



# Construction and Evaluation of a Deep Learning Model for Assessing Acne Vulgaris Using Clinical Images

Yin Yang · Lifang Guo · Qiuju Wu · Mengli Zhang · Rong Zeng ·  
Hui Ding · Huiying Zheng · Junxiang Xie · Yong Li · Yiping Ge ·  
Min Li · Tong Lin

Received: February 14, 2021 / Accepted: April 22, 2021 / Published online: May 18, 2021  
© The Author(s) 2021

## ABSTRACT

**Introduction:** Accurate assessment is the basis for the effective treatment of acne vulgaris. The goal of this study was to achieve standardised diagnosis and treatment based on a deep learning model that was developed according to the current Chinese Guidelines for the Management of Acne Vulgaris.

**Methods:** The first step was to divide each image of acne vulgaris into four regions. Each of these four regions of the same patient was then combined to form a complete facial region. The second step was to classify the images based

lesion type, in accordance with the current Chinese guidelines, and by treatment strategy adopted by experienced dermatologists. The final step was to evaluate the performance of the deep learning model in patients with acne vulgaris.

**Results:** The results showed that the average F1 value of the assessment model is 0.8 (optimum value = 1). The weighted kappa coefficient between the evaluation according to the artificial intelligence model and the evaluation by the attending dermatologists was 0.791 (95% confidence interval 0.671–0.910,  $P < 0.001$ ), indicating a high degree of consistency.

**Conclusions:** The assessment model based on deep learning and according to the Chinese guidelines had a slightly higher overall performance is comparable to that of the attending dermatologist.

**Keywords:** Acne vulgaris; Assessment model; Artificial intelligence; Deep learning

---

Y. Yang · L. Guo · Q. Wu · M. Zhang · R. Zeng ·  
H. Ding · H. Zheng · Y. Ge (✉) · M. Li (✉) ·  
T. Lin (✉)

Hospital of Skin Diseases and Institute of  
Dermatology, Chinese Academy of Medical Sciences  
and Peking Union Medical College, 12  
Jiangwangmiao Street, Nanjing 210042,  
People's Republic of China  
e-mail: gypfeng@hotmail.com

M. Li  
e-mail: limin@pumcderm.cams.cn

T. Lin  
e-mail: ddlin@hotmail.com

J. Xie · Y. Li  
Institute of Medical Information, Chinese Academy  
of Medical Sciences & Peking Union Medical  
College, 3 Yabao Road, Beijing 100020,  
People's Republic of China

### Key Summary Points

Effective and accurate assessment is the basis for the treatment of acne vulgaris.

We have developed a deep learning model for the evaluation of acne conditions in accordance with current Chinese evidence-based guidelines.

A high degree of consistency was found between the model and attending dermatologist-level treatment strategy.

The model was used to retrospectively assess ten cases of acne; patients who received treatment equal to or better than the recommended treatment regimen of the model were found to have received more efficacious treatment than those receiving treatments less than that recommended treatment regimen of the model.

Due to the advantages of deep learning, we will continue to obtain new data based on practical clinical applications of the model in order to constantly update the general performance of the model and to optimise it, making the model a more valuable tool in clinical practice.

## DIGITAL FEATURES

This article is published with digital features, including a summary slide, to facilitate understanding of the article. To view digital features for this article go to <https://doi.org/10.6084/m9.figshare.14464815>.

## INTRODUCTION

Acne vulgaris is a common inflammatory skin disease of the pilosebaceous unit that results from androgen-induced increased sebum production, altered keratinisation, inflammation

and bacterial colonisation of hair follicles by *Cutibacterium acnes*, also known as *Propionibacterium acnes* [1]. Clinical manifestations of acne take various forms, including comedoes, papules, pustules, nodules, cysts, among others, which occur mostly in adolescence but may persist into adulthood. Almost all adolescents between the ages of 15 and 17 contract the disease, and 5% of women and 3% of men aged 40–49 years continue to be affected [2–4]. The economic and psychological impact of acne is substantial. Total cost in terms of acne treatment and loss of productivity every year in the USA alone is more than 3 billion U.S. dollars [5], which imposes a heavy burden on the health-care system. Changes in physical appearance due to acne often results in depression and social isolation in adolescent patients [1, 6], and delayed or improper treatments may bring serious and irreversible damage to the physical and mental health of patients [7, 8]. Therefore, the treatment of acne vulgaris is of great significance to patients in particular and society in general.

