Identification the inflection points of wrinkle types in a large-scale population study of 431,321 subjects

Fudi Wang[1,2,#], Wen Sha[1,#], Siying Fu[2], Xiaoxue Mo[2], Xinyi Fu[2], Sijia Wang[1,*]

[1] CAS Key Laboratory of Computational Biology, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China;

[2] EVELAB INSIGHT（SINGAPORE）PTE.LTD.

[*] Corresponding author's email: wangsijia@picb.ac.cn

Abstract

The aging process affects human skin in various ways, but the extent and timing of these changes remain poorly understood. Previous studies have been limited by small sample sizes, impeding accurate measurement of skin aging rates. In this study, we analyzed eight wrinkle related skin aging traits (i.e., wrinkle under-eyes, wrinkle forehead, crow's feet, nasolabial folds, frown lines, marionette lines, lacrimal sulcus and total wrinkles) from a large sample of 431,321 subjects and developed machine learning models to estimate aging rates. Our results show that the rate of total wrinkle score inflects at age 24, indicating that the rate of wrinkle severity increases after this age. More precisely, our results indicate that the severity crow's feet and nasolabial folds starts to increase at an early age and then stabilizes at around 55 years old. In contrast, marionette lines, frown lines and lacrimal sulcus, tend to occur at a later age, around 30 years old, and also stabilize at an older age. Additionally, our findings suggest that the severity of under-eye wrinkles and lacrimal sulcus undergoes a significant increase at a younger age, around 24 years old, before stabilizing at around 40 years old.

Materials and Methods

Samples

The 431,321 participants enrolled in this study were healthy women aged from 18-70, living in China. The mean age was 45.33. Participants were required to remove make-up and finish the skin test with MeituEve M. Each participant was informed of the testing process before the study.

Phenotyping

We generated the wrinkle phenotypes with MeituEve M after facial cleaning (Figure 1). The device captures high resolution 2D facial images with three HD cameras and five spectral dimensions. Then the fine lines and creases on the skin were restored with a two-step deep learning network, involving a multi-task learning of facial segmentation and a novel detection solution[1].

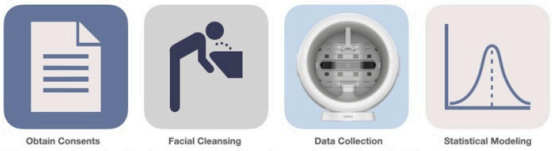


Figure 1. Study design of the inflection points of wrinkle types.

Statistical analysis

We employed two different methods (i.e., polynomial fitting and sigmoid function) to fit the model and calculated R^2 values. The best fitting curve was selected based on the highest R^2 value. The first-order derivative (slope) of the fitted curve was calculated, and the points where the slope began to change rapidly and where to reach its minimum were identified.

Results

The trend of aging varies across different wrinkle phenotypes.

We observed a significant correlation between age and wrinkle phenotypes (PCC: 0.43-0.86, P<0.001). As each wrinkle phenotype is regarded as a biological marker of aging, those findings provide further support for the reliability of wrinkle detection by MeituEve M.We grouped our samples by age, with each age group comprising a ten-year range. Subsequently, we calculated the average wrinkle score for each age group. Our results indicated that all types of wrinkle scores decreased with increasing age, indicating an exacerbation of wrinkles with age (Figure 2). Furthermore, rate of aging varies across different wrinkle type.

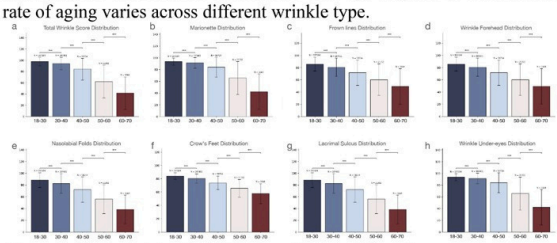


Figure 2. The bar plots of wrinkle phenotypes in different age groups

X axis is the age group and Y axis is the wrinkle score. The trend of aging varies across different wrinkle phenotypes. Wrinkle score was compared between two groups using t tests (*** P<0.001).

Regression models reveal the inflection point of wrinkle

We developed polynomial regression models for each type of wrinkle. By utilizing the first derivative, we obtained the age at which wrinkle scores decline rapidly (onset of aging) and the age at which wrinkle scores change most rapidly (slowing of aging). The blue curve in Figure 3 demonstrates trend of aging for each wrinkle phenotype, while the red curve represents the rate of wrinkle severity.

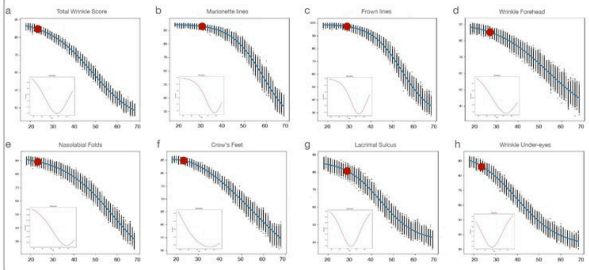


Figure 3. The trend of wrinkle phenotypes. The blue curve demonstrates the trend of aging, and the red curve represents its first derivative, which reflects the slope of wrinkle changes with age. The age point at which the absolute value of the slope exceeds 90% is defined as the onset of aging (red points), whereas the age corresponding to the lowest point of the slope represents the age at which the rate of aging begins to slow down.

Table 1. The inflection points of each wrinkle phenotype

Phenotypes	Model	R^2	Onset Age	Slowing Age	Pattern
Total Wrinkles	Sigmoid	0.99	24	46	-
Frown Lines	Sigmoid	0.99	28	52	
Marionette Lines	Sigmoid	0.99	28	58	😠
Forehead Wrinkles	Sigmoid	0.98	29	53	
Crow's Feet	Polynomial	0.99	23	55	
Nasolabial Fold	Polynomial	0.99	24	58	😊
Lacrimal Sulcus	Sigmoid	0.99	28	44	
Under-eye Wrinkles	Sigmoid	0.99	26	39	others

The results show that the onset and slowing of aging for total wrinkle is age 24 and 46, respectively (Table 1), indicating that the rate of wrinkle severity increases after 24 and slow down after 46. More precisely, our results indicate that the severity crow's feet and nasolabial folds starts to increase at an early age and then stabilizes at around 55 years old. In contrast, marionette lines, frown lines and lacrimal sulcus, tend to occur at a later age, around 30 years old, and also stabilize at an older age. Additionally, our findings suggest that the severity of under-eye wrinkles and lacrimal sulcus undergoes a significant increase at a younger age, around 24 years old, before stabilizing at around 40 years old.

Discussion

Our results indicate that the severity of smile-related expression lines, such as crow's feet and nasolabial folds, starts to increase at an early age. However, anger or sadness-related expression lines, such as marionette lines, frown lines, and lacrimal sulcus, tend to occur at a later age. There is still a need for additional investigation. In conclusion, our big data study provides strong and unprecedented evidence for inflection points of wrinkle manifestion, and underscores the importance of precision cosmetic prevention of different skin aging signs.

Conclusions

- We accurately quantified facial wrinkle phenotypes in a total of 431,321 samples with MeituEve M using deep-learning base algorithm.
- Based on big data analysis of 431,321 individuals, we identified significant variations in the trends of different wrinkle types with increasing age. By performing statistical modeling, we were able to identify the age inflection points of wrinkle phenotypes.
- Our findings have important implications for skin aging protection and skin rejuvenation strategies.

References

C. Qian, T.L., H. Jiang, Z. Wang, P. Wang, M. Guan. (2018). A Detection and Segmentation Architecture for Skin Lesion Segmentation on Dermoscopy Images. In International Conference on Medical Image Computing and Computer. (Granada, Spain).

Contact: S. Wang, Dermatogenomics Group, CAS Key Laboratory of Computational Biology, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China, E-mail: wangsijia@picb.ac.cn

SID 2024

View Abstract

Belowisaproofofyourabstract. Please review for accuracy as this is how your abstract will be printed.

CONTROL ID: 4106229

TITLE: Identifying Facial Regions and Aging Features Associated with Perceived Age: A Deep Learning-Based Facial Aging Assessment Method

AUTHORS (FIRST NAME, LAST NAME): FudiWang [1,2], BaolinChen [1], Zhiyang Li [1], Sijia Wang [2]

INSTITUTIONS (ALL): 1. EVELABINSIGHT(SINAOGE) TE.LTD.
2. CAS Key Laboratory of Computational Biology, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China.

ABSTRACT BODY:
Abstract Body: Facial aging features manifest with considerable inter-individual variability, leading some individuals to appear younger while others appear older. Classic experiments on perceived age rely on human assessment, which demands significant human resources. In this study, we assembled 160 evaluators to assess the perceived age of 3,186 subjects' faces. Evaluators were stratified by gender and divided into four age groups (18-30, 30-40, 40-50, 50-60), with 20 individuals per group. Then, we employed deep learning model (Short-Term Dense Concatenate network, STDC network) to simulate human age assessments. Results revealed that female evaluators yielded more accurate assessments compared to male (MAE: 2.68 vs. 2.77), with the most accurate evaluations observed for women aged 40-50 (MAE=2.59). Furthermore, deep learning was utilized to visualize facial regions most influential in perceived age. Facial regions were divided into five areas (i.e., forehead, cheeks, nose, eyes and mouth), with eyes exerting the greatest impact on facial aging compared to other areas (β=0.12, p=1.02e-04). Additionally, we quantified eight main facial aging features (i.e., pigmentation spots, nasolabial folds, forehead wrinkles, crow's feet, periorbital wrinkles, tear troughs, under-eye bags, glabella wrinkles and wrinkle under eyes) and investigated their respective effects on perceived age. We found that under-eye bags exhibited the highest weight (β=0.21, p=2.21e-3), while forehead wrinkles had the smallest impact (β=0.03, p=0.02). Thus, this study quantifies facial aging using deep learning model and identifies facial regions and aging features with the greatest influence on perceived age.

KEYWORDS: Aging,AI(ArtificialIntelligence),Imaging.

CURRENT PREFERRED CATEGORY: SystemsBiology, Advanced Technologies, and Imaging

CURRENT SECOND PREFERRED CATEGORY: Clinical Research - Epidemiology and Observational Research

PRIMARY MEMBERSHIP:
Fudi Wang : 7. Other

GENDER:
Fudi Wang : Female

Source of Support: Institution

Disclosure Agreement: Agree to all disclosure conditions.

Journal of Cosmetic Dermatology

WILEY

JCD Journal of Cosmetic Dermatology

ORIGINAL ARTICLE OPEN ACCESS

Transdermal Delivery of Baicalin Based on Bio- Vesicles and Its Efficacy in Antiaging of the Skin

Liang Chen[1,2] | Fudi Wang[3] | Xiaoyun Hu[1,2] | Nihong Li[1,2] | Ying Gao[4] | Fengfeng Xue[5] | Ling Xie[1,2] | Min Xie[1,2]

[1]Scientific Research Laboratory, Shanghai Le- Surely Biotechnology Co. Ltd, Shanghai, China | 2SASELOMO Research Institute and Biological Laboratory, Shanghai Chuanmei Industrial Co. Ltd, Shanghai, China | 3Evelab Insight (Singapore) Pte. Ltd, Singapore, Singapore | 4Zhejiang Moda Biotech Co. Ltd, Hangzhou, China | 5Nanomedicine and Intestinal Microecology Research Center, Shanghai Tenth People's Hospital, School of Medicine, Tongji University, Shanghai, China

Correspondence: Liang Chen (liang_chen13@fudan.edu.cn) | Fengfeng Xue (xuefengfeng@msn.com)

Received: 5 September 2024 | **Revised:** 25 December 2024 | **Accepted:** 24 January 2025

Funding: The authors received no specific funding for this work.

Keywords: antiaging | baicalin | bio- vesicle | skin physiology and cell culture | transdermal delivery

ABSTRACT

Objective: To develop a stable and efficient delivery system for baicalin, a flavonoid with potential antioxidant and antiaging properties, to overcome its limitations in solubility, stability, and skin permeability.

Methods: Baicalin was encapsulated using ATP synthase molecular motor technology into bio- vesicles derived from yeast/bacillus cell membranes, forming "motor baicalin" (MB). The liposome baicalin (LB), baicalin raw material (BRM), and bio- vesicles were used for comparison. The stability, transdermal penetration, and antioxidant activity of MB, LB, BRM, and bio- vesicles were evaluated through in vitro and in vivo tests.

Results: MB formed a stable core- shell structure, significantly enhancing the water solubility and long- term stability of baicalin. The tests confirmed superior transdermal penetration and antioxidant activity of MB, evidenced by increased expression of SOD, CAT, and GSH- Px enzymes and improved cell proliferation and migration. Clinical trials demonstrated significant reductions in wrinkle depth and improvements in skin elasticity.

Conclusion: This study presents a promising approach to improving the stability and transdermal delivery of baicalin. MB showcases potent antioxidant and antiaging properties, making it a valuable component in skincare products.

1 | Introduction

Skin aging, characterized by increased wrinkles, dullness, dryness, pigmentation, and loss of elasticity, is a complex biological process induced by intrinsic and extrinsic factors over time [1]. Generally, UV radiation, inflammation, glycation, and reactive oxygen species (ROS) are believed to be the primary contributors to skin aging [2]. Among these, ROS is considered to play a dominant role, as the other three factors ultimately accelerate skin aging by upregulating ROS levels in the skin [3, 4]. Although it is well known that low concentrations of ROS have positive effects on cell signaling, cell proliferation, differentiation, and autophagy [5], excessive ROS accumulated over the years can attack cells and tissues, leading to the loss of collagen and intercellular lipids and inducing skin aging. Consequently, understanding the relationships between ROS and skin aging phenotypes, as well as developing new and highly effective antioxidants are current research hotspots in cosmetic industry [6–8]. In recent years, numerous antioxidant ingredients have been developed and incorporated into cosmetics, such as Vitamin C (VC), Vitamin E (VE), ferulic acid, coenzyme Q10, and res- veratrol. However, the effectiveness of these antioxidants is

limited by their structural instability and low antioxidant efficiency [9–11]. As is well known, there are two levels of antioxidant systems in our skin: relatively inefficient nonenzymatic defenses and highly efficient enzymatic defenses. The nonenzymatic antioxidant system neutralizes ROS through compounds like VC, VE, and certain carotenoids, whereas the enzymatic antioxidant system primarily relies on activating the nuclear factor erythroid 2- related factor 2/antioxidant response element, (Nrf2/ARE) signaling pathway. This activation promotes the expression of powerful antioxidant enzymes, such as catalase (CAT), superoxide dismutase (SOD), and glutathione peroxidase (GSH- Px) [12]. Therefore, developing more advanced antioxidant ingredients necessitates finding a substance that activates and stabilizes the Nrf2/ARE pathway.

Baicalin, a flavonoid derived from Scutellaria baicalensis Georgi, exhibits a range of biological activities including anti-inflammatory, antineoplastic, and immunoregulatory properties. It has, thus, become a significant pharmacological component in the biomedical field [13, 14]. Recent studies have demonstrated that baicalin possesses excellent antioxidant properties by activating the Nrf2/ARE pathway, making it a promising antioxidant for use in cosmetics [15]. However, baicalin monomer is insoluble in water and polyol solvent, making it prone to precipitation from formulations, which results in product instability. Furthermore, the precipitated baicalin is susceptible to discoloration and off- odor under high temperature and light conditions. This not only compromises the efficacy of the products but also significantly affects the user experience. Moreover, BRM has been shown to have poor skin absorption, potentially leading to excessive local epidermal concentration and skin irritation. Therefore, in order to expand the application of baicalin in cosmetics, it is essential to address these aforementioned drawbacks first. In fact, recent advancements in transdermal drug delivery systems used in cosmetics can effectively resolve similar issues. For instance, liposome encapsulation technology not only stabilizes certain ingredients but also enhances their transdermal absorption, thereby improving the efficacy while reducing irritation.

In this study, we developed a novel baicalin delivery system based on our previous advancements in ATP synthase molecular motor technology [16, 17]. The ATP molecular motor, also known as ATP synthase, plays a crucial role in cellular energy metabolism, encompassing processes such as nerve conduction, muscle contraction, and substance transport. This molecular motor features a distinctive bio- vesicle structure that can encapsulate active molecules for effective delivery. Our previous research has demonstrated that the ATP molecular motor facilitates directional movement of the entire loading system with F1 rotating during ATP hydrolysis [18]. Leveraging the low pH active targeting capability of ATP molecular motors and the weak acidity characteristic of human skin tissue, we found that baicalin encapsulated within these bio- vesicles can effectively penetrate deep into the skin. This mechanism is essential for enhancing the transdermal delivery efficiency of this technology.

Specifically, BRM is encapsulated within bio- vesicles to create a “core- shell” structure. The resulting baicalin delivery system is designated as “motor baicalin” (MB). For comparison, the baicalin raw material (BRM), bio- vesicles (without baicalin), and liposome baicalin (named LB, a commonly used modi- fied technology in cosmetics) also proceeded. As a result, both in vitro and in vivo tests have shown that MB exhibits excellent transdermal properties along with antiaging and skin- repairing effects. This study presents a promising strategy to address challenges related to instability and transdermal delivery of active ingredients in cosmetic formulations.

2 | Materials and Methods

2.1 | Materials

Commercial baicalin (Cosmetic Grade; LogP: 0.1422; Solubility: nearly insoluble in water and polyol solvent; Stability: stable in dilute acid) was purchased from Shanxi Jiahe Biotechnology Co. Ltd. Retinol powder (≥ 95%) was obtained from Sigma- Aldrich Reagent Company. Trifluoroacetyl tripeptide- 2 powder was sourced from Shanghai Peptide Biotechnology Co. Ltd., PR China. Acetone (AR), isotonic phosphate buffer, and fluores- cein (AR) were acquired from Shanghai Titan Technology Co. Ltd. NaCl (AR), propylene glycol (AR), hydrogenated lecithin (AR), caprylic acid glyceride (AR), disodium EDTA (AR), DAPI solution, and other reagents were purchased from Greagent Company. SOD test kit, CAT test kit, GSH-Px test kit, hematox- ylin, Ponceau-Fuchsin, and aniline blue were purchased from Beyotime Biotechnology. Purified water (18 MΩ) was obtained using a Milli-Q system (Millipore). All the chemical reagents were utilized without further purification.

2.2 | Instruments

The MB or LB sample was sonicated by the bath type of sonication (KQ- 700DM) from Kunshan Ultrasonic Instrument Co. Ltd. (China) with the ultrasonic frequency of 40 KHz at 25°C. In these sonication conditions, no structural modifica- tions occurred during the sonication process. DLS and zeta potential were measured with a Malvern Nano- ZS90 (Britain). Transmission electron microscopy (TEM) was performed using an Ultra 55 and an FEI Tecnai 20 operating (America) at 200 KV. The test sample was dispersed in an ethanol/water mixture (7:3, v/v), sonicated for 5 min, and then dropped onto a copper grid for examination. HPLC measurements were conducted on a 1260 Infinity system (Agilent, America) with Elite C18 OSD- BP column. The confocal images were measured by a Nikon single- photon confocal microscope (NikonC1Si, Japan). The fluorescence spectra were conducted on a fluorescence mi- croscope (Leica Microsystems Model DMi8 manual, Germany). Stability was tested in an incubator with different temperatures (YiHeng Shanghai, DHP9402). pH values were measured by a pH meter (Mettler Toledo, S400-K).

2.3 |The Preparation of MB

The preparation of MB was conducted following a previously reported method with appropriate modifications, and the utilizing bio- vesicles derived from the cell membranes of yeast and *Bacillus* [18]. The detailed procedure is outlined as follows:

2.3.1 | Preparation of Baicalin Cyclodextrin Encapsulated Nanoparticles

Cyclodextrin (5 equal) is weighed and subjected to grinding in a colloid mill. Subsequently, baicalin (15 equal) is also weighed and added to the colloid mill. The mixture is thoroughly ground to ensure homogeneity. Following this, water (100 equal) is introduced into the mixture to completely dissolve the evenly mixed cyclodextrin–baicalin powder. The resulting solution undergoes homogenization using a high- pressure microjet homog- enizer at a pressure of 30 MPa. This homogenization process is repeated several times to achieve uniformity. Finally, the solu- tion is collected and subjected to freeze- drying at −80°C to re- move the solvent, achieving baicalin cyclodextrin encapsulated nanoparticles.

2.3.2 | Cultivation of *Bacillus subtilis* and Purification of the Vesicle Fraction

The freeze- dried *Bacillus subtilis* powder is dissolved in nutrient broth (NB) medium. The resultant bacterial suspension is transferred into sterile test tubes containing an appropriate volume of liquid culture medium and mixed thoroughly for cultivation purposes. Subsequently, the activated bacterial strain is introduced into a conical flask filled with NB medium and incubated under optimal shaking conditions for further use. Afterward, *Bacillus subtilis* cells (24 equal) are transferred into NB liquid medium supplemented with a designated amount of yeast powder (55 equal) for further cultivation. Upon completion of this cultiva- tion phase, the fermentation broth is collected and centrifuged to obtain bacterial sediment.

The bacterial sediment is then resuspended in a phosphate buffer solution at a defined volume ratio. This buffer solution containing the bacterial cells undergoes homogenization followed by treatment with high- pressure microjet homogenization at a pressure of 30 MPa. The resultant solution containing bacterial fragments is passed through a solid chromatography column to separate and purify the vesicle fraction.

2.3.3 | Preparation of the Target MB Product

A measured quantity of the vesicle fraction is diluted with phosphate buffer solution. An appropriate amount of baicalin cyclodextrin encapsulated nanoparticles is weighed and combined with the vesicle fraction. The mixture undergoes ultrasonic treatment using an ultrasonic instrument to ensure thorough mixing. The resulting liquid is then transferred to a high- pressure extruder, where nitrogen serves as the pressure source. The mixture is extruded through a polycarbonate membrane and collected via centrifugation. Following precipitation, the product is dried and ground to obtain the final target product.

2.4 | The Preparation of LB

A specific amount of propylene glycol (35.5%) and hydrogenated lecithin (7%) are weighed and stirred for 10 min at 75°C and 600 rpm until all particles dissolve in the liquid. Subsequently, an appropriate quantity of caprylic acid glyceride (4.5%) is added to this solution, which continues stirring for an additional 5 min under identical conditions until a homogeneous mixture is achieved. Next, a measured amount of baicalin (8.5%) is introduced into the mixture, followed by stirring for another 10 min at 75°C and 600 rpm until a uniform yellow liquid forms. Afterward, an ap- propriate amount of disodium EDTA (0.05%) is incorporated into the solution and stirred for 5 min at the same temperature until achieving a consistent yellow liquid. The resultant solution undergoes homogenization using a mi- crojet homogenizer across three cycles at a pressure of 2000 psi while maintaining discharge temperatures at 50°C during homogenization and cooling down to 15°C afterward. This process yields a slightly viscous yellow liquid. Finally, the solution centrifuged at 5000 rpm for 10 min to eliminate any insoluble solid impurities.

2.5 | Stability Test

The structure stability of MB was assessed by monitoring changes of zeta potential and particle size over a period of 90 days across temperatures: −20oC, 4oC, room temperature (RT), and 40oC through a Malvern Nano- ZS90. We also monitored the change of loading capacity and pH value within 90 days at dif- ferent temperatures of −20oC, 4 oC, RT, and 40oC by using HPLC and S400- K.

2.6 |Detection of Baicalin Loading Capacity in MB

The baicalin loading capacity was determined using high-performance liquid chromatography (HPLC), with a baicalin standard sample sourced from Aladdin, which has a purity of ≥ 98%. The conditions and samples used were as follows: an Elite C18 OSD- BP column maintained at 25°C, a flow rate of 1 mL/min, and a detection wavelength set at 280 nm. The mobile phase consisted of methanol and 0.2% of phosphoric acid in a ratio of 47:53. Acetone was used as the solvent for the standard curve samples, while pure water was used for the test samples. The calculation of standard curve equations (concentration range from 0 to 100 μg/mL) was y = 34.526x- 28.428 with the $R2$ value of 0.9987 (Figure S1).

