scientific reports

Check for updates

OPEN

Comparisons between Caucasian-validated and Chinese-validated photo-numeric scales for assessing facial wrinkles

Jun Yan Ng[1], Hongyu Zhou[1], Tianqi Li[1] & Fook Tim Chew[1,2]✉

Skin ageing results in wrinkling. In this study, we discuss four types of facial wrinkles: Crow's Feet wrinkles, forehead wrinkles, glabellar frown wrinkles, and nasolabial folds. These four phenotypes can be assessed either with a photo-numeric scale developed and validated on Caucasian skin (i.e., Caucasian scale) or with a photo-numeric scale developed and validated on Chinese skin (i.e., Chinese scale). As Caucasian and Chinese skin have inherent differences, the main objective of this study is to determine whether these inherent differences affect the suitability of evaluating facial wrinkles on ethnic Chinese skin with a Caucasian scale. Three trained assessors studied four types of wrinkles on the faces of 1,081 ethnic Chinese young adults from the Singapore/Malaysia Cross-sectional Genetics Epidemiology Study (SMCGES). We found that Caucasian scales and Chinese scales are concordant (Spearman's Rank Correlation (ρ) values: 0.53–0.80) and the level of agreement between the Caucasian scales and Chinese scales is moderately high (Cohen's Kappa (κ) values: 0.40–0.49). When tested on ethnic Chinese skin, both the Caucasian scale and the Chinese scale are largely consistent in showing presence or absence of a given facial wrinkle (Area under curve (AUC) values: 0.79–0.90). All assessors are highly internally consistent (Weighted Kappa (κ_w) values: 0.686–0.992). Our results build confidence that four types of facial wrinkles on ethnic Chinese faces can be assessed with Caucasian scales. To the best of our knowledge, Chinese scales for facial wrinkles beyond the four types discussed here have yet to be developed. Caucasian scales for these other facial wrinkles will also need to be tested for their suitability to be used on ethnic Chinese skin as and when more Chinese scales are developed.

Abbreviations

SMCGES	Singapore/Malaysia Cross-sectional Genetics Study
ISAAC	International Study of Asthma and Allergies in Childhood
IBM SPSS/PC	International Business Machines Corporation Statistical Package for Social Scientists Personal Computer (SPSS/PC)
EOS	Electro-Optical System
IS	Image stabilisation
USM	Ultrasonic motor
EF	Electronic focus
SD	Standard deviation
SGD	Singapore dollars
HDB (Public housing)	Public housing constructed by the Singapore Housing Development Board
BMI	Body mass index
cm	Centimetres
m^2	Squared metres
kg	Kilograms
ROC	Receiver operator characteristic
AUC	Area under curve

[1]Department of Biological Sciences, Faculty of Science, National University of Singapore, Singapore 117543, Singapore. [2]Allergy and Molecular Immunology Laboratory, Lee Hiok Kwee Functional Genomics Laboratories, Block S2, Level 5, 14 Science Drive 4, Lower Kent Ridge Road, Singapore 117543, Singapore. ✉email: dbscft@nus.edu.sg

Skin ageing is a complex process which presents itself in a wide-ranging span of phenotypes including wrinkles, sagging skin, changes in skin pigmentation, and photodamage.

In our recent systematic review of skin ageing phenotypes[1], we compiled a detailed list of 56 skin ageing phenotypes identified from published literature and reputable medical books.

While skin ageing phenotypes are evaluated using photo-numeric scales, the bulk of these phenotypes are evaluated on scales which are developed and validated on Caucasian skin (i.e., Caucasian scales). Meanwhile, our recently published meta-analysis of the risk factors of skin ageing found significant differences in skin ageing progression across races[2]. Importantly, it is known for some time that wrinkle onset and facial wrinkling rates differ between Caucasian skin and Chinese skin[3]. Moreover, Asian races are significantly associated with decreased Crow's Feet wrinkles, forehead wrinkles, and glabellar frown wrinkles[4]. A recent paper by Zhang *et al.* in 2017[5] which developed several scales (i.e., Chinese scales) for evaluating facial wrinkles on Chinese skin appears to reinforce the idea that Caucasian skin and Chinese skin are different.

However, despite the innate differences between Caucasian and Chinese skin, it is unclear whether Caucasian scales remain as suitable assessment tools to evaluate facial wrinkles on the ethnic Chinese skin. The availability of both a Caucasian scale and a Chinese scale for evaluating the same facial wrinkle presents us with an opportunity to investigate this question.

In this study, we evaluate four facial wrinkles (Crow's Feet wrinkles, forehead wrinkles, glabellar frown wrinkles, and nasolabial folds) on both a Caucasian scale and a Chinese scale and investigate how they compare with each other.

Methods

Participant recruitment

The Singapore/Malaysia Cross-sectional Genetics Epidemiology Study (SMCGES) participants were recruited for this study. These participants have been described in detail in our previous studies[6–8]. In brief, the SMCGES comprises of participants from Singapore and Malaysia and has been previously studied in epidemiological and genetic studies of allergic diseases.

Participant recruitment for the previous study (i.e., the epidemiological and genetic study) was conducted through emails and posters from the National University of Singapore, Singapore (2005 to 2023), Universiti Tunku Abdul Rahman (UTAR) Campus, Malaysia (2016 to 2018), and Sunway University, Malaysia (2019 and 2022). These participants were recruited on a walk-in and voluntary basis.

All the participants from collection drives held from 2011 to 2023 and consented to re-contact were invited to participate in this present study (i.e., the skin ageing study). A total of 10,248 participants were invited to participate in the skin ageing study, of which 3,365 completed the study. All 3,365 participants constitute our study population. The demographics of the respondents and non-respondents are largely the same and have been previously reported[9]. In this study, a representative subset of the 3,365 participants was selected for detailed analyses by three trained and blinded assessors. This subset comprises of 1,081 ethnic Chinese young adult participants (Table 1).

The study was conducted in accordance with the Declaration of Helsinki and Good Clinical Practices.

Survey data collection

The SMCGES participants have previously completed a set of investigator-administered, validated International Study of Asthma and Allergies in Childhood (ISAAC) questionnaires which collated data on their sociodemographic, personal lifestyles, and familial and personal medical history.

Following their completion of the ISAAC questionnaires, the SMCGES participants signed an informed consent form to participate in this present study (i.e., the skin ageing study). In this present study, the participants completed an investigator-administered skin ageing questionnaire with more personal lifestyle questions (e.g., usage of anti-ageing skincare creams, substances, or therapies).

Data on participation in skincare routine is not available. However, data is available for the following dichotomous question: 'Have you ever applied any anti-ageing creams, consumed any anti-ageing substances or undergone any anti-ageing therapies?' Participating in skincare routine and engaging in anti-ageing interventions share the same underlying motivation (i.e., taking the initiative to deliberately perform interventional acts which are believed to be beneficial to oneself, such as applying cream onto one's skin). Hence, participation in skincare routine is approximated with data on the usage of anti-ageing skincare creams, substances, or therapies.

Image data collection

Investigators acquire photographic documentation of all participants. Photographs are taken using the same camera (Canon EOS 6DII Body with a EF85 f/1.4 L IS USM lens) and tripod positioned one metre away from the participant. Photographs are taken with identical camera settings, lighting, and positioning at five angles – *en face*, 45° oblique, and 90° side profiles, in line with the standards of recent studies[10,11]. Photos of four facial expressions are taken under standardised conditions at each angle – at-rest, eyes closed, smiling, and with eyebrows raised, enabling the study of in-motion skin ageing phenotypes alike what is performed by Zhang *et al.* in 2017 [5]. In total, 20 photos are collected from each participant.

