Skin Research and Technology 2013; 0: 1–7
Printed in Singapore · All rights reserved
doi: 10.1111/srt.12073

Skin Research and Technology

SWIRL, a clinically validated, objective, and quantitative method for facial wrinkle assessment

Lily I. Jiang, Thomas J. Stephens and Robert Goodman

Thomas J. Stephens & Associates Inc., Carrollton, TX, USA

Background: Facial wrinkles are an undesirable feature caused by extrinsic photodamage and intrinsic aging process. Many cosmetic products and esthetic procedures strive to ameliorate the appearance of wrinkles. Currently the effects of those products and procedures on wrinkles are mainly evaluated by clinical grading, subjective self-assessment questionnaires, and optical profilometry of replica impressions. An objective and quantitative method is in demand.

Methods: Raking light optical profilometry was applied directly to facial photography to cast wrinkles as dark shadows. The resulting high-resolution digital images were analyzed using Image Pro software. A high-throughput method, Stephens Wrinkle Imaging using Raking Light (SWIRL), was developed to analyze the severity of wrinkles using photographs taken under the raking light condition. This method was applied to photographs taken from many panelists with a wide range of wrinkle severity scores. The result was compared with clinical grading scores to determine its validity. In addition, this method was applied to photographs taken from panelists before and after product usage to determine its sensitivity.

Results: Using the SWIRL method, multiple wrinkle parameters were quantitatively assessed, including wrinkle count, length, width, area, and relative depth. Those parameters correlated well with clinical grading scores, showing correlation coefficient (*r* value) of about 0.8 for all parameters. This result indicates that the SWIRL method is a valid method for analyzing wrinkle severity. When applied to a clinical study, the SWIRL method was sensitive enough to detect improvement after 8 weeks of product application.

Conclusions: The SWIRL method has been fully validated through clinical studies. It is accurate, objective, and quantitative. As multiple wrinkle parameters are analyzed simultaneously, it can provide more detailed information on how wrinkles change over time and therefore has the potential to shed light on the action and mechanism of antiwrinkle products.

Key words: wrinkle – raking light photography – quantitative analysis – photonumeric wrinkle scale

Accepted for publication 3 May 2013

Skin undergoes both intrinsic and extrinsic aging processes. Like all the other organs in the body, the skin progressively decreases in its functional capacity as we age (1). Chronic sun exposure especially to the UV wavelength accelerates damage to the epidermis and the dermis, resulting in the appearance of wrinkles, mottled pigmentation, and skin sagging (2, 3). There is a great consumer demand for products and procedures that improve the appearance of aging skin by helping the skin to maintain a more radiant and youthful look. Many researches have been conducted to understand the mechanism of skin aging and photodamage (2–4). Many research trials are conducted each year on medical devices, prescription products, and cosmetic products to evaluate their effectiveness in improving the texture, color, and contour features of aging skin (5, 6).

Typically the effects of cosmetic products are assessed by expert grading during clinical trials and wrinkle assessment is considered a primary endpoint of those trials. Several wrinkle severity scales have been published over the years and have been used as guidelines for clinical grading (7–11). In an effort to make wrinkle assessment less subjective, published wrinkle scales have verbal descriptors of the skin condition and/or photographs of varying wrinkle severity which are assigned to numeral scores. Although this approach adds objectivity to wrinkle assessments, it still has its limitation in that it requires training and experience to reach a level of consistency. In addition, every face displays a unique pattern of wrinkles; a grading score only reflects a combined impression of all features without differentiation of distinct parameters, such as wrinkle count, length, or depth.

Alternatively, wrinkles can be assessed using skin replica impressions and optical profilometry techniques following JCSS guidelines, a standardized method for analyzing digital images taken from small replica impression of the crow's foot area (12). Although a proven method, preparing quality silicone replicas can be quite challenging even for veteran clinicians. Making skin replicas is labor intensive, highly variable due to the procedure, and is typically done on a small area. Typical problems include replica ring positioning errors, air bubbles in the replica impression, and controlling the polymerization process. Slight variations in temperature, humidity, and body temperature can produce unsuitable replica impressions. Attempts to standardize the process of preparing quality replicas impressions can be quite frustrating.