2.7 | In Vitro Transdermal Release Experiment (Franz Diffusing Cells)

The test was conducted at 32°C using an isotonic phosphate buffer (pH 7.4) as the medium. When the amount of baicalin added was 2 mg/mL, the solution became clear and transparent, and no precipitation or stratification was observed after being stored at room temperature for 48 h. Three groups were established: the BRM group, the LB group, and the MB group, each with three parallel subgroups. The skin samples were ob- tained from the back of 1 kg New Zealand white rabbits (the related animal study was approved by the Ethics Committee on Institute of Basic Medicine and Cancer, Chinese Academy

of Sciences with the approval number 2023R0071). Following isoflurane respiratory anesthesia, the rabbits were humanely euthanized by injecting 1 mL of an appropriately concentrated potassium chloride solution. The back skin was harvested, with the hair and fat layers removed, thoroughly washed with water, divided into appropriate sizes, and stored in normal saline at 4°C for later use. The active solutions used in the diffusion cell were diluted with isotonic buffer. The concen- trations of BRM, MB, and LB were 150 μg/mL, 1 mg/mL, and 1.76 mg/mL, respectively (both equivalent to 150 μg/mL of BRM). For each experiment, 2 mL of the active solution was added to the donor cell. The receiving cell contained 8 mL of isotonic phosphate buffer (potassium chloride, sodium chlo- ride, pH 7.4). The skin was placed between the donor and re- ceiving cells and clamped in the transdermal instrument. At specified time intervals, 0.5 mL of the receiving cell solution was withdrawn and replaced with an equal volume of isotonic solution. After collecting samples at each time point, the solu- tions were analyzed using a UV spectrophotometer at 280 nm. The concentration of baicalin was calculated using a baicalin standard curve. The transmittance was then calculated to create the chart.

2.8 |Transdermal Penetration Test

The baicalin molecule itself lacks fluorescence, rendering it unsuitable for tracking via fluorescence methods. Therefore, in this study, baicalin was labeled with fluorescein isothiocyanate (FITC) to facilitate fluorescence tracking. The labeling proce- dure is outlined as follows (Figure S2): Appropriate amounts of baicalin were dissolved in an acetic acid–sodium acetate buffer.

Subsequently, tyramine (p-hydroxyphenethylamine) and so- dium cyanoboride were added successively, the mixture was then stirred in a water bath under dark conditions. FITC was introduced to the tyraminated baicalin using borate buffer as the labeling medium and allowed to react overnight. Denaturing vertical plate polyacrylamide gel electrophoresis was employed subsequently to detect and separate the FITC- labeled baicalin. Further purification involved separating the free baicalin from the labeled compound using a glucose gel column, followed by elution with Tris–HCl to obtain FITC-labeled baicalin.

The experiment utilized skin from 3-month- old miniature pigs.

Following hair removal, a scraper was used to remove the fat layer and a portion of the dermis. The skin was then cut into circular pieces, each approximately 3.14 cm2 in area. These cuticles were secured on the penetration device with the epidermis facing upward.

The MB actives, LB actives, and BRM sample were diluted to a concentration of 2‰ (2 mg/mL, based on the baicalin concentration) with PBS buffer at pH 7.4. Subsequently, 0.5 mL of each diluent was added to the penetration device and incubated in a dark environment for 1 h. Postincubation, the skin was removed and meticulously rinsed with water until no surface fluorescence discoloration was observed. The surface water was absorbed using absorbent paper, and the skin was slightly dried in the dark. The samples were then sectioned and observed under a fluorescence microscope to assess the penetration of fluorescence into the skin.

2.9 |Cell Uptake Assay

L929 fibroblasts with a density of 1 × 106 cells per well were seeded in dishes overnight and subsequently incubated with FITC- labeled BRM, FITC-labeled MB, and FITC-labeled LB at equivalent concentrations for 2, 4, and 8 h and then, stained with DAPI (P0131, BYT) at room temperature. The cellular uptake and intracellular distribution were observed by a Nikon single- photon confocal microscope. The FITC- labeled baicalin can reflect the "uptake effect" of cells on baicalin by observing the intracellular fluorescence intensity at specific excitation/emis- sion wavelengths. Fluorescence intensity was measured using a fluorescence enzyme-linked immunosorbent assay (ELISA) reader, and semiquantitative analysis of optical density intensity was performed using Image J.

2.10 |Anti- ROS Property

The anti- ROS detection was conducted using a 2′, 7′- dichlorofluorescein diacetate (DCFH-DA) stain assay. The

experiment was divided into six groups: Control, UVB or H2O2, MB (133.3 μg/mL, equal with BRM 20 μg/mL), LB (235 μg/mL, equal with BRM 20 μg/mL), BRM (20 μg/mL), and bio- vesicles (113.3 μg/mL, equal with MB without BRM). Human skin fibroblast (HSF) cells were plated in 12- well plates (2 × 105 cells per well) and treated with cell culture medium, MB, LB, BRM, and bio- vesicles for 24 h. The cells were then exposed to UVB (200

mJ/

cm2) or H2O2 (1.6 mM, for 30 min) to induce ROS except for the control group. Subsequently, the cells were incubated with 10 μM DCFH- DA for 20 min at 37°C. After washing with PBS, fluorescence images were visualized using a fluorescence microscope (Leica Microsystems Model DMi8 manual).

2.11 | The Enzymatic Activities of SOD, CAT, GSH- Px In Vitro

The enzymatic activities of SOD, CAT, and GSH-P x were tested in HSF cell lines. Hydrogen peroxide (H2O2) and UVB radiation were introduced as stimuli to establish a ROS damage model. The experiment was divided into six groups: Control, UVB or H2O2, MB (133.3 μg/mL), LB (235 μg/mL), BRM (20 μg/mL), and bio- vesicles (113.3 μg/mL). HSF cells were seeded into a six-well plate and incubated with the cell culture medium, MB, LB, BRM, and bio- vesicles for 24 h. Subsequently, the cells were exposed to UVB or H2O2, washed twice with PBS, and replenished with a complete culture medium. The cells were then continuously cultured at 37°C and 5% CO2 for another 24 h. Finally, the enzyme activities were assessed using SOD, CAT, and GSH-Px enzyme activity detection kits.

2.12 | MMP- 1, TIMP- 1 Expression, and Photo Damage Repair Assay

HSF cells were utilized in this experiment to assess their ability to express MMP- 1 and TIMP- 1. Initially, the cells were cultured to a specific density and then incubated with MB (133.3 μg/ mL), LB (235 μg/mL), BRM (20 μg/mL), bio- vesicles (113.3 μg/ mL), retinol (20 μg/mL), and trifluoroacetyl tripeptide- 2 (20 μg/

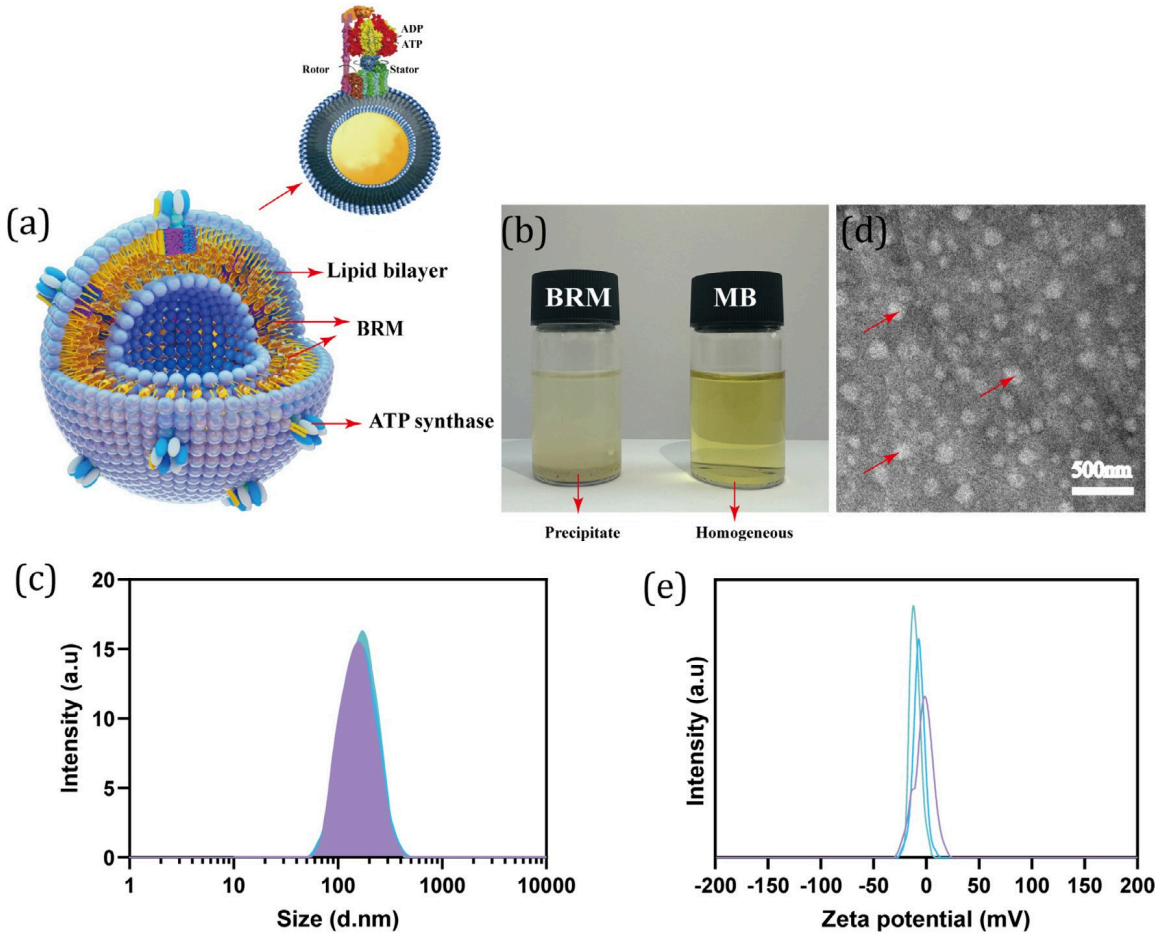


FIGURE 1 | Physical and chemical data of motor baicalin (MB). (a) The structure diagram of MB shows ATP synthase molecules on the surface of bio- vesicles. The yellow particles dispersed in the hydrophobic phospholipid bilayer present BRM. (b) The solubility of BRM and MB in water. (c) The particle size distributions of three batches of MB. (d) TEM image of MB, where the white circular hollows depict MB vesicles. (e) Zeta potential properties of three batches of MB.

TABLE 2 | Characteristics of baicalin delivery system of MB.

	Parameters	Batch 1	Batch 2	Batch 3	Average value
Particle size	Z- Average, d.nm	147.6	145.6	144.6	145.9
	PDI	0.21	0.17	0.17	0.18
Zeta potential	Zeta potential, mV	−2 .7 1	−2 .9 7	−4.54	−3 . 41
Loading capacity	Percentage, %	14 .7 0	15.20	14.80	14 .9 0

batches is approximately 14.90 wt% (Table 2), which exceeds the LB values (approximately 8.5 wt%). Moreover, the structural stability of MB is crucial for practical application. Consequently, we also tested its long- term stability over 90 days. Results demonstrated no significant changes in MB's pH values, baicalin load- ing capacity, particle size, or zeta potential values at different temperatures (Figure S3), further revealing the good long- term stability of MB.

3.2 | In Vitro Release Assessment

The in vitro drug release measurement is a crucial method for evaluating the efficacy of bio- vesicle delivery technology and estimating its in vivo performance qualitatively. In this study, we assessed the release kinetics of MB using the Franz dif- fusion cells method. For comparison, BRM and LB were also evaluated. As illustrated in Figure S4, a burst release effect was observed within the first 4 h of BRM application, with ap- proximately 60% of baicalin being released. This rapid release could lead to skin sensitivity due to local hyper- concentration. Indeed, high concentrations of BRM exhibited significant cyto- toxicity toward both HaCaT and HSF cell lines (Figure S5). In contrast, both MB and LB demonstrated slow- release behavior during this initial period, approximately 39% and 29% of bai- calin were released from MB and LB, respectively, within the first 4 h. These findings suggest that both MB and LB samples conform to a first- order release model, extending the release time of the active substance may effectively mitigate potential cytotoxicity and skin irritation risks. This phenomenon can be

attributed to the fact that when used as potent antioxidants or whitening agents in cosmetics, polyhydroxyl structures or acid anhydride forms of baicalin may compromise the skin's barrier function. When extending the observation period to 60 h, it was found that the release ratio for MB reached 88%, whereas LB only achieved 64%. This indicates that MB exhibits superior release efficiency compared to LB, which is also very crucial for their realization of antiaging efficacy.

3.3 | Transdermal Penetration and Endocytosis Evaluation

To investigate the potential efficacy of MB in cosmetics, an in vitro test was conducted to examine transdermal penetra- tion and endocytosis performance. For comparison, BRM and LB were also evaluated. As illustrated in Figures 2 and S7, BRM demonstrated notably poor transdermal permeability, with a penetration depth of only 264 ± 2 μm, remaining primarily on the skin's surface as evidenced by fluorescence imaging. In contrast, following modification with lipid and bio- vesicle tech- nology, LB and MB demonstrated enhanced skin penetration capabilities of 555 μm and 1347 μm, respectively, indicating a significant improvement in penetration. Unlike traditional lipo- some technology and static vesicle encapsulation methods, the primary advantage of ATP molecular motors lies in their ability to undergo rotational disturbance upon receiving biological signals. When the ATP molecular motor bio- vesicles contact cell membrane, this rotational disturbance further facilitates the fusion between biofilms and enhances the delivery efficiency of active molecules. Consequently, utilizing ATP molecular motor bio- vesicles to encapsulate baicalin markedly improves its skin penetration and cellular uptake ability, thus significantly amplifying its antiaging effects. The ability of active ingredients to penetrate the skin is cru- cial, but their effectiveness also largely depends on their ca- pacity to bind to cells. Therefore, we conducted an endocytosis experiment using L929 fibroblasts. Similar to the transdermal results, the BRM sample was not captured by the cells, owing to its poor permeability, as evidenced by the almost complete absence of fluorescence in the cytoplasmic, even after 8 h of coincubation. In contrast, the cells displayed a relatively strong uptake of LB and MB actives, especially the MB actives. Compared with the LB actives, a substantial amount of bai- calin was found in the cytoplasm of the MB. The remarkable uptake of MB is possibly derived from the ATP synthase units on the bio- vesicles, as ATP synthase can act as a molecular motor, driving the bio- vesicles that carry baicalin to integrate into specific skin cells.

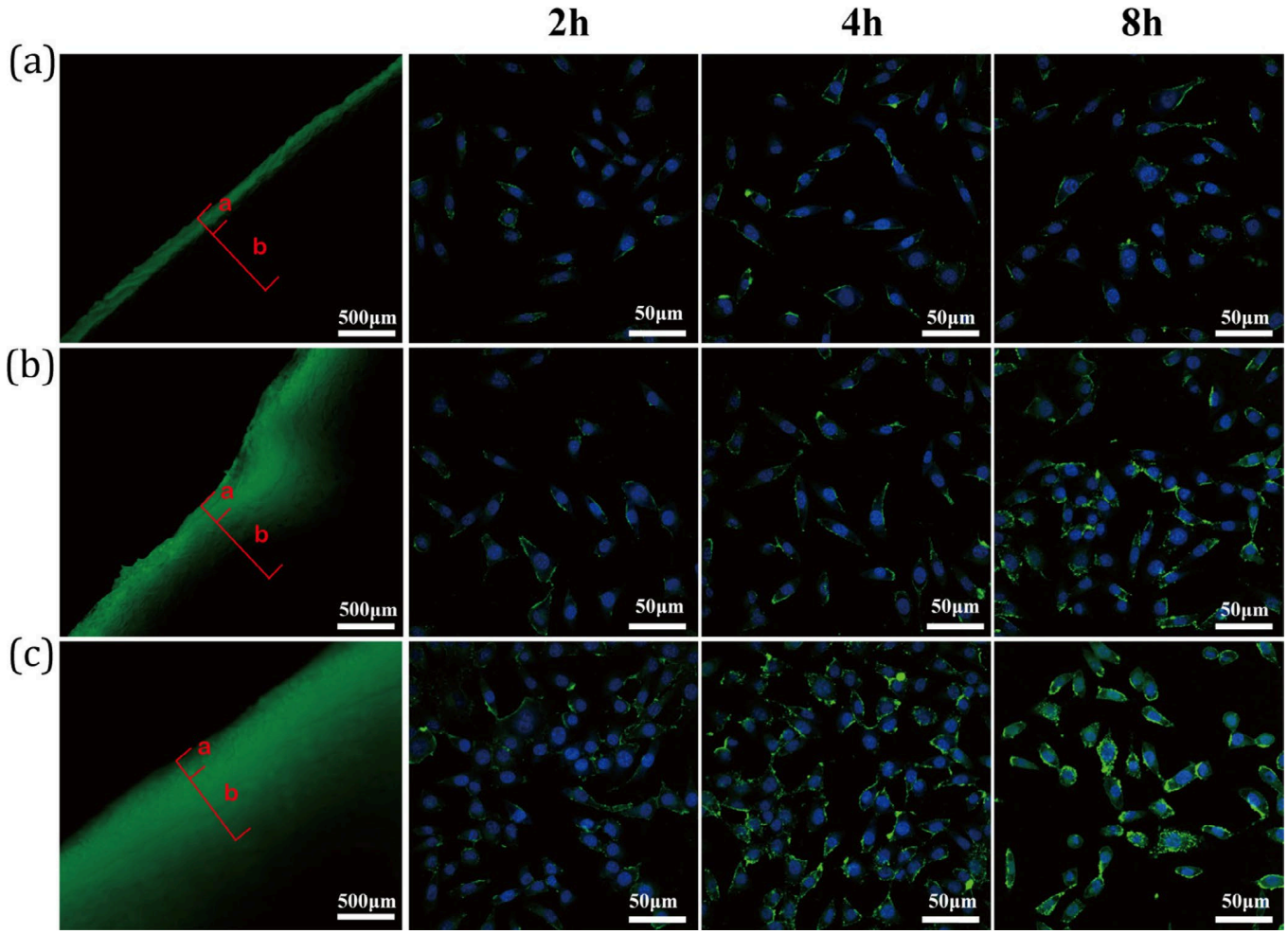


FIGURE 2 | Transdermal penetration and cellular uptake properties of (a) BRM, (b) LB, and (c) MB. In the figure, the red labels “a” and “b” denote the epidermis and dermis, respectively.

3.4 | Effect of MB Against Cell Damage Caused by Oxidative Stress

The CCK- 8 assay of bio- vesicles and baicalin- related actives in different concentrations confirmed that bio- vesicles and baicalin-related actives exhibited excellent bio- compatibility (Figure S5). The cell viability of MB, LB, BRM, and bio-

vesicles exposed to UVB and H2O2 is shown in Figure S6. The results exhibited that the antioxidant efficacy of MB was superior to LB and BRM, which can be attributed to the enhanced delivery efficiency of BRM by ATP synthase molecular motor technology. Then, we further investigated the relative activity changes of three key intracellular enzymes: SOD, CAT, and GSH- Px, which are strongly correlated with cellular endogenous antioxidant capacity. As shown in Figure 3, the activities of SOD, CAT, and GSH- Px were significantly reduced to less than 50% compared to the control cells ($p < 0.05$) when exposed to UVB or H2O2, indicating severe oxidative damage. Subsequently, treatment with baicalin- related actives markedly improved the expression of these enzymes, particularly in the MB group. The activities of SOD, CAT, and GSH-Px in the MB group increased to 101.5%, 122.9%, and 128.9% in

the UVB model and to 98.2%, 86.2%, and 121.3% in the H2O2
model, respectively ($p < 0.05$), compared to the blank control group, demonstrating the excellent antioxidant potential of MB actives. This result is consistent with the immunofluorescence test findings for UVB (Figure S8) and H2O2 (Figure S9) as ROS stimuli.

3.5 |Effect of MB on Antiaging Evaluation

Matrix metalloproteinase- 1 (MMP- 1) and tissue inhibitor of metalloproteinase- 1 (TIMP- 1) are a pair of enzymes in our skin that regulate the expression and degradation of collagen, play- ing a crucial role in skin aging. To explore the antiaging proper- ties of baicalin, we examined the activity changes of MMP- 1 and TIMP- 1 in baicalin- related actives. As shown in Figure 4a,b, when cells were incubated with test samples, all baicalin- related groups, including BRM, MB, and LB, demonstrated a significant inhibitory effect on the MMP- 1 gene compared to the blank con- trol group ($p < 0.05$), except for the bio- vesicles group. Among these, the MB active exhibited the most substantial inhibition, indicating excellent collagen protection potential. Similarly, MB actives also showed the most pronounced promotion of TIMP- 1 expression, further confirming their superior collagen- protective and antiaging capabilities. Furthermore, to elucidate the exceptional collagen- protective performance of MB more clearly, commonly used antiaging ingredients such as retinol and trifluoroacetyl tripeptide- 2 were evaluated for comparison purposes. The results disclosed that MB exhibits comparable or even superior inhibition of MMP- 1 while promoting TIMP- 1 activity relative to both retinol and trifluoroacetyl tripeptide- 2 (Figure S10). In addition, the protein levels of MMP- 1 and TIMP- 1 treated with BRM, MB, LB, and bio- vesicles were con- firmed by ELISA (Figure S11a,b). This suggests that MB- based ATP molecular motors enhance the delivery of baicalin effec- tively, bolstering collagen protection and antiaging properties.

