

Exploration of Anti-Redness Short-Term Evaluation Models for Cosmetics in Asian Females: Anti-Redness Cosmetics Evaluation Exploration

Qianxin Zhu¹, Daniela Zoellner², Yanwen Jiang³, Yuchen Qiu³, Ping Xu⁴, Yuanyuan Diao^{1,*}

¹ Beiersdorf Innovation Center Shanghai, Shanghai, China

² Research and Development, Beiersdorf AG, Hamburg, Germany

³ Shanghai China-Norm Quality Technical Service Co., Ltd., Shanghai, China

⁴ Department of Dermatology, Shuguang Hospital Affiliated to Shanghai University of Traditional Chinese Medicine, Shanghai, China

***Correspondence:** Yuanyuan Diao, Beiersdorf Innovation Center Shanghai, Shanghai 200233, China.

E-Mail: Harbor.Diao@Beiersdorf.com

Abstract

Objective: This study aimed to characterize two short-term erythema induction models for cosmetic anti-redness testing in Asian females.

Methods: In this single-blind monocentric study, 78 healthy Asian females aged 18–60 years were enrolled. Stage 1 characterized glycolic acid ([GA] 50% on face) and tape stripping ([TS] forearm) models. Stage 2 used a reference cosmetic product solely to assess model detectability, not to confirm product efficacy. Evaluations included 5-point clinical grading, erythema index, a* value, and imaging analysis; GA induction was performed under dermatologist supervision.

Results: In Stage 1, both models induced significant erythema versus untreated controls. In Stage 2, the GA model detected reference product-related redness reduction on clinical grading (4 h, $P = 0.009$; 24 h, $P = 0.020$) and VISIA-CR red-area analysis (1 h, 2 h, and 24 h, $P < 0.05$). The TS model showed numerical improvement on clinical grading at 2–4 h without significant treated versus untreated differences. No protocol-defined adverse events (AEs) were recorded through 24 h.

Conclusions: The GA model suits chemical irritation-related facial erythema; the TS model suits barrier-disruption erythema on body skin. Clinical grading and imaging were most sensitive for GA; TS was mainly supported by grading at 2–4 h. Both models provide a practical framework for matching anti-redness claims to erythema phenotype rather than definitive validation. Delayed irritation and post-inflammatory hyperpigmentation (PIH) beyond 24 h cannot be excluded.

Keywords: skin redness; glycolic acid induction; tape stripping; clinical grading; imaging analysis

Introduction

Acute skin redness, also known as “flush”, is an acute

vasodilation reaction that can be caused by various factors. In this manuscript, “skin redness” denotes the visually perceived cutaneous red appearance; “erythema”

Received: 13 August 2026; **Accepted:** 22 August 2026

DOI: 10.2738/SKIN.2026.0024

© The Author(s) 2026. This article is published by Higher Education Press on behalf of People's Medical Publishing House. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

denotes clinically or instrumentally assessed inflammatory redness with a physiological basis; and the “redness/erythema score” refers to the predefined 5-point clinical grading scale. Instrumental a^* value and erythema index (EI) are quantitative correlates of erythema intensity. External factors such as heat, cold, or sun exposure can accelerate microcirculation in the skin, and external irritants can weaken the skin barrier^[1]. Internal factors like a hyper-reactive skin, hormones, emotions, consumption of coffee or spices, and heredity (vasoreactive skin) also play a role^[1-3]. Skin aging makes the skin more sensitive to redness starting from the age of 25^[4]. Moreover, with the popularity of beauty treatments, some special procedures such as laser therapy, chemical peels, and intensive light therapy can also disrupt the skin barrier or inflame the skin, which may result in increased sensitivity and redness as a side effect^[5].

Redness can have a multi-faceted impact on individual's normal life, influencing social interactions, emotional well-being^[5], and daily life quality with physical discomfort like burning or itching^[6]. If acute redness were not relieved by suitable treatment, it may cause persistent facial redness and even chronic conditions like rosacea. Thus, managing redness often involves a combination of pharmacologic treatments and special cosmetic products to avoid irritation and provide symptom relief^[6].

In the field of cosmetics, to verify the acute anti-redness, soothing, and repairing efficacy of cosmetic products, certain induction methods are used to simulate different skin irritation reactions. These induction methods include mechanical induction (tape stripping [TS], shaving), chemical agent induction (sodium lauryl sulfate [SLS] and glycolic acid [GA] induction)^[7,8], and thermal induction. These methods can simulate different external redness cause factors and induce various degrees of redness to the treated area, but the induction results are not always consistent.

TS is usually used as the main method to induce skin barrier function disruption. Relevant research has revealed that the process of repeated TS results in a transient impairment of the skin's barrier function. This impairment initiates a cascade of histological alterations, notably the hyperplasia of epidermal cells and the infiltration of inflammatory cells. Such findings underscore the impact of TS on the skin's structural integrity and its subsequent

inflammatory response^[9]. Additionally, this method has been widely applied in some skin diseases, such as psoriasis^[10], atopic dermatitis^[11]. It can effectively simulate the skin barrier damage caused by the disease for comparative efficacy or related research on barrier recovery. In the realm of healthy skin, this method is more widely applied in the study of product formulations percutaneous absorption^[12]. We also found that relevant articles have used this method for the study of skin barrier function repair, in which the main comparison was made on the transepidermal water loss rate, and observations on skin erythema were also included, but the observations were made at relatively longer time points (3 days, 5 days)^[13]. In general, this technique is frequently employed to mimic the disruption of the skin barrier and monitor the subsequent restoration processes, particularly focusing on the transepidermal water loss rate as a key metric. Nevertheless, research that examines the resolution of immediate skin redness and erythema is comparatively scarce. To this end, leveraging the mechanism whereby TS induces skin epithelial cell shedding, which in turn stimulates the skin, disrupts the stratum corneum barrier, and triggers inflammatory responses, our study delves deeper into the practical application of this model. Specifically, it examines the model's utility in inducing acute redness to assess the soothing and reparative properties of various products.

In the chemical agent induction model with SLS^[7,8,14,15], the research on the induction of skin redness by chemical agents is well-founded in general applications. The most common evaluation method for redness recovery is the use of occlusive patches on the back or torso skin. It usually requires 24 h of occlusion before observing the success of induction. This method is relatively convenient, fast, and indeed a high-throughput method for formula screening. However, considering the effectiveness claims for future marketing of products, it is necessary to evaluate the acute anti-redness and soothing effects of products in a manner that is closer to the actual application sites and usage methods of consumers. Cumulative SLS irritation models have also been used to characterize barrier disruption and repair kinetics^[16].

To this end, this exploratory study aimed to (1) characterize two short-term erythema induction models (GA and TS), (2) identify the most sensitive evaluation endpoints

for each model, and (3) examine model consistency and detectability by using a reference product intervention in Stage 2. The primary objective was model exploration and characterization; the test product was used only as a reference intervention to determine whether the models could detect short-term anti-redness changes, not to confirm the efficacy of that specific product.

Materials and methods

Subject population

A total of 78 healthy Asian female subjects from Shanghai completed their assigned study protocol across two sequential stages (48 in Stage 1; 30 in Stage 2). Only female participants were included because facial GA induction and consumer anti-redness claims are most frequently evaluated in women in cosmetic practice, and sex-related differences in skin structure, barrier function, and vascular reactivity may confound initial model-establishment studies if both sexes were combined. Fitzpatrick skin phototypes were not systematically recorded in this exploratory study; therefore, phototype distribution cannot be reported, and this limitation is discussed below. Demographic characteristics of the final analytic cohorts are summarized in Table 1.

Table 1: Demographic characteristics of analyzed subjects.

	Stage 1	Stage 2
<i>N</i>	48	30
Age distribution	Mean ($\pm$ SD): 46.52 $\pm$ 7.27 y.o.	Mean ($\pm$ SD): 46.03 $\pm$ 9.05 y.o.
	Max: 60 y.o.	Max: 60 y.o.
	Min: 28 y.o.	Min: 20 y.o.

Abbreviations: SD, standard deviation; y.o., years old.

Participant flow

Stage 1 and Stage 2 enrolled separate subject cohorts. Stage 1 evaluated the GA and TS induction models in two independent, non-overlapping groups.

In the GA part, 27 subjects underwent facial GA induction after preliminary screening. Two subjects were excluded because post-induction facial erythema did not reach

$\geq$ grade 2 on the 5-point scale, leaving 25 subjects included in the GA analyses.

In the TS part, 49 subjects underwent forearm TS induction. Twenty-six subjects were excluded for the same reason, leaving 23 subjects included in the TS analyses.

In Stage 2, 42 subjects underwent bilateral induction screening after providing written informed consent and meeting preliminary eligibility criteria. Ten subjects were excluded because post-induction erythema did not meet the predefined endpoint ($\geq$ grade 2 and higher than the pre-induction score). The remaining 32 subjects entered the full study protocol. Two subjects subsequently withdrew for personal reasons, yielding 30 subjects in the final Stage 2 analyses. The same 30 subjects completed both the GA and TS assessments in Stage 2.

Missing data were handled by available-case analysis without imputation. No formal a priori sample-size calculation was performed; sample sizes were chosen for exploratory model characterization and feasibility assessment.

The entire study followed ethical principles. The two stages were conducted after undergoing ethical review and receiving approval with approval numbers of Frontier Science Ethics Committee (FSEC) 2023-124-01 and 2023-123-01 respectively, thus ensuring the ethical compliance of the study. All study details were explained and written informed consent was obtained before enrollment.