Evaluation of the severity of acne is directly related to its management strategy. At the present time, however, although many different types of methods and guidelines are available for grading the severity of acne, diagnosis and the choice of a treatment tend to be made based more on the treating physician's own clinical experiences [1], and while a long training period is often required to become a certified dermatologist, dermatologists do differ in terms of their clinical experience. Also contributing to management strategy is patient inaccessibility to dermatologists and the lack of awareness among patients of the need to seek medical treatment, both of which have contribute to inadequate treatment among acne patients [8, 9]. To address the problems mentioned above, our goal has been to establish a deep learning model for the evaluation of acne conditions in accordance with current evidence-based guidelines (Chinese Guidelines for the Management of Acne Vulgaris; [10]). Such a model could provide less experienced doctors, such as general practitioners and dermatology residents, with an objective, reliable and rapid assessment tool to reduce the incidence of

nonstandard treatment and facilitate follow-up observation of treatment effects. In addition, patients may be able to conveniently conduct self-evaluation assessments and seek medical advice in a timely manner. Therefore, the aim of this study on this deep learning-based model is to use ordinary clinical photos of different severity levels of acne vulgaris to train and finally establish a disease evaluation model.

## METHODS

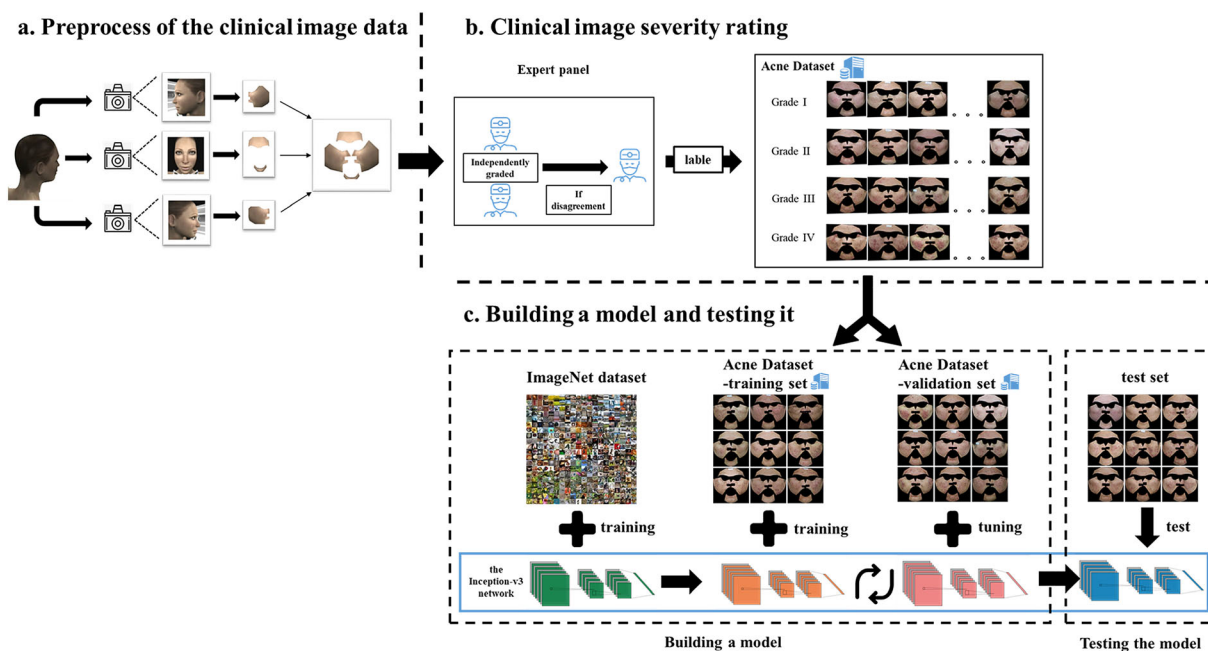
This study was approved by the Ethics Committee of the Hospital of Skin Diseases, Chinese Academy of Medical Sciences (IRB number: 2019-KY-005). Since all patient records were anonymous and privacy was removed through image preprocessing, no personal data were collected from existing patient records and no written informed consent was required. Patients were informed that the images may be used for scientific research and publication after privacy removal and approval by the ethics committee.