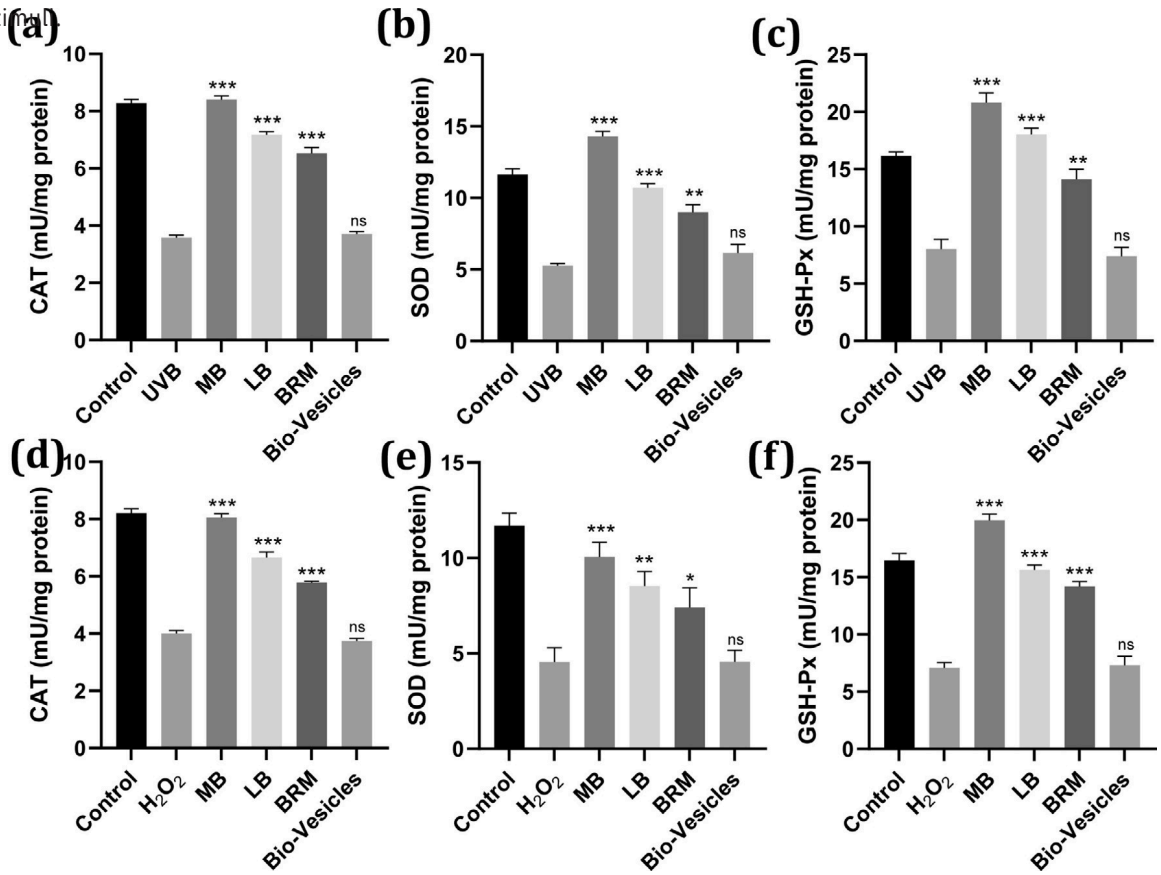


FIGURE 3 | Antioxidant capabilities of various baicalin- related actives and bio- vesicles. Panels (a, d) depict CAT, (b, e) show SOD, and (c, f) represent GSH- Px gene expression in HSF cells. Data are reported as mean values ± SD. The symbol * indicates values that are significantly different from those of the UVB or H2O2 groups (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

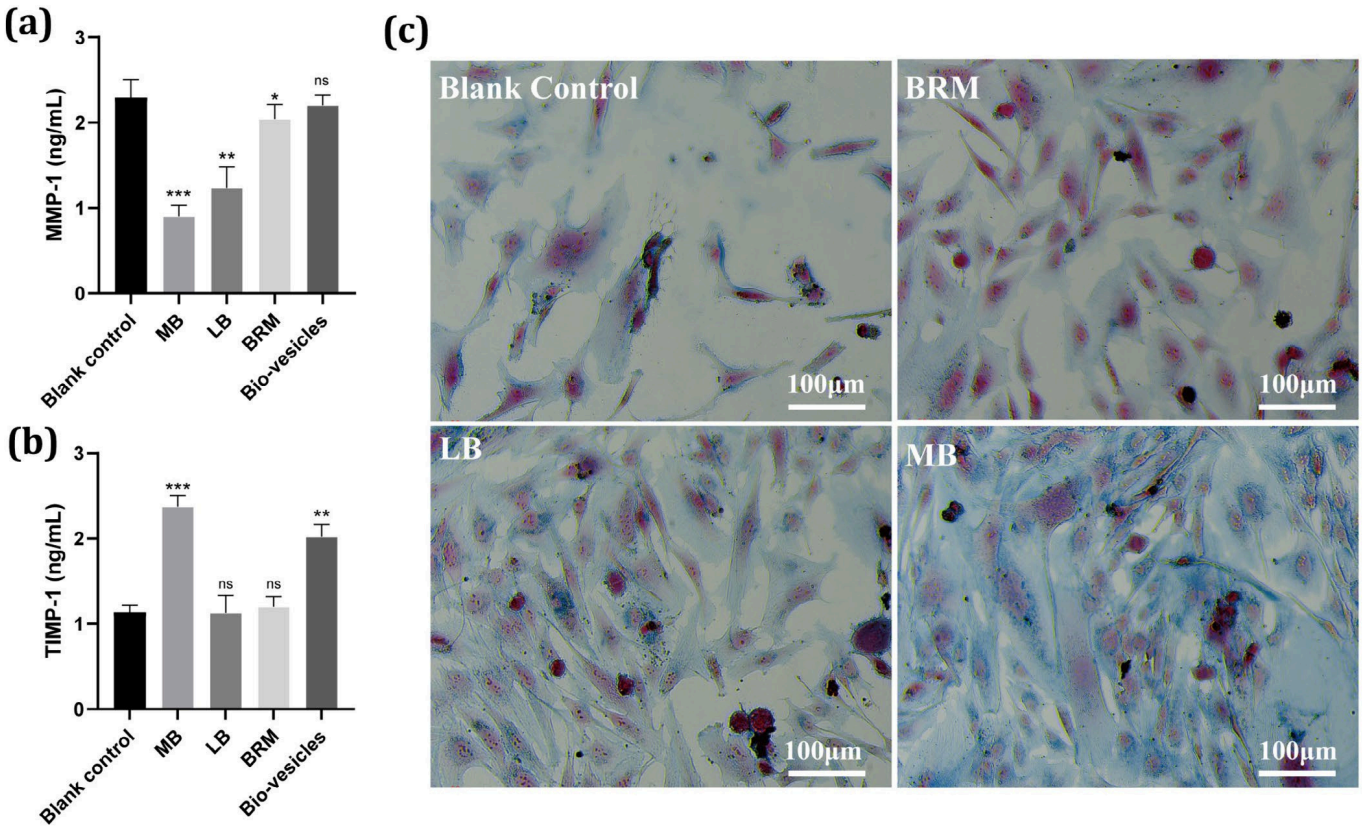


FIGURE 4 | Antiaging properties of various baicalin- related actives and bio- vesicles. (a, b) MMP- 1 and TIMP- 1 gene expression in HSF. Data are presented as mean values ± SD. An asterisk (*) indicates values that are significantly different from the blank control (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$). (c) Cell fibrous staining images of BRM, LB, and MB actives. The blue regions represent the fibrous structures of HSF, while the red regions denote the fibroblasts themselves.

3.6 | Effect of MB on Repairing of UV- Damaged Fibroblasts

Fibroblasts are a pivotal component of skin tissue, and the density, thickness, and orientation of fibrous structures within them predominantly reflect the cellular state and degree of aging. UV radiation can thin collagen fibers within cells, thereby accelerating cellular aging. In this study, we further investigated the reparative and promoting effects of baicalin on fibroblasts post- UV irradiation. Human skin fibroblasts were used in the experiment; for details, see the experimen- tal section. After UV exposure, the fibrous structures in the blank control group became thinner and exhibited abnormal orientation, indicating accelerated cellular aging due to UV irradiation (Figure 4c). Upon incubating UV- damaged cells with baicalin- related actives, a significant improvement in cell vitality was observed. The fibrous structures within the cells appeared thicker compared to the blank control. This suggests that all three baicalin- related actives have a reparative effect on photoaged human skin fibroblasts. Among them, the MB actives showed the most significant enhancement, evidenced by a large amount of blue collagen fibers generated around the cells. This may be attributed to the strong uptake effect of fi- broblasts on MB actives. In addition, the expression protein level of Col- I also confirmed that all three baicalin- related ac- tives have a reparative effect on photoaged HSF by increasing Col- I (Figure S11c).

3.7 |Effect of MB on Wound Healing

Enhanced fibroblast migration is crucial for effective skin healing and regeneration, particularly in aged or photodamaged skin. By facilitating faster wound closure, enhancing collagen production, and restoring skin structure and elasticity, fibroblast migration plays a significant role in overall skin health. The application of therapies designed to enhance fibroblast function presents a promising strategy for improving the ap- pearance and quality of aging skin. Therefore, the scratch assay was performed on HSF cells to evaluate the effects of three baicalin-related actives on cell spread in the wound area. As shown in Figure 5, the migration of HSF cells was calculated to be 53.8%, 79.8%, 59.2%, 52.5%, and 54.7% for control, MB, LB, BRM, and bio- vesicles group, respectively. Compared to BRM and LB actives, the MB actives showed a superior ability to pro- mote HSF cells move to scratch areas, indicating that MB could increase the migration ability of HSF cells. Therefore, baicalin can significantly enhance cell vitality, combined with its ability to repair fibers, suggesting its great potential in antiaging field.

3.8 |Clinical Research

Given the excellent antiaging potential demonstrated by MB actives in vitro, we conducted clinical research to further validate these findings. The in vivo instrumental evaluation results are

summarized in Figure 6. Compared to the baseline, the target group with 2% MB additives showed significant benefits in reducing wrinkles and improving elasticity. According to Antera 3D data, there was a notable reduction of 8.30%, 7.06%, 7.43%, and 4.27% in the depth of nasolabial folds, forehead wrinkles, under- eye wrinkles, and crow's feet, respectively, on Day 28. Additionally, the length of nasolabial folds, marionette lines, and under- eye wrinkles decreased by 3.46%, 9.59%, and 32.31%, respectively. These effects were significantly different from the placebo at both intervals. Furthermore, an increase of 10.92% in the *R*2 value, indicating improved skin elasticity, was ob- served in the 2% MB additives group on Day 28 compared to the placebo. Figures 7 and 8 also provide illustrative examples highlighting the antiwrinkle efficacy of the MB actives in the facial cream.

4 | Discussion

In this study, we developed a novel baicalin delivery system, named Motor Baicalin (MB), to address the challenges of insolubility, instability, and low skin permeability associated with baicalin. The MB system employs bio- vesicles derived from yeast and bacillus cell membranes to encapsulate baicalin, forming a core- shell structure that enhances its water solubility and stability. Our findings indicate that the MB system mark- edly enhances the bioavailability, antioxidant capacity, and antiaging properties of baicalin. This system outperforms both BRM and commonly used LB formulations while demonstrat- ing comparable efficacy to other previously established baicalin delivery systems [19–24].

The physicochemical characterization of MB revealed a particle size of approximately 145 nm with a low polydispersity index, indicating a homogeneous system. The zeta potential values suggested good safety and long- term stability, further confirmed by the absence of significant changes in pH, baicalin loading capacity, particle size, and zeta potential over 90 days. These properties are crucial for practical application, ensuring that MB can maintain its effectiveness over extended periods. Moreover, the in vitro release assessment demonstrated that MB exhibited a slower and more controlled release of baicalin compared to BRM and LB, resulting in higher cumulative release over 60 h. This slow- release behavior reduces the risk of local hyperconcentration and associated skin sensitivity, providing a safer and longer lasting efficacy. Additionally, transdermal penetration studies showed that MB significantly enhanced the penetration of baicalin into the deeper layers of the dermis compared to BRM and LB. This improved penetration is likely due to the ATP synthase units on the bio- vesicles, which facilitate active transfer and integration into specific skin cells [16, 25]. The antioxidant efficacy of MB was evaluated by measuring the activities of key intracellular enzymes (SOD, CAT, and GSH- Px) under oxidative stress conditions. The results indi- cated that MB significantly enhanced the activities of these enzymes, highlighting its strong antioxidant potential [26]. Moreover, MB exhibited superior antiaging properties by inhibiting MMP- 1 expression and promoting TIMP- 1 expression, suggesting a promising collagen protection ability [27]. These findings were corroborated by the improved cell vital- ity and fibrous structure observed in UV- damaged fibroblasts treated with MB. The wound-healing assay demonstrated that MB facilitated faster wound closure compared to BRM and LB, likely due to its superior cell bioavailability and its ability to stimulate cell pro- liferation and migration [28]. Clinical research further validated the antiaging potential of MB actives, showing significant reductions in wrinkles and improvements in skin elasticity in participants using a facial cream with 2% MB additives. These findings underscore the potential of MB as a highly effective cosmetic ingredient for antiaging and skin repairing.

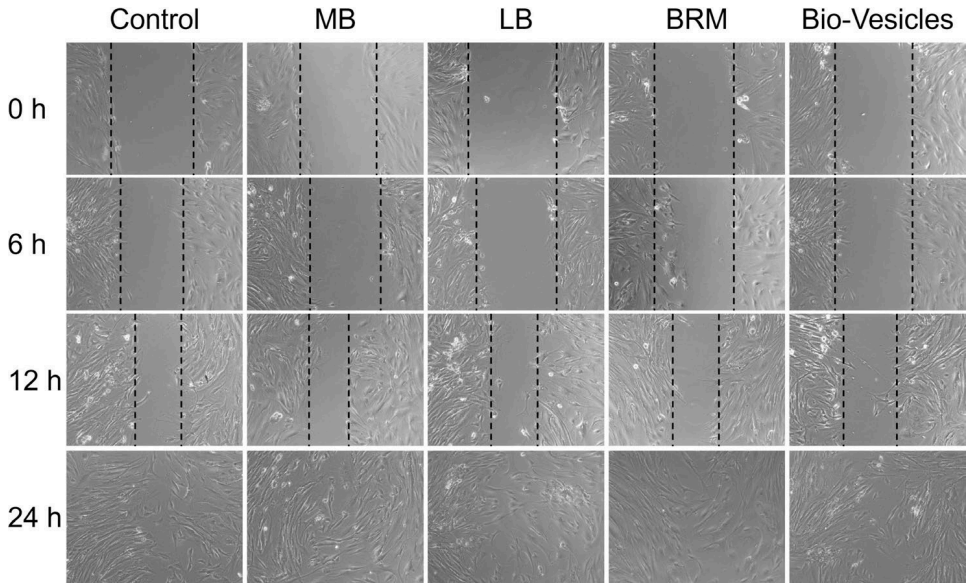


FIGURE 5 | The Effect of MB on wound closure. Images of scratch assays performed on HSF monolayers were treated with control, UVB, MB, LB, BRM, and bio- vesicles groups for 0, 12, 24, and 36 h. Scale bar: 100 μm. The results show that MB enhances the rate of wound closure.

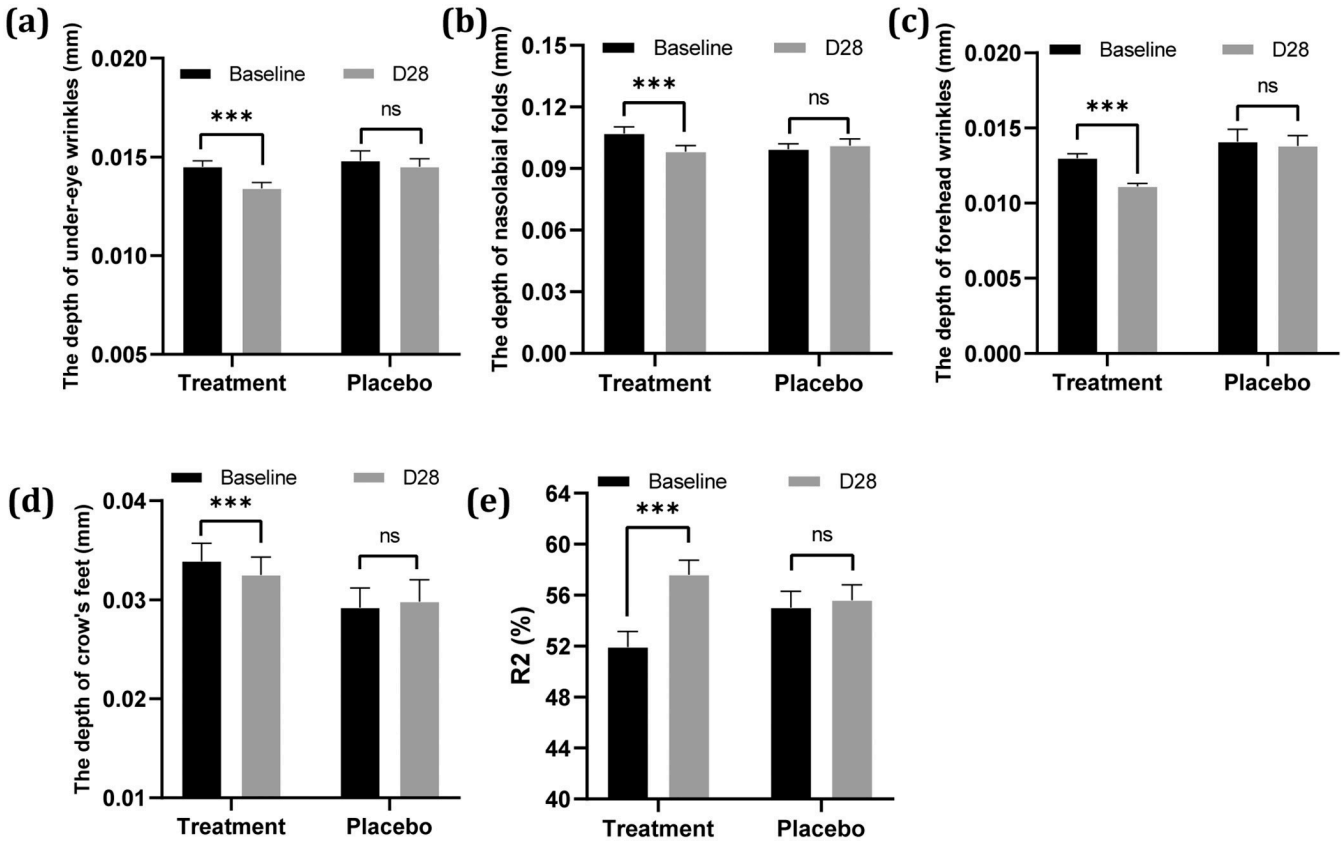


FIGURE 6 | Efficacy of the facial cream in reducing wrinkles and improving skin elasticity. (a) Reduction in under- eye wrinkles, (b) Reduction in nasolabial wrinkles, (c) Reduction in forehead wrinkles, (d) Reduction in crow's feet, (e) Improvement in skin elasticity) (***$p \leq 0.001$ vs. baseline).

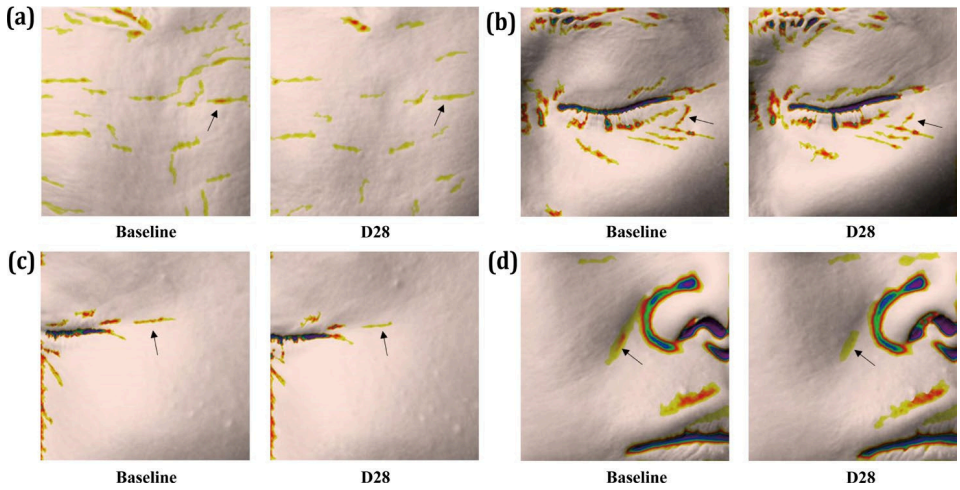


FIGURE 7 | The examples of in vivo antiaging efficacy of the eye cream. (a) Forehead wrinkles reduction, (b) Under- eye wrinkle reduction, (c) Crow's feet reduction, (d) Nasolabial folds reduction).

5 | Conclusions

Our study introduces a promising baicalin delivery system that effectively addresses the primary challenges of insolubility, instability, and low skin permeability. The MB system not only enhances the bioavailability and stability of baicalin but also boosts its antioxidant, antiaging, and wound- healing properties. Clinical findings further validate the efficacy of MB in reducing wrinkles and improving skin elasticity, making it a valuable addition to cosmetic formulations. Future research should explore the long- term effects of MB in clinical settings and investigate its potential applications in other skin- related conditions.

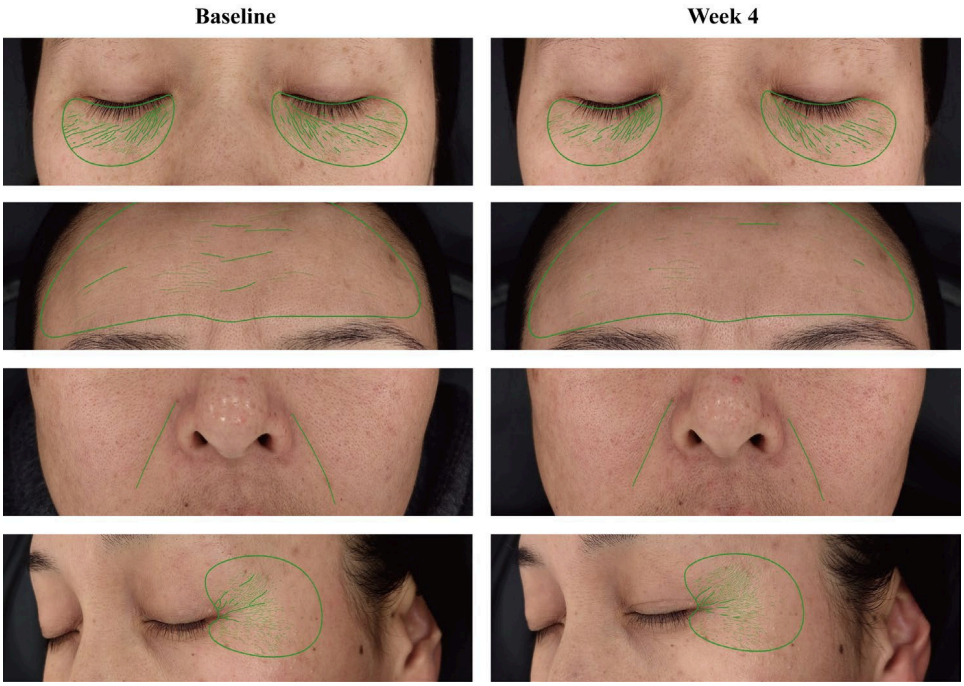


FIGURE 8 | Analysis of nasolabial wrinkles, under- eye wrinkles, forehead wrinkles, and crow's feet in participants following supplementation treatment. Representative Meitu images taken at baseline and after 4 weeks of supplementation.

Author Contributions

L.C., F.X., and Y.G. contributed to conceptualization. F.X. and Y.G. developed the methodology. L.C. and F.W. were responsible for the software. Y.G., F.W., and N.L. conducted the formal analysis. L.C. and F.W. took part in the investigation. L.X., M.X., and L.C. were responsible for resources. L.X., M.X., and L.C. were responsible for data curation. L.C. and F.W. contributed to writing – original draft preparation. L.C., F.W., and F.X. were responsible for writing – review and editing. L.C. and N.L. were involved in visualization. L.X., M.X., and L.C. were involved in supervision. All authors have read and agreed to the published version of the manuscript.

Ethics Statement

The study was conducted in accordance with the Declaration of Helsinki and approved by the Shanghai Ethics Committee (ER- SINH- 262441).

Consent

Informed consent was obtained from all subjects involved in the study.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. S. P. Wyles, J. D. Carruthers, P. Dashti, et al., "Cellular Senescence in Human Skin Aging: Leveraging Senotherapeutics," *Gerontology* 70, no. 1 (2024): 7–14.

2. T. M. Ansary, M. R. Hossain, K. Kamiya, M. Komine, and M. Ohtsuki, "Inflammatory Molecules Associated With Ultraviolet Radiation-Mediated Skin Aging," *International Journal of Molecular Sciences* 22, no. 8 (2021): 3974.

3. H. Pageon, "Reaction of Glycation and Human Skin: The Effects on the Skin and Its Components, Reconstructed Skin as a Model," *Pathologie Biologie* 58, no. 3 (2010): 226–231.

4. J. W. Shin, S. H. Kwon, J. Y. Choi, et al., "Molecular Mechanisms of Dermal Aging and Antiaging Approaches," *International Journal of Molecular Sciences* 20, no. 9 (2019): 2126.

5. I. Juranek, D. Nikitovic, D. Kouretas, A. W. Hayes, and A. M. Tsatsakis, "Biological Importance of Reactive Oxygen Species in Relation to Difficulties of Treating Pathologies Involving Oxidative Stress by Exogenous Antioxidants," *Food and Chemical Toxicology* 61 (2013): 240–247.

6. P. Rattanawiwatpong, R. Wanitphakdeedecha, A. Bumrungpert, and M. Maiprasert, "Anti- Aging and Brightening Effects of a Topical Treatment Containing Vitamin C, Vitamin E, and Raspberry Leaf Cell Culture Extract: A Split- Face, Randomized Controlled Trial," *Journal of Cosmetic Dermatology* 19, no. 3 (2020): 671–676.

7. J. R. Neves, S. Grether- Beck, J. Krutmann, et al., "Efficacy of a Topical Serum Containing L- Ascorbic Acid, Neohesperidin, Pycnogenol, Tocopherol, and Hyaluronic Acid in Relation to Skin Aging Signs," *Journal of Cosmetic Dermatology* 21, no. 10 (2022): 4462–4469.

8. X. L. He, N. Wang, X. Teng, et al., "Dendrobium Officinale Flowers' Topical Extracts Improve Skin Oxidative Stress and Aging," *Journal of Cosmetic Dermatology* 23, no. 5 (2024): 1891–1904.

9. F. H. Lin, J. Y. Lin, R. D. Gupta, et al., "Ferulic Acid Stabilizes a Solution of Vitamins C and E and Doubles Its Photoprotection of Skin," *Journal of Investigative Dermatology* 125, no. 4 (2005): 826–832.