Evaluation of skin ageing phenotypes

Skin ageing phenotypes are evaluated by three trained assessors. While we did not conduct a live validation, the evaluation and validation exercise were conducted shortly after the photographs were taken. Each of the three assessors used all 20 photos from each participant to aid the evaluation process. For each of the 1,081 participants, four kinds of facial wrinkles were evaluated using a total of ten photo-numeric scales. The ten

Characteristic	Participants evaluated ($n = 1,081$)
Age, mean (SD), y	26.15 ± 8.11[a]
Height, mean (SD), cm	166.03 ± 8.60[a]
Weight, mean (SD), kg	60.96 ± 12.60[a]
Age range, y	18 to 73
BMI, mean (SD), kg/m^2	22.02 ± 3.99 [a]
Sex[c]	
Male	420 (38.85%)[b, d]
Female	661 (61.15%)[b, d]
Race[c]	
Chinese	1,081 (100%)[b]
Malay	0 (0%)[b]
Indian	0 (0%)[b]
Others	0 (0%)[b]
Total monthly family income per capita	
Low	117 (10.82%)[b, d]
Moderate	275 (25.44%)[b, d]
High	226 (20.91%)[b, d]
Very high	446 (41.26%)[b, d]
Missing/Invalid	17 (1.57%)[b, d, e]
Housing	
Flat	664 (61.42%)[b, d]
Condominium/Private Apartment	320 (29.60%)[b, d]
Landed Property	76 (7.03%)[b, d]
Missing/Invalid	21 (1.94%)[b, d, e]

Table 1. Demographics of Singapore and Malaysia ethnic Chinese participants recruited from the Singapore/Malaysia Cross-sectional Genetics Epidemiology Study (SMCGES) for the current assessment. [a] The values after ± are standard deviation values. [b] Data are presented as number (percentage) of study participants. [c] Sex and race are self-reported. [d] Percentages may not total 100 due to rounding. [e] Missing/Invalid refers to responses that are either left blank or otherwise invalid. SD, standard deviation; y, years, cm, centimetres; kg, kilograms, BMI, body mass index; kg/m^2, kilograms per square metre.

photo-numeric scales are (1) a Caucasian scale for Crow's Feet wrinkles[12], (2) a Chinese scale for Crow's Feet wrinkles at-rest[5], (3) a Chinese scale for Crow's Feet wrinkles in-motion[5], (4) a Caucasian scale for forehead wrinkles[13], (5) a Chinese scale for forehead wrinkles at-rest[5], (6) a Chinese scale for forehead wrinkles in-motion[5], (7) a Caucasian scale for glabellar frown wrinkles[14], (8) a Chinese scale for glabellar frown wrinkles[5], (9) a Caucasian scale for nasolabial folds[15], and (10) a Chinese scale for nasolabial folds[5].

Each of the ten photo-numeric scales is composed of an ordered series of photos. The first photo shows the absence of the specific facial wrinkle studied by the scale. Every successive photo shows an increase in the depth and number of wrinkles in the target area as compared to the photo which comes immediately before it. The target area of the Crow's Feet wrinkles scale is the skin at the corner of the eyes, the target area of the forehead wrinkles scale is the skin of the forehead, the target area of the glabellar frown wrinkles is the skin between the eyebrows, and the target area of the nasolabial folds scale is the skin beside the nose.

More details on the validated photo-numeric scales used in our current assessment are reported in Table S1. More details on the facial wrinkling scales can also be found in our previous work[16].

Before phenotyping the bulk of the participants, photographs of 30 participants are randomly selected and openly discussed to reach a consensus among all three assessors. This is done to calibrate the assessment scores. After calibration, each of the three trained assessors grades the rest of the photographs independently from the other two assessors. Thus, other than a randomly selected handful of participants ($n = 30$) whose photographs are used for calibration purposes; all the 1,081 participants are assessed three times independently.

Statistical analysis

All the photo-numeric scales are standardised.

Two-tailed bivariate correlations for Spearman's Rank correlation (ρ), Cohen's Kappa (κ), and weighted kappa (κ_w) are calculated using Version 25 of the IBM Statistical Package for Social Scientists (SPSS/PC). Mean values are reported in Table 2 and S3 while raw values split by assessor are reported in Tables S4-S5. Weighted kappa coefficients (κ_w) for intra-assessor agreement are reported in Table S6. The strengths are interpreted as follows – 0.00-0.19: very weak, 0.20–0.39: weak, 0.40–0.59: moderate, 0.60–0.79: strong, and 0.80-1.00: very strong.

Bubble plots (Fig. 1 and S1) compare the mean scores given on the Caucasian scale and the Chinese scales for four wrinkling phenotypes: (i) Crow's Feet wrinkles, (ii) forehead wrinkles, (iii) glabellar frown wrinkles, and (iv) nasolabial folds. A linear relationship is assumed for computing the equation for the goodness of fit and the

Measurement		Phenotype			
		Crow's Feet wrinkles	Forehead wrinkles	Glabellar frown wrinkles	Nasolabial folds
Mean Spearman's Rank Correlation (ρ)	Value	0.53 ± 0.03	0.66 ± 0.02	0.78 ± 0.03	0.80 ± 0.01
	p-Value	1.00×10^{-80}	2.21×10^{-134}	9.24×10^{-224}	8.50×10^{-245}
Mean Cohen's Kappa (κ)	Value	0.42 ± 0.26	0.40 ± 0.32	0.40 ± 0.02	0.49 ± 0.20
	p-Value	2.28×10^{-7}	1.83×10^{-8}	4.17×10^{-94}	2.33×10^{-3}
Mean area under curve (AUC) of the Receiver Operator Characteristic (ROC) curve	When the Caucasian photo-numeric scale is the gold standard	0.82 ± 0.07	0.84 ± 0.09	0.90 ± 0.10	0.79 ± 0.05
	When the Chinese at-rest photo-numeric scale is the gold standard	0.61 ± 0.06	0.62 ± 0.04	0.84 ± 0.12	0.77 ± 0.04
Mean Coefficient of determination (R^2) for the equation for the goodness of fit		0.8749	0.8924	0.8748	0.8925
Equation for the goodness of fit		Chinese scale = 2.1024 × Caucasian scale + 0.0252	Chinese scale = 2.0424 × Caucasian scale + 0.0225	Chinese scale = 2.1862 × Caucasian scale + 0.0115	Chinese scale = 1.853 × Caucasian scale + 0.0119

Table 2. Comparison[1] between the caucasian photo-numeric scale and the Chinese at-rest photo-numeric scale for assessing four skin wrinkling phenotypes: (i) Crow's Feet wrinkles, (ii) forehead wrinkles, (iii) glabellar frown wrinkles, and (iv) nasolabial folds. [1] Means are calculated from three assessors.

corresponding coefficient of determination (R^2) value. The strength of the R^2 value is interpreted as follows – 0.00-0.19: very weak, 0.20–0.39: weak, 0.40–0.59: moderate, 0.60–0.79: strong, and 0.80-1.00: very strong.

The area under curve (AUC) of the Receiver Operator Characteristic (ROC) curve is computed using SPSS. Mean values are reported in Table 2 and S3 while raw values split by assessor are reported in Figures S2-S5. The strength of the AUC is interpreted as follows – 0.51–0.59: very weak, 0.60–0.69: weak, 0.70–0.79: fair, 0.80–0.89: strong, and 0.90-1.00: very strong.

Pair-wise comparisons were conducted using chi-square tests to evaluate the proportion of participants with facial wrinkles of different severity levels when stratified by age. Four facial wrinkles were evaluated in this manner: (i) Crow's Feet wrinkles (Figure S6i), forehead wrinkles (Figure S6ii), glabellar frown wrinkles (Figure S6iii), and nasolabial folds (Figure S6iv). Chi-square trend tests are performed to evaluate whether the changes in these proportions follow a significant trend with chronological age (Figures S6i-iv).