Applying raking light optical profilometry (RLOP) directly to facial photography is a novel, noninvasive, and direct method to assess facial wrinkle condition at multiple sites (13). The basic idea is that when the lighting is at a scant angle to the face of the subject, wrinkles are casted as dark shadows. Using this approach and Image Pro Plus software we developed a method, Stephens Wrinkle Imaging using Raking Light (SWIRL), which can analyze facial wrinkles in an automated way. Multiple parameters are quantified and reported regarding the number and dimensions of the wrinkles. In this study we report the functions and capability of the SWIRL method and the full clinical validation of this method.

Materials and Methods

Study design

These clinical trials were conducted in accordance with applicable Good Clinical Practice (GCP) regulations and guidelines and Institutional Review Board (IRB) regulations. All subjects were required to read and sign an inform consent form, HIPAA release form, and a photograph release form. Photographs were taken under raking light condition as detailed in the 'Image acquisition' section below. For validation of the SWIRL method, photographs were pooled from multiple trials to cover a range of wrinkle severity conditions. Data from a single study were used for determining the sensitivity of SWIRL method. The trial was conducted as a single-center efficacy study to test the effect of a product on wrinkle reduction in the periocular area. Clinical grading was performed at baseline and after 8 weeks of product usage. Raking light photographs were taken at both time points.

Clinical grading

Wrinkles and fine lines in the crow's feet area, under the eye area, forehead, and upper lip area were clinically evaluated by investigators on a 0–9 photographic scale. Score 0 indicates no fine lines or wrinkles; score 9 indicates numerous deep wrinkles and fine lines. Images used for the development and validation of SWIRL were pooled from multiple studies over a 2-year period.

Image acquisition

Subjects were instructed to clean their faces, remove all makeup and jewelry, and equilibrate to ambient conditions within the clinic prior to photography procedure. A subject's face was positioned at a head rest and a chin rest. Strobe lighting was placed above and at a scant angle to the side of the face. A Nikon D7000 camera was used to take the digital images. At returning visit, an image ghosting software developed by Stephens & Associates was used to align the positions of the face to match baseline images.

Image analysis

A proprietary macro developed based on Image Pro Plus v7 is used for image analysis. Briefly, an area of interest (AOI) is defined and wrinkles within the AOI are identified as dark shadows. The number of identified wrinkles is recorded. For each wrinkle, the length, width, and area are calibrated and calculated. The depth of a wrinkle is defined as how dark the wrinkle appears to be; therefore is expressed in relative unit. The measurements for each wrinkle are summed up and the final reported parameters represent the sum of all wrinkles within the AOI.

Statistical analysis
Correlation coefficient, r, was calculated using Pearson's method. r value of 0 indicates complete lack of correlation; r value of 1 indicates perfect linear relationship. The Wilcoxon signed rank sum test was used to determine the statistical significance of differences before and after product treatment.

Results

Crow's feet wrinkle severity grading atlas
During the past decade of clinical research, Stephens & Associates has accumulated a library of photographs displaying a variety of photodamaged skin conditions on the face. Using those images we compiled a photographic grading scale on wrinkle severity in the crow's feet area. As shown in Fig. 1, the wrinkle severity score increases as the number, length, and/or depth of the wrinkles increases. A representative image for each grading score is shown. Grading consistency among different investigators is achieved using this grading atlas as a reference.