Study procedure

In both stages, subjects first completed written informed consent and portrait-use authorization, underwent preliminary inclusion/exclusion screening, and then received standardized facial cleansing and equilibration before GA or TS induction. Post-induction reassessment confirmed that erythema met the predefined endpoint ($\geq$ grade 2 on the 5-point scale). Eligible subjects then underwent clinical, instrumental, and imaging assessments at the scheduled time points; Stage 2 additionally included reference-product application as described below.

In Stage 1, two groups were designed for GA and TS induction model. This stage consists of one visit, which

includes baseline evaluation of clinical grading, instrumental measurement, and image capture, followed by the model induction procedure. Clinical grading, instrumental measurements, and image captures were then conducted at 10 min, 30 min, 1 h, 2 h, and 4 h after induction.

In Stage 2, a reference cosmetic product containing Symstive™, hyaluronic acid, and panthenol^[17] was applied after induction to examine whether the models could detect short-term anti-redness changes; observation was extended to 24 h^[18]. The reference product was applied at about 0.3 g on the treated half-face and 2 mg/cm² on the treated forearm immediately after post-inducing assessments and was reapplied on-site before the 24 h assessment using the same method. The contralateral side served as a blank control. In Stage 1, the GA-treated half-face and TS one side of arm were assigned according to a computer-generated randomization list. In Stage 2, the side receiving the reference product was likewise assigned according to a computer-generated randomization list (after bilateral induction). During study conduct, only the staff who applied the induction agent and/or the reference product were aware of treatment allocation. Clinical graders, instrumental technicians, image analysts, and data analysts remained blinded to the randomization assignment throughout evaluation and analysis (single-blind design with respect to assessors and analysts). Subjects could not be fully blinded because the contralateral side remained a blank control.

During GA application, a dermatologist provided continuous oversight. Upon occurrence of intolerance, subjects were guided to use a neutralizing solution under dermatologist supervision. Monitoring continued until symptoms resolved, or until subjects sought medical care, with ongoing follow-up thereafter. Detailed safety procedures for the GA model are described in the “Glycolic acid induction: safety procedures” subsection. The randomization list was generated before study start and kept by staff not involved in clinical grading or instrumental assessment; allocation was concealed until the assigned side was prepared for product application.

Glycolic acid induction: safety procedures

Because 50% GA was applied to the face, standardized safety procedures were implemented under continuous

dermatologist supervision. Before induction, subjects were screened for contraindications to facial acid application, including active facial dermatoses, recent chemical peel or laser procedures, known hypersensitivity to alpha-hydroxy acids, and pregnancy or lactation.

During induction, 0.2–0.3 g of commercially available 50% GA was applied to a limited facial area (half-face nasolabial fold) for approximately 8–10 min until clinical erythema reached $\geq$ grade 2, as assessed by the same dermatologist. The potential of hydrogen (pH) of the formulation was not recorded in the study file; the commercial preparation has since been discontinued and is no longer available for retrospective pH measurement, so pH cannot be reported. Application was stopped immediately if subjects reported intolerable burning/stinging, if clinical erythema exceeded grade 4 on the 5-point scale, if vesicles or erosions appeared, or at the dermatologist’s discretion. The matched neutralizing solution supplied with the same commercial product was applied as needed under medical supervision.

After induction, subjects were instructed to avoid additional exfoliating or irritating products on the test area, minimize sun exposure, and use gentle cleansing and basic moisturization as recommended by the dermatologist until local reactions resolved. All adverse events, prolonged irritation, discomfort, and any signs suggestive of post-inflammatory hyperpigmentation (PIH) were recorded at each visit through 24 h (Stage 2) and during follow-up when applicable. Subjective discomfort was monitored by dermatologist inquiry during procedures but was not formally graded. Stage 2 included a 24 h follow-up visit with on-site observation.

Stopping criteria included intolerable subjective symptoms despite neutralization, clinical erythema grade $\geq$ 4 persisting with worsening trend, blistering, suspected allergic reaction, or any event requiring medical referral. No induction was continued once stopping criteria were met.

Clinical grading

The skin sign involved in this study is redness/erythema, which is graded by verified expert before and after model induction at various times (Table 2) using a 5-point rating scale (grade 0 means clear, grade 1 means very slight,

Table 2: Evaluation timepoints & test areas.

Stage 1			
Model	N	Test areas	Evaluation timepoint
Glycolic acid induction model	25	Face	- T0 (baseline which is before induction) - Timm (10-min after induction) - T30 min (30-min after induction) - T1h (1-h after induction) - T2h (2-h after induction) - T4h (4-h after induction)
Tape stripping induction model	23	Forearms	- T0 (baseline which is before induction) - Timm (10-min after induction) - T30 min (30-min after induction) - T1h (1-h after induction) - T2h (2-h after induction) - T4h (4-h after induction)
Stage 2			
Model	N	Test areas	Evaluation timepoint
Glycolic acid model with product intervention	30	Face	- T01 (baseline which is before induction) - T02 (10-min after induction) - T30 min (30-min after product single application) - T1h (1-h after product single application) - T2h (2-h after product single application) - T4h (4-h after product single application) - T24h (24-h after product repeat application)
Tape stripping model with product intervention	30	Forearm	- T01 (baseline which is before induction) - T02 (10-min after induction) - T30 min (30-min after product single application) - T1h (1-h after product single application) - T2h (2-h after product single application) - T4h (4-h after product single application) - T24h (24-h after product repeat application)

grade 2 means slight, grade 3 means moderate, grade 4 means severe) (Fig. 1). The higher the score, the redder it gets^[19,20].

Instrumental measurement

The skin redness (a^* value) is measured by Spectrophotometer CM-26dG (Konica Minolta, Japan) under Specular Component Excluded (SCE) mode (light mode of surface reflection removing), and the erythema index is measured by Mexameter MX18 (Courage and Khazaka, Germany)^[21,22]. All measurements were conducted by trained technician before and after model induction at various times (Fig. 2 and Table 2). The higher the value, the redder it gets. Objective quantification of erythema using colorimetric/ a^* measurements is established in photobiology and irritation research^[23].

Standard photos capture and imaging analysis

The standard facial photo is captured by VISIA-CR Version 5.0 (Canfield, America) on front, left and right sides under standard and cross-polarized light source on face^[19] and the arm photo captured by arm capture frame under the CANON EOS 6D MARK II camera, with the LED170 set of illumination at a color temperature of 5,600 K and the illuminance set to a maximum of 2,200 lux for optimal lighting effect; the shooting on the inner forearms under cross-polarized light source operated by trained technician before and after model induction at various times (Table 2). Some parameters of red area, regarding the percentage of red area on face and redness (a^* value) on face and forearms are analyzed by the image analysis program (Image-Pro Plus, Version 7.0)^[24]. The higher the value, the redder it gets.



Figure 1: Atlas of skin redness/erythema severity 5-point scale provided by Shanghai China-norm Quality Technical Service Co., Ltd.

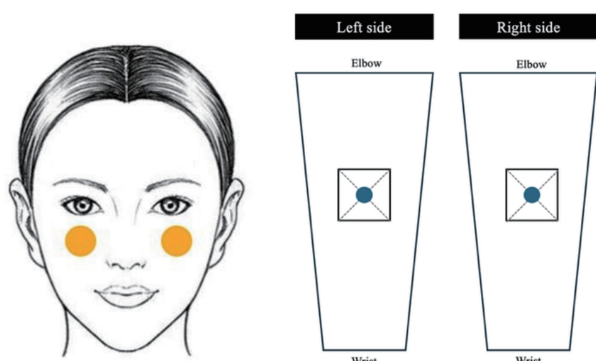


Figure 2: Schematic diagram of the instrumental measurement area, where the orange dots are for facial measurement positioning and the blue dots are for the tape stripping area as well as the measurement positioning in the designated areas of the forearm.

Redness induction procedure

This study includes two induced redness models. The GA model was designed to induce redness primarily associated with chemical peeling, whereas the TS model was designed to induce redness through physical stimulation and subsequent skin barrier disruption, and the methods applied in the two different stages vary. Endpoint assessment was performed by the same dermatologist for all subjects.

In Stage 1, for the **GA model**, 0.2–0.3 g of commercially available 50% GA^[25] was applied on one side half face (nasolabial fold) according to randomization table. The timing was recorded until subjects' half face turned

red (minimum grade 2; 5-point scale; approximately 8–10 minutes after GA application), with the primary goal of observing whether the induction can reach the preset level and monitoring the skin's natural recovery process. For the **TS model**, one side of forearm was stripped according to randomization table on a 3 cm × 3 cm area with D-Squame D100 tape; each tape was pressed for 10 s (timed) before removal to make it fit the skin more completely and evenly, then stripping was repeated until erythema ≥ grade 2 or a predefined maximum of 20 strips. In Stage 1, the exact number of strips in each subject was not archived. In Stage 1, the contralateral forearm served as an untreated control without TS, allowing us to explore the duration required for skin self-recovery after induction.

In Stage 2, the GA model was applied to both sides of the face; after induction assessments, the reference product (0.3 g) was randomly applied to one half-face, and the contralateral side remained a blank control. Similarly, both forearms underwent TS induction, and the number of strips required to reach the induction endpoint was recorded for each forearm ($n = 30$ subjects; both forearms induced): the mean was 11.43 ± 2.84 strips (range 7–18), and no subject reached the maximum of 20 strips. The reference product (2 mg/cm^2) was then randomly applied to one forearm, with the contralateral forearm remaining a blank control, to examine assay detectability rather than to confirm product efficacy.

Descriptive statistics

The descriptive statistics included the number of valid samples (N), the mean (Mean), the standard error (SE), and the mean change from the induction.

