Using artificial intelligence to compute Severity of Alopecia Tool scores

To the Editor: Alopecia areata (AA) is a complex autoimmune condition that causes nonscaring hair loss. A systemic review of the epidemiology of AA indicated a similar worldwide lifetime incidence of around 2%.[1] The Severity of Alopecia Tool (SALT) was developed to provide a standardized scale for assessing the severity of individual AA cases and is used for clinical practice and research.[2] Existing methods for assigning SALT scores are cumbersome, time-consuming, and inconsistent among raters.[3]

We have developed an artificial intelligence (AI) system for objectively quantifying AA and assigning SALT scores. The AI system computes SALT scores by analyzing 4 images (1 for each quadrant) of a person's head. The system has 2 main parts. Part 1 uses a deep learning model (Unet with EfficientNetB4 as base model) whose input is a head image and whose outputs are a scalp segmentation (with outputs of 0 or 1) and a heat map (the result of a regression task) showing the percentage of hair loss on that image. The model was trained on 823 head images labelled by trained annotators and verified by doctors using our annotation system. The severity of hair loss is divided into 4 levels: 0; 0.33; 0.66 and 1. The annotation of hair loss was done by area. In part 2, the system compiles the results of the 4 images to compute a single SALT score (the percentage of hair loss) for the entire head. Details are in the Supplementary File, available via Mendeley at https://data.mendeley.com/datasets/kv4658vwgs/1.

We evaluated 188 images containing 47 sets of 4 views (left, right, top, back) for each subject. All subjects are adults, have alopecia aerata and are from the Philadelphia region. The images were taken during clinic visits using an iPad. Each image was independently scored by a human investigator (a veteran U.S. dermatologist) and by the AI system, with the score representing the percentage of hair loss in that image (see Fig 1). We then used intraclass correlation coefficient (ICC) to measure accuracy between scores by the doctor and the AI. We note that based on a 95% confidence interval, ICC scores between 0.75 and 0.9 indicate good correlation, and scores greater than 0.90 indicate excellent correlation (see Fig 2).[4] We obtained an ICC score of 0.97 for SALT scoring, and ICC scores of 0.96, 0.97, 0.95, and 0.92 for left, right, top, and back view images.

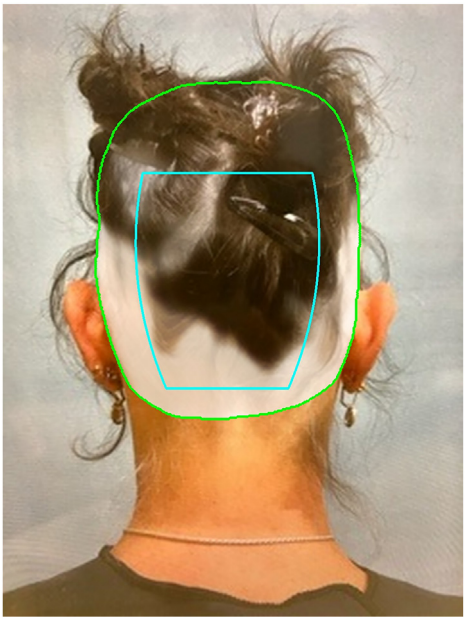

Fig 1. Dermatologist-estimated hair loss is 33%. AI analysis: Calculated heatmap (*whiter* - more severe), scalp segmentation (*green*) and identified quadrant (*blue*). AI-estimated hair loss is 15%. *AI*, Artificial intelligence.

Results were not affected by skin tone, skin color, or hair length. Most cases where the doctor and AI scores differed involved severe alopecia or patients with new, white and short hair (doctor scored 100% hair loss but AI scored less, error is about 10%).

Our data supports a potential role for AI techniques in assigning SALT scores. For the next step, we are working on a 3D model for scalp area which will give a much more accurate hair loss percentage.

Hang Nguyen, PhD,[a,b] *Lea Gazeau,*[a,b] *and Jonathan Wolfe, MD*[b,c]

From the Torus AI, Ramonville-Saint-Agne, France[a]*; BelleTorus Corporation, Cambridge, Massachusetts*[b]*; and Division of Dermatology, Jefferson-Einstein Montgomery Medical Center, East Norriton, Pennsylvania.*[c]

Funding sources: The project was supported by BelleTorus Corp. and Torus AI.