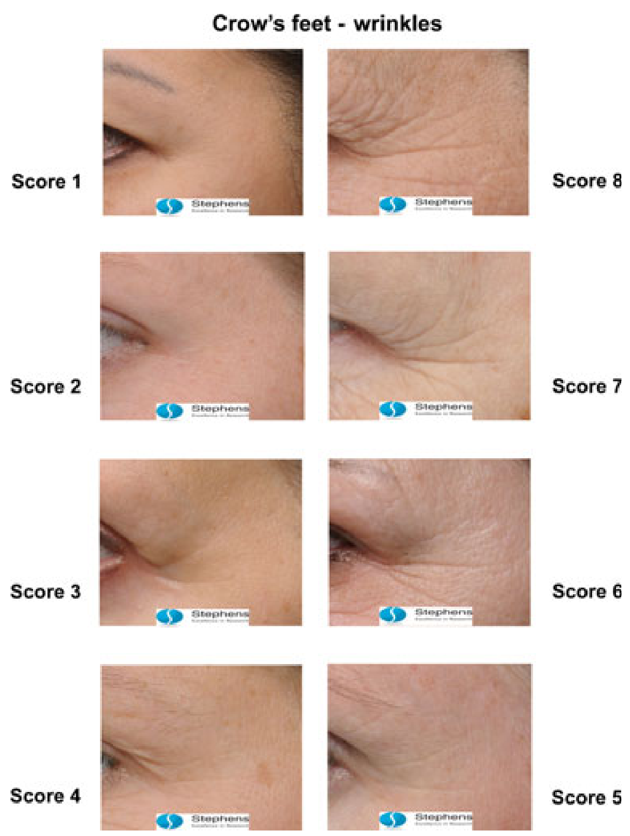

Fig. 1. Stephens wrinkle grading atlas for the crow's feet area. Wrinkles in the crow's feet area of female panelists were graded on a scale of 0–9; 0 = no wrinkles, 9 = numerous deep coarse wrinkles. Representative photographs for grading score 1–8 are selected from Stephen's photo library.

SWIRL, an objective and quantitative analysis for wrinkles
To objectively and quantitatively analyze the features of facial wrinkles we developed a raking light photography procedure to better capture the wrinkles and fine lines on the face (13). Strobe lighting was placed above and beside subject's face at a scant angle. Under such lighting conditions wrinkles and fine lines on the face are casted as shadows of varying shades. The photographs were then analyzed using Image Pro Plus software. Photographs taken under raking light condition from our image library provided us various examples and conditions to modify and improve our analysis method during the development process. As a result, a method was developed which we call SWIRL, for Stephens Wrinkle Imaging using Raking Light.

As shown in Fig. 2, an AOI near the crow's feet area was defined for analysis. The SWIRL method identified wrinkles as dark shadows in the defined area and produced an image overlay revealing the identified wrinkles (Fig. 2). Five wrinkle parameters were analyzed and reported (Table 1). In addition to the number of wrinkles, the length, width, and area of wrinkles were calibrated to millimeter or sq.

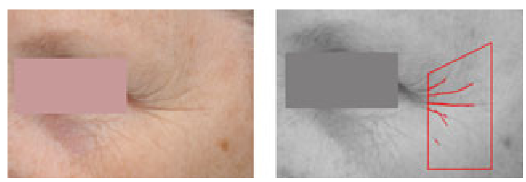

Fig. 2. Example of SWIRL analysis for crow's feet wrinkles. A digital color image taken under raking light condition was used for SWIRL analysis. An area of interest (AOI) near the crow's feet area was selected (red trapezoid). Wrinkles within the AOI were identified and highlighted in red.

TABLE 1. Wrinkle parameters analyzed through SWIRL method. The depth is assessed using arbitrary unit (a.u.) and higher value indicates deeper wrinkle

Parameter	
Number	8
Length (mm)	58.14
Width (mm)	6.70
Area (mm^2)	20.45
Depth (a.u.)	158

millimeter unit. The depth of the wrinkles was reported in a relative unit with higher value indicating deeper depth.

The SWIRL method has a built-in automation process. Multiple AOIs and multiple images can be analyzed automatically. The process time is less than 1 min per image.

Clinical validation of SWIRL: part I. wrinkle severity
To validate the reliability and robustness of the SWIRL method, we compared the results from SWIRL analysis to clinical grading scores. From our image library, a total of 40 raking light images with various grading scores were randomly selected. Those images were subjected to SWIRL analysis to obtain wrinkle parameters. Correlation coefficient (*r*) was calculated between each parameter and the clinical grading score.