Differential statistics

All data difference tests were conducted using two-tailed tests, and the significance level was set at $\alpha = 0.05$. The difference statistics included the comparisons of time-point differences and between treatment differences. Before the analysis, the data distribution test was carried out, and the statistical analysis method was selected according to the distribution results. When data passed normality (all normality of the distributions is confirmed for the same parameter), a parametric test of paired T test was used to detect significant difference among time points; When data fails normality (if one or more normality of the distributions for the same parameters is rejected), a non-parametric test of Wilcoxon signed rank was used to detect significant difference between timepoints.

No primary endpoint or primary assessment time point was prespecified. Multiplicity was not formally adjusted; P values are presented descriptively for this exploratory analysis, and mixed-effects models were not applied, paired tests and Wilcoxon rank test were used at each

time point. In Stage 2, treated versus untreated comparisons used change from induction. Mean differences of change with 95% confidence interval (CIs) and absolute post-induction, pre-application between side levels (comparability only) are reported in the Results.

Results

Model establishment conformity in STAGE 1 —induction model exploration

Significant findings are summarized of Stage 1 in Table 3 and Table 4.

Clinical grading results of redness/erythema

GA model and **TS model** showed the effect on inducing skin redness/erythema.

Both **GA model** and **TS model**, an increased value indicates higher severity of redness/erythema on facial skin and forearm skin, it was observed that the redness/erythema severity after **GA model** and **TS model** inducing at all time-points are significantly increased ($P < 0.05$) compared to that of T0. **TS model** recovered faster within 2 h since there is no significant difference vs. baseline after 2 h but also delivers a significantly stronger effect ($P < 0.05$) on inducing redness compared with the untreated control. (Fig. 3 and Fig. 4).

Table 3: Summary of P value between time-points in Stage 1.

Attributes	Model	Comparison with T0				
		Timm	T30 min	T1h	T2h	T4h
Redness/erythema by clinical assessment	GA	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
	TS	< 0.001	< 0.001	< 0.001	0.002	0.004
a* value by CM26dG SCE mode	GA	0.993	0.345	0.367	0.950	0.379
	TS	< 0.001	< 0.001	0.083	0.076	0.040
EI by mexameter MX18	GA	0.952	0.093	0.996	0.606	0.172
	TS	< 0.001	0.024	0.864	0.739	0.893
a* value by photo analysis	GA	0.008	< 0.001	< 0.001	0.022	0.396
	TS	< 0.001	< 0.001	0.002	0.514	0.030
Area ratio of redness by photo analysis	GA	0.003	0.003	< 0.001	0.057	0.634

Abbreviations: GA, glycolic acid; TS, tape stripping; SCE, specular component excluded; CM26dG, spectrophotometer CM-26dG; EI, erythema index; a* value, redness chromaticity value.

Table 4: Summary of *P* value between treatments in Stage 1.

Attributes	Model	Comparison with untreated				
		Timm	T30 min	T1h	T2h	T4h
Redness/erythema by clinical assessment	TS	< 0.001	< 0.001	< 0.001	0.004	0.004
a* value by CM26dG SCE mode	TS	< 0.001	< 0.001	0.058	0.170	0.176
EI by mexameter MX18	TS	< 0.001	< 0.001	0.005	0.270	0.599
a* value by photo analysis	TS	< 0.001	< 0.001	< 0.001	0.810	0.392

Abbreviations: TS, tape stripping; SCE, specular component excluded; CM26dG, spectrophotometer CM-26dG; MX18, Mexameter MX18; EI, erythema index; a* value, redness chromaticity value.

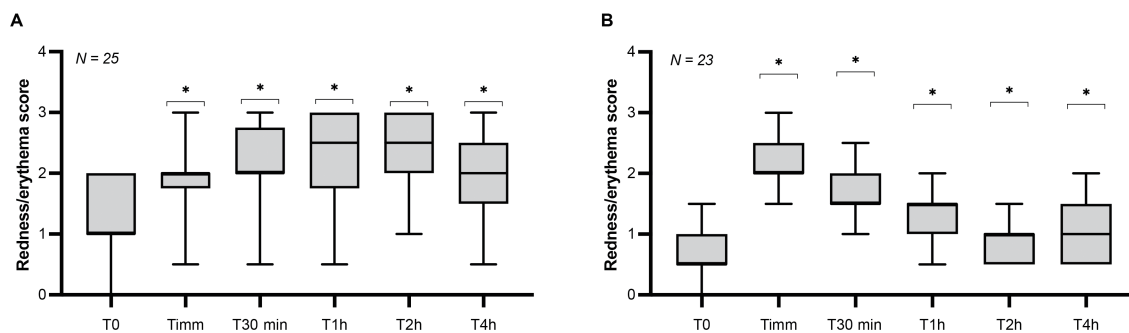


Figure 3: Clinical scoring of redness/erythema using 5-point scale refer to the details in the “Clinical grading” subsection (grade 0 means clear, grade 4 means severe).

(A) GA: time-points comparison of redness/erythema score. (B) TS: time-points comparison of redness/erythema score. On the figure, the clinical scoring of redness (mean, median, maximum and minimum) in each time-points. Subsequent times that are significantly different ($P < 0.05$ at 95% CI) from T0 are marked with an asterisk “*”. *N* means valid sample size. Abbreviations: GA, glycolic acid; TS, tape stripping. CI, confidence interval.

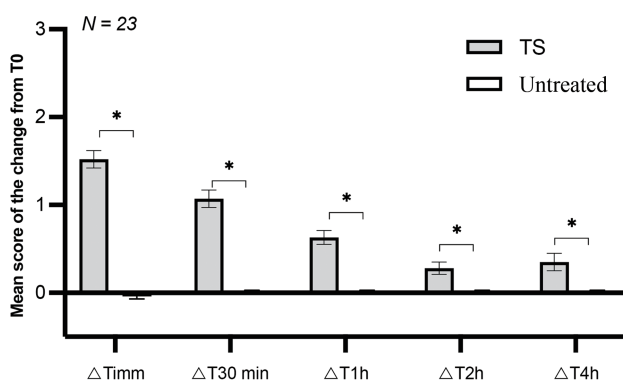


Figure 4: Clinical scoring of redness/erythema using 5-point scale refer to the details in the “Clinical grading” subsection (grade 0 means clear, grade 4 means severe).

Abbreviations: TS, tape stripping. CI, confidence interval.

On the figure, the clinical scoring of redness/erythema (mean with standard error bar) in each time-points adjusted by T0. The model showed that the difference was statistically significant ($P < 0.05$, 95% CI) compared with the untreated area; significant points are marked with an asterisk (*). *N* means valid sample size.

Instrumental results of spectrophotometer CM-26dG (a* value, SCE mode)

GA model and **TS model** showed the visual effect on inducing skin redness.

Both **GA model** and **TS model**, an increased value indicates higher value of redness (a* value) on facial skin and forearm skin, it was observed that the redness value after **GA model** inducing at Timm, T30 min, T1h and T2h compared to that of T0, but there is no significant difference. The redness value after **TS model** inducing at Timm, T30 min and T4h is significantly increased ($P < 0.05$) compared to that of T0. **TS model** delivers a significantly stronger effect ($P < 0.05$) on inducing redness at Timm and T30 min compared with the untreated control. (Fig. 5 and Fig. 6).

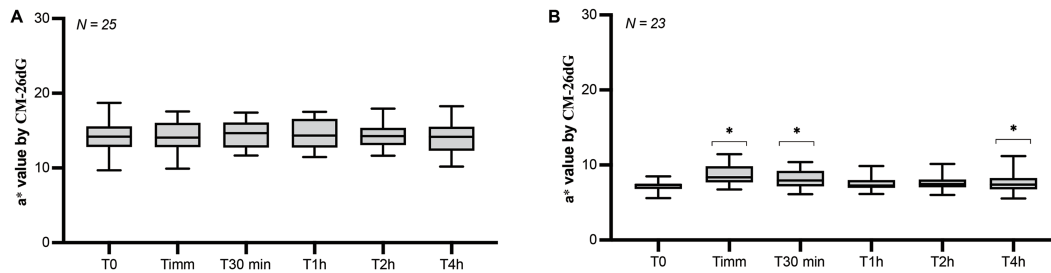


Figure 5: Instrumental measurement of a* value (mean, median, maximum and minimum) in each time-points. Subsequent times that are significantly different ($P < 0.05$ at 95% CI) from T0 are marked with an asterisk “*” on the figure.

(A) GA: time-points comparison of a* value. (B) TS: time-points comparison of a* value. *N* means valid sample size.

Abbreviations: GA, glycolic acid; TS, tape stripping; a* value, redness chromaticity value. CI, confidence interval.

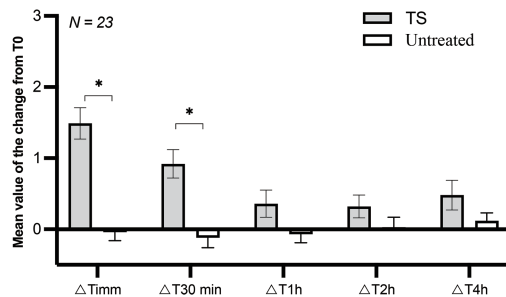


Figure 6: Instrumental measurement of a* value (mean with standard error bar) in each time-points adjusted by T0.

Abbreviations: TS, tape stripping; a* value, redness chromaticity value. CI, confidence interval. The model showed that the difference was statistically significant ($P < 0.05$, 95% CI) compared with the untreated area; significant points are marked with an asterisk (*). *N* means valid sample size.

Instrumental results of mexameter MX18 (EI, erythema index)

GA model and TS model showed the effect on inducing erythema.

Both **GA model** and **TS model**, an increased value indicates higher value of EI on facial skin and forearm skin, it was observed that the EI after **GA model** inducing at T30 min and T2h compared to that of T0, but there is no significant difference; and the EI after **TS model** inducing at T1min and T30 min is significantly increased ($P < 0.05$) compared to that of T0. **TS model** delivers a significantly stronger effect ($P < 0.05$) on inducing erythema at T1min, T30 min and T1h compared with the untreated control. (Fig. 7 and Fig. 8).