Participants were also stratified according to whether they practice skincare routines. Another set of pair-wise comparisons was also conducted in both stratified groups to evaluate the proportion of participants with facial wrinkles of different severity levels when further stratified by age (Figures S7-S10). Chi-square trend tests are performed to evaluate whether the changes in these proportions follow a significant trend with chronological age in participants who practice skincare routines and in participants who do not practice skincare routines (Figures S7-S10).

Results

Participant demographics

The 1,081 participants evaluated in this current study present a representative overview of the epidemiology of skin ageing in the Singapore ethnic Chinese young adult population. The same 1,081 participants have been previously evaluated for aspects of photo-ageing in our earlier work[9]. There are more females ($n = 661$, 61.15%) than males in our current study. On average, participants are 26.15 ± 8.11 years old, 166.03 ± 8.60 cm in height, 60.96 ± 12.60 kg in weight, and have a BMI of 22.02 ± 3.99 kg/m^2. Our study population consists predominantly of young adults aged 21 to 30. The youngest participant is 18 years' old, and the oldest participant is 73 years' old (Figure S11). Most participants have a total monthly family income per capita of ≥ 6,000 Singapore dollars (SGD) ($n = 446$, 41.26%). Most participants stay in HDB public housing ($n = 664$, 61.42%) (Table 1).

Despite being a relatively young study group, we report that the way facial wrinkles manifest in one age group is significantly different from another age group (Figures S6i-iv). We treat participants aged 18 to 20 years old as the reference group. When measured on the Crow's Feet wrinkles scale by Tsukahara *et al.*[12], participants aged 21 to 25 years old (p-value $= 3.41 \times 10^{-15}$), participants aged 26 to 30 years old (p-value $= 4.21 \times 10^{-10}$), participants aged 31 to 35 years old (p-value $= 3.31 \times 10^{-18}$), participants aged 36 to 40 years old (p-value $= 1.32 \times 10^{-17}$), and participants aged above 40 years old (p-value $= 2.63 \times 10^{-121}$) exhibit significantly more Crow's Feet wrinkles as compared to the reference group (Figure S6i). The severity of Crow's Feet wrinkles also increases progressively with age (Chi-square trend test p-value < 0.001) (Figure S6i).

Similar observations are made for a different facial wrinkle: forehead wrinkles. Forehead wrinkles are assessed through a photo-numeric scale by Flynn *et al.*[13]. Participants aged 21 to 25 years old (p-value $= 6.04 \times 10^{-10}$), participants aged 26 to 30 years old (p-value $= 8.71 \times 10^{-7}$), participants aged 31 to 35 years old (p-value $= 3.16 \times 10^{-12}$), participants aged 36 to 40 years old (p-value $= 4.75 \times 10^{-14}$), and participants aged above 40 years old (p-value $= 2.88 \times 10^{-55}$) exhibit significantly more photo-ageing as compared to the reference group (Figure S6ii). The severity of forehead wrinkles also increases progressively with age (Chi-square trend test p-value < 0.001) (Figure S6ii).

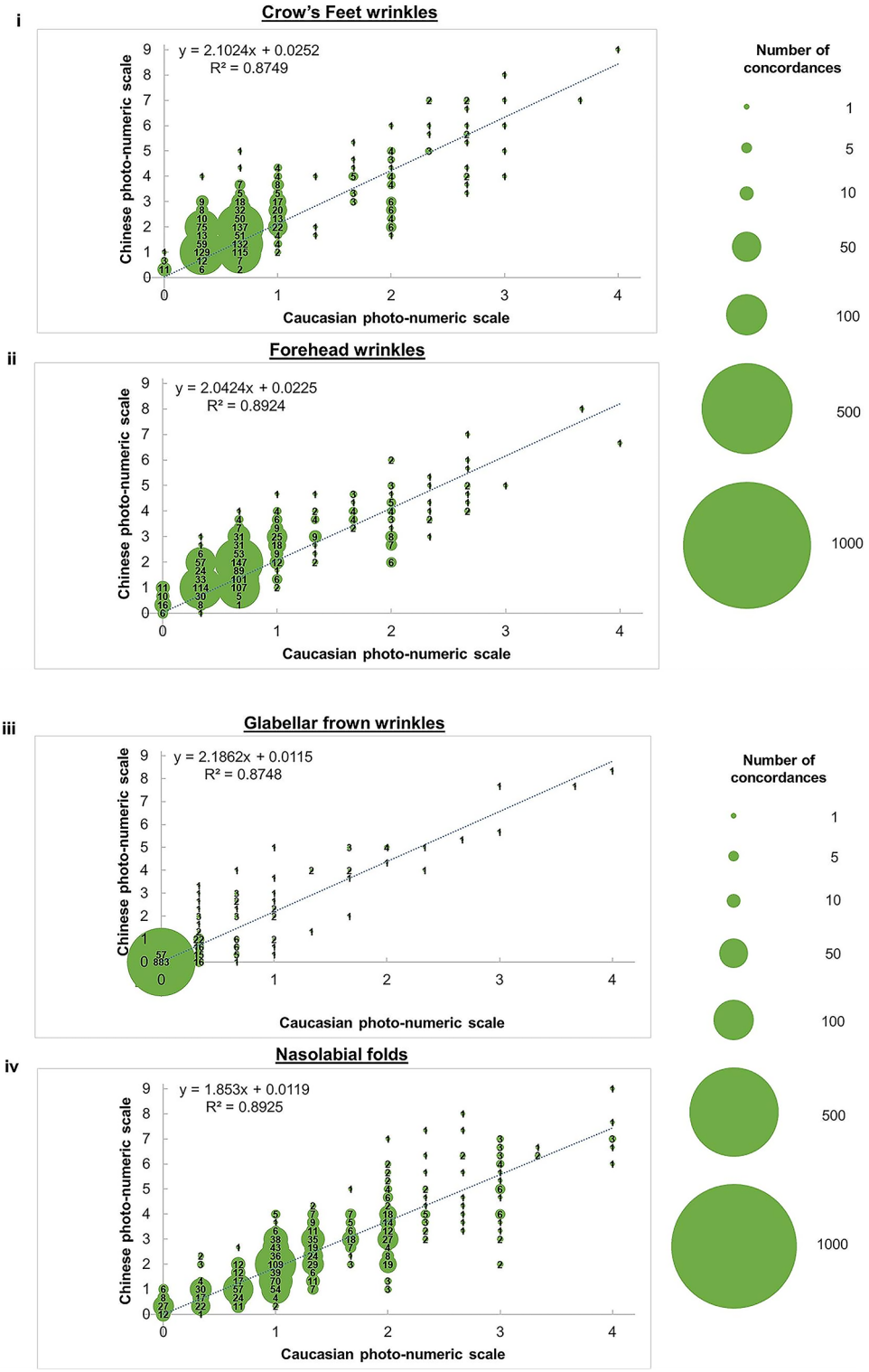

Fig. 1. **i**, **ii** Bubble plots. Compare the Caucasian photo-numeric scale and the Chinese at-rest photo-numeric scale for the same phenotype (**i**) Crow's Feet wrinkles, and (**ii**) forehead wrinkles. Means scores calculated from three assessors on the Chinese photo-numeric scale were plotted against mean scores calculated from the same three assessors on the Caucasian photo-numeric scale. Larger circles indicate greater concordance between the two scales. Numbers in the circles are the number of concordances. The sample size of each plot is 1,081 participants. (**iii**), (**iv**) Bubble plots. Compare the Caucasian photo-numeric scale and the Chinese at-rest photo-numeric scale for the same phenotype (**iii**) glabellar frown wrinkles, and (**iv**) nasolabial folds. Means scores calculated from three assessors on the Chinese photo-numeric scale were plotted against mean scores calculated from the same three assessors on the Caucasian photo-numeric scale. Larger circles indicate greater concordance between the two scales. Numbers in the circles are the number of concordances. The sample size of each plot is 1,081 participants.