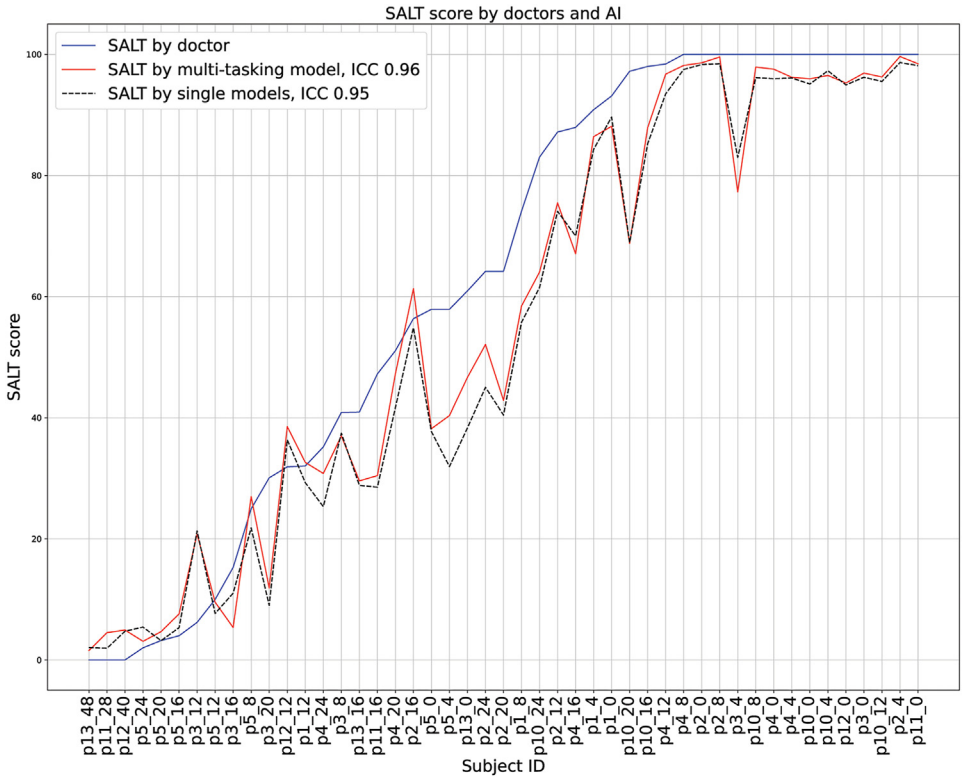

Fig 2. SALT scores for patients by dermatologist (in *blue*) and by AI (*red*). The horizontal axis represents individual patients. *AI*, Artificial intelligence; *SALT*, Severity of Alopecia Tool.

Patient consent: Consent for the publication of recognizable patient photographs or other identifiable material was obtained by the authors and included at the time of article submission to the journal stating that all patients gave consent with the understanding that this information may be publicly available.

IRB approval status: Not applicable.

Key words: hair-loss regression; multi-task learning; SALT; scalp segmentation; Severity of Alopecia Tool.

Correspondence to: Hang Nguyen, PhD, BelleTorus Corporation, 245 First St Riverview II, 18th Floor, Cambridge, MA 02142

E-mail: hangntt@belle.ai

X handle: @belletorus

Conflicts of interest

Dr Hang Nguyen, Lea Gazeau and Dr Wolfe are shareholders of BelleTorus.

REFERENCES

1. Villasante Fricke AC, Miteva M. Epidemiology and burden of alopecia areata: a systematic review. *Clin Cosmet Investig Dermatol*. 2015;8:397-403.
2. Olsen EA, Hordinsky MK, Price VH, et al. Alopecia areata investigational assessment guidelines — part II. National alopecia areata foundation. *J Am Acad Dermatol*. 2004;51:440-447.
3. Bernardis E, Castelo-Soccio L. Quantifying alopecia areata via texture analysis to automate the SALT score computation. *J Invest Dermatol Symp Proc*. 2018;19(1):S34-S40.
4. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155-163. Erratum in: J Chiropr Med, 16(4), 346. PMID: 27330520; PMCID: PMC4913118.

https://doi.org/10.1016/j.jdin.2024.04.003