Imaging analysis results of redness (a* value)

GA model and TS model showed the effect on inducing skin redness.

Both **GA model** and **TS model**, an increased value indicates higher severity of redness (a* value) on facial skin and forearm skin, it was observed that the redness after **GA model** inducing at T1min, T30 min, T1h and T2h are

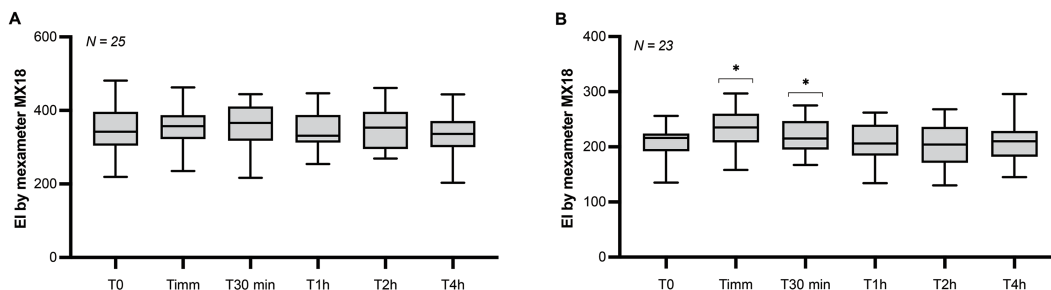


Figure 7: Instrumental measurement of EI (mean, median, maximum and minimum) in each time-points.

(A) GA: time-points comparison of EI. (B) TS: time-points comparison of EI. Subsequent times that are significantly different ($P < 0.05$ at 95% CI) from T0 are marked with an asterisk “*” on the figure. *N* means valid sample size.

Abbreviations: GA, glycolic acid; TS, tape stripping; EI, erythema index. CI, confidence interval.

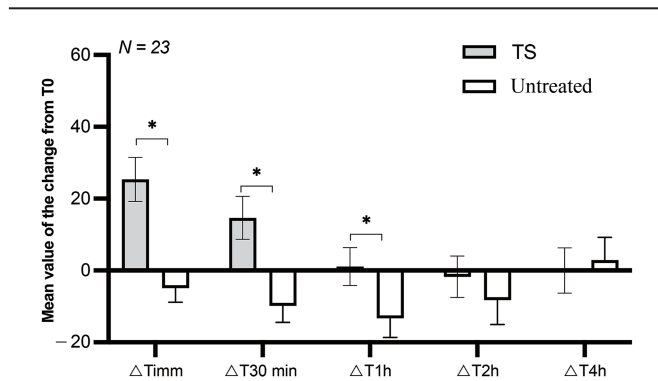


Figure 8: Instrumental measurement of EI (mean with standard error bar) in each time-points adjusted by T0.

Abbreviations: TS, tape stripping; EI, erythema index. CI, confidence interval.

The model showed that the difference was statistically significant ($P < 0.05$, 95% CI) compared with the untreated area; significant points are marked with an asterisk (*). *N* means valid sample size.

significantly increased ($P < 0.05$) compared to that of T0. and the redness after **TS model** inducing at Timm, T30 min, T1h and T4h are significantly increased ($P < 0.05$) compared to that of T0. **TS model** delivers a significantly stronger effect ($P < 0.05$) on inducing redness at Timm, T30 min and T1h compared with the untreated control. (Fig. 9 and Fig. 10).

Imaging analysis results of red area ratio% (on face)

Photographs of **GA model** showed the effect on inducing skin redness.

GA model, an increased value indicates higher value of red area ratio% on facial skin, it was observed that the red area after **GA model** induction at Timm, T30 min

and T1h are significantly increased ($P < 0.05$) compared to that of T0. (Fig. 11).

Model product intervention in Stage 2 – detectability assessment using a reference product

Significant findings of Stage 2 are summarized in Table 5 and Table 6.

Post-induction, pre-application between-side differences are summarized in Table 7 (mean, SE, 95% CI, and P value). Clinical erythema grades were comparable between sides for both models. For the GA model, Mexameter EI was higher on the subsequently treated side than on the untreated side (mean difference 14.70, approximately 95% CI ± 13.73 , $P = 0.045$); other instrumental and imaging endpoints did not differ significantly. A nonsignificant P value was not interpreted as evidence of equivalence. The EI imbalance was considered when interpreting GA instrumental endpoints. Mean differences of change from induction (treated versus untreated) with 95% CIs are reported in Table 8.

Clinical grading results of redness/erythema

The Stage 2 reference product intervention was used to examine treated versus untreated detectability under the tested conditions, rather than to confirm reference-product efficacy or formally validate either model.

The results showed that under **GA model** the reference product delivers a significantly stronger effect ($P < 0.05$)

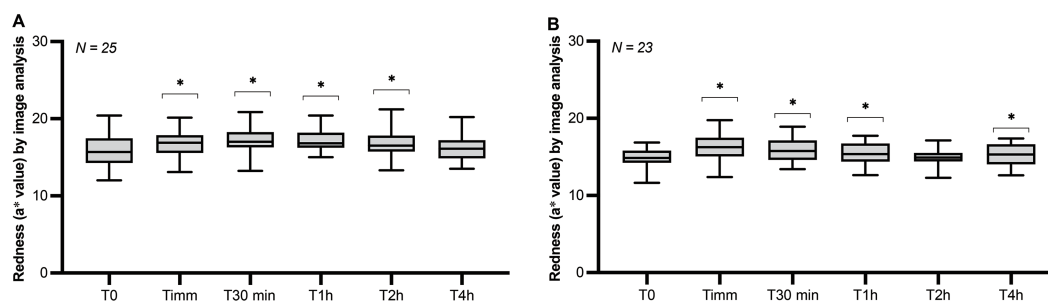


Figure 9: Imaging analysis of redness a^* value (mean, median, maximum and minimum) in each time-points. Subsequent times that are significantly different ($P < 0.05$ at 95% CI) from T0 are marked with an asterisk “*” on the figure.

(A) GA: time-points comparison of redness (a^* value). (B) TS: time-points comparison of redness (a^* value). *N* means valid sample size.

Abbreviations: GA, glycolic acid; TS, tape stripping; a^* value, redness chromaticity value. CI, confidence interval.

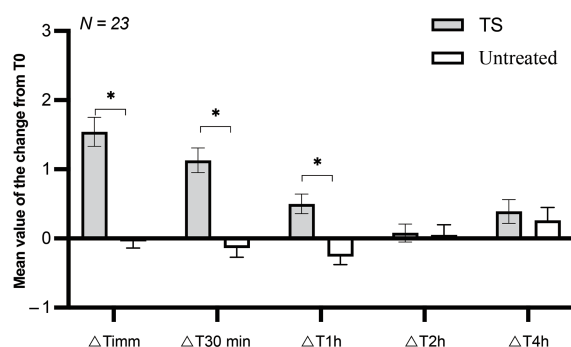


Figure 10: Imaging analysis of redness a* value (mean with standard error bar) in each time-points adjusted by T0.

Abbreviations: TS, tape stripping; a* value, redness chromaticity value. CI, confidence interval. The model showed that the difference was statistically significant ($P < 0.05$, 95% CI) compared with the untreated area; significant points are marked with an asterisk (*). *N* means valid sample size.

on decreasing redness/erythema at T4h and T24h compared with the untreated control; The TS model showed numerical improvement at T2h and T4h, without a statistically significant treated versus untreated difference. (Fig. 12 and Fig. 13).

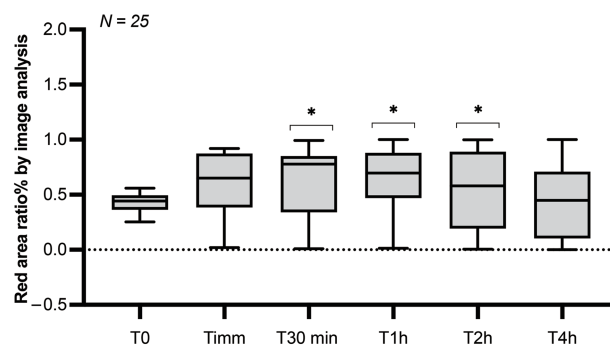


Figure 11: Imaging analysis of red area ratio% value (mean, median, maximum and minimum) in each time-points.

CI, confidence interval.

Subsequent times that are significantly different ($P < 0.05$ at 95% CI) from T0 are marked with an asterisk "*" on the figure. *N* means valid sample size.

Instrumental results of spectrophotometer CM-26dG (a* value, SCE mode)

The models showed trends consistent with detectability of short-term anti-redness changes using the reference product.

Table 5: Summary of *P* value between time-points in Stage 2.

Attributes	Model	Treatments	Comparison with Timm after induction				
			T30 min	T1h	T2h	T4h	T24h
Redness/erythema by clinical assessment	GA	Treated	0.005	0.000	0.000	0.000	0.000
		Untreated	0.002	0.001	0.000	0.001	0.000
	TS	Treated	0.000	0.000	0.000	0.000	0.000
		Untreated	0.000	0.000	0.000	0.000	0.000
a* value by CM26dG SCE mode	GA	Treated	0.744	0.375	0.283	0.298	0.951
		Untreated	0.757	0.905	0.355	0.521	0.640
	TS	Treated	0.004	0.000	0.000	0.000	0.036
		Untreated	0.000	0.006	0.001	0.032	0.011
EI by mexameter MX18	GA	Treated	0.004	0.000	0.000	0.000	0.003
		Untreated	0.002	0.003	0.000	0.006	0.002
	TS	Treated	0.001	0.002	0.000	0.006	0.022
		Untreated	0.000	0.017	0.000	0.003	0.007
a* value by photo analysis	GA	Treated	0.301	0.020	0.056	0.097	0.004
		Untreated	0.689	0.029	0.098	0.109	0.023
Area ratio of redness by photo analysis	GA	Treated	0.432	0.039	0.013	0.032	0.003
		Untreated	0.934	0.278	0.252	0.048	0.190

Abbreviations: GA, glycolic acid; TS, tape stripping; SCE, specular component excluded; CM26dG, spectrophotometer CM-26dG; MX18, Mexameter MX18; EI, erythema index; a* value, redness chromaticity value.