The design and workflow of the study is shown schematically in Fig. 1.

### Clinical image data collection and pretreatment

**Clinical image data collection** For this study we collected 5871 clinical photographs of 1957 patients who visited the Hospital of Skin Diseases, Chinese Academy of Medical Sciences from 2004 to 2016. Basic parameters of the data were as follows: (1) the pictures were taken by two single-lens reflex cameras (FinePix S9500, Fujifilm, Tokyo, Japan and Canon model EOS 800D, Canon Corp., Tokyo Japan, respectively), and the images were 2 million by 20 million pixels; (2) the clinical photographs contained all the lesion areas. All images used in this study were processed for privacy masking.

**Preprocessing of the clinical image data** Facial photographs of each acne patient were preprocessed by following procedures. First, 68 human face markers were used to mark the photographs. According to the face markers, the two eyes in each person’s facial image were placed on a horizontal line by rotation, and then the image size was adjusted to ensure that the distance between the two pupils was 800



**Fig. 1** Procedure for development of an evaluation model for acne vulgaris based on deep learning. **a** Preprocessing of the clinical image data, **b** severity rating of clinical images, **c** building and subsequently testing the model

pixels. The purpose of this procedure was to ensure that each face was kept horizontal and zoomed standardly. Then, in order to avoid interference from the eyes, nose and mouth area, each face was divided into different regions, as shown in Fig. 1a, and the four regions of the same patient (i.e. the forehead, lower jaw, left side of the face and right side of the face) were subsequently combined to form a complete facial region. Thus, the entire face area of the patient was fully visible in the two-dimensional photograph. Finally, due to the lack of sufficient training data and data imbalance, we utilised the ImageDataGenerator in the neural network library Keras as an image augmentation technique to increase the size of the training set. It should be noted that all the above-mentioned steps ran automatically without human involvement.

### Rating of acne severity on clinical images

The processed clinical images were independently graded by two experienced dermatologists, and a third dermatologists who had more experience was consulted in the case of

disagreement. The three dermatologists were blinded—and had no access to—the deep-learning predictions, as shown in Fig. 1b. According to the Chinese guidelines for the management of acne vulgaris [10], the images were classified into four grades based on the type of the clinical features and ensued treatment strategy, as shown in Table 1. The three dermatologists entered their consensus directly into the dataset which we named Acne Dataset. The Acne Dataset was divided into two separate directories following the division of 80% and 20% for training and validation, respectively.

### Model construction

The model was trained using the Inception-v3 network (a deep learning-based classification model), as shown in Fig. 1c. When training the network, we first used pretrained network parameters in the ImageNet Dataset (including 1.28 million images and 1000 objects) because these pretrained networks have preserved the shallow features of the image, which helps to improve classification accuracy. Then, based on the transfer learning method, we used the Acne

**Table 1** Classification criteria for clinical image severity

Clinical features and treatment strategy	Severity rating			
	Grade I	Grade II	Grade III	Grade IV
Clinical manifestations	Some comedones with no more than one small inflammatory lesions	Some papules with no more than a few pustules only, no nodular	Some pustules, but no more than one small nodular	Some nodules and cysts
Treatment strategy	Topical retinoids	Topical retinoids + benzoyl peroxide ± topical antibiotic or benzoyl peroxide + topical antibiotics	Oral antibiotics + topical retinoids ± benzoyl peroxide ± topical antibiotics	Oral isotretinoin ± benzoyl peroxide/topical antibiotics. Patients with strong inflammatory response may take oral antibiotics + benzoyl peroxide/topical antibiotics, followed by oral isotretinoin

Dataset-training set to train the network for learning the high-level semantic features of the image. In the experiment, we used a learning rate of 0.001 for training and used the cross-entropy loss function and the Adam optimizer. The Acne Dataset-validation set was later used to provide an unbiased evaluation of a model fit on the training dataset while tuning model hyperparameters for better performance. The model was trained and validated on a server (Intel® Xeon® Processor E5: 2.10 GHz, 32 GB RAM, 1080 GTX GPU; Intel Corp., Santa Clara, CA, USA).