Next, we look at glabellar frown wrinkles as assessed through the photo-numeric scale by Honeck *et al.*[14]. Participants aged 21 to 25 years old (p-value = 1.28×10^{-4}), participants aged 26 to 30 years old (p-value = 6.28×10^{-5}), participants aged 31 to 35 years old (p-value = 6.00×10^{-6}), participants aged 36 to 40 years old (p-value = 1.35×10^{-8}), and participants aged above 40 years old (p-value = 6.81×10^{-79}) exhibit significantly more severe glabellar frown wrinkles as compared to the reference group (Figure S6iii). Glabellar frown wrinkles also steadily increase with age (Chi-square trend test p-value < 0.001) (Figure S6iii).

Lastly, we analysed nasolabial folds using a photo-numeric scale by Narins *et al.*[15]. Participants aged 21 to 25 years old (p-value = 6.45×10^{-15}), participants aged 26 to 30 years old (p-value = 1.98×10^{-9}), participants aged 31 to 35 years old (p-value = 2.86×10^{-16}), participants aged 36 to 40 years old (p-value = 2.91×10^{-23}), and participants aged above 40 years old (p-value = 1.71×10^{-92}) exhibit significantly more severe nasolabial folds as compared to the reference group (Figure S6iv). Nasolabial folds also steadily increase with age (Chi-square trend test p-value < 0.001) (Figure S6iv).

Comparisons between the Caucasian scales and Chinese scales

Concordance between Chinese scales

Currently, there are not many photo-numeric scales evaluating facial wrinkles which are designed specifically for Chinese skin. As of writing, there are only six photo-numeric scales designed to evaluate four kinds of facial wrinkles on Chinese skin. The six photo-numeric scales are (1) a scale for Crow's Feet wrinkles at-rest[5], (2) a scale for Crow's Feet wrinkles in-motion[5], (3) a scale for forehead wrinkles at-rest[5], (4) a scale for forehead wrinkles in-motion[5], (5) a scale for glabellar frown wrinkles[5], and (6) a scale for nasolabial folds[5].

Here, we focus on the two Chinese scales which evaluate Crow's Feet wrinkles at-rest and in-motion. Both Chinese scales behave similarly (Figure S1, Tables S2, S3 and S7) and we selected the Chinese at-rest Crow's Feet wrinkles scale for further study.

Similarly, there are two Chinese scales for forehead wrinkles (Figure S1, Tables S2, S3 and S7): forehead wrinkles at-rest and in-motion; we selected the Chinese at-rest forehead wrinkles scale for further study because it has a simple relationship with the Caucasian forehead wrinkles scale.

Facial features more common in the Chinese as opposed to Caucasians include epicanthal folds, a narrower palpebral fissure, a less apparent upper eyelid crease, a wider intercanthal distance, an upward lateral canthal tilt[17], floating forehead or the calves, and specific landmarks in the periorbicular area. These facial features could also be evaluated in the future if they are found to undergo changes as one's age increases.

Level of agreement among assessors

To test the agreement between two gradings of the same image by the same assessor, the weighted kappa for intra-assessor was calculated for each of the three assessors and for all the Caucasian scales and Chinese scales. Grades given by each of the three assessors have high inter-assessor consistency (Tables S4 and S5). Weighted kappa coefficients (κ_w) for intra-assessor range from strong (0.686) to very strong (0.992) (Table S6).

Concordance and level of agreement between the Caucasian scales and Asian scales

Using the mean evaluation of three assessors, we found that Caucasian scales and Chinese scales are moderate-to-very-strongly concordant in evaluating facial wrinkles in the ethnic Chinese. Caucasian scales and Chinese scales are moderately concordant in evaluating Crow's Feet wrinkles (Spearman $\rho = 0.53 \pm 0.03$, p-value = 1.00×10^{-80}), strongly concordant in evaluating both forehead wrinkles (Spearman $\rho = 0.66 \pm 0.02$, p-value = 2.21×10^{-134}) and glabellar frown wrinkles (Spearman $\rho = 0.78 \pm 0.03$, p-value = 9.24×10^{-224}), and very strongly concordant in evaluating nasolabial folds (Spearman $\rho = 0.80 \pm 0.01$, p-value = 8.50×10^{-245}) (Table 2).

The level of agreement between Caucasian scales and Chinese scales is moderately high (κ ranges from 0.40 to 0.49) for all four facial wrinkling phenotypes. Caucasian scales and Chinese scales have moderate levels of inter-scale agreement in evaluating all four facial wrinkling phenotypes: Crow's Feet wrinkles ($\kappa = 0.42 \pm 0.26$, p-value = 2.28×10^{-7}), forehead wrinkles ($\kappa = 0.40 \pm 0.32$, p-value = 1.83×10^{-8}), glabellar frown wrinkles ($\kappa = 0.40 \pm 0.02$, p-value = 4.17×10^{-94}), and nasolabial folds ($\kappa = 0.49 \pm 0.20$, p-value = 2.33×10^{-3}) (Table 2).

Goodness of fit between the Caucasian scales and Asian scales

The Caucasian scales are 5-point scales (Grades 0–4) in which the severity of the phenotype is clearly reinforced with each photo. In contrast, the Chinese scales are 10-point scales (Grades 0–9) with a smoother gradation in the severity of the phenotypes presented (Fig. 1). The Chinese scales therefore allow facial wrinkling phenotypes to be studied at a greater resolution.

We determined that the relationship between the Caucasian scale and the Chinese scale largely follows a 1:2 ratio in which an increase of one grade on the Caucasian scale approximately corresponds to an increase of two successive grades on the Chinese scale. For instance, most participants graded as Grade 0 (out of 4) on the Caucasian scale are graded as either Grade 0 or 1 (out of 9) on the Chinese scale. In detail, the Chinese scale to Caucasian scale ratio is approximately 1:2.1024 ($R^2 = 0.8749$) for Crow's Feet wrinkles, 1:2.0424 ($R^2 = 0.8924$) for forehead wrinkles, 1:2.1862 ($R^2 = 0.8748$) for glabellar frown wrinkles, and 1:1.853 ($R^2 = 0.8925$) for nasolabial folds (Fig. 1). The coefficient of determination (R^2) for the equation for the goodness of fit is very strong (0.8748 to 0.8925) for all four facial wrinkling phenotypes. Thus, the Chinese scale enables a stricter characterisation of wrinkle progression on the ethnic Chinese skin.

Area under curve (AUC) between the Caucasian scales and Asian scales

When the Caucasian scale is taken as the gold standard, the Chinese scale has a strong predictive accuracy in discerning the presence of Crow's Feet wrinkles (AUC = 0.82 ± 0.07) and forehead wrinkles (AUC = 0.84 ± 0.09). Glabellar frown wrinkles can be very strongly discerned (AUC = 0.90 ± 0.10) while the presence or absence of

nasolabial folds are fairly strongly distinguishable apart from each other (AUC = 0.79 ± 0.05) (Table 2). In other words, when the Caucasian scale is taken as the gold standard, the Chinese scale also correctly identifies the presence or absence of a facial wrinkle in most instances. This evidence suggest that Caucasian scales are suitable assessment tools to evaluate Crow's Feet wrinkles, forehead wrinkles, glabellar frown wrinkles, and nasolabial folds on the ethnic Chinese facial skin.

In comparison, when the Chinese scale is taken as the gold standard, the Caucasian scale is less capable in distinguishing between the presence and absence of Crow's Feet wrinkles (AUC = 0.61 ± 0.06) and forehead wrinkles (AUC = 0.62 ± 0.04). Glabellar frown wrinkles can still be strongly discerned (AUC = 0.84 ± 0.12) and the presence or absence of nasolabial folds can still be distinguishable apart fairly well (AUC = 0.77 ± 0.04) (Table 2).