10. M. A. Santos, F. N. Franco, C. A. Caldeira, et al., "Antioxidant Effect of Resveratrol: Change in MAPK Cell Signaling Pathway During the Aging Process," *Archives of Gerontology and Geriatrics* 92 (2021): 104266.

11. S. V. E. Silva, M. C. Gallia, J. R. D. D. Luz, et al., "Antioxidant Effect of Coenzyme Q10 in the Prevention of Oxidative Stress in Arsenic- Treated CHO- K1 Cells and Possible Participation of Zinc as a

Pro- Oxidant Agent," *Nutrients* 14, no. 16 (2022): 3265, https:// doi. org/ 10. 3390/ nu141 63265 .

12. I. Bellezza, I. Giambanco, A. Minelli, and R. Donato, "Nrf2- Keap1 Signaling in Oxidative and Reductive Stress," *Biochimica et Biophysica Acta, Molecular Cell Research* 1865, no. 5 (2018): 721–733.

13. L. Li, H. Bao, J. Wu, et al., "Baicalin Is Anti- Inflammatory in Cigarette Smoke- Induced Inflammatory Models In Vivo and In Vitro: A Possible Role for HDAC2 Activity," *International Immunopharmacology* 13, no. 1 (2012): 15–22.

14. W. Min, X. Liu, Q. Qian, et al., "Effects of Baicalin Against UVA-Induced Photoaging in Skin Fibroblasts," *American Journal of Chinese Medicine* 42, no. 3 (2014): 709–727.

15. M. Bao, M. Liang, X. Sun, et al., "Baicalin Alleviates LPS- Induced Oxidative Stress via NF- kappaB and Nrf2- HO1 Signaling Pathways in IPEC- J2 Cells," *Frontiers in Veterinary Science* 8 (2021): 808233.

16. Y. Gao, J. Zhang, J. Pan, et al., "F(O)F1- ATP Synthase Molecular Motor Biosensor for miRNA Detection of Colon Cancer," *Life Sciences* 319 (2023): 121527.

17. W. Hong, B. Lou, Y. Gao, et al., "Tumor Microenvironment Responded Naturally Extracted F(O)F1- ATPase Loaded Chromatophores for Antitumor Therapy," *International Journal of Biological Macromolecules* 230 (2023): 123127.

18. Q. Yang, X. Zhou, B. Lou, N. Zheng, J. Chen, and G. Yang, "An FOF1-ATPase Motor- Embedded Chromatophore as a Nanorobot for Overcoming Biological Barriers and Targeting Acidic Tumor Sites," *Acta Biomaterialia* 179 (2024): 207–219.

19. M. Manconi, M. L. Manca, C. Caddeo, et al., "Nanodesign of New Self- Assembling Core- Shell Gellan- Transfersomes Loading Baicalin and In Vivo Evaluation of Repair Response in Skin," *Nanomedicine* 14, no. 2 (2018): 569–579.

20. M. L. Manca, S. Mir- Palomo, C. Caddeo, et al., "Sorbitol- Penetration Enhancer Containing Vesicles Loaded With Baicalin for the Protection and Regeneration of Skin Injured by Oxidative Stress and UV Radiation," *International Journal of Pharmaceutics* 555 (2019): 175–183.

21. R. V. Contri, L. A. Fiel, N. Alnasif, A. R. Pohlmann, S. S. Guterres, and M. Schäfer- Korting, "Skin Penetration and Dermal Tolerability of Acrylic Nanocapsules: Influence of the Surface Charge and a Chitosan Gel Used as Vehicle," *International Journal of Pharmaceutics* 507, no. 1–2 (2016): 12–20.

22. J. Niu, M. Yuan, Z. Zhang, et al., "Hyaluronic Acid Micelles for Promoting the Skin Permeation and Deposition of Curcumin," *International Journal of Nanomedicine* 17 (2022): 4009–4022.

23. Y. Zhang, K. Zhang, T. Guo, Y. Li, C. Zhu, and N. Feng, "Transdermal Baicalin Delivery Using Diethylene Glycol Monoethyl Ether-Mediated Cubic Phase Gel," *International Journal of Pharmaceutics* 479, no. 1 (2015): 219–226.

24. S. Mir- Palomo, A. Nácher, O. Díez- Sales, et al., "Inhibition of Skin Inflammation by Baicalin Ultradeformable Vesicles," *International Journal of Pharmaceutics* 511, no. 1 (2016): 23–29.

25. A. Guinart, M. Korpidou, D. Doellerer, et al., "Synthetic Molecular Motor Activates Drug Delivery From Polymersomes," *Proceedings of the National Academy of Sciences of the United States of America* 120, no. 27 (2023): e2301279120.

26. S. Hu, P. Wang, J. Ke, et al., "Protective Effect of Peucedanum Praeruptorum Dunn Extract on Oxidative Damage of LLC- PK1 Cells Induced by H2O2," *Experimental and Therapeutic Medicine* 26, no. 5 (2023): 517.

27. K. C. Song, T. S. Chang, H. J. Lee, J. H. Kim, J. H. Park, and G. S. Hwang, "Processed *Panax ginseng*, Sun Ginseng Increases Type I Collagen by Regulating MMP- 1 and TIMP- 1 Expression in Human Dermal Fibroblasts," *Journal of Ginseng Research* 36, no. 1 (2012): 61–67.

28. A. Grada, M. Otero- Vinas, F. Prieto- Castrillo, Z. Obagi, and V. Falanga, "Research Techniques Made Simple: Analysis of Collective Cell Migration Using the Wound Healing Assay," *Journal of Investigative Dermatology* 137, no. 2 (2017): e11–e16.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

中国科学院
CHINESE ACADEMY OF SCIENCES

EVELAB INSIGHT

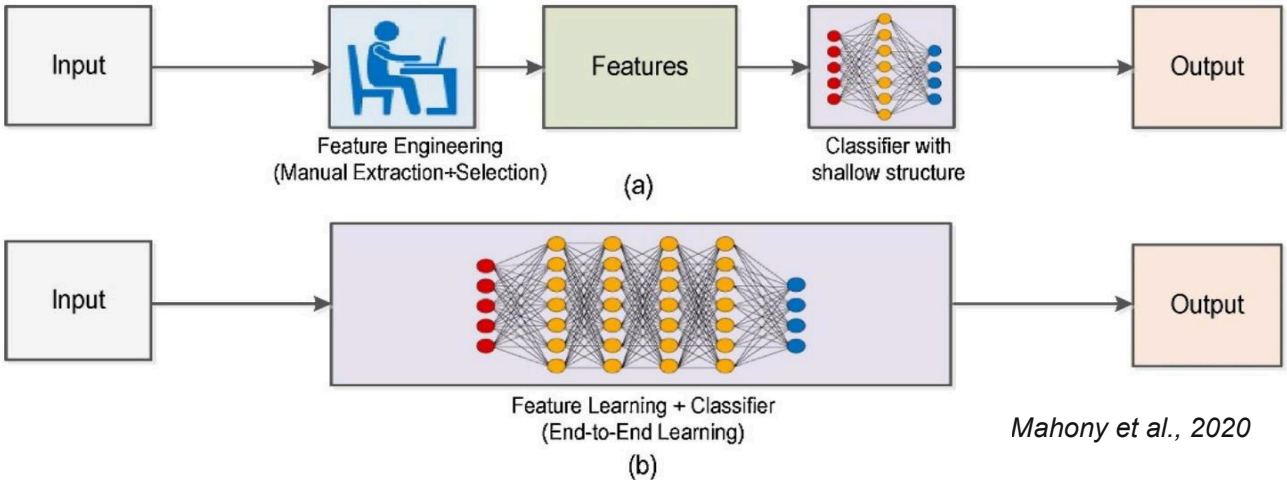
Quantifying facial skin aging signs by deep learning-based algorithm

Sijia Wang

Key Laboratory of Computational Biology
Chinese Academy of Science

Automatic Detection of Skin Aging Signs

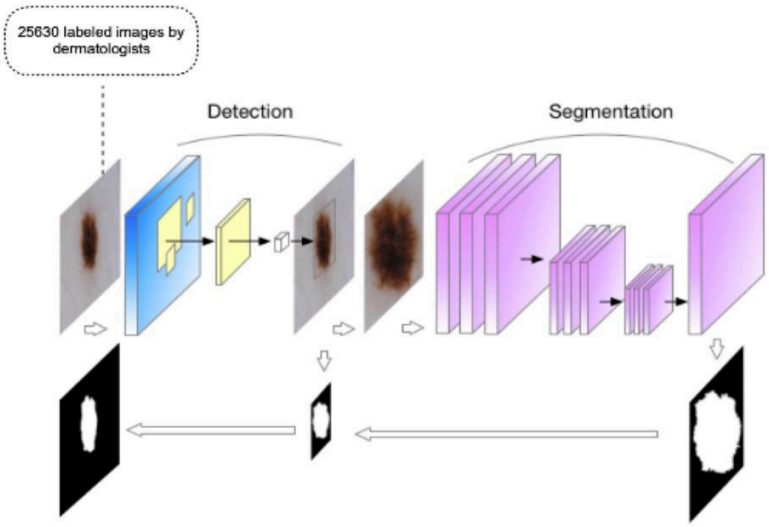
Traditional Computer Vision vs. Deep Learning



Mahony et al., 2020

- ☐ Need large-scale training dataset
- ☐ End-to-end learning
- ☐ Deep learning networks
- ☐ Save manpower in feature extraction
- ☐ Higher accuracy

Deep Learning-based Algorithm to Quantify Skin Aging Signs



Won 2018 ISIC Challenge in skin lesion detection

- Two-stage design
- 25,000+ labeled images by dermatologists
- Data augmentation to further increase training dataset
- Multi-task learning network
- VAE to enhance network performance

Deep Learning-based Algorithm Implemented in a Device

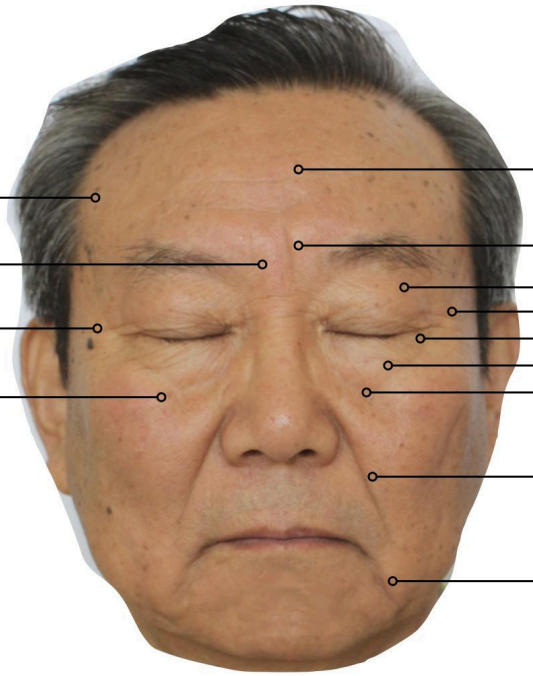
Devices based on conventional computer vision methods

Perceived age

Wrinkle (Front view)

Wrinkle (Side view)

Pigmentation spots



Device based on deep learning-based methods (MeituEve)

Wrinkle forehead

Frown lines

Laxity eyelids

Crow's feet

Wrinkle under eyes

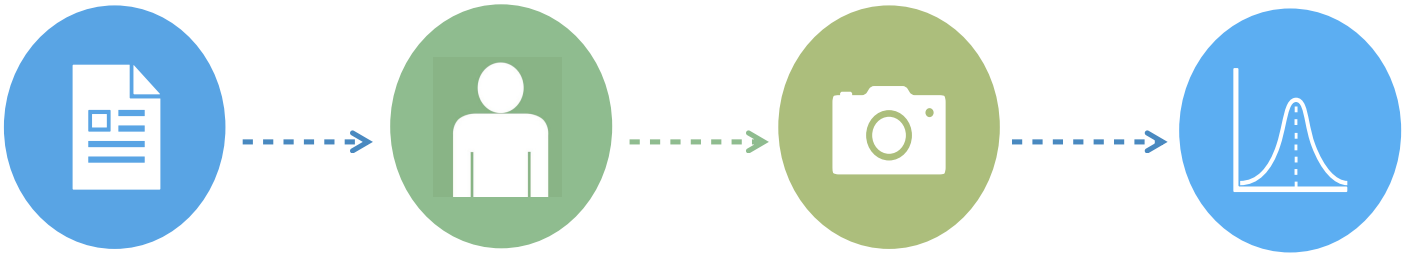
Pouch

Lacrimal sulcus

Nasolabial folds

Marionette lines

Performance Evaluation



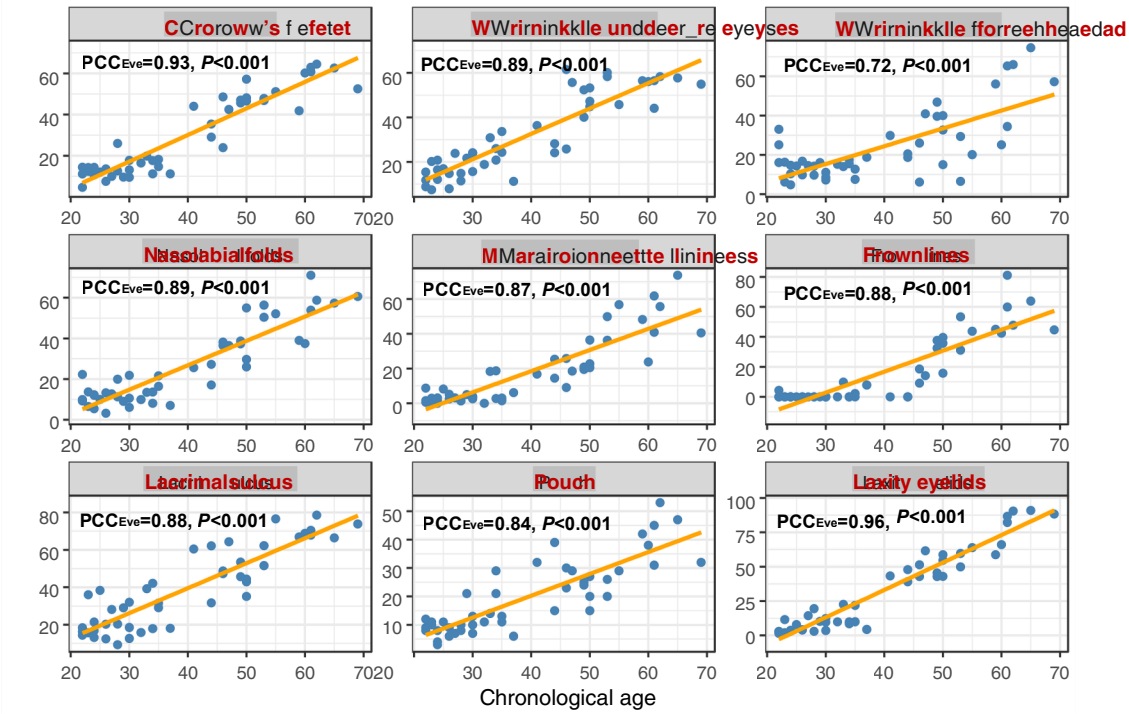
Obtain Consents
(N = 46, Age: 22-69)

Remove makeup
Fix hair

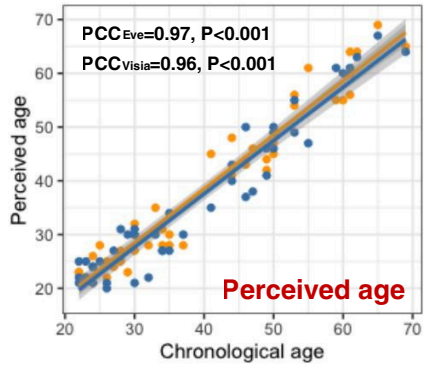
Collect facial images
(Eve and Visia)

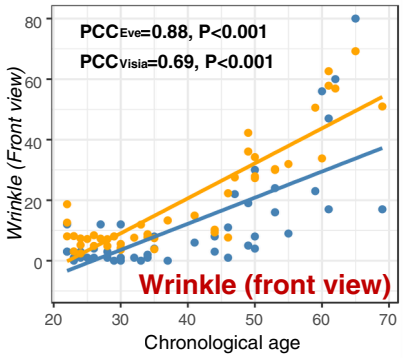
Data cleaning
Data analysis

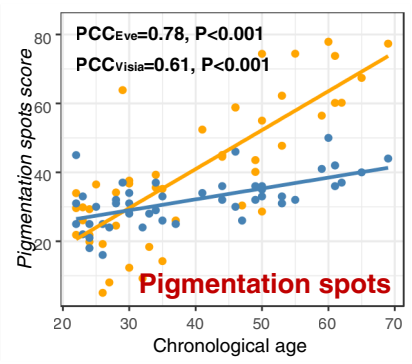
Skin Aging Related Phenotypes Highly Correlated with Age

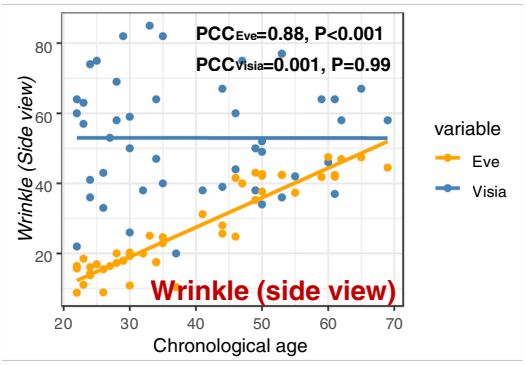


Comparison of Eve and VISIA on overlapped phenotypes

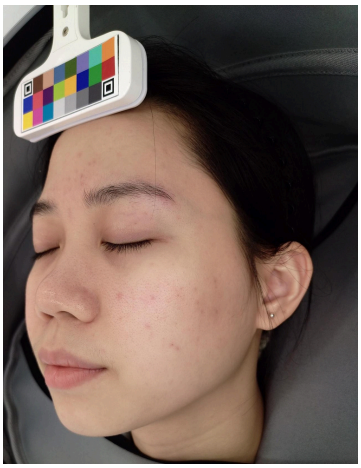


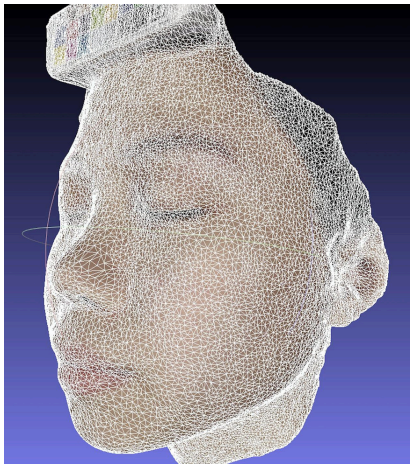


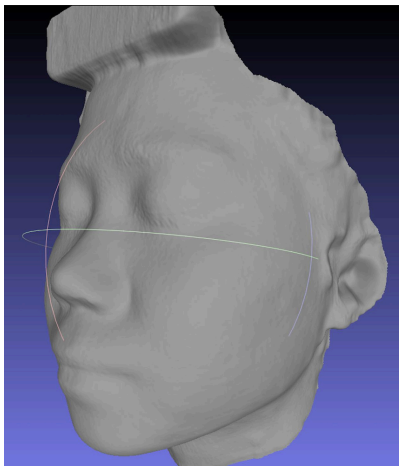




MeituEve V: 3D Imaging Technology Implemented

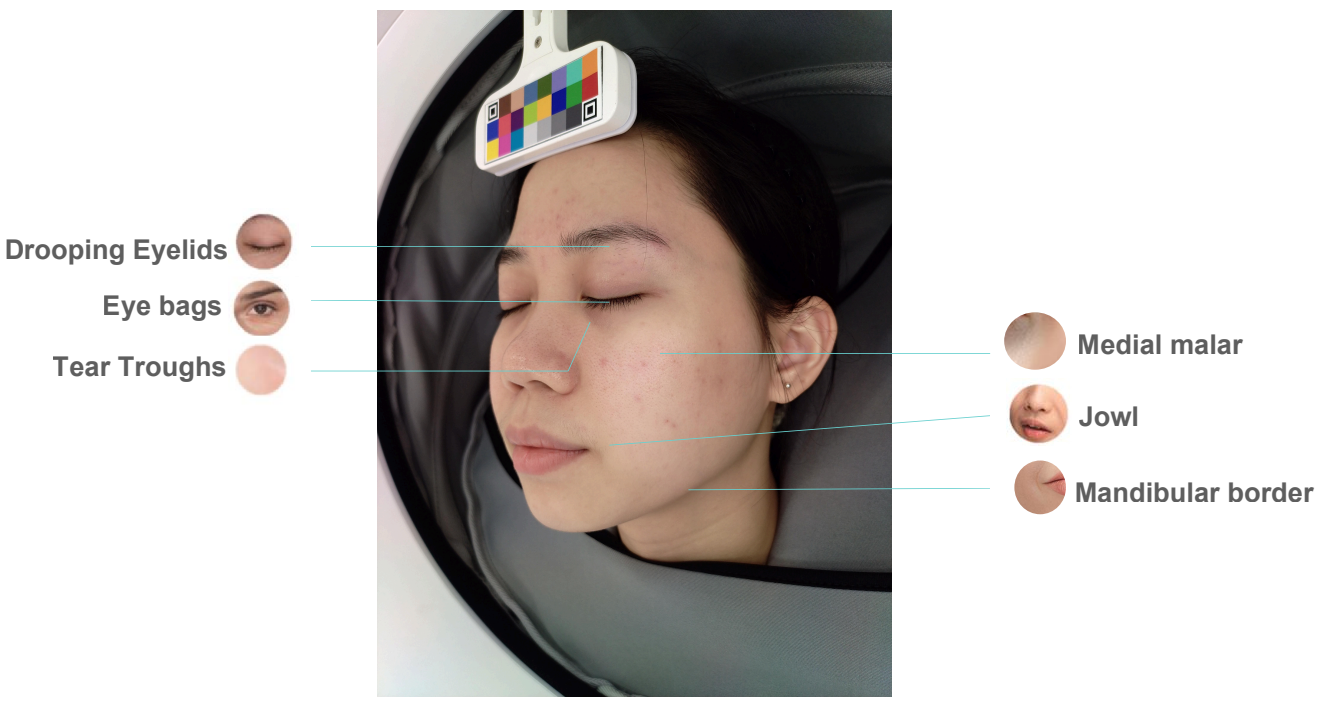




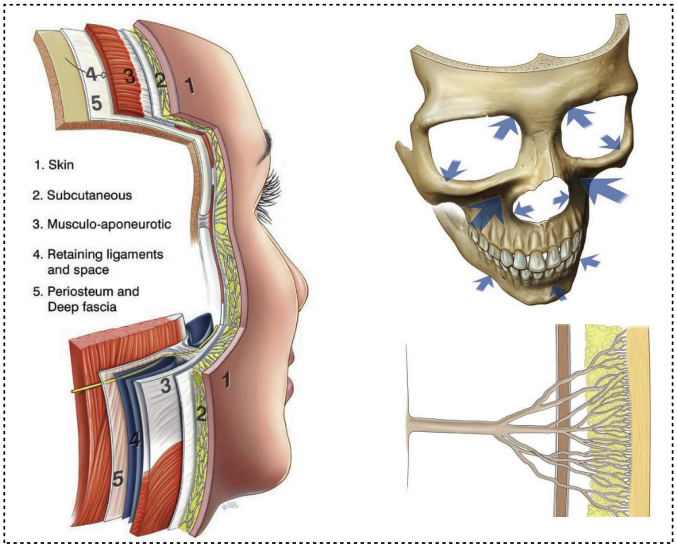


- Five precision 3D cameras
- Collecting millions of point clouds of human face
- Accuracy level of 0.5mm
- Facial reconstruction completed within 1 second

MeituEve V: 3D Image-based Skin Aging Signs



MeituEve V: Why 3D?



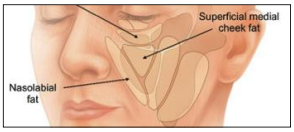
Aesthetic surgery of the face, 2013

MeituEve V: Why 3D?