Table 6: Summary of *P* value between treatments in Stage 2.

Attributes	Model	Comparison with untreated				
		T30 min	T1h	T2h	T4h	T24h
Redness/erythema by clinical assessment	GA	1.000	0.157	0.132	0.009	0.020
	TS	0.114	0.426	0.783	0.807	0.294
a* value by CM26dG SCE mode	GA	0.589	0.482	0.932	0.575	0.696
	TS	0.548	0.390	0.165	0.158	0.622
EI by mexameter MX18	GA	0.993	0.102	0.359	0.180	0.748
	TS	0.336	0.321	0.487	0.737	0.912
a* value by photo analysis	GA	0.278	0.904	0.410	0.786	0.317
Area ratio of redness by photo analysis	GA	0.129	0.021	0.007	0.157	0.006

Abbreviations: GA, glycolic acid; TS, tape stripping; SCE, specular component excluded; CM26dG, spectrophotometer CM-26dG; MX18, Mexameter MX18; EI, erythema index; a* value, redness chromaticity value.

Table 7: Treated versus untreated differences immediately after induction (pre-application).

Attributes	Model	Mean diff (T–U)	95% CI	<i>P</i> value
Redness/erythema by clinical assessment	GA	0.00	–0.11 to 0.11	1.000
	TS	–0.07	–0.24 to 0.10	0.412
a* value by CM26dG SCE mode	GA	0.26	–0.36 to 0.88	0.426
	TS	–0.16	–0.62 to 0.30	0.504
EI by mexameter MX18	GA	14.70	0.97 to 28.43	0.045
	TS	–1.63	–15.27 to 12.01	0.816
a* value by photo analysis	GA	0.12	–0.20 to 0.44	0.465
Area ratio of redness by photo analysis	GA	0.02	–0.02 to 0.06	0.328

Abbreviations: GA, glycolic acid; TS, tape stripping; T–U, treated minus untreated; diff, difference; CI, confidence interval; SCE, specular component excluded; CM26dG, spectrophotometer CM-26dG; MX18, Mexameter MX18; EI, erythema index; a* value, redness chromaticity value.

The results showed that under both **GA model** and **TS model** the reference product show the better effect on decreasing redness at T1h, T2h, T4h and T24h compared with the untreated control, but there is no significant difference. (Fig. 14 and Fig. 15).

Instrumental results of mexameter MX18 (EI, erythema index)

The models showed trends consistent with detectability of short-term anti-redness changes using the reference product.

The results showed that both under **GA model** and **TS model** the reference product showed the better effect on decreasing erythema at Timm (except GA), T1h, T2h,

T4h and T24h compared with the untreated control, but there is no significant difference. (Fig. 16 and Fig. 17).

Imaging analysis results of redness (a* value, on face)

The models showed trends consistent with detectability of short-term anti-redness changes using the reference product.

The results showed that under **GA model** the reference product showed the better effect on decreasing redness at T30 min, T1h, T2h, T4h and T24h compared with the untreated control, but there is no significant difference. (Fig. 18 and Fig. 19).

Table 8: Treated versus untreated differences immediately after application (T30 min, T1h, T2h, T4h and T24h).

Time-point	Attributes	Model	Mean diff (T-U)	95% CI	P
T30 min-after induction	Redness/erythema by clinical assessment	GA	0	-0.07 to 0.07	1.000
		TS	0.13	-0.04 to 0.31	0.114
	a* value by CM26dG SCE mode	GA	0.16	-0.45 to 0.78	0.589
		TS	0.17	-0.4 to 0.73	0.548
	EI by mexameter MX18	GA	-0.07	-15.13 to 15	0.993
		TS	-8.53	-26.36 to 9.29	0.336
	a* value by photo analysis	GA	-0.09	-0.27 to 0.08	0.278
	Area ratio of redness by photo analysis	GA	-0.02	-0.05 to 0.01	0.129
T1h-after induction	Redness/erythema by clinical assessment	GA	-0.07	-0.16 to 0.03	0.157
		TS	0.07	-0.13 to 0.26	0.426
	a* value by CM26dG SCE mode	GA	-0.26	-1 to 0.48	0.482
		TS	-0.24	-0.8 to 0.32	0.39
	EI by mexameter MX18	GA	-14.5	-32.07 to 3.07	0.102
		TS	-12.53	-37.93 to 12.87	0.321
	a* value by photo analysis	GA	-0.01	-0.24 to 0.21	0.904
	Area ratio of redness by photo analysis	GA	-0.04	-0.08 to -0.01	0.021
T2h-after induction	Redness/erythema by clinical assessment	GA	-0.08	-0.19 to 0.03	0.132
		TS	-0.03	-0.24 to 0.17	0.783
	a* value by CM26dG SCE mode	GA	-0.04	-0.89 to 0.82	0.932
		TS	-0.39	-0.95 to 0.17	0.165
	EI by mexameter MX18	GA	-7.03	-22.47 to 8.4	0.359
		TS	-4.83	-18.88 to 9.21	0.487
	a* value by photo analysis	GA	-0.11	-0.36 to 0.15	0.41
	Area ratio of redness by photo analysis	GA	-0.06	-0.09 to -0.02	0.007
T4h-after induction	Redness/erythema by clinical assessment	GA	-0.22	-0.37 to -0.06	0.009
		TS	-0.02	-0.25 to 0.22	0.807
	a* value by CM26dG SCE mode	GA	-0.2	-0.92 to 0.52	0.575
		TS	-0.39	-0.95 to 0.16	0.158
	EI by mexameter MX18	GA	-14.23	-35.42 to 6.95	0.18
		TS	-2.53	-17.8 to 12.74	0.737
	a* value by photo analysis	GA	-0.05	-0.4 to 0.31	0.786
	Area ratio of redness by photo analysis	GA	-0.03	-0.07 to 0.01	0.157
T24h-after induction	Redness/erythema by clinical assessment	GA	-0.15	-0.27 to -0.03	0.02
		TS	0.13	-0.1 to 0.37	0.294
	a* value by CM26dG SCE mode	GA	-0.16	-0.98 to 0.66	0.696
		TS	0.14	-0.44 to 0.73	0.622
	EI by mexameter MX18	GA	-2.9	-21.17 to 15.37	0.748
		TS	-1.23	-23.84 to 21.37	0.912
	a* value by photo analysis	GA	-0.13	-0.4 to 0.13	0.317
	Area ratio of redness by photo analysis	GA	-0.07	-0.11 to -0.02	0.006

Abbreviations: GA, glycolic acid; TS, tape stripping; T-U, treated minus untreated; diff, difference; CI, confidence interval; SCE, specular component excluded; CM26dG, spectrophotometer CM-26dG; MX18, Mexameter MX18; EI, erythema index; a* value, redness chromaticity value.

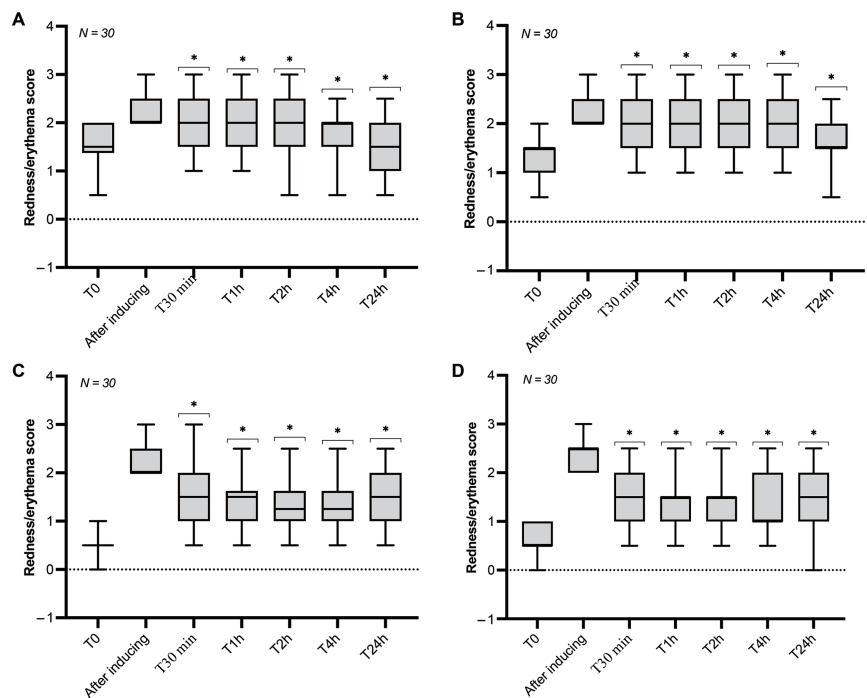


Figure 12: Clinical scoring of redness/erythema using 5-point scale refer to the details in the “Clinical grading” subsection (grade 0 means clear, grade 4 means severe).