Discussion

Participant demographics

Singapore being a racially diverse country, enables the recruitment of participants of diverse races. However, an unequal proportion and sampling probability exists among some races, with only ethnic Chinese remaining as the major racial group with a large enough sample size. We therefore focused only on the ethnic Chinese, from which, we obtained a subset of 1,081 participants to analyse in our current study. Future analyses and description of skin ageing in other races (e.g., Malays and Indians) and dual-race participants will be possible when the number of participants of the above races grows sufficiently large through progressive, annual recruitment drives to expand the Singapore/Malaysia Cross-sectional Genetics Epidemiology Study (SMCGES).

The participants assessed in the current study largely have no or slight wrinkles (Fig. 1 and S1, Table S2). Hence, the correlation between the Caucasian and Chinese photo-numeric scales in our current assessment should be interpreted in the context of a group whose facial wrinkles are just beginning to emerge.

A facial wrinkle scale has two broad objectives. First, it must be able to detect the presence or absence of a wrinkle. Secondly, the series of photos used to construct the scale should capture a smooth and gradual increase in the severity of the facial wrinkle evaluated by the scale.

As most of the participants in this study have no wrinkles or have barely visible wrinkles, we seized the opportunity to rigorously test the ability of the Caucasian and Chinese facial wrinkle scales to meet the first objective. Our study also partially satisfied the second objective as we have a sizable number of participants with moderately severe wrinkles (Fig. 1 and S1, Table S2). Ongoing efforts to expand the SMCGES through annual recruitment drives will add more participants with severe wrinkles, enabling us to confirm our findings for higher wrinkle grades. Future work to increase the number and diversity of assessors could also allow us to better confirm our findings.

Comparing Caucasian scales and Chinese scales

Many researchers have attempted to understand how age affects the skin by designing ways to quantify skin ageing, one of which is a photo-numeric scale. Photo-numeric scales are considerably better than previously available tools for evaluating skin ageing[18]. Thus, photo-numeric scales became an area of intense research interest[19–22]. This eventually led to the proliferation of many scales developed and validated on Caucasian skin, with each scale specialising in a specific type of wrinkle or a specific form of skin ageing[13,15,20,23]. Meanwhile, multiple publications also began reporting that differences exist in skin ageing progression across races[3]. In our recent meta-analysis, we found that these differences across races are significant[2]. The timing that the collective community realises that people of different races have inherently different skins roughly coincided with the timing that ethnic-specific scales (including Chinese ones) to describe skin ageing in different ethnic groups start gaining research interest[18,24–26].

The main objective of this study is to determine whether the inherent differences between Chinese and Caucasian skin affect the suitability of evaluating facial wrinkles on the ethnic Chinese skin with a Caucasian scale. To achieve our objective, we must address two related questions.

Firstly, we must understand how similar and different it is to evaluate the same facial wrinkle (e.g., Crow's Feet wrinkles) on the ethnic Chinese skin with a Chinese scale and a Caucasian scale.

Secondly, we must identify the motivations to create Chinese wrinkling scales. Specifically, we wonder whether the recent interest in developing Chinese scales is the outcome of an increased collective understanding that different races have different skin, or whether the interest in developing Chinese scales is because of irreconcilable differences between Caucasian and Chinese skin that made it challenging, if not impossible, to use a Caucasian scale to evaluate Chinese skin.

To figure out whether it is possible and practical to evaluate Chinese skin on a Caucasian scale, we graded many Chinese participants on both a Caucasian scale and a Chinese scale and compared the results.

We closely consulted a systematic review of clinical scales by Dobos *et al.*[27] to identify the best scales to use to evaluate the four facial wrinkling phenotypes. In this systematic review of clinical scales, the researchers studied multiple skin ageing phenotypes from the Skin Ageing Atlas by Bazin *et al.*[28]. The researchers also evaluated a 9-point forehead wrinkle scale, a 7-point glabellar frown wrinkles scale, and an 8-point nasolabial folds scale by Bazin *et al.*[28]. In total, 111 skin ageing instruments were subjected to a methodological appraisal by the research team, using parameters such as internal consistency, reliability, measurement error, content validity, structural validity, hypothesis testing, cross-cultural validity, criterion validity, and responsiveness. The scales by Bazin *et al.* 2007 were evaluated to be 'excellent' only for two parameters: content validity and criterion validity.

Dobos *et al.* (2015) also appraised the methodology of other skin ageing instruments for evaluating facial wrinkles. For instance, the 5-point upper face ageing scale by Flynn *et al.*[13] is positively appraised for its interrater reliability, intrarater reliability, hypothesis testing, and validity. The 5-point glabellar wrinkles scale[14] by Honeck *et al.* (2003) and the nasolabial folds scale by Narins *et al.*[15] are both positively rated for reliability.

Lastly, the 5-point Crow's Feet wrinkles scale[12] by Tsukahara *et al.* (2000) received positive ratings for content validity. Overall, Dobos *et al.*[27] concluded that the 5-point upper face ageing scale[13] by Flynn *et al.* (2012), the 5-point glabellar wrinkles scale[14] by Honeck *et al.* (2003), the 5-point Crow's Feet wrinkles scale[12] by Tsukahara *et al.* (2000), and the 5-point nasolabial folds scale[15] by Narins *et al.* (2012) are 'higher performing scales' that outperform the others. Thus, we conducted our facial wrinkle assessments using the upper face ageing scale by Flynn *et al.*[13], the glabellar wrinkles scale[14] by Honeck *et al.* (2003), the nasolabial folds scale by Narins *et al.*[15], and the Crow's Feet wrinkles scale[12] by Tsukahara *et al.* (2000).

We found that the Caucasian scale had strong or very strong correlations (Spearman $\rho \geq 0.66$) with the Chinese scale for forehead wrinkles, glabellar frown wrinkles, and nasolabial folds, and moderately strong correlations for Crow's Feet wrinkles (Spearman $\rho = 0.53$) (Table 2). We also found that these high correlations persist across time (strong/very strong intra-assessor grading consistency: 0.686–0.992) (Table S6), persist across assessors (moderately high inter-scalar reliability: 0.40–0.49 (Table 2)), and have a high inter-assessor intra-scalar consistency (Tables S4 and S5)). We interpret these results to mean that using either a Caucasian scale or a Chinese scale to evaluate Chinese skin gives largely the same results and these results remain valid across multiple assessors and across two time points. Although there are currently no gold standards for grading facial wrinkles, the moderately high inter-scalar reliability suggests that Caucasian scales are a valid comparative measurement to Chinese scales for evaluating ageing on the Chinese skin.

We interpret that our results are indicative that the facial wrinkles on Chinese skin can be evaluated equally well on the Caucasian scales and the Chinese scales. As the Caucasian skin is known in the literature to wrinkle more than the skin of other races, the design of the Caucasian facial wrinkle scales tends to emphasise capturing wrinkles on the face. We have shown that these high-performing Caucasian facial wrinkle scales are suitable for evaluating facial wrinkles on Chinese skin. Taking into consideration its design objectives, at first glance, the Caucasian scales could arguably fare better than the respective Chinese facial wrinkle scales in capturing subtle wrinkles on Chinese skin. However, the Chinese facial wrinkle scales made up for this apparent shortcoming by being 10-point scales (Grades 0–9) with a smoother gradation in the severity of the phenotypes presented (Fig. 1, Table S2) as opposed to the 5-point (Grades 0–4) Caucasian facial wrinkle scales. In doing so, the Chinese scales allow facial wrinkling phenotypes to be studied at a greater resolution. Overall, we found two things. Firstly, both the Caucasian scales and the Chinese scales have their own strengths. Secondly, when the Caucasian scales and the Chinese scales are used to assess the same wrinkle, both scales consistently deliver similar results.