### Testing the model

Using the same preprocessing method as the Acne Dataset, a test set of 40 preprocessed images were obtained, as shown in Fig. 1c. Three attending dermatologists and three dermatology residents were invited to classify the photos of the test set. Each attending dermatologist and dermatology resident who did not participate in the labelling independently completed the evaluation. The evaluation results of the three attending dermatologists were voted on, and those on which more than two attending dermatologists agreed were considered to be the results of the evaluation from the attending dermatologist. The same method was also used for the selection of evaluation results at the dermatology resident level.

### Statistical indicators for evaluating the model classification performance

This study used the F1 value as a classification evaluation indicator. The F1 value is the harmonic mean of the accuracy and the recall rate. When the accuracy and recall rate are high, the F1 value will also be high accordingly. The F1 value reaches the optimum value (i.e. with a perfect accuracy and recall rate) at 1, and the worst value at 0.

This study also used Kendall's coefficient of concordance (Kendall's W) and its test to evaluate the consistency of the evaluation results among the three attending dermatologists. The same method was applied to the evaluation

**Table 2** Basic information of the clinical data and datasets

Data sets	Severity rating			
	Grade I	Grade II	Grade III	Grade IV
Training sets	532	537	336	161
Validation sets	133	134	84	40
Test sets	10	10	10	10
Total	675	681	430	211

results among the dermatology residents. The linearly weighted kappa coefficient and its test were used to evaluate the consistency of the model-based evaluation results and the dermatologist-based evaluation results. A value of  $> 0.75$  indicates high consistency, a value between 0.75 and 0.65 indicates moderate consistency and a value  $< 0.65$  indicates low consistency. The above test was carried out in SPSS version 24.0 (SPSS IBM Corp., Armonk, NY, USA).

## RESULTS

### Basic information of the clinical data and datasets

In total, 1957 processed clinical images were included in the Acne Dataset, of which 1565 images were used as training sets and 392 images were used as validation sets. In addition, 40 processed clinical images were used as test sets (Table 2).

### Network model test results

Test results for the test set The results show that the average F1 value of the assessment model is 0.8 (Table 3).

Comparison of model evaluation with evaluation of dermatologists This study used Kendall's W to analyse the consistency of the diagnosis of disease severity in 40 acne patients.



**Table 3** Test results for the test set

Grade I		Grade II		Grade III		Grade IV		Average		
Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall	F1 <sup>a</sup>
0.78	0.7	0.75	0.9	0.78	0.7	0.9	0.9	0.80	0.80	0.80

<sup>a</sup> F1 value is the harmonic mean of the accuracy and the recall rate. The F1 value will be high when the accuracy and recall rate are high, with a value of 1 (optimum value) indicating perfect accuracy and recall rate; 0 is the worst value

The results showed that Kendall's coefficient of concordance of the evaluations among the three attending dermatologists was 0.938 ( $P < 0.001$ ), which indicated a high consistency. Similarly, Kendall's coefficient of concordance among the three dermatology residents was 0.907 ( $P < 0.001$ ), indicating that they also showed a high consistency regarding their diagnosis.

The linearly weighted kappa coefficient was used to analyze the consistency between the dermatologists and the model for the diagnosis of disease severity in 40 pairs of acne patients. The results showed that the attending dermatologists and the dermatology residents had a consistent diagnosis for 25 subjects and inconsistent for 15 subjects. The weighted kappa coefficient was 0.650 (95% confidence interval [CI] 0.515–0.784,  $P < 0.001$ ), indicating lower consistency. The artificial intelligence (AI) model and the attending dermatologist-level weighted kappa coefficient was 0.791 (95% CI 0.671–0.910,  $P < 0.001$ ), indicating a high degree of consistency. The AI model and dermatology resident-level weighted kappa coefficient was 0.629 (95% CI 0.502–0.756,  $P < 0.001$ ), indicating a lower consistency than that of the attending dermatologist-level comparison.