As a control, we first assessed the correlation between skin color and wrinkle grading scores (Fig. 3A). No correlation was expected between the color of the skin and the severity of

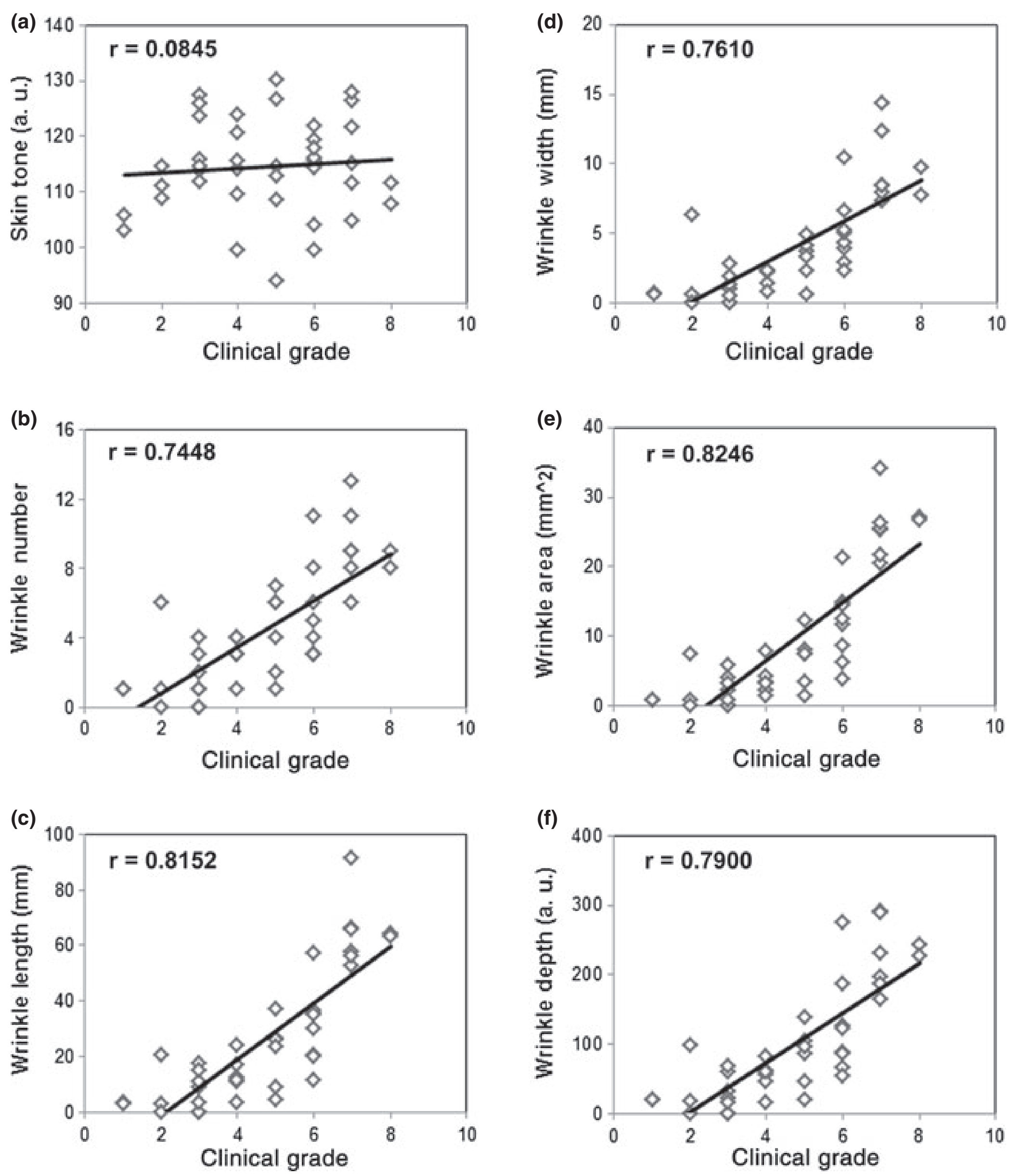

Fig. 3. Correlation analysis of wrinkle parameters and clinical grading. A total of 40 images with various wrinkle grading scores in the crow's feet area were analyzed using SWIRL method. The correlation coefficient (r) between the clinical grading scores and each parameter, skin tone (a), winkle number (b), total wrinkle length (c), total wrinkle width (d), total wrinkle area (e), and total wrinkle depth (f) was calculated and shown in the upper left corner of each graph. Each diamond represents an image, black trace is the trend line derived from linear regression.

wrinkles in the crow's feet area. As predicted, the correlation coefficient was very poor between these two features ($r = 0.08$). On the other hand, all the wrinkle parameters highly correlated with clinical grading score, displaying a correlation coefficient around 0.8 (Fig. 3b–f). Among all the parameters, the area of wrinkles best correlated with clinical grading score, likely because the area parameter takes into account both wrinkle length and width.