As it has been reported in the literature that Asian races are significantly associated with decreased Crow's Feet wrinkles, forehead wrinkles, and glabellar frown wrinkles[4], we needed a reliable and validated way to report fine and barely noticeable wrinkles on the Chinese face. We found that using the Caucasian scale, most ($n = 983$, 91%) of our participants are identified to have a mean score between Grade 0 (i.e., no wrinkles) and Grade 1 (Fig. 1i, Table S2). This is consistent with the literature about Asians having decreased Crow's Feet wrinkles[4]. When the same participants were scored on the Chinese scale, we observe a wider spread in which participants were sorted into Grades 1, 2, 3, and 4 (out of 9) (Fig. 1i, Table S2). A very similar distribution can be seen for forehead wrinkles (Fig. 1ii, Table S2) and nasolabial folds (Fig. 1iv, Table S2). Most ($n = 883$, 82%) of our participants do not have glabellar frown wrinkles (Fig. 1iii, Table S2) and our observation follows the expected results in the literature about fewer glabellar frown wrinkles in Asians[4].

We can draw two conclusions from our findings. Firstly, our results are consistent with the literature that Asian skin has decreased Crow's Feet wrinkles, forehead wrinkles, and glabellar frown wrinkles. Additionally, we contribute to and strengthen the literature by showing that ethnic Chinese skin from Singapore and Malaysia also follows this pattern.

Secondly, our analyses support the idea that the inherent differences between Caucasian and Chinese skin do not compromise the suitability to use a Caucasian scale to evaluate Chinese skin. Rather, it is our increased collective understanding that different races have different skin (such as how Asian skin has decreased Crow's Feet wrinkles, forehead wrinkles, and glabellar frown wrinkles) that led to the development of Chinese scales. These Chinese scales are in turn, better designed scales to capture fine and barely noticeable wrinkles on the ethnic Chinese skin.

The same study[4] also identified that Asian races are significantly associated with decreased perioral wrinkles. At present, Chinese scales for perioral wrinkles have yet to be developed. Future work in this area can proceed when a Chinese scale for perioral wrinkles is developed. Such a scale should then be compared against an existing Caucasian scale for perioral wrinkles to investigate whether this newly developed Chinese scale for perioral wrinkles can capture fine and barely noticeable perioral wrinkles better than the Caucasian scale does on Chinese skin.

Beneficial effects of skincare routine on facial wrinkling

We have previously analysed the effects of anti-ageing interventions on skin sagging (eyebags, droopy eyelids, low eyebrow positioning)[29]. We wondered whether skincare routines bring about similar effects for facial wrinkling phenotypes. We found that practicing skincare routines have beneficial effects for Crow's Feet wrinkles, forehead wrinkles, glabellar frown wrinkles, and nasolabial folds.

Firstly, we used the Caucasian scale by Tsukahara *et al.*[12], the Chinese at-rest scale by Zhang *et al.*[5] and the Chinese in-motion scale by Zhang *et al.*[5] to evaluate Crow's Feet wrinkles. We found that among participants who practice skincare routines, Crow's Feet wrinkles do not significantly increase in severity from 26 to 40 years old (Chi-square test p-value ≥ 0.0902) (Figure S7i). In contrast, among participants who do not practice skincare routines, Crow's Feet wrinkles steadily increase in severity from 26 to 40 years old (Chi-square test p-value ≤ 0.05) (Figure S7ii). Earlier in this paper, we established that when the Caucasian scales and the Chinese scales are used to assess the same wrinkle, both scales consistently deliver similar results. Here, we also found that when Crow's Feet wrinkles are assessed using the Chinese at-rest scale, people who practice skincare routines

also do not see a significant increase in the severity of Crow's Feet wrinkles from 26 to 40 years old (Chi-square test p-value ≥ 0.0672) (Figure S7iii). In comparison, people who do not practice skincare routines have Crow's Feet wrinkles that become steadily more severe from 26 to 40 years old (Chi-square test p-value ≤ 0.05) (Figure S7iv). The same effects are seen for the Chinese in-motion scale, but in a narrower age band (26 to 35 years old) (Figures S7v-vi).

Secondly, we used the Caucasian scale by Flynn *et al.*13, the Chinese at-rest scale by Zhang *et al.*[5] and the Chinese in-motion scale by Zhang *et al.*[5] to evaluate forehead wrinkles. We found that among participants who practice skincare routines, forehead wrinkles do not significantly increase in severity from 18 to 40 years old (Chi-square test p-value ≥ 0.204) (Figure S8i). In contrast, among participants who do not practice skincare routines, forehead wrinkles steadily increase in severity from 18 to 40 years old (Chi-square test p-value < 0.001) (Figure S8ii). Like the case with Crow's Feet wrinkles, we also found that when forehead wrinkles are assessed using the Chinese at-rest scale, people who practice skincare routines also do not see a significant increase in the severity of forehead wrinkles from 18 to 40 years old (Chi-square test p-value ≥ 0.642) (Figure S8iii). In comparison, people who do not practice skincare routines have forehead wrinkles that become steadily more severe from 18 to 40 years old (Chi-square test p-value ≤ 0.05) (Figure S8iv). When assessed using the Chinese in-motion scale, forehead wrinkles among participants who practice skincare routines maintained at a constant severity level from 18 to 30 years old (Chi-square test p-value ≥ 0.0577) (Figure S8v), and maintained at a constant, higher severity level from 31 to 40 years old (Chi-square test p-value ≥ 0.0708) (Figure S8v). In contrast, forehead wrinkles among participants who do not practice skincare routines become steadily more severe from 18 to 40 years old (Chi-square test p-value ≤ 0.05) (Figure S8vi).

Thirdly, we used the Caucasian scale by Honeck *et al.*14 and the Chinese scale by Zhang *et al.*[5] to evaluate glabellar frown wrinkles. We found that while glabellar frown wrinkles do not significantly increase in severity from 26 to 40 years old for people who do not practice skincare routines (Chi-square test p-value ≥ 0.248) (Figure S9ii), this youthfulness persists longer from 18 to 40 years old for people who practice skincare routines (Chi-square test p-value ≥ 0.362) (Figure S9i). We see similar effects when glabellar frown wrinkles are graded on the Chinese photo-numeric scale for glabellar frown wrinkles, but the youthfulness persisted in a narrower age band (26 to 35 years old) (Figure S9iv). When graded on the Chinese scale for glabellar frown wrinkles, while these facial wrinkles do not significantly increase in severity from 26 to 35 years old for people who do not practice skincare routines (Chi-square test p-value ≥ 0.201) (Figure S9iv), this youthfulness persists longer from 18 to 40 years old for people who practice skincare routines (Chi-square test p-value ≥ 0.400) (Figure S9iii).

Lastly, we used the Caucasian scale by Narins *et al.*[15] and the Chinese scale by Zhang *et al.*5 to evaluate nasolabial folds. On first glance, we found that the development of nasolabial folds across different age groups, when evaluated on the Caucasian scale by Narins *et al.*[15], is similar regardless of whether one practices skincare routines (Figures S10i-ii). The benefits of practicing skincare routines become apparent when the Chinese scale by Zhang *et al.*[5] is used, as this 10-point scale allows facial wrinkling phenotypes to be studied at a greater resolution. People who practice skincare routines do not see a significant increase in the severity of nasolabial folds from 26 to 40 years old (Chi-square test p-value ≥ 0.0814) (Figure S10iii). In comparison, people who do not practice skincare routines have nasolabial folds that become steadily more severe from 18 to 40 years old (Chi-square test p-value ≤ 0.05) (Figure S10iv).