## DISCUSSION

Medical image recognition is a cross-disciplinary field involving clinical medicine, mathematical image processing, pattern recognition, machine learning, among other areas of knowledge. The main research aspects include medical image classification, lesion location

and segmentation and three-dimensional reconstruction and visualisation. Traditional medical image recognition techniques include pixel-level image processing and mathematical modelling, both of which are based on specific recognition rules. Some studies have evaluated the severity of acne. Chang and Liao [11] used the support vector machine (SVM) classifier and feature extraction to divide acne vulgaris into comedones, papules, pustules, etc. Malik et al. [12] used a similar method to divide acne vulgaris into comedones, papules, etc., and then classified acne into mild, mild to moderate, moderate to severe and severe according to the established scoring rules. Patwardhan et al. [13] extracted features from VISIA-CR images of acne to calculate the number of inflammatory lesions, noninflammatory skin lesions, erythema, and post-acne pigmentation. These studies were based on specific acquisition equipment or traditional image recognition methods. It is necessary to design image features according to the color or texture of acne and then select the classifier for training and classification. These kinds of artificially designed image features are easily affected by the illumination environment and imaging quality and cannot effectively describe the appearance characteristics of acne, which limits the clinical application.

In recent years, thanks to the development of computers and the expansion of datasets, deep learning models have been widely used in the field of image classification and detection. With the development of machine learning, scholars are beginning to use convolutional neural networks to train medical images, such as magnetic resonance imaging (MRI) images, computed tomography (CT) images,

microscopic images and clinical photographs, to achieve higher recognition rates. For acne research, Shen et al. [14] used the 16-layer Visual Geometry Group (VGG16) model to classify acne patients into seven categories: those with normal skin, whiteheads, blackheads, pimples, pustules, cysts and nodules. The same patient can be output to more than one class of results. Zhao and Spoelstra [15] used the 152-layer residual neural network (ResNet-152) to classify acne into five levels: 1-clear, 2-almost clear, 3-mild, 4-moderate, and 5-severe, which was compared with 11 dermatologists on the same acne patients. Seité [16] developed an artificial intelligence algorithm for smartphones to determine the severity of facial acne using the Global Acne Severity Scale for Europe.

In view of previous studies and the practical application of the evaluation of the acne severity, the model presented in this study was designed with following characteristics:

1. A grading standard with operability. Not only the quantity and type of skin lesions but also the standard treatment plan were used as the grading standard. Each severity level corresponded to an established standard treatment regimen, which could assist the physician to formulate a clinical treatment plan and provide patients with standardised and targeted precision treatment. In this study, the assessment of acne severity and the corresponding treatment recommendations were in accordance with current Chinese guidelines [10], which were formulated by 30 expert dermatologists based on user feedback of the previous guidelines, research progress on acne in China and abroad, as well as experts' experience. These guidelines have a practical guiding and normative role in diagnosis and treatment. In addition, this model can re-evaluate the severity of patients after treatment to determine whether the current treatment regimen should be changed. For instance, when a patient's grading is downgraded from Grade III to Grade I after treatment, his/her treatment can thus be changed from the original oral antibiotics to topical retinoids.
2. Full-section photos of multi-angle stitching. The overall degree of disease cannot be fully described by a single facial photograph. In this study, we used a multi-angled face to remove the area near the eyes, nostrils, and lips that did not have skin lesions. Splicing the dataset not only reflects the skin lesions in the entire face of the patient but also minimises the privacy exposure of patient data.
3. Voting method to process data. The average classification was not processed by orderly classification data, such as severity classification, but by a more accurate voting method.
4. Training using the Inception-v3 model. Inception-v3 is a widely used image recognition model that achieves better performance in image recognition. Compared with VGGNet and ResNet, it is more suitable to our task because of its unique Inception architecture, in which multiscale convolutions are performed in parallel and the convolution results for each branch are further concatenated. Additionally, VGGNet requires more computational cost and ResNet performs poorly in identifying subtle object differences.