(A) GA: time-points comparison of redness/erythema score. (B) GA: time-points comparison of redness/erythema score (untreated). (C) TS: time-points comparison of redness/erythema score. (D) TS: time-points comparison of redness/erythema score (untreated). On the figure, the clinical scoring of redness (mean, median, maximum and minimum) in each time-points. Subsequent times that are significantly different ($P < 0.05$ at 95% CI) from the time after induction are marked with an asterisk “*”. *N* means valid sample size. Abbreviations: GA, glycolic acid; TS, tape stripping. CI, confidence interval.

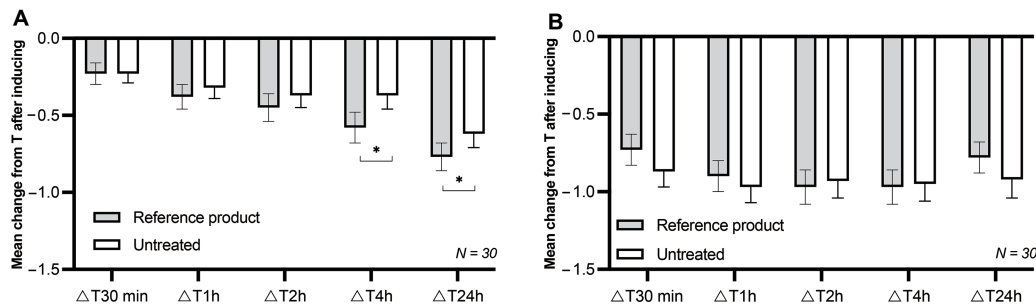


Figure 13: Clinical scoring of redness/erythema using 5-point scale refer to the details in the “Clinical grading” subsection (grade 0 means clear, grade 4 means severe).

(A) GA: treatments comparison of redness/erythema score. (B) TS: treatments comparison of redness/erythema score. On the figure, the clinical scoring of redness (mean with standard error bar) in each time-points adjusted by the time after induction. The model showed that the difference was statistically significant ($P < 0.05$, 95% CI) compared with the untreated area; significant points are marked with an asterisk (*). *N* means valid sample size. Abbreviations: GA, glycolic acid; TS, tape stripping; CI, confidence interval.

Imaging analysis results of the percentage of red area (on face)

The Stage 2 reference-product intervention supported short-term detectability mainly for the GA model on

selected endpoints (clinical grading and facial red-area imaging); the TS model showed mainly numerical trends. These findings support exploratory model characterization rather than definitive efficacy confirmation of the reference product or formal validation of either model.

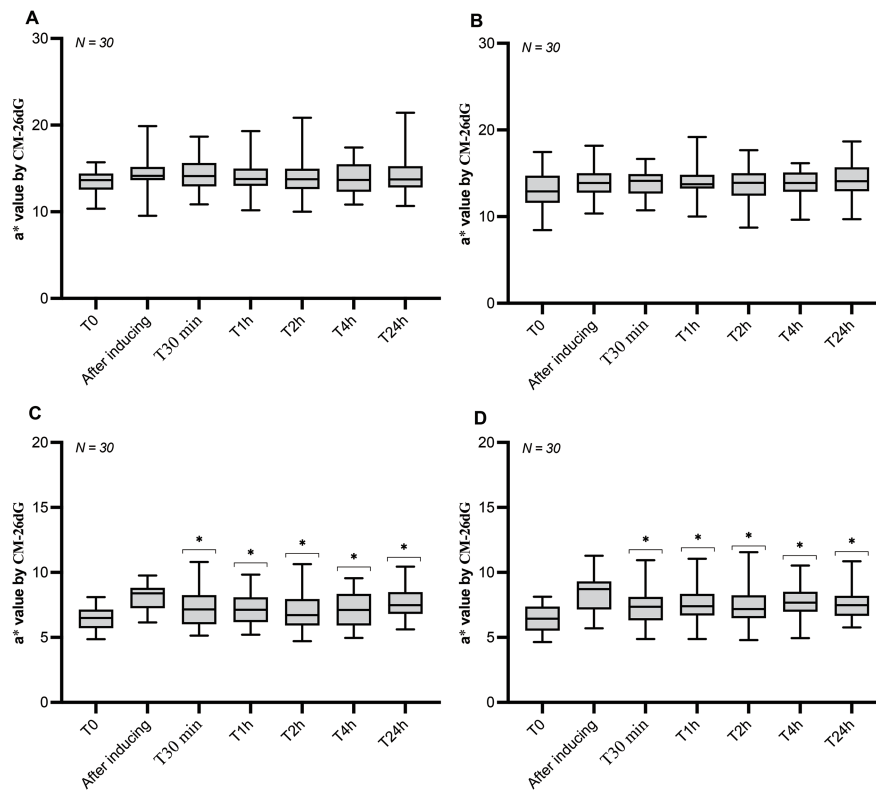


Figure 14: Instrumental measurement of a^* value (mean, median, maximum and minimum) in each time-point.

(A) GA: time-points comparison of a^* value. (B) GA: time-points comparison of a^* value (untreated). (C) TS: time-points comparison of a^* value. (D) TS: time-points comparison of a^* value (untreated). Subsequent times that are significantly different ($P < 0.05$ at 95% CI) from the time after induction are marked with an asterisk "*" on the figure. N means valid sample size. Abbreviations: GA, glycolic acid; TS, tape stripping; a^* value, redness chromaticity value. CI, confidence interval.

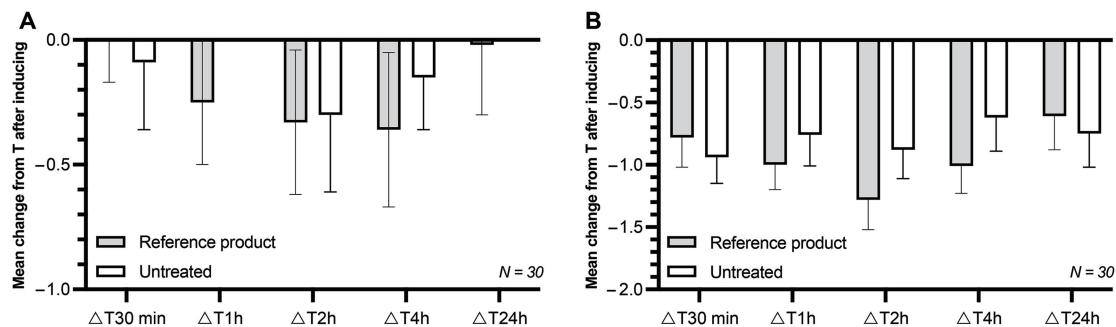


Figure 15: Instrumental measurement of a^* value (mean with standard error bar) in each time-points adjusted by the time after induction.

(A) treatments comparison of a^* value. (B) TS: treatments comparison of a^* value. The model showed that the difference was statistically significant ($P < 0.05$, 95% CI) compared with the untreated area; significant points are marked with an asterisk (*). N means valid sample size.

Abbreviations: TS, tape stripping; CI, confidence interval.

The results showed that under **GA model** the reference product delivers a significantly stronger effect ($P < 0.05$) on decreasing redness at T1h, T2h and T24h compared with the untreated control (Fig. 20, Fig. 21 and Fig. 22).

Safety outcomes

During Stage 1 and Stage 2, all GA induction procedures were conducted under dermatologist supervision. No

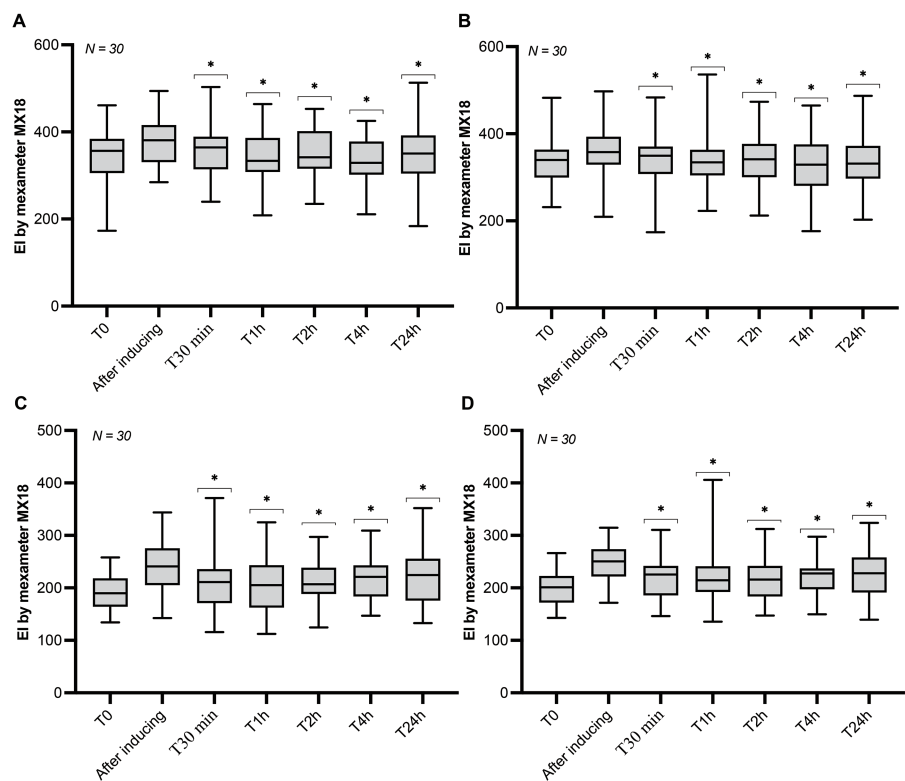


Figure 16: Instrumental measurement of EI (mean, median, maximum and minimum) in each time-points.

(A) GA: time-points comparison of EI. (B) GA: time-points comparison of EI (untreated). (C) TS: time-points comparison of EI. (D) TS: time-points comparison of EI (untreated). Subsequent times that are significantly different ($P < 0.05$ at 95% CI) from the time after induction are marked with an asterisk "*" on the figure. N means valid sample size. Abbreviations: GA, glycolic acid; TS, tape stripping; EI, erythema index. CI, confidence interval.