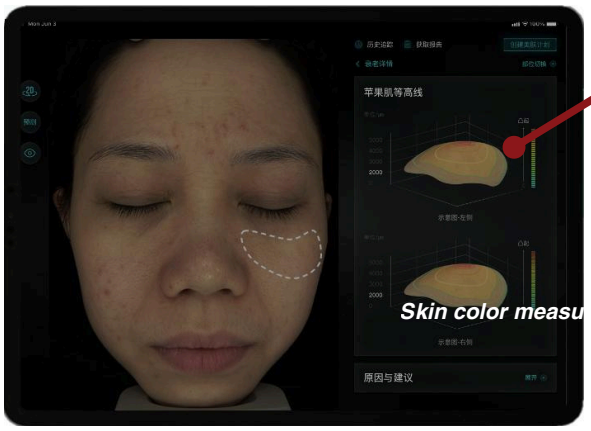
Severity of nasolabial fold associated with the aging change of facial contour structure

JPRAS, 2018

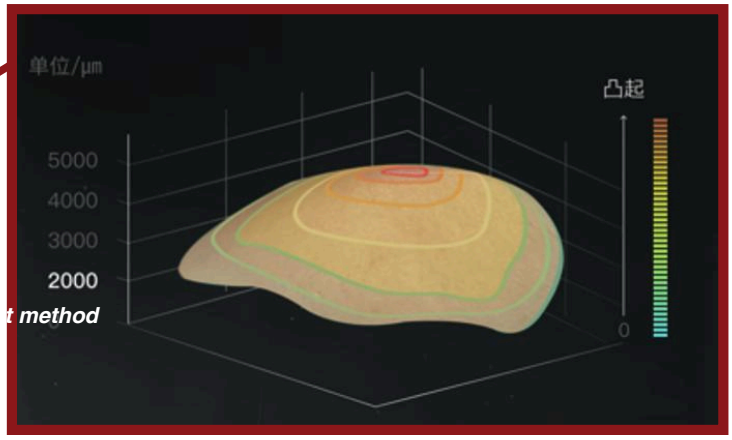
Nasolabial fold



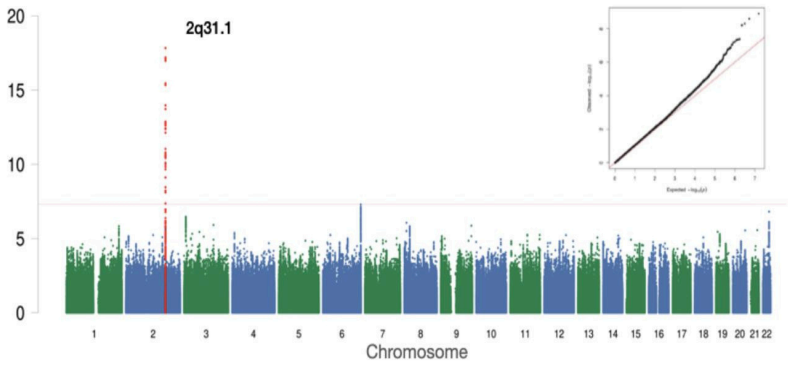
Facial contour structure

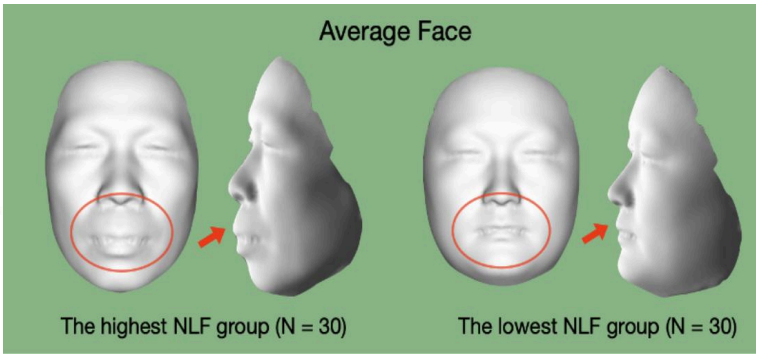


Skin color measu t method



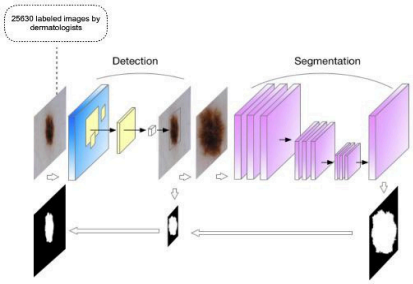
Nasolabial Fold (NLF) vs. Facial Contour Structure



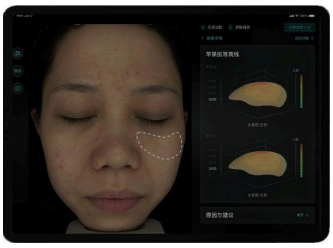


- Genetic variants at 2q31.1 associated with NLF
- Same gene was found to affect facial morphology in the middle face
- Mendelian randomization found evidence for a causal relationship
- Facial morphology Skin aging (NLF)

Conclusions



- **Deep learning-based algorithm** works well in detecting skin aging signs from the facial images



- Implementation of **3D images** to define more skin aging related phenotypes

IFSCC 2025 full paper (2025-866)

Deep Learning Analysis of Perceived Facial Aging and Influential Features Across Evaluator Groups

Fudi Wang [1,2], Siying Fu 2, Baolin Chen 2, Zhiyang LI 2, Eagle Lee 2, Sijia Wang 1*,

[1] Shanghai Institute of Nutrition and Health, Chinese Academy of Sciences, Shanghai, China; 2 EVELAB INSIGHT (SINGAPORE) PTE. LTD., Singapore, Singapore

1. Introduction

Facial aging has long been recognized as an important biomarker of aging and overall health [1-6]. It is closely linked to numerous age-related diseases, including cardiovascular conditions, metabolic disorders, and neurodegenerative syndromes[7-12]. As a visible and socially relevant indicator, facial aging has received sustained attention from both scientific communities and the public.

Over the years, various approaches have been proposed to quantitatively assess facial aging. Some methods focus on localized features—such as wrinkles, pigmentation, or skin texture—while others rely on perceptual experiments, where evaluators estimate overall facial age based on visual appearance[13-16]. Among these, perceived age assessments have emerged as a widely accepted and effective method for quantifying global facial aging, due to their strong correlation with biological age and health outcomes[1, 17-19].

However, a major limitation of existing perceptual studies is the lack of consideration for individual-level variability among assessors. Demographic attributes such as the evaluator's own age and gender may introduce systematic biases into perceived age ratings, potentially confounding the interpretation of facial aging estimates.

In this study, we address this gap by leveraging a crowdsourced evaluation framework to systematically examine how individual characteristics of assessors influence perceived facial aging outcomes. By modeling these effects, we aim to simulate the perceptual process more objectively and remove potential bias sources.

Subsequently, we develop deep learning and statistical models to replicate human perceptual judgments of age in a more standardized and reproducible manner. Finally, we identify and quantify the facial features most strongly associated with perceived aging, providing insights into the visual cues that drive human judgments of facial age. The overall workflow of this study, from data collection to bias modeling, deep learning simulation, and feature analysis, is summarized in Figure 1.

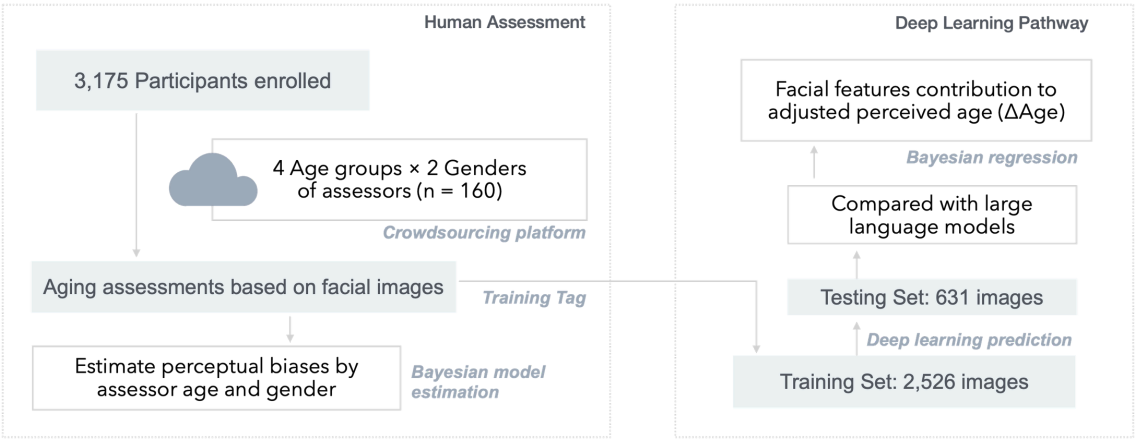


Figure 1. Study Design and Flowchart. The study begins with a large-scale crowdsourced experiment, where assessors rated perceived age for 3,157 facial images. Bayesian modeling was then applied to quantify and remove demographic biases introduced by assessor age and gender. A deep learning model (STDC2-FLD-HR) was trained and compared against large language models (LLMs) to simulate perceptual judgments. The adjusted perceived age (ΔAge) was derived by subtracting chronological age from bias-corrected estimates. Finally, Bayesian regression was used to identify which facial features contributed most significantly to perceived aging.

2. Materials and Methods 2.1 Study Population

The Jidong cohort (JD) is a community-based, long-term observational cohort study to evaluate health related risk factors[20]. The baseline data were collected from 2013 to 2014 in the Staff Hospital, Jidong Oilfield Branch, China. Approval was obtained from the Ethics Committee of Kailuan General Hospital of Tangshan City and the Medical Ethics Committee, Staff Hospital, Jidong Oilfield Branch, China National Petroleum Corporation in July, 2013 (approval No. 2013 YILUNZI1). In this study, 3,157 individuals (1,712 men and 1,445 women, aged 18-87) have been enrolled after excluding individuals who were unable or unwilling to participate. Written informed consent was obtained from all participants. The facial images were collected in the Staff Hospital for further analysis.

2.2 Crowdsourced Perceived Age Assessment

In this study, we recruited 152 assessors through a crowdsourcing platform, evenly distributed across four age groups (21–30, 31–40, 41–50, and 51–60 years) with equal representation of male and female participants (19 men and 19 women per group). Each assessor was tasked with estimating the perceived age via facial images. In addition, assessors provided localized aging assessments based on cropped facial regions of facial images. The evaluated features included under-eye bags, glabellar lines, nasolabial folds, tear troughs, forehead wrinkles, frown lines, crow's feet and marionette lines[14].

2.3 Deep Learning Model Training

We employed deep learning models to simulate human age estimation processes[21]. The performance of three different network architectures was evaluated: STDC2-FLD-HR and STDC2-FLD (both using a CNN-based backbone), Pvt_v2-FLD (Transformer-based backbone), and Vmamba-FLD (VSSM-based backbone)[22-24]. Model training and evaluation were conducted on an NVIDIA 3090 GPU. To enhance the segmentation accuracy of wrinkle regions, STDC2-FLD-HR incorporated feature maps from Stage 1 into the decoder phase, whereas the other models only integrated features from Stage 2 (Figure 2)[25]. For the decoder, we adopted the flexible and lightweight FLD module from PP-LiteSeg to ensure efficient and accurate feature decoding.

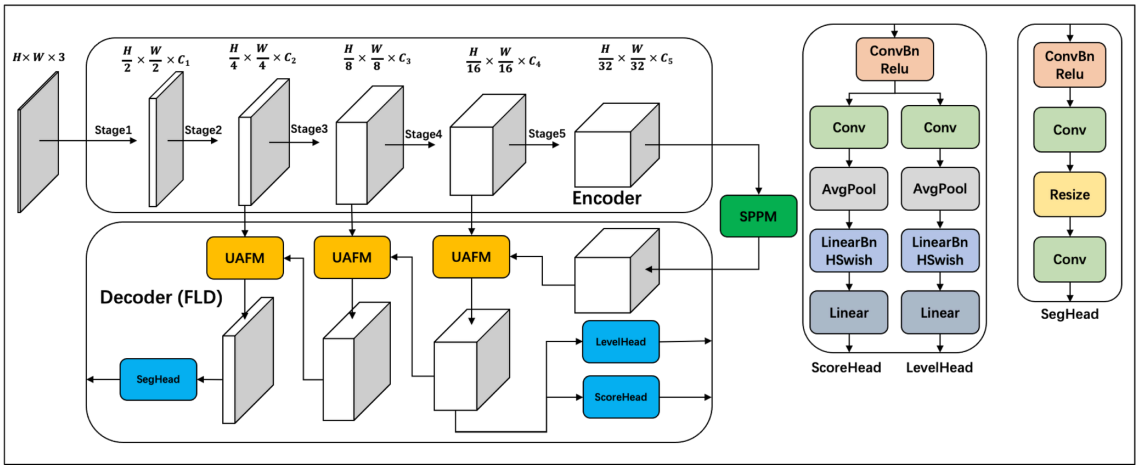


Figure 2. Schematic illustration of the deep learning architecture used for perceived age prediction. Themodel is basedona Short-TermDenseConcatenate(STDC2) backbonefor feature extraction, followed by a lightweight FLD decoder adapted from PP-LiteSeg. High-resolution feature maps from early encoder stages are integrated during decoding (in STDC2-FLD-HR) to enhance fine-grained wrinkle and skin texture representation.

2.4 Deep Learning Visualization

To visualize the information captured by the deep learning models, we first extracted downsampled feature maps and performed Principal Component Analysis (PCA) to reduce their dimensionality. The top three principal components were then mapped to the RGB channels to generate interpretable feature visualizations. Additionally, the low-resolution feature maps were upsampled to higher resolutions to provide clearer and more detailed visual representations.

2.5 Bayesian Rating Model Construction

To account for subjective biases in perceived age ratings, we constructed a hierarchical Bayesian model. The model assumes that evaluators from different age groups and genders may systematically over- or under-estimate the perceived age of facial images. These latent biases are captured as random effects and inferred through Bayesian estimation. Specifically, y_{ij} given by evaluator j to subject i is decomposed into a global mean, a the observed score

subject-specific fixed effect, and a rater-specific bias term that varies by age and gender group. Random noise is added to model unexplained variance.

The rating bias β_{ij} is modeled as the sum of random effects from the evaluator's age group and gender:

where $\gamma_{age[j]} \sim \mathcal{N}(0, \sigma_{age}^2)$ and $\gamma_{gender[j]} \sim \mathcal{N}(0, \sigma_{gender}^2)$ represent the random effects associated with the evaluator's age group and gender, respectively.

3. Results

3.1 Sample Characteristics

A total of 3,157 facial images were included in this study, representing 1,712 males and 1,445 females aged between 18 and 87 years (mean age = 48.12). These images were used as stimuli in the perceived age evaluation experiments. Initially, 152 assessors were recruited to perform the age assessments. The distribution of perceived age ratings was examined across all age and gender groups and found to approximately follow a normal distribution (P>0.05). Table 1 shows the standard deviations of perceived age ratings within each subgroup, compared to the combined results from the full evaluator samples. The results indicate a high degree of consistency across subgroups, with coefficient of variation (CV) values ranging between 0.11 and 0.13. mean absolute error (MAE) values remained relatively stable, generally between 7.16 and 7.50 years, suggesting that assessors provided similarly accurate age estimations across each subgroup.

Table 1. Summary of perceived age rating variability stratified by assessor groups

Gender 1	Age Group 2	Mean (years)	SD (years)	CV (%)	MAE
Male	20-30	48.67	9.84	0.11	7.27
	30-40	48.66	9.84	0.12	7.40
	40-50	48.80	9.85	0.12	7.44
	50-60	49.13	10.27	0.11	7.50
Female	20-30	47.92	10.11	0.13	7.16
	30-40	48.66	9.91	0.12	7.26
	40-50	48.30	10.26	0.12	7.25
	50-60	49.00	10.17	0.12	7.50
Total	20-30	48.27	9.92	0.12	7.21
	30-40	48.64	9.83	0.12	7.33
	40-50	48.53	9.99	0.12	7.35
	50-60	49.05	10.17	0.12	7.50

[1] The gender of assessors; 2 The age group of assessors. The table reports the mean perceived age, standard deviation (SD), coefficient of variation (CV, %) and mean absolute error (MAE) for each sub-group. Lower CV values indicate greater consistency among assessors within the corresponding group.

Table 2 summarizes the inter-assessor consistency in perceived age ratings using intraclass correlation coefficients (ICC) across different demographic groups. The total ICC across all assessors was 0.72, indicating substantial agreement. Subgroup analyses revealed similar reliability across age groups and genders, with slightly higher ICC observed in the 50–60 age

group (ICC = 0.74, 95% CI: [0.73, 0.75]). These results suggest that perceived age assessments were stable across demographic subgroups.

Table 2. Intraclass Correlation Coefficients (ICC) for perceived age ratings

Group	ICC1	F-value	p-value	95% CI
Total	0.72	399.80	<0.001	[0.71, 0.73]
20–30	0.72	101.01	<0.001	[0.71, 0.73]
30–40	0.72	99.51	<0.001	[0.71, 0.73]
40–50	0.72	103.35	<0.001	[0.71, 0.73]
50–60	0.74	113.29	<0.001	[0.73, 0.75]
Male	0.72	202.55	<0.001	[0.71, 0.73]
Female	0.72	202.58	<0.001	[0.71, 0.73]

[1] Intraclass correlation coefficients (ICC) reflecting inter-assessor reliability across demographic subgroups. ICC values were computed using a two-way random effects model for absolute agreement. Higher ICC indicates stronger consistency among assessors.

3.2 Perceptual Biases by Assessor Demographics

To examine potential subjective biases in perceived age assessments, we applied a hierarchical Bayesian model to estimate group-level deviations based on assessor age and gender. As shown in Table 3, the estimated effects (γ) for different assessor age groups demonstrated a clear trend: older assessors tended to give higher perceived age ratings. In particular, the 50–60 age group showed a positive bias (posterior mean = 0.41), while younger groups had slightly negative or near-zero estimates from -0.01 to -0.37.

Regarding gender, female assessors on average provided slightly lower perceived age ratings, with a posterior bias estimate of -0.34 compared to male assessors (0.003). The directional trends suggest potential demographic influences on perceptual judgments.

Table 3. Posterior summaries of age and gender group-level bias effects

	Mean	SD	MeanMCSE[1]	SDMCSE
$\gamma_{age[20-30]}$	-0.37	1.03	0.09	0.07
$\gamma_{age[30-40]}$	-0.01	1.03	0.09	0.07
$\gamma_{age[40-50]}$	-0.12	1.03	0.09	0.07
$\gamma_{age[50-60]}$	0.41	1.03	0.09	0.07
$\gamma_{gender[Male]}$	0.003	1.45	0.10	0.07
$\gamma_{gender[Female]}$	-0.34	1.45	0.10	0.07

[1] MCSE: Monte Carlo standard errors. Results are derived from a hierarchical Bayesia

3.3 Deep Learning Model Performance in Simulating Human Perception

To assess the ability of deep learning models to simulate human perceptual judgment of facial age, we evaluated the performance of four model architectures using the Pearson correlation coefficient (PCC) between predicted and human-assessed perceived age, stratified by assessor age and gender groups. As shown in Table 4, all models achieved statistically significant correlations with human ratings ($p < 0.001$ across all subgroups), confirming their predictive validity.

Among the models tested, STDC2-FLD-HR consistently achieved the highest performance, with PCC values ranging from 0.94 to 0.95 across all gender and age combinations, demonstrating superior ability to approximate human perception. This was followed by STDC2-FLD (PCC = 0.91–0.94), Pvt_v2-FLD (PCC = 0.86–0.88), and Vmamba-FLD (PCC = 0.83–0.87). STDC2-FLD-HR was used in further analysis.

Table 4. Pearson correlation coefficients (PCC) between model-predicted and human-perceived ages across assessor gender and age groups.

Gender	**Female**				**Male**			
Age Group	**20-30**	**30-40**	**40-50**	**50-60**	**20-30**	**30-40**	**40-50**	**50-60**
STDC2-FLD-HR	0.95[1]	0.95	0.94	0.94	0.94	0.94	0.94	0.94
STDC2-FLD	0.94	0.94	0.94	0.94	0.94	0.93	0.93	0.91
Pvt_v2-FLD	0.88	0.87	0.87	0.86	0.86	0.86	0.86	0.86
Vmamba-FLD	0.87	0.87	0.87	0.87	0.86	0.86	0.86	0.83

[1]Values in the table represent the Pearson correlation coefficients (PCC) between predicted ages from deep learning models and perceived ages assessed by human raters. Higher PCC values indicate stronger alignment between model predictions and human perception. All reported correlations are statistically significant ($p<0.001$) .

To further evaluate the capability of deep learning models in simulating human-like age perception, we compared the performance of a visual model (STDC2-FLD-HR) with two large language models (Qwen and LLaMA) prompted to perform perceived age estimation from facial images.

Performance was evaluated based on mean absolute error (MAE), Pearson correlation coefficient (PCC), and linear regression R^2 between predicted and true chronological age. The STDC2-FLD-HR model demonstrated superior performance across all metrics, achieving the lowest MAE (5.56), highest PCC (0.87), and R^2 (0.76). In contrast, Qwen achieved moderate performance, with an MAE of 9.895, PCC of 0.779, and R^2 of 0.608, suggesting that while it captured some meaningful patterns, its estimations were less precise and less aligned with

true age distributions. LLaMA showed significantly poorer performance, with an MAE of 11.045, PCC of 0.392, and R^2 of 0.154, reflecting limited capacity to extract and interpret visual age-related cues.

3.4 Drivers of Perceived Aging After Bias Adjustment

To isolate intrinsic facial aging effects from potential demographic biases, we first adjusted the perceived age scores predicted by the deep learning model using a Bayesian correction model accounting for assessor age and gender. The resulting metric, referred to as ΔAge (delta age), represents the deviation between perceived age and actual chronological age. We then examined the relationship between ΔAge and a set of facial aging features using a Bayesian regression model. As shown in Table 5, the most influential features contributing to higher ΔAge values were nasolabial folds (β = 0.09), eye bags (β = 0.08), and pigmentation spots (β = 0.08), suggesting these features were most strongly associated with an older per- ceived appearance. These findings highlight specific regional markers that drive perceived facial aging, independent of demographic rater effects.

Table 5. Facial feature effects on perceived age deviation (ΔAge)

Aging Features	β1	Aging Features	β
Nasolabial_fold	0.086	Wrinkle around eyes	0.054
Eye bags	0.083	Tear_troughs	0.035
Pigmentation spots	0.079	Wrinkle forehead	0.031
Crows' feet	0.064	Eye sagging	0.007
Marionette lines	0.060	Frown lines	0.004

Bayesian regression coefficients (β) for facial aging features contributing to ΔAge (Perceived age – chronological age). ΔAge represents perceived age deviation from chronological age after demographic bias correction. Higher β values indicate stronger influence on perceived aging.

4. Discussion

This study sought to better understand the process of human facial age perception and improve its objectivity by modeling demographic influences and leveraging deep learning technologies. Building on the notion that facial aging is a socially and biologically relevant biomarker, we first confirmed that perceived age assessments can be systematically influenced by assessor characteristics—particularly age and gender. Through a large-scale crowdsourced experiment and hierarchical Bayesian modeling, we identified measurable perceptual biases associated with these demographic variables.

We developed a visual deep learning model (STDC2-FLD-HR) that demonstrated strong alignment with human ratings while maintaining high consistency across gender and age

groups. This model significantly outperformed large language models (LLMs) when tasked with estimating perceived age from facial images. Moreover, by removing assessor-level bias and analyzing the remaining discrepancy between predicted and actual chronological age (ΔAge), we identified several facial features—such as nasolabial folds, eye bags, and pigmentation—that strongly contribute to the faical aging. These features may serve as reliable visual markers for future studies of skin aging.

Despite these promising findings, our study has several limitations. First, the dataset and assessors were limited to Han Chinese populations, without accounting for ethnic diversity. Given that facial aging patterns and perceptual tendencies vary significantly across populations, future work will aim to expand the dataset to include participants of different ethnic backgrounds. Second, our analysis of language model performance was restricted to two models (Qwen and LLaMA) available before January 2025. As the field of multimodal AI rapidly evolves, subsequent versions of LLMs may offer improved performance and should be evaluated in future research.
In addition, while our current sample includes over 3,000 individuals, further expansion in both sample size and demographic diversity will enhance the robustness and generalizability of our findings. Incorporating cross-cultural evaluations and larger datasets will be critical for developing universally reliable tools for facial age estimation.