This study performed a statistical analysis of the evaluation results of three attending dermatologists and three dermatology residents, confirming that the dermatologists at the same clinical level of treatment had more consistency. The evaluations by our model have a strong consistency with the evaluations of the attending dermatologists, indicating that the model can achieve the same level of assessment of acne severity as an attending dermatologist with more experience in diagnosis and treatment.

The model is highly practical since it is based on photographs taken by SLR cameras, which are commonly used in everyday practice. It can be applied to the teaching of dermatology and can assist primary dermatologists, general practitioners and physicians in developing treatment plans and help acne patients understand the severity of their disease. Of course, the

treatment and management regimen of the disease is not fixed. It should be noted that:

1. The purpose of this research model is to provide an objective assessment of the severity of acne vulgaris in patients and to propose basic treatment recommendations. However, there is a far way to go before such an AI system could actually make clinical decisions in the place/absence of human input.
2. In order to fully embody the principle of individualised treatment, dermatologists should make choices according to the actual conditions of patients based on, for example, medical history, contraindications of drugs, etc.

There are a number of shortcomings to this study. First, since only East Asian people with Fitzpatrick skin type III and IV were included in the dataset, we adopted the Chinese Guidelines for the Management of Acne Vulgaris as the sole grading standard. In future studies, we are going to use guidelines [17–20], datasets and evaluation methods [21] from different ethnic groups to train the model, with the aim to expand the application range of the model. Secondly, the team will conduct a retrospective or prospective clinical study of the model's graded treatment plan to optimise present the treatment regimen of the study model. Furthermore, patients' metadata, such as patients' medical history and other clinical information in addition to image data will be added to the model. Later, a standardised acne vulgaris severity assessment model applied to smart-phones and other modes will be developed to promote the implementation of precise, accessible, individualised medical treatment. Thirdly, this study focused on the overall evaluation of the disease without classifying and quantifying specific skin lesions, which is an aim of our future research. Moreover, due to the advantages of deep learning, we will continue to obtain new data based on clinical practical application, so as to constantly update the generalisation performance and optimise the model, making it a more valuable tool in clinical practice.

## CONCLUSIONS

In this study, clinical images were used to construct a deep learning model to evaluate acne. The model had a slightly higher overall performance than and is comparable to evaluations carried out by attending dermatologists. We will continue to verify the performance of this model with more data.

## ACKNOWLEDGMENTS

We thank the participants of the study. We are grateful to Nanjing suoyousuoyi Information Technology Co., Ltd for assisting in building the model.

**Funding.** The work was financially supported by CAMS Innovation Fund for Medical Sciences (CIFMS-2017-I2M-1-017, CIFMS-2018-I2M-AI-018). Primary Research & Development Plan of Jiangsu Province (BE2017808). The Fundamental Research Funds for the Central Universities (3332018123). The Rapid Service Fee was funded by the authors.

**Disclosures.** Yin Yang, Lifang Guo, Qiuju Wu, Mengli Zhang, Rong Zeng, Hui Ding, Huiying Zheng, Junxiang Xie, Yong Li, Yiping Ge, MD, Min Li, Tong Lin have nothing to disclose.

**Compliance with Ethics Guidelines.** This study was approved by the Ethics Committee of the Hospital of Skin Diseases, Chinese Academy of Medical Sciences (IRB number: 2019-KY-005). Since all patient records were anonymous and privacy was removed through image preprocessing, no personal data were collected from existing patient records and no written informed consent was required. Patients were informed that the images may be used for scientific research and paper publication after privacy removal and approval by the ethics committee.

**Authorship.** All named authors meet the International Committee of Medical Journal



Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

**Authorship Contributions.** All authors were involved in the conceptual design of this manuscript, drafting and development, and agreement to publish.

**Data availability.** The datasets generated analysed during the current study are available from the corresponding author on reasonable request.