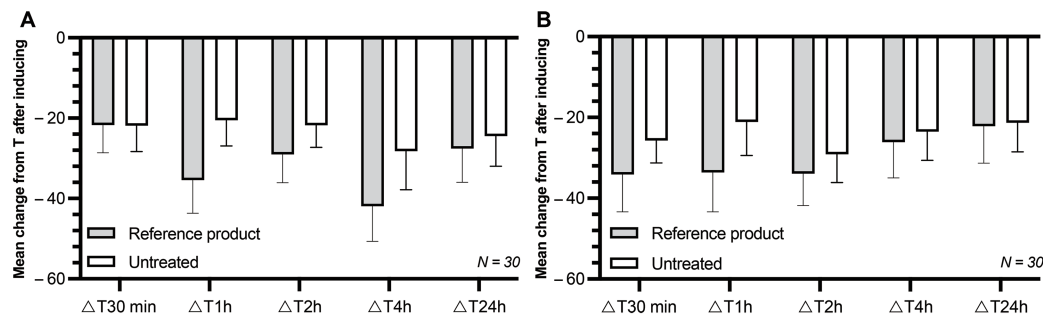


Figure 17: Instrumental measurement of EI (mean with standard error bar) in each time-points adjusted by the time after induction.

(A) treatments comparison of EI. (B) TS: treatments comparison of EI. The model showed that the difference was statistically significant ($P < 0.05$, 95% CI) compared with the untreated area; significant points are marked with an asterisk (*). N means valid sample size. Abbreviations: EI, erythema index; TS, tape stripping; CI, confidence interval.

protocol-defined AEs were recorded during scheduled visits through 24 h. Expected transient induction-related erythema was managed in accordance with the safety protocol (the "Glycolic acid induction: safety procedures" subsection). No subject required

medical referral. No clinically apparent PIH was observed at the 24 h visit; delayed irritation and PIH beyond 24 h cannot be fully excluded. These findings support the feasibility of the GA model under the pre-defined safety and stopping criteria. Discomfort was

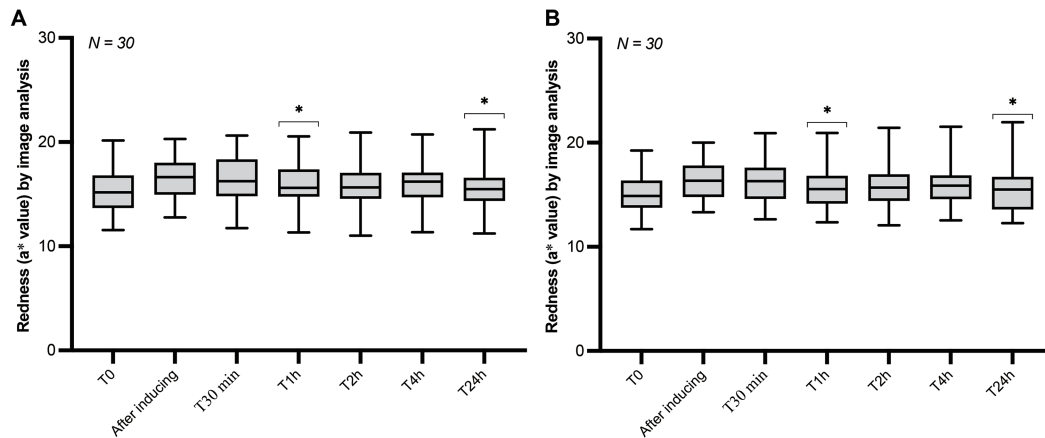


Figure 18: Imaging analysis of redness a^* value (mean, median, maximum and minimum) in each time-points.

(A) GA: time-points comparison of redness (a^* value). (B) GA: time-points comparison of redness (a^* value) (untreated). Subsequent times that are significantly different ($P < 0.05$ at 95% CI) from the time after induction are marked with an asterisk "*" on the figure. N means valid sample size.

Abbreviations: GA, glycolic acid; CI, confidence interval; a^* value, redness chromaticity value.

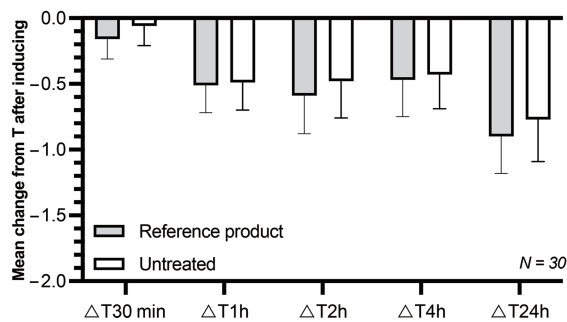


Figure 19: Imaging analysis of redness a^* value (mean with standard error bar) in each time-points adjusted by the time after induction.

The model showed that the difference was statistically significant ($P < 0.05$, 95% CI) compared with the untreated area; significant points are marked with an asterisk (*). N means valid sample size. CI, confidence interval; a^* value, redness chromaticity value.

monitored by dermatologist inquiry but not formally graded.

Discussion

This exploratory clinical model study addressed a practical need in cosmetic anti-redness testing: how to induce erythema in a site-relevant manner and which endpoints best capture short-term recovery. Stage 1 showed that GA induction on the face and TS induction on the forearm produced measurable erythema with different recovery kinetics, supporting their use as complementary rather than interchangeable models.

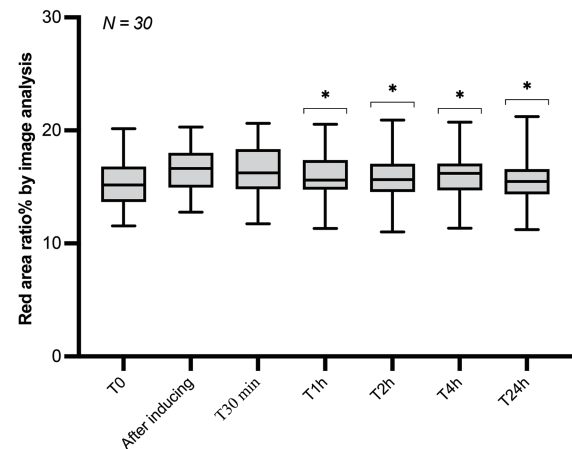


Figure 20: Imaging analysis of the percentage of red area value (mean, median, maximum and minimum) in each time-points.

Subsequent times that are significantly different ($P < 0.05$ at 95% CI) from the time after induction are marked with an asterisk "*" on the figure. N means valid sample size. CI, confidence interval.

Given that both GA and TS induction can cause skin irritation, equal induction intensity between treated sides is important in Stage 2, and predefined maximum induction thresholds should be maintained. Continuous dermatologist oversight, neutralization, and explicit stopping criteria (the "Glycolic acid induction: safety procedures" subsection) are essential when using 50% GA on the face.

Redness is rarely homogeneous across the test area, as

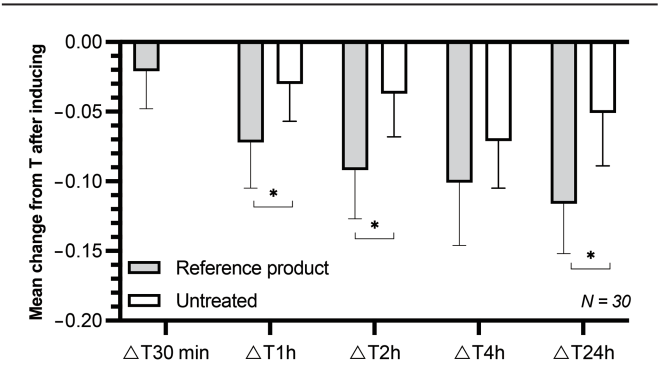


Figure 21: Imaging analysis of red area ratio% value (mean with standard error bar) in each time-points adjusted by the time after induction.

The model showed that the difference was statistically significant ($P < 0.05$, 95% CI) compared with the untreated area; significant points are marked with an asterisk (*). *N* means valid sample size. CI, confidence interval.

shown in Figure 21. Depending on the device, point-based instrumental probes may therefore underestimate spatial erythema, whereas whole-area clinical grading and imaging analysis capture spatial variation more effectively. For model-based testing, endpoint selection should match the erythema phenotype and application site rather than relying on a single universal readout. Accordingly, this study provides a practical framework for selecting induction models and endpoints when evaluating cosmetic anti-redness claims. Although 50% GA represents a relatively strong controlled chemical challenge compared with many everyday exposures, the resulting short-term facial erythema phenotype remains relevant to consumer-perceivable redness and may therefore inform daily anti-redness claim testing when results are interpreted in an exploratory, model-based

context. Extrapolation to real-world medical aesthetic settings may be more limited, because post-procedure erythema can vary with procedure type, energy/dose, and aftercare, which were not simulated here. The TS forearm model better approximates barrier-disruption scenarios but is not a direct facial-use surrogate.

Statistical significance should not be equated with clinical meaningfulness. No minimal clinically important difference (MCID) was prespecified for the 5-point clinical erythema grade or for instrumental a^*/EI values. Consumer-relevant improvement in cosmetic anti-redness testing is often inferred from concordant clinical grading and imaging changes rather than from a single instrumental threshold. We therefore emphasize consistency across endpoints, effect estimates with 95% confidence intervals, and model detectability, rather than claiming a clinically validated efficacy margin for the reference product.