5. Conclusion

This study demonstrated that facial age perception is influenced by both skin-related features and assessor demographics such as age and gender. By correcting for these biases through Bayesian modeling and applying a deep learning framework, we achieved accurate and objective facial age estimation. Our model outperformed language-based models and revealed that features like nasolabial folds, pigmentation spots, and crow’s feet play key roles in perceived aging. Future work will expand to more diverse populations and AI models to enhance generalizability.

Reference

1. Liu,F., et al., *The MC1R Gene and Youthful Looks.* Curr Biol, 2016. **26**(9): p. 1213-20.
2. Bonfante, B., et al., *A GWAS in Latin Americans identifies novel face shape loci, implicating VPS13B and a Denisovan introgressed region in facial variation.* Sci Adv, 2021. **7**(6).
3. Xiong, Y., et al., *Prevalence and associated factors of metabolic syndrome in Chinese middle-aged and elderlypopulation:anational cross-sectional study.* Aging Male, 2021. **24**(1): p. 148-159.
4. Gurovich, Y., et al., *Identifying facialphenotypesofgeneticdisordersusing deep learning.* Nat Med, 2019. **25**(1): p. 60-64.
5. Xiong, Z., et al., *Combininggenome-wideassociationstudieshighlightnovel loci involved in human facialvariation.* Nat Commun, 2022. **13**(1): p. 7832.
6. Yu, Z., et al., *Thermalfacialimageanalysesrevealquantitativehallmarks of aging and metabolic diseases.* Cell Metab, 2024. **36**(7): p. 1482-1493 e7.
7. Peng, Q., et al., *3Dfacialimaging:anovelapproachformetabolicabnormalities risk profiling.* Sci China Life Sci, 2025.
8. Cox-Brinkman, J., et al., *Three-dimensional face shape in Fabry disease.* Eur J Hum Genet, 2007. **15**(5): p. 535-42.
9. Kim, M.K., et al.,*Associations ofVariabilityin BloodPressure, Glucoseand Cholesterol Concentrations, and Body Mass Index With Mortality and Cardiovascular Outcomes in the General Population.* Circulation, 2018. **138**(23): p. 2627-2637.
10. Meng, T., et al., *Identifying Facial Features and Predicting Patients of Acromegaly Using Three-Dimensional Imaging Techniques and Machine Learning.* Front Endocrinol (Lausanne), 2020. **11**: p. 492.
11. Lin, S., et al., *Feasibility of using deep learning to detect coronary artery disease based on facial photo.* Eur Heart J, 2020. **41**(46): p. 4400-4411.
12. Kong, X., et al., *Facial recognition for disease diagnosis using a deep learning convolutional neural network: a systematic review and meta-analysis.* Postgrad Med J, 2024. **100**(1189): p. 796-810.
13. Liu, Y., et al., *Genome-wide scan identified genetic variants associated with skin aging in a Chinese female population.* J Dermatol Sci, 2019. **96**(1): p. 42-49.
14. Vierkotter, A., et al., *The SCINEXA: a novel, validated score to simultaneously assess and differentiate between intrinsic and extrinsic skin ageing.* J Dermatol Sci, 2009. **53**(3): p. 207-11.
15. Saffari, P.S., et al., *Facial Aging in Thyroid Eye Disease: Quantification by Artificial Intelligence.* J Craniofac Surg, 2025.
16. Estler, A., et al., *Quantification of Facial Fat Compartment Variations: A Three-Dimensional Morphometric Analysis of the Cheek.* Plast Reconstr Surg, 2023. **152**(4): p. 617e-627e.
17. Sun, N., et al.,*Self-perception ofaging and perceived medical discrimination.* J Am Geriatr Soc, 2023. **71**(10): p. 3049-3058.
18. Nkengne, A., et al., *Influenceoffacial skin attributes on the perceived age of Caucasian women.* J Eur Acad Dermatol Venereol, 2008. **22**(8): p. 982-91.
19. Flament, F., et al., *Changesin facial signs due to age and their respective weights on the perception of age and skin plumpness among differently aged Korean women.* Skin Res Technol, 2021. **27**(4): p. 526-536.

20. Wang, F., et al., *A Genome-Wide Scan on Individual Typology Angle Found Variants at SLC24A2 Associated with Skin Color Variation in Chinese Populations.* J Invest Dermatol, 2022. **142**(4): p. 1223-1227 e14.
21. Fan, M.a.L., et al., *Rethinking BiSeNet For Real-time Semantic Segmentation*, in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2021. p. 9711-9720. Wang, W., et
22. al., *Pyramid Vision Transformer: A Versatile Backbone for Dense Prediction without Convolutions*, in *IEEE/CVF international conference on computer vision*. 2021: Montreal, QC, Canada. p. 568-578. Wang, W., E. Xie, and X. Li, *PVT v2: Improved baselines with pyramid vision transformer.*
23. Computational Visual Media, 2022. **8:3**: p. 415-424. Bian, J., M. Feng, and W. Dong, *Locally Aware Visual State Space for Small Defect Segmentation in Complex Component Images.* IEEE
24. Transactions on Industrial Informatics, 2025: p. 1-12. Shreve, M., R. Bala, and W. Wu, *Region-wise Modeling of Facial Skin Age using Deep CNNs*. 2019 16th International Conference on
25. Machine Vision Applications (MVA). 2019.

LB1126

A decade-long comparative analysis of skin disease trends in the United States and globally using the global burden of disease database

I Draw [1], LS Shqair , O Alani , D Alkurdi and Z Schwager 1 Dermatology, University of Louisville, Louisville, Kentucky, United States, 2 Dermatology, Icahn School of Medicine at Mount Sinai, New York, New York, United States and 3 Lahey Hospital & Medical Center, Burlington, Massachusetts, United States

Skin diseases impact the quality of life and strain healthcare systems. Studying long-term trends helps guide prevention and treatment options. Our study aims to analyze 2011e2021 US and global skin disease trends using the Global Burden of Disease Database, showing heightened disparities and potential causes. We extracted prevalence and incidence data for 15 conditions, including acne vulgaris, atopic dermatitis (AD), bacterial skin infections (BSI), decubitus ulcers (DU), scabies, viral skin diseases (VSD), and pyoderma. Changes were compared between US and global populations, with confidence intervals (95% CI) used to assess significance. Several conditions showed a greater burden in the US. BSIs rose to 0.1511 (95% CI: 0.13-0.17), nearly double the global rate (0.0806, 95% CI: 0.0773-0.0839). Cellulitis also increased (0.1022, 95% CI: 0.08-0.12) compared to global levels (0.0855, 95% CI: 0.0773-0.0943). DUs (0.2242, 95% CI: 0.16-0.27) and pyoderma (0.2440, 95% CI: 0.23-0.26) were significantly higher than global rates (0.1121, 95% CI: 0.0891-0.1330 and 0.0803, 95% CI: 0.0768-0.0838), reflecting greater challenges in wound care. Some conditions declined in the US. Scabies fell to 0.0966 (95% CI: 0.0293 to 0.0415), a sharper drop than globally (0.0293, 95% CI: 0.0415 to 0.0174), likely due to improved hygiene. VSD followed a similar trend (0.0246, 95% CI: 0.0289 to 0.0239 in the US vs. 0.0263, 95% CI: 0.0289 to 0.0239 globally). AD declined in both (0.0563, 95% CI: 0.07 to 0.04 in the US, 0.0680, 95% CI: 0.0756 to 0.0610 globally). Acne vulgaris remained stable (0.00046, 95% CI: 0.01146 to 0.01048). Our analysis shows differing US and global skin disease trends, with rising infections and chronic wounds in the US and a decline in scabies and AD. Our results underscore the need for better infection control and further research into healthcare disparities and environmental factors.

LB1127

Epidemiologic trends of cellulitis in the US: National and state-level insights (2010e2021)

H Akbarialiabad [1], CG Bunick [2] and A Grada 1 [3] Department of Dermatology, University of Utah Health, Salt Lake City, Utah, United States, 2 Department of Dermatology, Yale University, New Haven, Connecticut, United States and 3 Department of Dermatology, Case Western Reserve University, Cleveland, Ohio, United States

Background: Cellulitis is a common bacterial skin infection affecting the dermis and subcutaneous tissues, characterized by erythema, edema, and tenderness. This study evaluates the epidemiologic trends and burden of cellulitis in the United States across demographic groups and states from 2010 to 2021. Methods: In this population-based study, we utilized established methods from the Global Burden of Diseases (2021) framework to quantify age-standardized rates for incidence (ASIR), prevalence (ASPR), and disability-adjusted life years or DALY (ASDR) per 100,000 in the US. This analysis integrates national surveys, medical records, and claims databases using Bayesian meta-regression modeling (DisMod-MR 2.1). All estimates were reported with 95% uncertainty intervals (UIs). Results: In 2021, there were 10.75 incident cases of cellulitis (95% UI: 10.15e11.37), reflecting an 18.53% increase from 2010. The ASIR increased from 2,664.7 per 100,000 (95% UI:2,524.2e2,815.7) in 2010 to 2,800.3 per 100,000 (95% UI: 2,648.9e2,964.4) in 2021. Males had a slightly higher inci- dence than females, and incidence rates increased progressively with age. The highest ASIR was observed in adults aged 80e84 years (7,547.5 per 100,000; 95% UI: 6,372.5e8,732.6). The highest incidence rates were reported in Louisiana, New York, Alabama, and Mississippi, while the lowest were observed in Minnesota, Colorado, and North Dakota. Conclusion: The burden of cellulitis in the U.S. significantly increased from 2010 to 2021, with a higher incidence in older adults, males, and specific states.

LB1128

Evaluation of imaging-based methods for facial aging detection

F Wang [1], B Chen [2] and Z Li 1 [2] Chinese Academy of Sciences Shanghai Institute of Nutrition and Health, Shanghai, Shanghai, China and 2 Xiamen Meitueve Technology CO., Ltd, Xiamen, China

Introduction: Quantifying facial aging is essential in dermatology for studying age-related changes and assessing the effectiveness of skincare products. However, facial aging detection remains challenging due to the absence of standardized benchmarks and unified testing protocols, making it difficult to compare different methods fairly. In this study, we evaluate the performance of commonly used imaging-based facial aging detection methods, focusing on accuracy,stability,andsensitivity.Byhighlightingtheirstrengthsandlimitations,thisresearch aims to provide a comprehensive review of facial aging detection approaches to help re- searchers select the most suitable method for their studies. Methods: A dataset of 2,000 in- dividuals was analyzed under standardized conditions using imaging systems (Visia CR 5.0, Antera 3D, EVE V). To compare the performance of the methods, we calculated correlations with age, consistency across repeated measurements, and detection of subtle changes post-intervention. Human assessments were used as the gold standard for measuring accuracy. Results: EVE V demonstrated the highest correlation with age, with results exceeding 0.82, outperforming Visia CR 5.0 (average 0.73) and Antera 3D (average 0.77). Compared to hu- man assessments as the gold standard, the accuracy adjustment for EVE V averaged 0.75, while Visia CR 5.0 was 0.68, and Antera 3D was 0.63. Antera 3D showed superior sensitivity, detecting over 20% changes in localized skin features. In terms of stability, EVE V achieved the highest score at 0.93, followed by Visia CR 5.0 at 0.88, and Antera 3D at 0.80. Discussion and Conclusion: Antera 3D's high sensitivity makes it ideal for analyzing localized skin features, while EVE V and Visia CR 5.0 excels in accuracy and stability, making it well-suited for research on age-related skin changes. These findings offer valuable insights for selecting the most appropriate imaging systems for clinical dermatology and efficacy testing.

LB1129

Bridging the gap in dermatologic care for afro-textured hair: Addressing disparities through collaboration and education

A Sataray-Rodriguez [1], D Oladinni and HF Hassan 1 A T Still University School of Osteopathic Medicine in Arizona, Mesa, Arizona, United States, 2 Northeast Ohio Medical University College of Medicine, Rootstown, Ohio, United States and 3 University of Nevada Reno, Reno, Nevada, United States

Hair and scalp health are vital components of dermatologic care, yet patients with Afro-textured hair face persistent disparities in diagnosis and treatment. These inequities stem from limited representation of dermatologists and insufficient training on Afro-textured hair in medical education. Patients often experience inadequate care and report discomfort addressing hair and scalp concerns with racially incongruent providers, further burdened by the need to explain cultural hair practices. Meanwhile, dermatologists cite a lack of confi- dence and minimal training on Afro-textured hair as barriers to providing equitable care. Hairstylists, often serve as primary assessors of hair and scalp health but rarely collaborate with dermatologists. Initiatives like The Black Hair Curriculum and The S.T.R.A.N.D. Network highlight the potential of integrating culturally relevant education into medical training and fostering partnerships between hairstylists and healthcare professionals. These efforts aim to improve diagnostic accuracy, patient-provider relationships, and health outcomes for patients. This paper explores the root causes of disparities in dermatologic care for Afro-textured hair, evaluates current solutions, and proposes systemic changes to center Afro-textured hair in medical education. By addressing knowledge gaps and leveraging community partner- ships, we can dismantle hair care inequities and create a healthcare system that prioritizes inclusivity and cultural competence.

Genome-wide association study of the nasolabial fold identified novel variants related to genes that also affect facial morphology

Fudi Wang[1,#], Yuepu Zhao[1,#], Siyuan Du1, Jiarui Li1, Xiaoyun Hu1, Zhiyang Li2, Sijia Wang[*]

[1] CASKey Laboratory ofComputational Biology, ShanghaiInstituteofNutritionandHealth,UniversityofChinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China;

[2] EVELAB INSIGHT (SINGAPORE) PTE. LTD.

[*] Corresponding author's email: wangsijia@picb.ac.cn

Abstract

Thenasolabial fold (NLF) is one of the most notable phenotypes of facial aging for aesthetic physicians. Previous studies on the three-dimensional structures of the NLF have clarified that it was protentially induced by morphological changes in the superficial musculoaponeurotic system (SMAS) architecture. Reportedly the NLF has a reasonable heritability, but the underlying genetic factors have not yet been identified. In this study, we developed a deep learning-based method to measure the NLF and performed the largest genome-wide association study (GWAS) of the NLF to date in 11,033 Han Chinese. A novel locus at 2q31.1 was significantly associated with the NLF (P=7.18×10-19). Interestingly, the genes near this locus have both been reported to play important roles in facial morphology. By further examining the 3D facial images of the same individuals with the NLF measurement, we found that the middle face (nose-mouth region) morphology was strongly associated with the NLF (PCC=0.62, *P*=0.03). Moreover, mendelian randomization suggests a morphology-to-NLF causal relationship (β=2.14, *P*=4.41×10-6), consistent with results from previous anatomical studies. In conclusion, we expanded the knowledge of the NLF development and verified its relationship with facialmorphology from a genetic perspective.

Materials and Methods

Samples

Thediscovery dataset included a total of 2,921 individuals of Chinese origin from the National Survey of Physical Traits cohort (NSPT). The replication dataset included a total of 8,112 individuals from two independent cohorts: the Jidong cohort (JD, N=5,201), Taizhou Longitudinal cohort (TZL, N=2,911). All participants provided written informed consent. **Genotyping** For both the NSPT and JD cohorts, genomic DNA was extracted from blood samples and genotyped using the Illumina Infinium Global Screening Array (GSA) consisting of about 710,000 SNPs. For TZL, all individuals were genotyped using the Illumina HumanOmni- ZhongHua-8 chip, which interrogates 894,517 SNPs. Imputation was performed using the 1000 Genomes Project Phase 3 as the reference panel. SNPs with an imputation quality score of less than 0.6, MAF less than 0.01 or a missing rate more than 0.01 were excluded from further analyses. Finally, 8,039,700 SNPs passed quality control and were used for the analyses in this study. **Phenotyping** We generated the nasolabial score from high resolution 2D facial images, with a two-step deep learning network, involving a multi-task learning of facial segmentation and a novel detection solution1. Among 8,122 individuals in the NSPT and JD cohorts, we also collected the 3D facial data using the 2-pod 3dMDface camera system. Each 3D facial images contained 13 landmarks for the quantification of facial morphology.

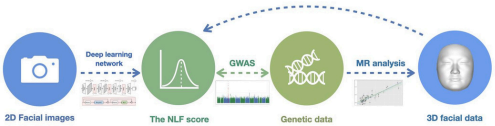


Figure 1. Study design of the nasolabial fold.

Statistical analysis

GWASswere separately conducted in NSPT, JD and TZL on the NLF using linear regression with an additive genetic model adjusted for covariates (i.e., sex, age, BMI and the first five genetic principal components). GWAS summary statistics of the three cohorts were combined using inverse variance fixed-effect meta-analysis. We performed causal inference using 2 sample mendelian randomization (MR) method, and selected instruments with P < 5×10-8 and a LD clumping threshold of 0.001.

Results

Characteristics of the study population

Wedivided thepopulationsinto twogroups (i.e., the milder NLF group and the severer NLF group) according to the median NLF score. Males had a significantly lower NLF score than females. The severer NLF group has significantly increased age and lower BMI than the milder group.

Table 1 Baseline characteristics of the individuals

	NSPT (N = 2,921)			TZL (N=2,962)			JD (N=5,102)		
	Milder group (N=1,460)	Severer group (N=1,460)	*P**	Milder group (N=1,481)	Severer group (N=1,481)	*P*	Milder group (N=2,551)	Severer group (N=2,551)	*P*
Age, mean (sd), years	42.31 (11.84)	57.44 (8.87)	1.68×10^{-21}	51.07 (8.69)	61.45 (7.17)	1.08×10^{-5}	36.67 (8.50)	53.45 (11.85)	3.52×10^{-7}
Female, N (%)	889 (47.90)	967 (52.10)	2.38×10^{-3}	899 (60.70)	970 (65.50)	0.01	1,056 (41.39)	1080 (42.34)	2.73×10^{-3}
The NLF, mean (sd)	80.77 (11.19)	30.95 (17.77)	-	56.38 (14.81)	25.71 (8.06)	-	88.45 (5.62)	55.26 (20.58)	-

*Baseline characteristics were presented as means (standard deviation [SD]) for quantitative traits and as numbers (percentages) for categorical variables. Characteristics were compared between the lower group and the higher group using t tests for continuous variables and tests for categorical variables.

GWAS identified variants at 2q31.1 associated with the NLF

Themeta-analysis identified a totalof 6 SNPsat 2q31.1 reached the genome-wide significance (P < 5×10-8) and being consistent in three cohorts, all at least reaching the nominal significance (P<0.05) and with betas of the same direction. No genome-wide significant association signal was observed after conditioning on the lead SNP rs6729788.

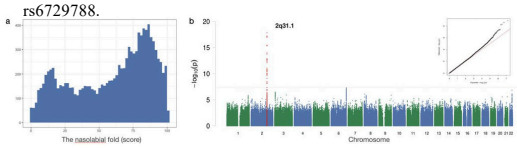


Figure 2. GWASs for the NLF in NSPT, JD and TZL. a) Histogram of the NLF. b) Manhattan plot of the meta-analysis results for the NLF from the GWASs (N=11,033). The −log10 (P values) for association were plotted for all SNPs according to their physical positions (genome-build GRCh38.p13). The red line was corresponding to the threshold for genome-wide significance (P=5×10-8).

Table 2. SNPs showing significantly association with the NLF (P<5×10-8)

							NSPT (N = 2,921)				TZL (N = 2,962)				JD (N = 5,102)				META	
SNP	Gene	CHR	MB	EA	OA	EAF	Beta	SE	P	EAF	Beta	SE	P	EAF	Beta	SE	P	Beta	P
rs6729788	*HOXD1*	2	177.09	T	G	0.34	-3.15	0.54	5.09×10^{-9}	0.34	-2.45	0.49	7.32×10^{-7}	0.34	-1.29	0.34	1.38×10^{-4}	-2.07	1.49×10^{-18}
rs6755093	*HOXD1*	2	177.12	A	G	0.33	-3.03	0.53	1.55×10^{-8}	0.34	-2.49	0.49	4.20×10^{-7}	0.34	-1.19	0.34	4.72×10^{-4}	-2.03	6.34×10^{-18}
rs1524082	*HOXD1*	2	177.11	A	G	0.33	-3.02	0.53	1.70×10^{-8}	0.34	-2.50	0.49	3.93×10^{-7}	0.34	-1.18	0.34	4.94×10^{-4}	-2.03	8.12×10^{-18}
rs7577205	*HOXD1*	2	177.09	G	C	0.38	-2.99	0.53	1.77×10^{-8}	0.39	-2.13	0.48	9.11×10^{-6}	0.40	-1.10	0.33	8.56×10^{-4}	-1.86	4.62×10^{-16}
rs970797	*HOXD1*	2	177.11	T	G	0.34	-2.98	0.53	1.70×10^{-8}	0.34	-2.52	0.49	2.93×10^{-7}	0.35	-1.15	0.34	6.74×10^{-4}	-2.01	1.07×10^{-17}
rs4893873	*HOXD1*	2	177.09	A	G	0.38	-2.95	0.53	2.57×10^{-8}	0.39	-2.14	0.48	8.06×10^{-6}	0.40	-1.07	0.33	1.18×10^{-3}	-1.87	3.51×10^{-16}

Functional annotation analysis of rs6729788

Thesignificant SNPs at 2q31.1 were locatednear HOXD1 andMTX2 genes. The derived T alleleof theleadSNPrs6729788 increasedthe severe level of the NLF. Functional annotation analysis of rs6729788 identified 8 interacting genes from antp homeobox family (HOXD, Figure 3a), which were involved in human development and facial morphology2. Epigenetic markers indicated a regulatory function of rs6729788 (Figure 3b). GTEx database showed the SNP functioned as eQTL of HOXD4 in heart ventricle (β=0.26, P = 2.44×10-5).

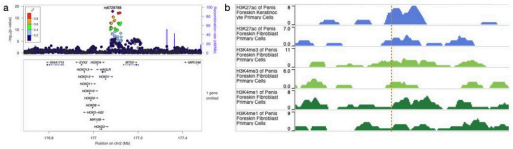


Figure 3. Functional annotation analysis of rs6729788. Regional association plot for the associated region at 2q31.1. b) Epigenetic markers H3k27ac, H3K4me1 and H3Kme3 at 2q31.3 region. The dotted line showed the position of rs6729788.

Facial morphology is potentially causal for the NLF

Since we also collected the 3Dfacial data in 8,122 individuals in the NSPT and JD cohorts, we performed MR analysis to test the potential causal relationship between facial morphology and the NLF. We chose the 13 most commonly used facial landmarks and calculated the Euclidean distance between the landmark pairs, which was then used to define three facial segments (i.e., upper face, middle face and lower face; Figure 4a). We extracted the first PC of each segment for causal analysis. MR analysis suggests that middle face morphology has a causal effect on the NLF (Table 3). Average face showed a significant difference in the middle face morphology between the highest and the lowest NLF score groups.

Figure 4. Mendelian randomization analysis of the NLF. a) Landmarks and Euclidean distance on three facial segments. b) Average face of highest and the lowest NLF score group.

Table 3. Relationship between facial morphology and the NLF in both directions.

Direction	Morphology → NLF			NLF → Morphology		
Facial region	Beta	SE	P value	Beta	SE	P value
Upper face	1.09	0.96	0.26	0.04	0.03	0.11
Middle face	2.14	0.47	4.41×10^{-6}	0.04	0.01	0.20
Lower face	0.98	0.57	0.08	0.04	0.03	0.25

Conclusions

1. Anovel signal at 2q31.1 near *HOXD1* was associated with the NLF.
2. The lead SNP rs6729788 at the 2q31.1 was interacted with 8 genes from antp homeobox family (HOXD), which were involved in the human development and facial morphology.
3. Middle face morphology was potentially causal for the NLF. Compared with the normal face, maxillary protrusion significantly aggravated the NLF.

References

1.C. Qian, T.L., H. Jiang, Z. Wang, P. Wang, M. Guan. (2018). A Detection and Segmentation Architecture for Skin Lesion Segmentation on Dermoscopy Images. In International Conference on Medical Image Computing and Computer. (Granada, Spain).