**Open Access.** This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

## REFERENCES

- Williams HC, Dellavalle RP, Garner S. Acne vulgaris. *Lancet*. 2012;379(9813):361–72.
- Cunliffe WJ, Gould DJ. Prevalence of facial acne vulgaris in late adolescence and in adults. *Br Med J*. 1979;1(6171):1109–10.
- Rademaker M, Garioch JJ, Simpson NB. Acne in schoolchildren: no longer a concern for dermatologists. *BMJ*. 1989;298(6682):1217–9.
- Law MP, Chuh AA, Lee A, Molinari N. Acne prevalence and beyond: acne disability and its predictive factors among Chinese late adolescents in Hong Kong. *Clin Exp Dermatol*. 2010;35(1):16–21.
- Bhate K, Williams HC. Epidemiology of acne vulgaris. *Br J Dermatol*. 2013;168(3):474–85.
- Kubota Y, Shirahige Y, Nakai K, Katsura J, Moriue T, Yoneda K. Community-based epidemiological study of psychosocial effects of acne in Japanese adolescents. *J Dermatol*. 2010;37(7):617–22.
- Inglese MJ, Fleischer AB, Feldman SR, Balkrishnan R. The pharmacoeconomics of acne treatment: where are we heading. *J Dermatol Treat*. 2008;19(1):27–37.
- Gieler U, Gieler T, Kupfer JP. Acne and quality of life—impact and management. *J Eur Acad Dermatol Venereol*. 2015;29(Suppl 4):12–4.
- Haider A, Mamdani M, Shaw JC, Alter DA, Shear NH. Socioeconomic status influences care of patients with acne in Ontario, Canada. *J Am Acad Dermatol*. 2006;54(2):331–5.
- Acne Group, Combination of Traditional and Western Medicine Dermatology, Acne Group, Chinese Society of Dermatology, Acne Group, Chinese Dermatologist Association, et al. Chinese Guidelines for the Management of Acne Vulgaris: 2019 Update#. *Int J Dermatol Vener*. 2019;2(3):129–38.
- Chang C, Liao H. Automatic facial skin defects detection and recognition system. In: 2011 Fifth International Conference on Genetic and Evolutionary Computing; 2011. p. 260–3. <https://doi.org/10.1109/ICGEC.2011.67>.
- Malik AS, Ramali R, Hani AFM et al. Digital assessment of facial acne vulgaris. In: 2014 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings; 2014. p. 546–50. <https://doi.org/10.1109/I2MTC.2014.68608042014>.
- Patwardhan SV, Kaczvinsky Jr, Joa JF, Canfield D. Auto-classification of acne lesions using multimodal imaging. *J Drugs Dermatol*. 2013;12(7):746–756.
- Shen X, Zhang J, Yan C, Zhou H. An automatic diagnosis method of facial acne vulgaris based on convolutional neural network. *Sci Rep*. 2018;8(1):5839.
- Zhao THZ, Zhang H, Spoelstra J. A computer vision application for assessing facial acne severity from selfie images. arXiv:1907.07901. 2019.

16. Seité S, Khammari A, Benzaquen M, Moyal D, Dréno B. Development and accuracy of an artificial intelligence algorithm for acne grading from smartphone photographs. *Exp Dermatol*. 2019;28(11):1252–7.
17. Hayashi N, Akamatsu H, Iwatsuki K, et al. Japanese Dermatological Association guidelines: guidelines for the treatment of acne vulgaris 2017. *J Dermatol*. 2018;45(8):898–935.
18. Zaenglein AL, Pathy AL, Schlosser BJ, et al. Guidelines of care for the management of acne vulgaris. *J Am Acad Dermatol*. 2016;74(5):945–73.
19. Asai Y, Baibergenova A, Dutil M, et al. Management of acne: Canadian clinical practice guideline. *CMAJ*. 2016;188(2):118–26.
20. Nast A, Dréno B, Bettoli V, et al. European evidence-based (S3) guideline for the treatment of acne—update 2016-short version. *J Eur Acad Dermatol Venereol*. 2016;30(8):1261–8.
21. Bernardis E, Shou H, Barbieri JS, et al. Development and initial validation of a multidimensional acne global grading system integrating primary lesions and secondary changes. *JAMA Dermatol*. 2020;156(3):296–302.