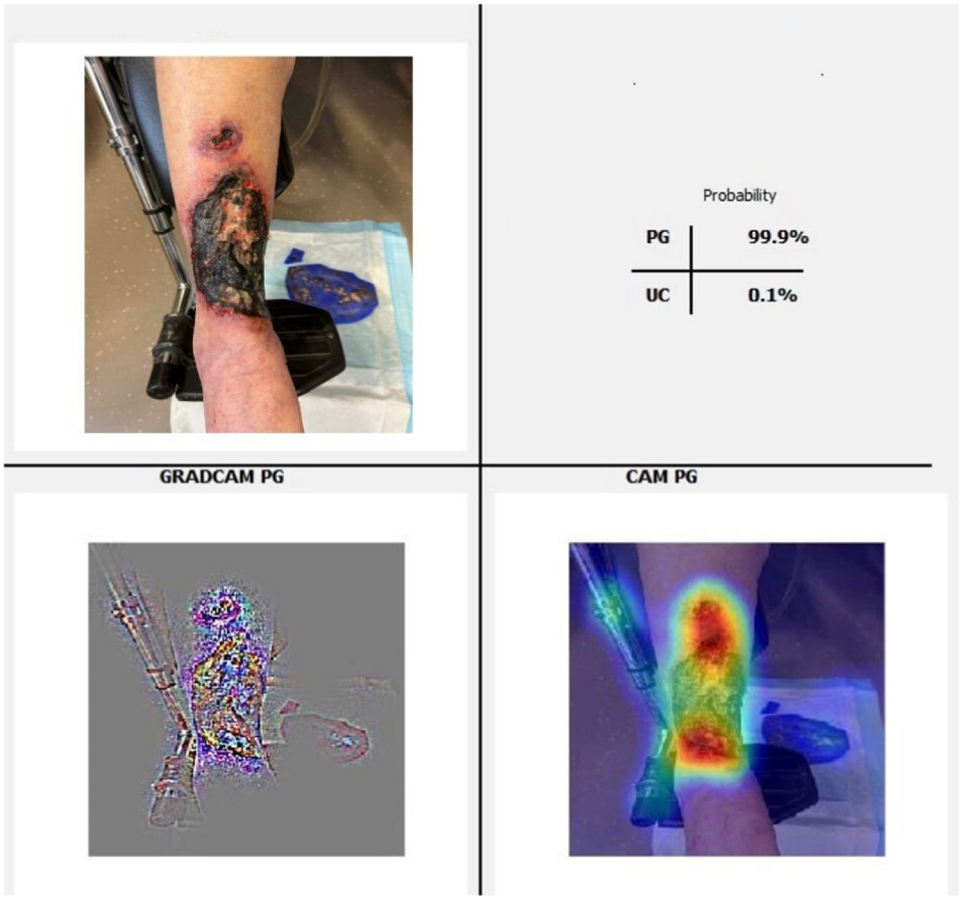

Fig 1. Application of the machine learning algorithm to photo taken at time of initial presentation to dermatology clinic. *CAM PG*, PG class activation map; *GradCAM PG*, gradient-weighted PG CAM; *UC*, ulcus cruris (ie, conventional leg ulcer).

a recently developed deep convolutional neural network trained to distinguish PG from venous ulcers. After the internal rescaling to 300 × 300 pixels, the image was run through the model. The neural network reported the probability for a PG within <1 second, together with the class activation map (ie, the regions of the images that matter most for the diagnosis), and the most similar images from the training data set. Her photograph reported a PG probability of > 99% (Fig 1). A countercase of a basal cell carcinoma originally misdiagnosed as PG was included to test the ability of the algorithm to exclude PG when appropriate. Photographs of this case's initial presentation reported a low-PG probability of 4.7 and 17.9, supporting the competence of the algorithm to make this distinction. Accordingly, our patient's cyclosporine was increased to 4 mg/kg, and the patient began to show improvement. Six months after the initial diagnosis, she continued to improve on 5 mg of prednisone every other day and 3 mg/kg of cyclosporine. At 6 months, the ulcerated plaque had decreased in size, no necrotic tissue was observed, and portions of the ulcer had reepithelialized (Fig 2).

DISCUSSION

Despite recent revisions and improvements in diagnostic criteria, a definitive diagnosis of PG remains challenging and is a diagnosis of exclusion.[3] In this case, and in similar cases with lack of obvious underlying condition, and lack of characteristic appearance or clinical history, the diagnosis is particularly difficult. Applying available diagnostic PG criteria to this case was unhelpful. On presentation to our dermatology clinic, this case initially met zero major and zero minor diagnostic criteria proposed by Haag et al.[3] It failed to meet the criteria established by Delphi consensus and scored 3 points out of the 10 needed to meet the high likelihood of PG criteria using the PARACELSUS method.[3]

AI-powered machine learning models used in dermatologic diagnosis are becoming increasingly common.[4] Although the majority of these models are used in categorization of melanocytic lesions,