Overall, practicing skincare routines delays the onset or severity of facial wrinkling phenotypes. However, as the four facial wrinkles (Crow's Feet wrinkles, forehead wrinkles, glabellar frown wrinkles, nasolabial folds) have individual developmental timelines, the beneficial effects of practicing skincare routines on these four phenotypes become the most apparent at different age ranges. All three scales (Caucasian scale, Chinese at-rest scale, Chinese in-motion scale) agree that practicing skincare routines are associated with delays in the onset and progression of Crow's Feet wrinkles from 26 to 35 years old, delays in the onset and progression of forehead wrinkles from 18 to 30 years old, and delays in the onset and progression of forehead wrinkles from 31 to 40 years old. Both the Caucasian scale and the Chinese scale agree that practicing skincare routines is associated with an extended delay in the progression of glabellar frown wrinkles by eight years from 26 to 40 years old to 18 to 40 years old. Practicing skincare routines is also associated with delays in the onset and progression of nasolabial folds from 26 to 40 years old.

Our results show that the appropriate age range for studying some facial wrinkling phenotypes, such as Crow's Feet wrinkles can be as young as 18 to 40 years old. We also show evidence that practicing skincare routines is associated with beneficial effects even when such skincare routines are practiced in people as young as 18 years old.

A potential limitation is that the data collected does not distinguish between anti-ageing interventions and skincare routines. Instead, we have shown that taking the initiative in performing deliberate interventional acts on the skin, such as applying facial creams, have significant benefits to facial wrinkling phenotypes.

Conclusion

In conclusion, our three trained assessors independently evaluated four types of facial wrinkles on 1,081 ethnic Chinese young adults: Crow's Feet wrinkles, forehead wrinkles, glabellar frown wrinkles, and nasolabial folds. All assessors arrive at the same consensus that the Caucasian scales and Chinese scales are concordant and the level of agreement between the Caucasian scales and Chinese scales is moderately high. When tested on ethnic Chinese skin, the Caucasian scale and the Chinese scale are aligned in most instances on whether a given facial wrinkle is present or absent. We found that Caucasian scales for Crow's Feet wrinkles, forehead wrinkles, glabellar frown wrinkles, and nasolabial folds are suitable assessment tools for evaluating facial wrinkles on the ethnic Chinese skin. We contribute to the literature by showing that Crow's Feet wrinkles, forehead wrinkles, glabellar frown wrinkles, and nasolabial folds on ethnic Chinese faces can be assessed with Caucasian scales. A facial wrinkle scale must be able to detect the presence or absence of a wrinkle and must also be able to capture a

smooth and gradual increase in the severity of the same wrinkle. As the 1,081 ethnic Chinese young adults have no wrinkles or slight wrinkles, we seized the opportunity to rigorously test the ability of the facial wrinkle scales in meeting the former objective. The latter objective can be better met when more participants with moderately severe and severe wrinkles are recruited to our study through annual recruitment drives.

Currently, Chinese scales for facial wrinkles beyond the four types discussed here have yet to be developed. Caucasian scales for facial wrinkles beyond these four types will also need to be tested for their suitability to be used on ethnic Chinese skin as and when more Chinese scales are developed.

Data availability

All data used and included in this study are available from the corresponding author (F.T.C.).

Received: 13 October 2023; Accepted: 5 November 2024
Published online: 16 November 2024

References

1. Ng, J. Y. & Chew, F. T. A systematic review of skin ageing genes: gene pleiotropy and genes on the chromosomal band 16q24. 3 may drive skin ageing. *Sci. Rep.* 1–23 (2022). https://doi.org/10.1038/s41598-022-17443-1
2. Wong, Q. Y. A. & Chew, F. T. Defining skin aging and its risk factors: a systematic review and meta-analysis. *Sci .Rep.***11** (1):1–13 https://doi.org/10.1038/s41598-021-01573-z (2021).
3. Nouveau-Richard, S. et al. Skin ageing: a comparison between Chinese and European populations: a pilot study. *J. Dermatol. Sci.* **40** (3), 187–193 (2005).
4. Goodman, G. J., Armour, K. S., Kolodziejczyk, J. K., Santangelo, S. & Gallagher, C. J. Comparison of self-reported signs of facial ageing among Caucasian women in Australia versus those in the USA, the UK and Canada. *Australas J Dermatol.* **59**(2),108–17. (2017). https://doi.org/10.1111/ajd.12637
5. Zhang, J., Hou, W., Feng, S., Chen, X. & Wang, H. Classification of facial wrinkles among Chinese women. *J. Biomed. Res.* **31** (2), 108–115 (2017).
6. Wong, Q. Y. A. et al. An updated prevalence of asthma, its phenotypes, and the identification of the potential asthma risk factors among young Chinese adults recruited in Singapore. *World Allergy Organ J.* **16** (3), 100757. https://doi.org/10.1016/j.waojou.2023.100757 (2023).
7. Lim, J. J. et al. An update on the prevalence, chronicity, and severity of atopic dermatitis and the associated epidemiological risk factors in the Singapore/Malaysia Chinese young adult population: A detailed description of the Singapore/Malaysia Cross-Sectional Genetics. *World Allergy Organ. J.* **15** (12), 100722 https://doi.org/10.1016/j.waojou.2022.100722 (2022).
8. Wong, Q. Y. A. et al. Allergic rhinitis in Chinese young adults from the Singapore/Malaysia cross-sectional genetics epidemiology study (SMCGES) cohort: prevalence, patterns, and epidemiology of allergic rhinitis. *World Allergy Organ J.* **15** (10), 100704 https://doi.org/10.1016/j.waojou.2022.100704 (2022).
9. Ng, J. Y. & Chew, F. T. Comparisons between Caucasian-validated photo-numeric scales and Korean-validated photo-numeric scales for photo-ageing. Insights from the Singapore/Malaysia cross-sectional genetics epidemiology study (SMCGES) cohort. *Ski. Res. Technol.* **30** (2), 1–11 https://doi.org/10.1111/srt.13637 (2024).
10. Chen, Y. et al. Accepted article A genome-wide association study identifies novel gene associations with facial skin wrinkling and mole count in Latin Americans. *Br. J. Dermatol.* 0–2 p. https://doi.org/10.1111/bjd.20436 (2021).
11. Ayer, J. et al. A photonumeric scale for the assessment of atrophic facial photodamage. *Br. J. Dermatol.* **178** (5), 1190–1195. https://doi.org/10.1111/bjd.16331 (2018).
12. Tsukahara, K. et al. A photographic scale for the assessment of human facial wrinkles. *J. Cosmet. Sci.* **51** (2), 127–139 (2000).
13. Flynn, T. C. et al. Validated assessment scales for the upper face. *Dermatol. Surg.*. **38** (2ptII), 309–19. https://doi.org/10.1111/j.1524-4725.2011.02248.x (2012).
14. Honeck, P., Weiss, C., Sterry, W. & Rzany, B. Reproducibility of a four-point clinical severity score for glabellar frown lines. *Br. J. Dermatol.* **149** (2), 306–10. https://doi.org/10.1046/j.1365-2133.2003.05436.x (2003).
15. Narins, R. S. et al. Validated assessment scales for the lower face. *Dermatol. Surg.* **38** (2ptII), 333–42. https://doi.org/10.1111/j.1524-4725.2011.02247.x (2012).
16. Ng, J. Y., Zhou, H., Li, T. & Chew, F. T. Comparisons between wrinkles and photo-ageing detected and self-reported by the participant or identified by trained assessors reveal insights from Chinese individuals in the Singapore/Malaysia Cross-sectional Genetics Epidemiology Study (SMCGES) cohort. *J. Physiol. Anthropol.* **43** (1), 1–22. https://doi.org/10.1186/s40101-024-00361-8 (2024).
17. Kiranantawat, K., Suhk, J. H. & Nguyen, A. H. The Asian eyelid: relevant anatomy. *Semin. Plast. Surg.* **29** (3), 158–64. https://doi.org/10.1055/s-0035-1556852 (2015).
18. Bernois, A. et al. A photographic scale for the evaluation of facial skin aging in Indian women. *Eur. J. Dermatol.* **21** (5), 700–4 https://doi.org/10.1684/ejd.2011.1462(2011).
19. Guinot, C. et al. Relative contribution of intrinsic vs extrinsic factors to skin aging as determined by a validated skin age score. *Arch. Dermatol.* **138** (11) https://doi.org/10.1001/archderm.138.11.1454 (2002).
20. Rzany, B. et al. Validated composite assessment scales for the global face. *Dermatol. Surg.* **38** (2 PART 2), 294–308 (2012).
21. Allerhand, M. et al. Skin ageing and oxidative stress in a narrow-age cohort of older adults. *Eur. Geriatr. Med.* **2** (3), 140–4. https://doi.org/10.1016/j.eurger.2011.04.009 (2011).
22. Vierkötter, A. et al. A novel, validated score to simultaneously assess and differentiate between intrinsic and extrinsic skin ageing. *J. Dermatol. Sci.* **53** (3), 207–11. https://doi.org/10.1016/j.jdermsci.2008.10.001 (2009).
23. Sattler, G. et al. Validated assessment scale for neck volume. *Dermatol. Surg.* **38** (2ptII), 343–50 https://doi.org/10.1111/j.1524-4725.2011.02253.x (2012).
24. Tsukahara, K. et al. Determination of age-related changes in the morphological structure (sagging) of the human cheek using a photonumeric scale and three-dimensional surface parameters. *Int. J. Cosmet. Sci.* **22** (4), 247–58. https://doi.org/10.1046/j.1467-2494.2000.00033.x (2000).
25. Sen, S. et al. A clinical rating scale for the assessment of facial aging in Indian population. *Indian. J. Dermatol. Venereol. Leprol.* **82** (2), 151–61 https://doi.org/10.4103/0378-6323.174369 (2016).
26. Jin Ho, C. et al. Cutaneous photodamage in koreans. *J. Dermatol.* **28** (11), 614–616 (2001).
27. Dobos, G., Lichterfeld, A., Blume-Peytavi, U. & Kottner, J. Evaluation of skin ageing: a systematic review of clinical scales. *Br. J. Dermatol.* **172** (5), 1249–61 https://doi.org/10.1111/bjd.13509 (2015).
28. Bazin, R. & Doublet, É. Skin Aging Atlas Volume 1: Caucasian Type. (2007).
29. Ng, J. Y. & Chew, F. T. Comparisons between eyebags, droopy eyelids, and eyebrow positioning identified by photo-numeric scales or identified by written descriptive scales: insights from the Singapore/Malaysia cross-sectional genetics epidemiology study (SMCGES) cohort. *Ski. Res. Technol.* **30** (2), 1–12 https://doi.org/10.1111/srt.13620 (2024).