This study has several limitations. First, Fitzpatrick skin phototypes were not systematically recorded; therefore, phototype distribution cannot be reported and findings may not generalize across all skin phototypes^[7,8]. Second, only female subjects were enrolled. Although acceptable for an exploratory cosmetic study, this limits generalizability, as sex-related differences in skin thickness, barrier function, sebum composition, and cutaneous vascular reactivity may influence induction and recovery. Third, all assessments were conducted within a single season, and seasonal differences in temperature, humidity, and baseline barrier status may affect erythema responses^[15]. Fourth, Stage 2 used a single reference product in a

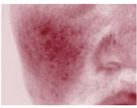
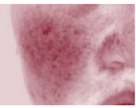
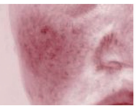
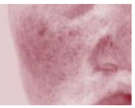
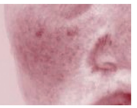
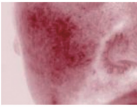
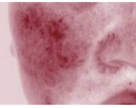
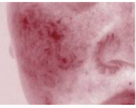
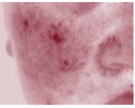
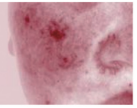
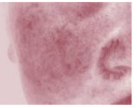
	After induction	T30 min	T1h	T2h	T4h	T24h
Active treatment for GA model						
Untreated for GA model						

Figure 22: Effective case photo of the percentage of red area (on face).

Abbreviations: GA, glycolic acid.

within-subject design and was not intended to confirm product efficacy. Fifth, the TS model was performed on the forearm whereas many anti-redness claims relate to the face, and the GA model used a high-concentration facial acid challenge that requires strict medical oversight. Future studies should enroll more diverse cohorts (including males and predefined Fitzpatrick subgroups), conduct multi-season testing, extend PIH follow-up, and further evaluate model performance with multiple reference interventions and longer safety follow-up. In addition, analyses were exploratory without multiplicity adjustment; the design was single-blind; GA amount varied within 0.2–0.3 g; safety follow-up was limited to 24 h; and the exact pH of the 50% GA preparation was not captured in the original documentation and could not be measured retrospectively because the product has been discontinued, which limits full chemical characterization of the induction challenge.

Conclusions

In summary, this exploratory study characterized two short-term anti-redness evaluation models for cosmetic testing. Stage 1 showed that GA (face) and TS (forearm) induction can elicit measurable erythema with different recovery kinetics. Stage 2 suggested that selected endpoints, particularly clinical grading and facial red-area imaging in the GA model can detect short-term treated-versus-untreated differences, whereas the TS model was supported mainly by numerical trends. These findings support feasibility and model selection rather than definitive validation of either model or confirmation of product efficacy. Because follow-up was limited to 24 h, we conclude only that no protocol-defined AEs were recorded during the short observation period; delayed irritation and PIH cannot be excluded.

Model selection should match the intended erythema phenotype (chemical irritation-related facial erythema for GA; barrier-disruption-related body erythema for TS), as detailed in the Discussion.

Author contributions

Yuanyuan Diao contributed to conceptualization, design, reviewing, and editing. Qianxin Zhu contributed to conceptualization, design, literature search and organization,

reviewing, and editing. Daniela Zoellner contributed to reviewing and editing. Yanwen Jiang contributed to conceptualization, design, data analysis and interpretation, literature search and organization. Yuchen Qiu contributed to data collection, data analysis and interpretation, literature search and organization. Ping Xu was responsible for safety control of the experimental process and provided professional dermatological expertise. All authors reviewed, edited, and approved the final version of the manuscript and agreed to its publication.

Funding information

This research received no external funding.

Acknowledgements

We express our sincere gratitude to the editors and reviewers for their constructive comments and to the journal and its publisher for their professional support.

Ethics statement

This study was conducted in accordance with the Declaration of Helsinki. This study was approved by the Institutional Review Board (IRB) of Ethics Committee, Shanghai China-norm Quality Technical Service Co., Ltd. (Approval Nos. FSEC 2023-124-01 and FSEC 2023-123-01).

All figures are either original to this work or properly attributed. No original images have been previously published, and all images comply with research integrity standards without inappropriate manipulation or misrepresentation.

Patient consent statement

Written informed consent was obtained from all individual participants prior to enrollment, including consent for publication of de-identified clinical data and accompanying images.

Data availability statement

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

AI statement

The authors used GPT-5.5, online version as a language-editing tool to improve grammar, clarity, and readability. The AI tool was not used for data analysis, data generation, image processing, or scientific content creation. All outputs were carefully reviewed and verified by the authors, who take full responsibility for the final content.

Conflicts of interest

The authors declare no conflicts of interest.

References

- Carle T, Gan D, Kalahasti G. Topical skin compositions for treating rosacea and skin redness : US202016792751(P). US Patent 11,090,352 B2. July 25, 2024.
- Raj N, Voegeli R, Rawlings AV, et al. A fundamental investigation into aspects of the physiology and biochemistry of the stratum corneum in subjects with sensitive skin. *Int J Cosmet Sci*. 2017;39(1):2-10.
- Aries MF, Poigny S. *Theobroma cacao* bean extract to fight against rosacea and skin redness : FR1853865 (P). French Patent 3,080,767 B1. July 25, 2024.
- National Center for Biotechnology Information. PubChem patent summary for WO-2019/211470-A1: extract of shells of *Theobroma cacao* beans for controlling rosacea and skin redness. NCBI PubChem Website. Accessed August 24, 2026.
- Cho BS, Yi YW, Kim KI. Composition for alleviating facial redness, comprising stem cell-derived exosomes as active ingredient: US202017104731(P). US Patent 2,0210,077,379 A1. July 25, 2024.
- Weber TM, Ceilley RI, Buerger A, et al. Skin tolerance, efficacy, and quality of life of patients with red facial skin using a skin care regimen containing Licochalcone a. *Journal of Cosmetic Dermatology*. 2006;5(3):227-232.
- Foy V, Weinkauff R, Whittle E, Basketter DA. Ethnic variation in the skin irritation response. *Contact Dermatitis*. 2001;45(6):346-349.
- Robinson MK. Racial differences in acute and cumulative skin irritation responses between Caucasian and Asian populations. *Contact Dermatitis*. 2000;42(3):134-143.
- Gerritsen MJ, Erp PEV, Vlijmen-Willems IMV, Lenders LTM, van de Kerkhof PCM. Repeated tape stripping of normal skin: a histological assessment and comparison with events seen in psoriasis. *Archives of Dermatological Research*. 1994;286(8):455-461.
- Wang HQ, Baran U, Wang RK. *In vivo* blood flow imaging of inflammatory human skin induced by tape stripping using optical microangiography. *J Biophotonics*. 2014;8(3):265-272.
- Tanaka M, Zhen YX, Tagami H. Normal recovery of the stratum corneum barrier function following damage induced by tape stripping in patients with atopic dermatitis. *Br J Dermatol*. 1997;136(6):966-967.
- Weigmann HJ, Lademann J, Meffert H, Schaefer H, Sterry W. Determination of the horny layer profile by tape stripping in combination with optical spectroscopy in the visible range as a prerequisite to quantify percutaneous absorption. *Skin Pharmacology & Applied Skin Physiology*. 1999;12(1-2):34-45.
- Visscher MO, Hoath SB, Conroy E, Wickett RR. Effect of semipermeable membranes on skin barrier repair following tape stripping. *Archives of Dermatological Research*. 2001;293(10):491-499.
- Pedersen LK, Held E, Johansen JD, Agner T. Short-term effects of alcohol-based disinfectant and detergent on skin irritation. *Contact Dermatitis*. 2005;52(2):82-87.
- Berardesca E, Vignoli GP, Distanti F, Brizzi P, Rabbiosi G. Effects of water temperature on surfactant-induced skin irritation. *Contact Dermatitis*. 1995;32(2):83-87.
- Wilhelm KP, Freitag G, Wolff HH. Surfactant-induced skin irritation and skin repair: evaluation of a cumulative human irritation model by noninvasive techniques. *J Am Acad Dermatol*. 1994;31(6):981-987.
- Maramaldi G, Togni S, Pagin I, et al. Soothing and anti-itch effect of quercetin phytosome in human subjects: a single-blind study. *Clin Cosmet Investig Dermatol*. 2016;9(1):55-62.
- Dreher F, Walde P, Luisi PL, Elsner P. Human skin irritation studies of a lecithin microemulsion gel and of lecithin liposomes. *Skin Pharmacol*. 1996;9(2):124-129.
- Basketter D, Reynolds F, Rowson M, Talbot C, Whittle E. Visual assessment of human skin irritation: a sensitive and reproducible tool. *Contact Dermatitis*. 1997;37(5):218-220.
- Kawai K. *Is visual assessment of human skin irritation reliable and reproducible?* Environmental Dermatology: The Official Journal of the Japanese Society for Contact Dermatitis; 2001.
- Stücker M, Hoffmann M, Altmeyer P. Instrumental evaluation of retinoid-induced skin irritation. *Skin Research and Technology*. 2002;8(2):133-140.
- Healy ZR, Dinkova-Kostova AT, Wehage SL, Thompson RE, Fahey JW, Talalay P. Precise determination of the erythema response of human skin to ultraviolet radiation and quantification of effects of protectors. *Photodermatol Photoimmunol Photomed*. 2009;25(1):45-50.
- Harrison GI, Young AR. Ultraviolet radiation-induced erythema in human skin. *Methods*. 2002;28(1):14-19.
- Abdlaty R. Hyperspectral imaging and data analysis of skin erythema post radiation therapy treatment [PhD thesis]. McMaster University; 2016.
- Okano Y, Abe Y, Masaki H, Santhanam U, Ichihashi M, Funasaka Y. Biological effects of glycolic acid on dermal matrix metabolism mediated by dermal fibroblasts and epidermal keratinocytes. *Exp Dermatol*. 2003;12:57-63.