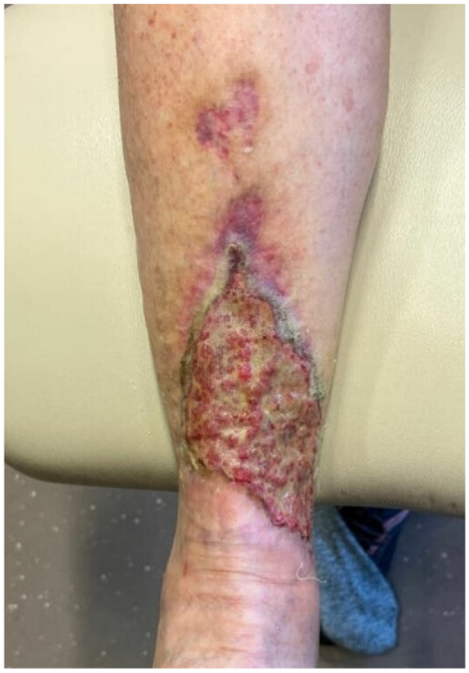

Fig 2. Recent photograph demonstrating decrease in size, resolution of necrotic tissue, and reepithelialization of portions of the wound.

models assisting in the diagnosis of inflammatory skin disease, such as palmoplantar psoriasis,[5] atopic dermatitis,[6] rosacea,[7] and infectious diseases, such as onychomycosis, have also been developed.[8] Apart from clinical photographs, models analyzing dermatoscopic photos, dermatopathology images, and confocal microscopy images have also been created.[9]

A deep convolutional neural network capable of analyzing wound photographs to facilitate PG diagnosis was recently developed, based on a data set that comprised 491 photographs of PG and conventional leg ulcers, primarily chronic venous ulcers. It outperformed dermatologists viewing photographs, with a significantly higher sensitivity to diagnosing PG (sensitivity 97% vs 72.7%).[2] In our case, this model improved diagnostic confidence and allowed for more aggressive and successful therapy.

This case highlights the potential global effect that accurate and efficient machine learning algorithms may have on heavily visual fields, such as dermatology. A potential pitfall in this case in particular, and in dermatology-focused machine learning applications in general, is the lack of diversity in clinical image repositories used to train algorithms. In addition, restricting the differential diagnosis to venous ulcers or PG is indeed a major limitation of this algorithm. However, Birkner et al[2] are working on further algorithm development and revisions to include more diagnostic options and provide public availability thereafter. With a high matching of PG and probability exceeding 99%, the present case shows potential for the use of this and similar algorithms in clinical practice in the future. With increasing diversity in datasets, similar machine learning algorithms will become increasingly generalizable and can easily be applied to broad populations worldwide. Because artificial intelligence-powered machine learning models continue to proliferate, it is important that they are tested and used in real-world cases.

Conflicts of interest

None disclosed.

REFERENCES

1. Kooper-Johnson S, Weber M, Eichtadt S, Nguyen BM. Limited commercialization of artificial intelligence technology in the diagnosis of pigmented lesions despite explosive growth in published programs. Preprint. Posted online February 23, 2023. *J Am Acad Dermatol*. 2023. https://doi.org/10.1016/j.jaad.2023.02.020. S0190-9622(23)00273-00276.
2. Birkner M, Schalk J, von den Driesch P, Schultz ES. Computer--assisted differential diagnosis of pyoderma gangrenosum and venous ulcers with deep neural networks. *J Clin Med*. 2022;11(23):7103. https://doi.org/10.3390/jcm11237103
3. Haag C, Hansen T, Hajar T, et al. Comparison of three diagnostic frameworks for pyoderma gangrenosum. *J Invest Dermatol*. 2021;141(1):59-63. https://doi.org/10.1016/j.jid.2020.04.019
4. De A, Sarda A, Gupta S, Das S. Use of artificial intelligence in dermatology. *Indian J Dermatol*. 2020;65(5):352-357. https://doi.org/10.4103/ijd.IJD_418_20
5. Paik K, Kim BR, Youn SW. Evaluation of the area subscore of the Palmoplantar Pustulosis Area and Severity Index using an attention U-net deep learning algorithm. *J Dermatol*. 2023;50(6):787-792. https://doi.org/10.1111/1346-8138.16752
6. 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. https://doi.org/10.21037/atm.2020.04.39
7. Binol H, Plotner A, Sopkovich J, Kaffenberger B, Niazi MKK, Gurcan MN. Ros-NET: a deep convolutional neural network for automatic identification of rosacea lesions. *Skin Res Technol*. 2020;26(3):413-421. https://doi.org/10.1111/srt.12817
8. Zhu X, Zheng B, Cai W, et al. Deep learning-based diagnosis models for onychomycosis in dermoscopy. *Mycoses*. 2022;65(4):466-472. https://doi.org/10.1111/myc.13427
9. Malciu AM, Lupu M, Voiculescu VM. Artificial intelligence-based approaches to reflectance confocal microscopy image analysis in dermatology. *J Clin Med*. 2022;11(2):429. https://doi.org/10.3390/jcm11020429