2.Huang Y, Li D, Qiao L, Liu Y, Peng Q, Wu S, Xu S, Jin L, Wang S, Tang K, Grünewald S. A genome-wide association study of facial morphology identifies novel genetic loci in Han Chinese. J Genet Genomics. 2021 Mar 20;48(3):198-207. doi: 10.1016/j.jgg.2020.10.004.

Contact: S. Wang, Dermatogenomics Group, CAS Key Laboratory of Computational Biology, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China, E-mail: wangsijia@picb.ac.cn

GWASsofthe NasolabialFold IdentifiedVariants RelatedtoGenes that AlsoAffect Facial Morphology

Journal of Investigative Dermatology (2025) ■,■—■; doi:10.1016/j.jid.2025.09.382

TO THE EDITOR The nasolabial fold (NLF) is a prominent dermatological phenotype of the aging midface. Previous anatomical studies have clarified that the NLF is potentially induced by the aging changes in the su- perficial musculoaponeurotic system ar- chitecture, cutaneous ligament, midface musculature and fat compartments, and craniofacial skeleton (Coleman and Grover, 2006; Cotofana et al, 2016). The NLF plays a critical role in facial es- thetics and is a key feature in perceived aging (Mayes et al, 2010). It also predicts certain craniofacial abnormalities linked to genetic factors (Chaaal 2018).

Methods for quantifying the NLF include clinical grading scales, image-based analyses, and device-based technologies (Hamer et al, 2018). These approaches have improved precision but still highlight variability in accuracy and reproducibility. Although the NLF has been studied extensively, the underlying genetic factors have not yet been iden- tified, presumably because of the limitations of previous phenotyping methodologies and the limited sample size.

In this study, we developed a deep learning—based method to measure the NLF and performed the largest GWAS of the NLF to date in 10,235 Han Chinese. The discovery sample included a total of 3091 individuals from National Survey of Physical Traits cohort. Studies in this cohort were approved by the Ethics Committees of Fudan University (14117) and the Shanghai Institutes for Biological Sciences (ER-SIBS-261410). The replication sample included a total of 7144 individuals from 2 cohorts: Taizhou Longitudinal cohort of Chinese origin (n = 2911) and Jidong cohort (n = 4233), which were conducted with the approval of the Ethics Committee of Fudan University (Ethics Approval no. 85) (Shanghai Ethics Committee of Kailuan General Hospital of Tangshan City, Jidong Oilfield Branch, National Petroleum Corporation in July, 2013 (approval number 2013 YILUNZI1). All participants provided written informed consent. Phenotyping details are pro- vided in Supplementary Materials and Methods. In brief, we generated the NLF score (0–100) from high-resolution 2-dimensional facial images, with a 2-step deep learning network, involving a multitask learning of facialsegmentationandadetection solution. Significant differences were observed among the 3 cohort in terms of age(P=3.87×10-286),sexdistribution(P=1.45×10-50),andtheNLF score (P < .001) (Table 1). Among 7324 individuals in the National Survey of Physical Traits and Jidong cohorts, we also collected the 3-dimensional facial data using the 2-pod 3dMDface camera system. Each 3-dimensional facial images contained 21 landmarks for the quantification of facial morphology. The deep learning—based NLF evaluation system demonstrated strong correlation with human assessment in 2 independent cohorts (Pearson correlation coefficient for National Survey of Physical Traits cohort = – 0.82, Pearson correlation coefficient for Taizhou Longitudinal cohort = –0.62) (Supplementary Figure S1).

Distributions of the NLF were statistically not normal (Shapiro—Wilk test, P < .05). Z-transformed NLF score was used in the subsequent genetic association analysis (Figure 1a). The meta-analysis identified a total of 29 SNPs at 2q31.1, reaching the genome-wide significance(P<5×10– 8) (Figure1band Supplementary Table S1) and being consistent in 3 cohorts, all at least reaching the nominal significance (P < .05)andwithbetasofthesamedirection. No genome-wide significant association signal was observed after conditioning on the lead SNP rs6729788 (multiple mapping analysis). Further prioritized rs6729788 as the most likely causal variant within this locus (posterior probability = 0.66) (Supplementary TableS2).

Functional annotation analysis of rs6729788 identified 8 interacting genes from antp homeobox family (*HOXD*) (Figure 1c), and these genes are highly enriched in reproductive system tissues, kidney tissues, adipose tissues, and skin tissues (Supplementary Figure S3). Genes near this locus have both been reported to play important roles in facial morphology, particularly among East Asian populations (Huang et al, 2021; Wu et al, 2019). Moreover, *Hox* genes exhibit pronounced age-dependent expression changes in dermal fibroblasts (Ko et al, 2024), potentially influencing NLF severity by promoting matrix metalloproteinase 2—mediated collagen degradation and altering superficial musculoaponeurotic system integrity. Epigenetic markers indicated a regulatory function of rs6729788 (Figure 1d). The derived T allele of the lead SNP rs6729788 increased the severe level of theNLF(β =–2.00,P=1.94×10– 15) (Figure 1e and Supplementary Figure S4). This allele is low frequent in East Asian populations but more common in European populations (Figure 1e), consistent with reports of more severe NLFs in Europeans (Vashi et al, 2016)

To further investigate the relationship between facial morphology and the NLF, we calculated the Euclidean distances between 21 landmark points

Abbreviation: NLF, nasolabial fold

Acceptedmanuscriptpublishedonline XXX;corrected proofpublishedonlineXXX

Table1.Baseline Characteristicsofthe Individuals

Characteristics	NSPT (n = 3091)	TZL (n = 2911)	JD(n = 4233)	*P*-Value
Age,y,mean(SD)	50.00(12.87)	56.30(9.50)	45.19(13.34)	3.87×10^{-286}
Female,n(%)	1942(62.83)	1845(63.38)	2047 (48.36)	1.45×10^{-50}
TheNLF,mean(SD)	55.19(29.09)	36.21(23.60)	71.75 (22.23)	<.001

Abbreviations:JD,Jidong;NLF,nasolabialfold;NSPT,National Survey of Physical Traits;TZL, Taizhou longitudinal.
Baselinecharacteristicsarepresentedasmeans(SD)forquantitative traits and as numbers (percentages)forcategoricalvariables.*P*-valueswerecalculated using the Kruskal—Wallistestfor continuous variables and chi-square tests for categorical variables.

derived from 3-dimensional facial data to measure facial morphology features. We then analyzed the correlation be- tween these Euclidean distances and the NLF, finding that the morphology of the lower facial region is significantly asso- ciated with the NLF (Supplementary Table S3). In addition, SNP-based causal inference test analysis suggested that the SNP influences the NLF through its effect on the Euclidean distances representing facial morphology features. This result remained significant after applying false discovery rate correction (adjusted *P* for causal inference test < .05)(SupplementaryTableS4), consistent with findings from previous anatomical studies (Sandulescu et al, 2019). To further validate the robust- ness of the result, we conducted coloc- alization and 2-sample Mendelian randomization analyses for the facial traits with significant causal inference test results. These analyses revealed that only a subset of traits (ie, AlR_ExR, ChR_ExR, EnR_EnL, and EnR_ExL) showed strong evidence of sharing a causal variant with NLF (Supplementary Tables S4—6) and exerted significant causal effects on NLF severity. Average face showed a significant difference in the lower-face morphology between the highest and the lowest NLF score groups (n = 60). Notably, the highest NLF group showed a longer middle face with prograding mouth (Supplementary Figure S5).

In conclusion, by applying a deep learning—based phenotyping approach, we identified a genome-wide significant association between the *HOXD* locus and NLF severity, highlighting a genetic link between facial morphology and skin aging. One limitation of our study is the lack of direct experimental validation for the identified genetic locus. Although functional analysis suggests regulatory roles for nearby *HOXD* genes, causal mechanisms remain unconfirmed. In addition, because our analysis was conducted exclusively in Han Chinese individuals, the generalizability of the findings to other ancestral populations remains unknown. Future studies incorporating multiethnic cohorts, CRISPR-based perturbation in skin cell models, and single-cell transcriptomic data from

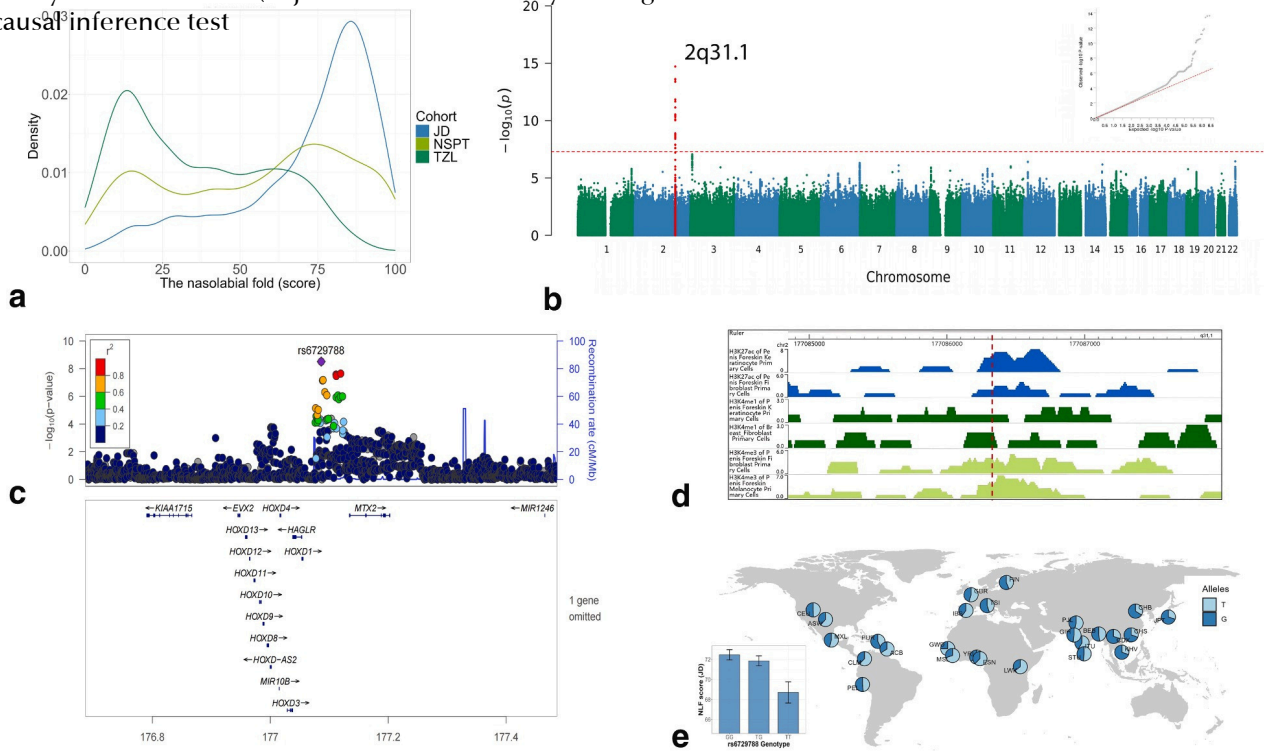


Figure 1. A meta-analysis of GWAS for the NLF identified a signal at 2q31.1. (**a**) Histogram of the NLF in the discovery cohort (NSPT) and replication cohorts (JDandTZL).(**b**)Manhattanplotofthemeta-analysisresultsfortheNLFfromtheGWASs(n=10,235).The−log10 *P*-valuesforassociationwereplottedforall SNPs according to their physical positions (genome-build GRCh37). The red line was corresponding to the threshold for genome-wide significance ($P = 5 \times 10^{-8}$).(**c**)Regionalassociationplotfortheassociatedregionat2q31.1.Colorfrombluetoredrepresentsincreasinglinkagedisequilibrium(r2)withrs6729788. (**d**) Epigenetic markers H3k27ac, H3K4me1, and H3Kme3 at 2q31.1 region. The dotted line showed the position of rs6729788. (**e**) Geography of allele frequency on rs6729788. The effect allele T on rs6729788 showed an increasing effect on more severe nasolabial folds. Data are from the 1000 Genomes Project (phase 3) through Ensembl Genome Browser, including 26 populations across 5 continental groups (n = 2504 individuals in total; sample size per population: 60—113). JD, Jidong; NLF, nasolabial fold; NSPT, National Survey of Physical Traits; TZL, Taizhou longitudinal.

humanand mouseskintoclarify functional consequences are needed.

ETHICS STATEMENT

This study was approved by the Ethics Committees of Fudan University (approval numbers 14117 and 85); Shanghai Institutes for Biological Sciences (ER-SIBS-261410); and the Ethics Committee of Kailuan General Hospital of Tangshan City, Jidong Oilfield Branch, National Petroleum Corporation (approval number 2013 YILUNZI1). Written informed consent was obtained from all participants.

DATA AVAILABILITY STATEMENT

The GWAS summary statistics are available from the National Omics Data Encyclopedia (http:// www.biosino.org/node/analysis/detail/OEZ00021305). Data usage must be in full compliance with the Regulations on Management of Human Genetic Resources in China. Individual genotype and phenotype data cannot be shared owing to Institutional Review Board restrictions on privacy concerns. Other relevant data sup- porting the key findings of this study are available within the letter and supplementary materials or from the corresponding author on reasonable request. The deep learning—based nasolabial fold scoring model developed in this study has been deployed as a cloud-based API: https://u.ss5.xyz/ nlfscore.

KEY WORDS

Facialmorphology; GWASs; Nasolabial fold

ORCIDs

Fudi, Wang: https://orcid.org/0000-0002-0208-3343
Zhao, Yuepu: https://orcid.org/0000-0002-4179-7533
Siyuan, Du: https://orcid.org/0000-0003-1602-1669
Jiarui, Li: https://orcid.org/0000-0003-0581-5850
Zhiyang, Li: https://orcid.org/0009-0007-7686-0700
Yonghao, Luo: https://orcid.org/0009-0000-1510-0790
Siying, 0009-5341-1333Fu: https://orcid.org/0009-

Wang, Sijia: https://orcid.org/0000-0001-6961-7867

CONFLICT OF INTEREST

The authors state no conflict of interest.

ACKNOWLEDGMENTS

This work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (grant number XDB38020400 to SW), National Natural Science Foundation of China (grant number 32325013), CAS Project for Young Scientists in Basic Research (grant number YSBR- 077), and Shanghai Science and Technology Commission Excellent Academic Leaders Program (22XD1424700) and was supported by the Human Phenome Data Center of Fudan University.

AUTHOR CONTRIBUTIONS

Conceptualization: SW; Data Curation: FW, YZ, SD,JL,ZL;FormalAnalysis: FW,YZ; Funding Acquisition: SW; Investigation: FW, YZ; Methodology:FW,YZ,YL; Project Administration:SW;Resources:SW, ZL,SF; Software:FW,YZ,YL;Supervision: SW;Validation:FW,YZ,SW;Visualization: FW,YZ, SW; Writing - Original Draft Preparation: FW, YZ, SW; Writing - Review and Editing: FW, YZ, SD, JL, ZL, YL, SF, SW

FudiWang1,2,3,YuepuZhao1,3 , SiyuanDu1 ,JiaruiLi1 ,ZhiyangLi2 , YonghaoLuo2,SiyingFu2 and Sijia Wang1,*

[1]CAS Key Laboratory of Computational Biology, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai,China;and2 EveLabInsight (Singapore), Singapore, Singapore [3]

Theseauthors contributedequallytothis work.

[*]Corresponding authore-mail:wangsijia@sinh.ac.cn

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of thepaperatwww.jidonline.org,andat 10.1016/j.jid.2025.09.382.

REFERENCES

Cha S, Lim JE, Park AY, Do JH, Lee SW, Shin C, et al. Identification of five novel genetic loci related to facial morphology by genome-wide association studies. BMC Genomics 2018;19:481.

Coleman SR, Grover R. The anatomy of the aging face: volume loss and changes in 3-dimensional topography. Aesthet Surg J 2006;26:S4—9.

Cotofana S, Fratila AA, Schenck TL, Redka-Swoboda W, Zilinsky I, Pavicic T. The anatomy of the aging face: a review. Facial Plast Surg 2016;32:253—60.

Hamer MA, Pardo LM, Jacobs LC, Deelen J, Uitterlinden AG, Slagboom E, et al. Facial wrinkles in Europeans: a genome-wide association study. J Invest Dermatol 2018;138:1877—80.

Huang Y, Li D, Qiao L, Liu Y, Peng Q, Wu S, et al. A genome-wide association study of facial morphology identifies novel genetic loci in Han Chinese. J Genet Genomics 2021;48:198—207.

Ko D, Mun S, Kim M, Nho YH, Lee DG, Kang S, et al. A glance into the destiny of transcriptomic activity, embodied by the HOX genes, in neonatal and aging dermal cells. Adv Biol (Weinh) 2024;8:e2300325.

Mayes AE, Murray PG, Gunn DA, Tomlin CC, Catt SD, Wen YB, et al. Ageing appearance in China: biophysical profile of facial skin and its relationship to perceived age. J Eur Acad Dermatol Venereol 2010;24:341—8.

Sandulescu T, Franzmann M, Jast J, Blaurock-Sandulescu T, Spilker L, Klein C, et al. Facial fold and crease development: a new morphological approach and classification. Clin Anat 2019;32:573—84.

Vashi NA, de Castro Maymone MB, Kundu RV. Aging differences in ethnic skin. J Clin Aesthet Dermatol 2016;9:31—8.

Wu W, Zhai G, Xu Z, Hou B, Liu D, Liu T, et al. Whole-exome sequencing identified four loci influencing craniofacial morphology in northern Han Chinese. Hum Genet 2019;138:601—11.

SUPPLEMENTARY MATERIALS AND METHODS

Study populations

The Jidong cohort. The Jidong cohort is a community-based, long-term observational cohort study to evaluate health-related risk factors. The baseline data were collected from 2013 to 2014 in the Staff Hospital, Jidong Oilfield Branch. Approval was obtained from the Ethics Committee of Kailuan General Hospital of Tangshan City and the Medical Ethics Committee, Staff Hospital, Jidong Oilfield Branch, China National Petroleum Corporation in July 2013 (approval number 2013 YILUNZI1). The approval had been renewed in 2018. To date, 9078 individuals aged >18years have been enrolled after excluding individuals who were unable or unwilling to participate. Written informed consent was obtained from all participants. This study included a total of 4233 individuals (2176 males and 2047 females, aged 20—80 years) who paid the return visit in 2018. The facial photograph and blood samples were collected in the Staff Hospital at the same time.

The National survey of physical traits cohort. TheNationalSurveyofPhysical Traits cohort is a population-based prospective cohort study of Han Chinese individuals recruited from 3 sites (ie, Taizhou, Nanning, and Zhengzhou) between 2015 and 2018, designed to explore environmental and genetic factors associated with physical traits and diseases. The National Survey of Physical Traits cohort study was conducted with the official approval of the Ethics Committees of Fudan University (14117) and the Shanghai Institutes for Biological Sciences (ER-SIBS-261410). All individuals provided written informed consent. Portrait photos of 3091 individuals (1149 males and 1942 females, aged 17—83 years) were collected in accordance with phenotyping standard operating procedure in this study. ***The Taizhou longitudinal cohort.*** The Taizhou longitudinal cohort study is a long-term observational cohort study to explore the environmental and genetic risk factors for common and non-communicable diseases. This research program was conducted with the approval of the Ethics Committee of Fudan approval 85). The detailed characteristics have been described before (Wang et al, 2009). Our replication set included 2911 health Han Chinese with portrait photos (1039 males and 1869 females), ranging from ages 31 to 87 years.

Facial data collection

All participants were asked not to take part in vigorous exercise an hour before their study visit, not to wear make-up, and to refrain from alcohol and to- bacco use 24 hours before the visit. Facial photographs were taken in a standardized space with a stabilized LED light source to ensure uniform illumination across subjects. Besides, all participants wore a shawl to help give consistent light illumination. A Canon 70D digital camera (lens: Canon EF 40 mm f/2.8) was used for all par- ticipants without the flash. The facial photograph for each participant consisted of a frontal facial shot with the eyes closed and no facial expression. The resolution of photographs was 300 dpi.

Phenotyping

Nasolabial folds (NLFs) are bilateral skin creases that extend from the sides of the nose (alae nasi) to the corners of the mouth. Traditional evaluation of the NLFs was performed using the SCINEXAscale(Vierkötteretal,2009).To better quantify NLFs, this study adopted the following 3 main steps to train the algorithm for NLF phenotyping:

1. Facial landmark detection: Facial keypoint detection algorithms were used to locate the regions of interest for NLFs. The green boxed region in the figure indicates the periocular wrinkle area (Supplementary Figure S6a).
2. Independentmodelsforsegmentation and score prediction: Independent segmentation and score prediction models are trained for NLF. These models including the following: (i) U-Net—like structure: the overall architecture is based on a U-Net, with modifications in the connection methods between feature layers (Supplementary Figure S6a).
   (ii) IBNmoduleintegrationinencoder: the convolutional layers in the encoder were combined with IBN modules (Pan et al, 2018), leveraging domain adaptation properties to enhance compatibility with input images across different skin tones and lighting conditions. It consisted of 5 downsampling stages, each constructed using IBN-b convolution blocks as the fundamental unit. The IBN-b blocks were repeated 2, 3, 4, 4, and 3 times sequentially across stages, progressively increasing and then stabilizing the feature extraction depth. This design enables hierarchical feature learning while maintaining computational efficiency in deeper layers.
   (iii)Decoder: This employed bilinear interpolation to upsample lower-resolution feature maps, followed by concatenation with corresponding higher-resolution skip connections from the encoder. Each merged feature map was then processed through 2 consecutive Conv-BN-ReLU blocks (convolution → batch normalization → ReLU activation) to refine the upsampled features.
   (v) Multiscalefeaturefusionfor score prediction: features from the intermediate layers of the decoder are fused to improve the accuracy and stability of score predictions (Supplementary Figure S6b).
3. Loss functions: segmentation loss: On the basis of BASNet (Qin et al, 2019), this consisted the following 3 components:
   (i) Weighted dice loss: an image-level loss function derived from dice loss (Sudre et al, 2017), making it particularly effective for imbalanced datasets where target regions occupy small areas. It optimized the score by introducing a weight parameter w∈[0,1] to balance precision and recall, with a logarithmic function applied to amplify gradient values. A constant C prevents division by zero:

$$L_{weighed\_dice} = -\log\frac{2w\sum_{i=1}^{H}\sum_{j=1}^{W}S(x_{i,j})G(x_{i,j})+C}{w\sum_{i=1}^{H}\sum_{j=1}^{W}S(x_{i,j})+(1-w)\sum_{i=1}^{H}\sum_{j=1}^{W}G(x_{i,j})+C}$$

(ii) StructuralSimilarityIndexMeasure (SSIM) loss: a patch-level loss function that calculates structural similarity for patches of size K × K between segmentation results and ground truth. It is originally designed for image quality evaluation, with particular emphasis on structural feature retention. For a given patch, the SSIM loss is computed as follows:

$$L_{ssim\_patch} = -\log\left(\frac{(2\mu_s\mu_g+C_1)(2\sigma_{sg}+C_2)}{(\mu_s^2+\mu_g^2+C_1)(\sigma_s^2+\sigma_g^2+C_2)}\right)$$

(iii) BinaryCross-Entropy(BCE)loss:a pixel-level loss function employing cross-entropy computation at each image pixel location. It is the most widely used loss in binary segmentation, which is computed as follows:

$$L_{bce} = -\frac{1}{W\times H}\sum_{i=1}^{W}\sum_{j=1}^{H}\left[G(x_{i,j})\log S(x_{i,j})+(1-G(x_{i,j}))\log(1-S(x_{i,j}))\right]$$

Score loss: standard regression lossis used:

$$L_{score} = \text{Smooth}_{l1}(y_s - y_g)$$

4. Evaluationmetrics:Toassesssegmentation performance, we adopted multiple pixel-level validation metrics. Specifically, the F-score (0.75) was used to summarize the harmonic mean of precision and recall, offering a balanced measure of model accuracy in delineating the NLF region. The mean absolute error between predicted NLF severity scores and expert-assigned wrinkle grades was 0.79, indicating strong agreement with clinical evaluation. These metricscollectivelydemonstrate the robustness and clinical relevance of our deep learning framework.