Acknowledgements
All validated photo-numeric scales used in the analysis were collected from publications in Table S1. The photo-numeric scales used in the current assessment are as follows: Forehead wrinkles are assessed using the forehead wrinkles at-rest scale by Zhang *et al.*[5], the forehead wrinkles in-motion scale by Zhang *et al.*[5], and the upper face ageing scale by Flynn *et al.*[13]. Glabellar frown wrinkles are assessed using the glabellar frown wrinkles scale by Zhang *et al.*[5] and the glabellar wrinkles scale by Honeck *et al.*[14]. Crow's Feet wrinkles are assessed using the Crow's Feet wrinkles at-rest scale by Zhang *et al.*[5], the Crow's Feet wrinkles in-motion scale by Zhang *et al.*[5], and the Crow's Feet wrinkles scale by Tsukahara *et al.*[12]. Nasolabial folds are assessed using the nasolabial folds scale by Zhang *et al.*[5], and the nasolabial folds scale by Narins *et al.*[15]. Permission to use the scales by Flynn *et al.*[13] and Zhang *et al.*[5] have been obtained. As the glabellar wrinkles scale by Honeck *et al.*[14], the nasolabial folds scale by Narins *et al.*[15], and the Crow's Feet wrinkles scale by Tsukahara *et al.*[12] are not under any copyright, permission to use them was not required. We extend our sincerest gratitude to all participants for their contributions to this study.

Author contributions
F.T.C. conceived and supervised the current review study. J.Y.N. conducted the literature review, analysed and interpreted the data, and wrote the manuscript. H.Y.Z and T.Q.L analysed the data. J.Y.N, H.Y.Z., and T.Q.L. assisted in recruiting study participants and data collation. All authors read and approved the final manuscript.

Funding
F.T.C. received grants from the National University of Singapore (N-154-000-038-001 (E-154-00-0017-01); C141-000-077-001 (E-141-00-0096-01)), Singapore Ministry of Education Academic Research Fund (R-154-000-191-112; R-154-000-404-112; R-154-000-553-112; R-154-000-565-112; R-154-000-630-112; R-154-000-A08-592; R-154-000-A27-597; R-154-000-A91-592; R-154-000-A95-592; R154-000-B99-114), Biomedical Research Council (BMRC) (Singapore) (BMRC/01/1/21/18/077; BMRC/04/1/21/19/315; BMRC/APG2013/108), Singapore Immunology Network (SIgN-06-006; SIgN-08-020), National Medical Research Council (NMRC) (Singapore) (NMRC/1150/2008; OFIRG20nov-0033; MOH-001636 (OFLCG23may-0038, A-8002641-00-00)), National Research Foundation (NRF) (Singapore) (NRF-MP-2020-0004), Singapore Food Agency (SFA) (SFS_RND_SUFP_001_04; W22W3D0006), Singapore's Economic Development Board (EDB) (A-8002576-00-00), and the Agency for Science Technology and Research (A*STAR) (Singapore) (H17/01/a0/008; and APG2013/108). The funding agencies had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Declarations

Ethics approval and consent
This study was conducted in accordance with the principles of the Declaration of Helsinki and Good Clinical Practices, and in compliance with local regulatory requirements. The cross-sectional studies in Singapore were conducted on the National University of Singapore (NUS) campus annually between 2005 and 2023 with the approval of the Institutional Review Board (Reference Code: NUS-07-023, NUS-09-256, NUS-10-445, NUS-13-075, NUS-14-150, and NUS-18-036) and by the Helsinki declaration, of which, participants between 2011 and 2022 participated in the current skin ageing study (Reference Code: NUS-2020-495). The cross-sectional studies in Malaysia were held in Universiti Tunku Abdul Rahman (UTAR) Campus and Sunway University. Ethical approval was granted from the Scientific and Ethical Review Committee of UTAR (Reference Code: U/SERC/03/2016) and the Sunway University Research Ethics Committee (Reference Code: SUREC 2019/029 and SUREC 2022/049). Before the data collection, all participants involved signed an informed consent form.

Consent for publication
All authors have read and consented to the publication of this manuscript.

Competing interests
F.T.C reports grants from the National University of Singapore, Singapore Ministry of Education Academic Research Fund, Singapore Immunology Network, National Medical Research Council (NMRC) (Singapore), Biomedical Research Council (BMRC) (Singapore), National Research Foundation (NRF) (Singapore), Singapore Food Agency (SFA), Singapore's Economic Development Board (EDB), and the Agency for Science Technology and Research (A*STAR) (Singapore), during the conduct of the study; and consulting fees from Sime Darby Technology Centre; First Resources Ltd; Genting Plantation, Olam International, Musim Mas, and Syngenta Crop Protection, outside the submitted work. The other authors declare no other competing interests. This research is supported by the National Research Foundation Singapore under its Open Fund-Large Collaborative Grant (MOH-001636) (A-8002641-00-00) and administered by the Singapore Ministry of Health's National Medical Research Council.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-78945-8.

Correspondence and requests for materials should be addressed to F.T.C.