5. Experimentalsetup:Weconducted training and validation using the Jidong dataset comprising 4233 facial images. The dataset was randomly split into 3809 images for training and 424 for validation (90—10% ratio). For model input, bilateral NLF regions were cropped to 256 × 128 pixels. Following the skin aging atlas protocol (Bazin and Flament, 2010), all images were clinically graded from 100 (mild) to 0 (severe) on the basis of NLF depth, the base of the nose, and the corner of lips. The segmentation masks were annotated to align with the NLF's central axis, defined as the deepest fold trajectory. We used a weight of w = 0.4 and C = 1e-7 in the weighted dice loss and set K = 11 in the SSIM loss. The network was optimized using Adam optimizer (Kingma and Ba, 2014[1]) withhyperparametersas follows: learning rate lr = 1e-3, betas = (0.9, 0.999), eps = 1e-8, and weight decay = 1e-5. Training converged after 500 epochs (batch size = 64) as monitored by validation loss plateau.

[1] Kingma DP, Adam BJ.Amethod for stochastic optimization. arXiv 2014.

Genotyping

For both Jidong and National Survey of Physical Traits cohorts, genomic DNA was extracted from blood samples us- ing the MagPure Blood DNA KF Kit. All samples were genotyped using the Illumina Infinium Global Screening Array consisting of about 710,000 SNPs. We implemented exclusion criteria for quality control using PLINK (version 1.9) (Chang et al, 2015). In detail, we excluded participants with >5% missing data, the duplicated par- ticipants, and participants who failed the X-chromosome sex concordance check or had ethnic information incompatible with their genetic information. We excluded SNPs that had more than 2% missing data, those with a minor allele frequency of less than 1%, and the ones that failed Hardy—Weinberg deviation test ($P<1\times10^{-5}$).Imputationwasperformed using the 1000 Genomes Project phase 3 as the reference panel. The chip genotype data were first phased using SHAPEIT3 (Delaneau et al, 2019), and IMPUTE2 (Howie et al, 2009) was then used to impute genotypes. SNPs with an imputation quality score (INFO) <0.6, minor allele frequency <0.01, or a missing rate ≥0.01 were removed. Finally, 6,036,700 SNPs passed quality control and were used for further analyses.

For Taizhou longitudinal cohort, blood samples were collected, and DNA was extracted. Allsampleswere genotyped using the Illumina HumanOmniZhongHua-8 chip, which analysed further, a total

interrogates 894,517 SNPs. After quality controlwithPLINK(version1.9),the genotype data were phased using SHAPEIT and were imputed using IMPUTE2 with the 1000 Genomes Project phase 3.

Statistical analysis ***GWASs and meta-analysis.*** GWASs were separately conducted in National Survey of Physical Traits and 2 replication cohorts (Taizhou longitudinal and Jidong cohorts) on the Z-transformed NLF score using software package PLINK. Genetic principal components were calculated from the chip genotype data. Additive allele effects were tested in linear models adjusted for covariates (age, sex, sampling locations, and the top 5 genomic principal components). GWAS outputs were combined using inverse variance fixed-effect meta-analysis as implemented in METAL (Willer et al, 2010). *P*-values ≤ 5×10− 8 fromthemeta-analysiswere considered as genome-wide significant. The inflation factor was estimated close to 1.0(λ =0.995)andnotfurtherconsidered. GWAS results were visualized using Manhattan plots, Q–Q plots, Regional plots, and association plots were generated using HaploView and LocusZoom.

Causal inference test analysis. The causal inference test was performed with the R package "cit" (version 2.2.2) (Chalupka et al, 2014). Briefly, the causal inference test requires the following 4 criteria to be met for a valid causal inference: (i) SNP and the NLF are associated, (ii) SNP is associated with facial morphology feature after adjusting for the NLF, (iii) facial morphology feature is associated with the NLF after adjusting for SNP, and (v) SNP is independent of the NLF after adjusting for facial morphology feature. The maximum *P*-value among the 4testswasreportedasthePCIT value. ResultswithPCIT value<0.05were considered statistically significant.

Colocalization analysis. Colocalization analysis was performed using a Bayesian approach in the R package 'coloc' (version 5.1.0) (Giambartolomei et al, 2014). The method evaluates whether 2 association signals share a common causal variant. Genetic association summary statistics for all facial Euclidean distance traits were obtained from a meta-GWAS, conducted in the same cohort as the NLF GWAS described earlier. Analyses were conducted within ±250 kb of the lead SNP, using default prior probabilities. Posterior probabilities were calculated for 5 hypotheses, with PP.H4. abf ≥ 0.9 considered strong evidence for colocalization.

Two-sample Mendelian randomization analysis. TheMendelianrandomization analysis was conducted using the 'TwoSampleMR' package (version 0.5.4) (Hemani et al, 2018) to explore the association between the NLF and facial morphology. Given the limited number of SNPs thatmettheGWASsignificance threshold of 5e-8, we relaxed the threshold to 5e-5. Multiple complementary approaches(Mendelianrandomization Egger,inversevarianceweighted, weighted median, simple mode, and weighted mode) were used to estimate the effectsoftheexposureandoutcome. Heterogeneity was assessed using Cochran's *Q* statistic, and horizontal pleiotropy was evaluated using the Mendelian randomizationPRESSOglobaltest (Verbanck et al, 2018), with $P < .05$ indicating significant pleiotropy.

Fine-mapping analysis. To prioritize likely causal variants within the genome-wide significant locus at 2q31.1, we performed statistical fine mapping using the PAINTOR (version 3.0) algorithm. We selected all SNPs located within ±50 kb of the lead SNP rs6729788 and included their Z-scores from GWAS summary statistics. Analyses were performed under models assuming either 1 or 2 causal variants. Posterior probabilities for each SNP were calculated to assess their likelihood of being causal.

Functional analysis. The gene expression heatmap of signal SNPs was visualized using FUMA (Watanabe et al, 2017) on the basis of Genotype-Tissue Expression data. Functional annotation was then visualized using the WashU Epigenome Browser (Li et al, 2022) where we displayed histone modification peaks (H3K27ac, H3K4me1, H3K4me3) for normal human foreskin fibroblasts and keratinocytes, obtained from the ENCODE database. These genomic elements are indicative of regulatory regions, providing insights into the functional roles of the SNPs in gene expression and facial morphology.

Facial traits and average face

Three-dimensional facialscans were captured using the dual-lens 3dMDface system. The raw data were preprocessed using MeshMonk (White et al, 2019) following the following steps: (i) the scans were preprocessed by manually marking the nasal tip as the origin, (ii) normalization of all face using an anthropometric mask template with 7906 dense landmarks, (iii) rigid registration based on the Iterative Closest Point algorithm, (iv) nonrigid registration based on the Thin-Plate Spline algorithm, and (v) the samples were aligned to a mean face using generalized Procrustes analysis. Each mapped image was checked, and any abnormal images with poor quality were deleted.

We further extracted facial morphological features on the basis of 3-dimensional facial data. In this study, we focused on 119 facial traits as paired distance between 21 previously defined facial landmarks (Xiong et al, 2019) (Supplementary Figure S7). To visualize the relationship between the NLF and facial morphology, principal component analysis clustering was performed on the Euclidean distances in the nasolabial region. On the basis of the first principal component, the top 30 and bottom 30 samples were selected to represent the largest and smallest phenotypes. The average face for each group was generated by calculating the mean X, Y, Z coordinates from the 7906 facial landmarks of these selected samples.

SUPPLEMENTARY REFERENCES

Bazin R, Flament F. Skin aging atlas. In: Asian type2. Paris: Med'Com; 2010.

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience 2015;4. s13742-015-0047-8:7.

Delaneau O, Zagury JF, Robinson MR, Marchini JL, Dermitzakis ET. Accurate, scalable and integrative haplotype estimation. Nat Commun 2019;10:5436.

Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian test for colocalisation between pairs of geneticassociationstudiesusingsummary statistics. PLoS Genet 2014;10:e1004383.

Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. eLife 2018;7: e34408.

Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the

next generation of genome-wide association studies. PLoS Genet 2009;5:e1000529.

Li D, Purushotham D, Harrison JK, Hsu S, Zhuo X, Fan C, et al. WashU epigenome browser update 2022. Nucleic Acids Res 2022;50:W774—81.

Millstein J, Chen GK, Breton CV. Cit: hypothesis testing software for mediation analysis in genomic applications. Bioinformatics 2016;32:2364—5.

Pan X, Luo P, Shi J, Tang X. Two at once: enhancing learning and generalization capacities via ibn-net. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y, editors. Proceedings of the European conference on computer vision (ECCV). Cham, Switerzland: Springer; 2018. p. 484—500.

Qin X, Zhang Z, Huang C, Gao C, Dehghan M, Jagersand M. Basnet: boundary-aware salient object detection. Long Beach, CA: Paper presented at: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition; 15—20 June 2019.

Sudre CH,LiW,VercauterenT,OurselinS,Jorge Cardoso M. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. Deep Learn Med Image Anal Multimodal Learn Clin Decis Support (2017) 2017;2017:240—8.

Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases [published correction appears in Nat Genet 2018;50:1196]. Nat Genet 2018;50:693—8.

Vierkö¨tter A, Ranft U, Krä¨mer U, Sugiri D, Reimann V, Krutmann J. The SCINEXA: a novel, validated score to simultaneously assess and differentiate between intrinsic and extrinsic skin ageing. J Dermatol Sci 2009;53:207—11.

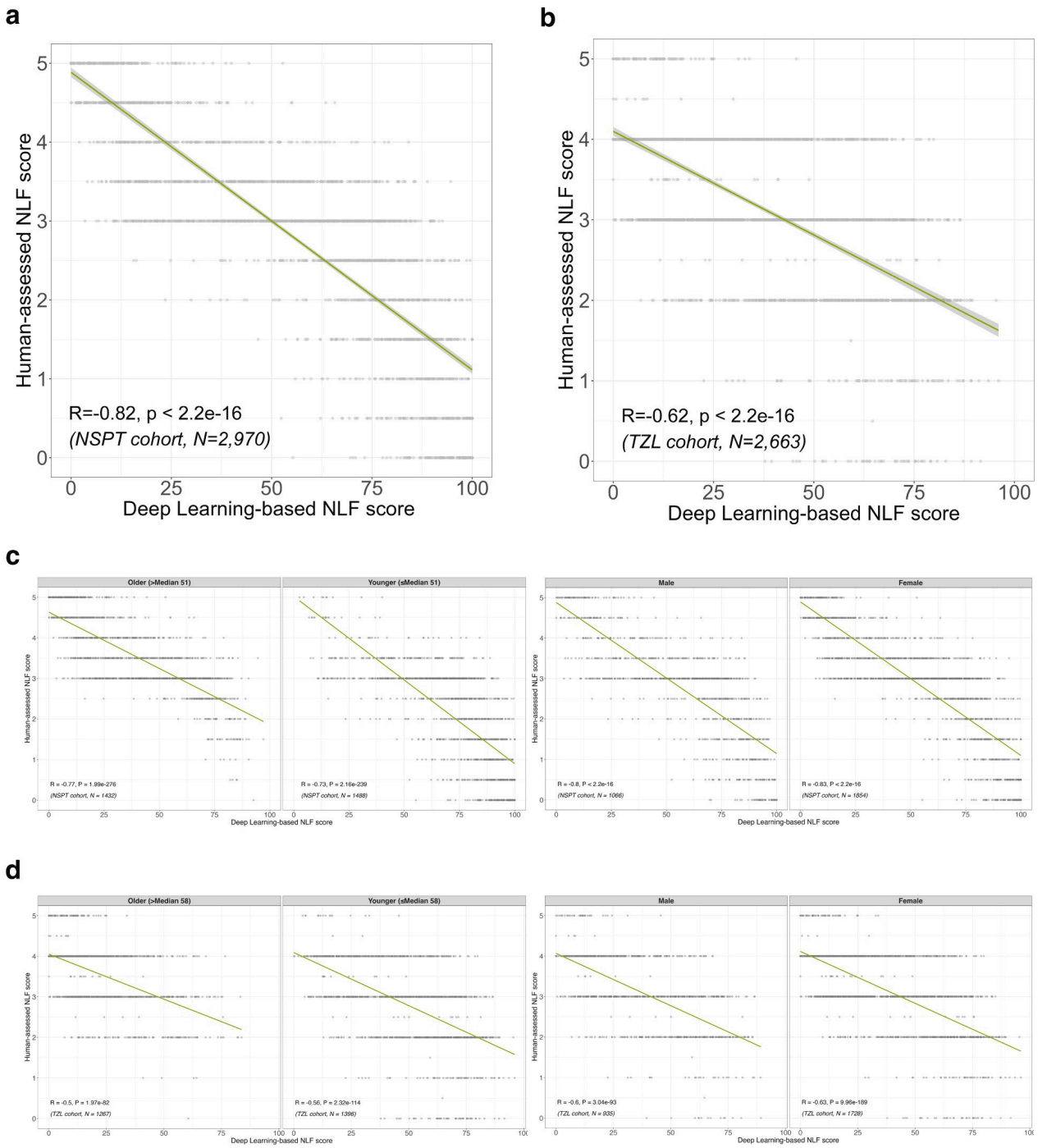
Wang X, Lu M, Qian J, Yang Y, Li S, Lu D, et al. Rationales, design and recruitment of the Taizhou Longitudinal Study. BMC Public Health 2009;9:223.

Watanabe K, Taskesen E, Van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun 2017;8:1826.

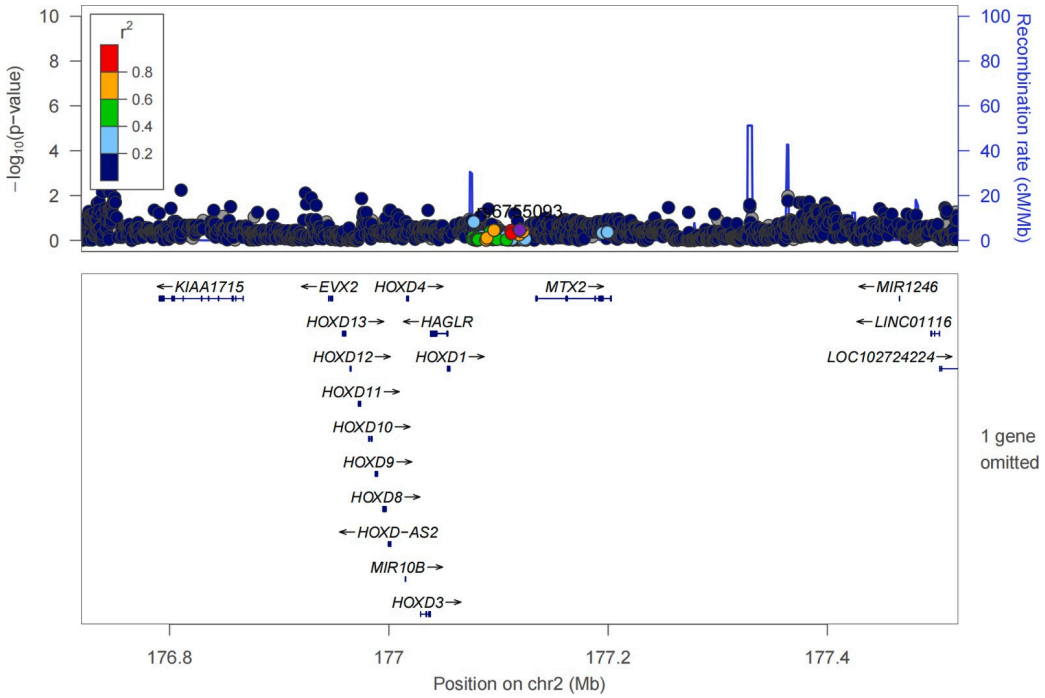
White JD, Ortega-Castrilló´n A, Matthews H, Zaidi AA, Ekrami O, Snyders J, et al. MeshMonk: open-source large-scale intensive 3D phenotyping. Sci Rep 2019;9:6085.

Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26:2190—1.

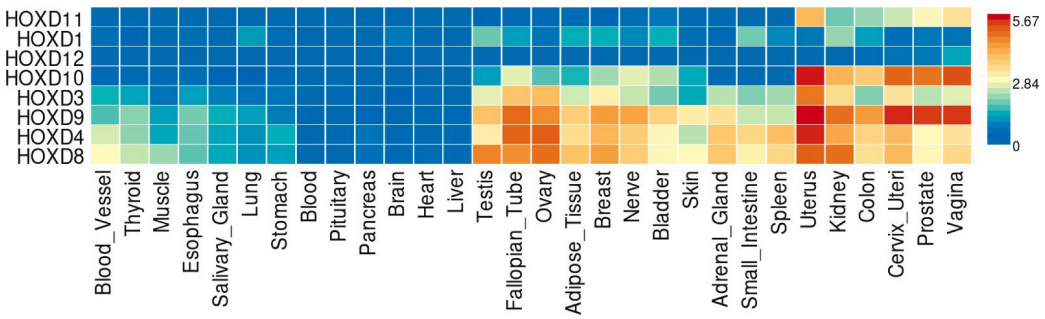
Xiong Z, Dankova G, Howe LJ, Lee MK, Hysi PG, De Jong MA, et al. Novel genetic loci affecting facial shape variation in humans. eLife 2019;8:e49898.



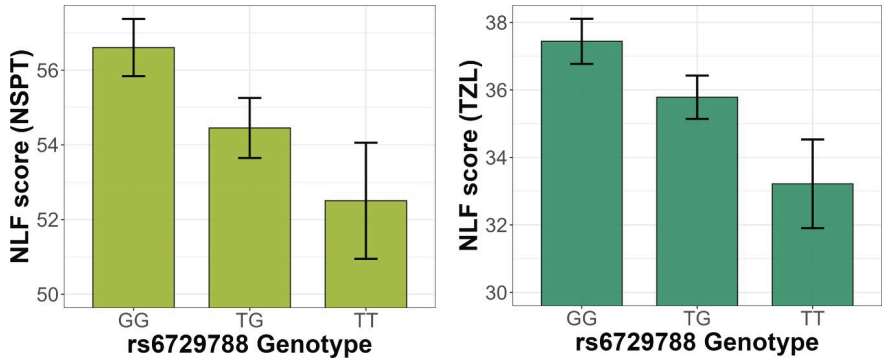
Supplementary Figure S1. Pearson correlation of NLF between human assessment and deep-learning evaluating system in NSPT and TZL. X-axis indicates NLF scores fromthedeep-learningsystem,andY-axisindicatesNLFscoresfromhumanassessment.Thecorrelationcoefficient(R)isprovidedforeachcohortto indicate thestrengthoftherelationshipbetweenthe2methods.(**a,b**)ResultsforNSPTandTZL.(**c,d**)Resultsstratifiedbyageandsexin NSPTandTZL, respectively. NLF, nasolabial fold; NSPT, National Survey of Physical Traits; TZL, Taizhou longitudinal.



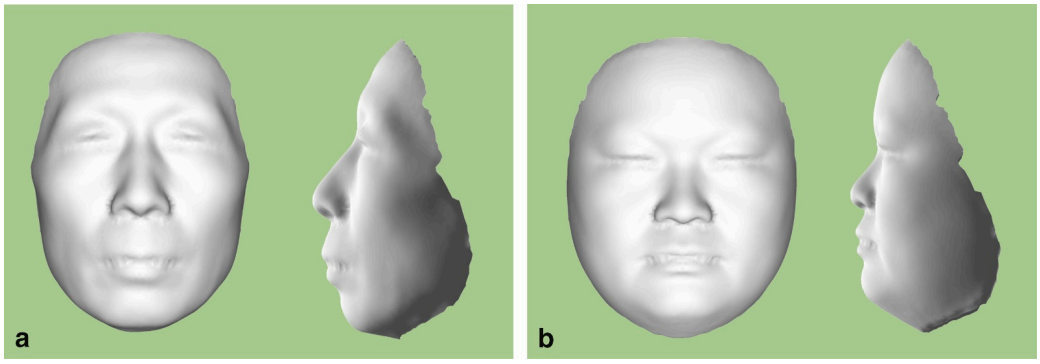
Supplementary Figure S2. Functional annotation analysisat2q31.1. Regional association plot for the associated regionat2q31.1 after controlling the leading SNP rs6729788.



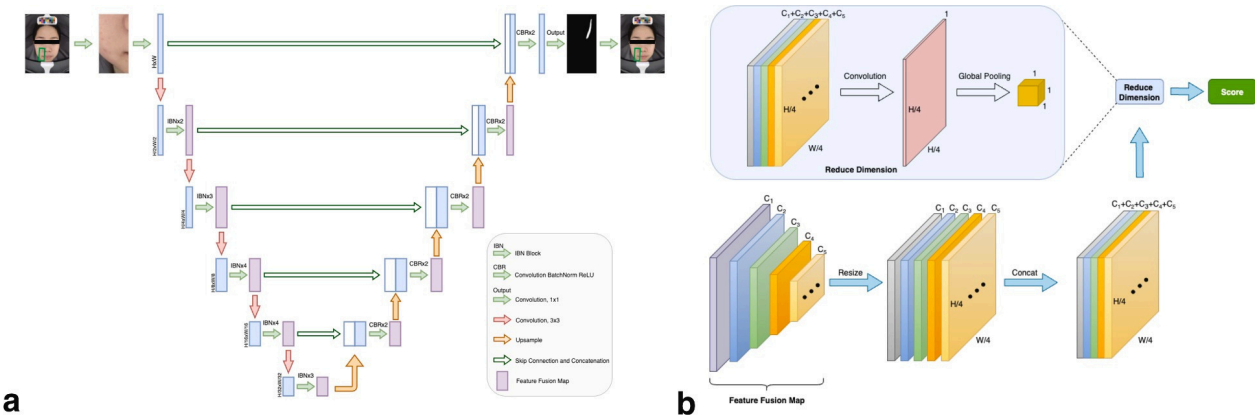
Supplementary FigureS3.Heatmapshowingtheexpressionofgenesnearrs6729788indifferenttissuesfrom GTExdata. The rowsrepresent the genes,andthe columnsrepresent differenttissuetypes. Geneexpressionisshownonablue-to-redscale, withblueindicatinglowexpressionandred indicatinghigh expression,highlightingthevariationacrosstissues. GTEx,Genotype-TissueExpression.



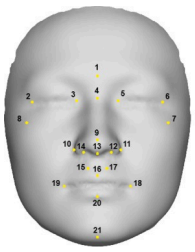
Supplementary Figure S4. Distribution of NLF scores by genotype groups of rs6729788. Bar plots show the NLF scores in the (**a**) NSPT cohort and the (**b**) TZL cohort. Each bar represents the mean NLF score for a genotype group, with error bars indicating the SEM. NLF, nasolabial fold; NSPT, National Survey of Physical Traits; TZL, Taizhou longitudinal.



Supplementary Figure S5. Average facial morphology differences by NLF severity in the NSPT cohort. Panels show the (**a**) highest and (**b**) lowest NLF score groups. NLF, nasolabial fold; NSPT, National Survey of Physical Traits.



SupplementaryFigureS6. Diagramofthenasolabialfolddetectionalgorithm.(**a**) Deeplearningframework. (**b**) Featuremapsschematicdiagram. The individual shownisavolunteer participantinour studywhoprovidedwritteninformedconsentforthe deidentified imagetobeusedinpublication.



No.	Name	Abbr.	No.	Name	Abbr.	No.	Name	Abbr.
1	Glabella	G	8	Zygion Left	ZyL	15	Philtrum Left	PhL
2	Exocanthion Left	ExL	9	Pronasale	Prn	16	Labiale Superius	Ls
3	Endocanthion Left	EnL	10	Alare Ledt	AIL	17	Philtrum Right	PhR
4	Nasion	N	11	Alare Right	AIR	18	Cheilion Right	ChR
5	Endocanthion Right	EnR	12	Subnasale Right	SnR	19	Cheilion Left	ChL
6	Exocanthion Right	ExR	13	Subnasale	Sn	20	Labiale Inferius	Li
7	Zygion Right	ZyR	14	Subnasale Left	SnL	21	Gnathion	Gn

SupplementaryFigure S7.Adetailedoverviewofthe21facialanatomicallandmarks,includingtheirpreciselocationsanddefinitions. The leftpanel provides a schematic representationofthehumanfacewiththepreciselocationsofthe21facialanatomicallandmarks.The rightpanelpresentsadetaileddescriptionof each landmark,includingitsnameandabbreviation.Abbrdenotes abbreviations,andNo denotesnumber.