Clinical validation of SWIRL: part II. detection of improvements

To validate the sensitivity of the SWIRL method, we applied it to a clinical study to determine if this method could detect any improvement in wrinkle severity after product usage. The clinical study was designed to assess the efficacy of a product in wrinkle reduction in the periocular area after 8 weeks of product usage. Significant improvement in wrinkle severity at the crow's feet area after 8 weeks of product usage was observed through clinical grading (Table 2). Raking light photographs were taken at baseline and after 8 weeks of product application. Those photographs were analyzed using the SWIRL method and the results are shown in Table 2.

As shown in Table 2, the SWIRL method detected improvements in all the wrinkle parameters as captured by the raking light photographs. An example of great improvement is shown in Fig. 4. The total number and severity of wrinkles were visibly reduced after product usage. Among the five wrinkle parameters, the statistical significance of the improvements varied as assessed by Wilcoxon signed rank sum test. The modest change in the wrinkle area parameter is likely because smaller, shorter, and shallower wrinkles disappear first after treatment yet their relative contribution to the area parameter is smaller in comparison that to the wrinkle count or length parameter.

TABLE 2. Statistical analysis of wrinkle parameters before and after product use (n = 26). A clinical study was designed to assess the efficacy of a product in wrinkle reduction in the periocular area after 8 weeks of product usage. Clinical grading was done and raking light photographs were taken at baseline and after 8 weeks of product application. Average values of clinical grading score and of the five wrinkle parameters determined by SWIRL at baseline and at week 8 are listed in table. % change is calculated as the difference between week 8 and baseline values normalized by baseline values for each parameter. P value is calculated using Wilcoxon Signed Rank Sum Test and bold number indicates statistically significant score (P < 0.05)

Parameter	Baseline	Week 8	% change	*P* value
Number	20.4	17.1	−16.0	**0.044**
Length (mm)	117.9	101.6	−13.6	**0.045**
Width (mm)	26.6	22.4	−15.6	0.089
Area (mm^2)	57.4	50.5	−12.1	0.128
Depth (a.u.)	577.7	461.3	−20.1	**0.014**
Clinical grading	4.92	4.52	−8.2	**<0.001**

Wrinkle analysis in other areas of the face

The SWIRL method can be easily applied to identify and analyze wrinkles in other areas beyond the crow's feet. Fig. 5 shows examples of using the SWIRL method to analyze wrinkles on areas of the forehead, under the eye, and upper lip. Similar to analysis of the crow's feet area, an overlay image with outlines of identified wrinkles can be generated for each area (Fig. 5). Five wrinkle parameters are reported

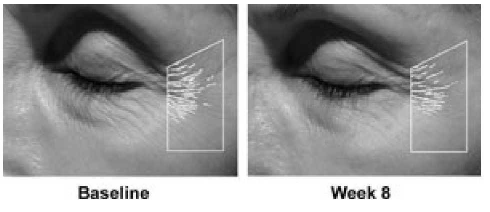

Fig. 4. Example of great improvement in wrinkle reduction after 8 weeks of product use through SWIRL analysis. SWIRL method was used to identify wrinkles in the crow's feet area at baseline and after 8 weeks of product use.

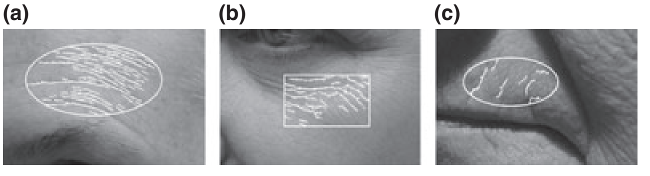

Fig. 5. Example of SWIRL analysis for wrinkles in the area of forehead (a), under the eye (b), and upper lip (c). AOI in the shape of oval or rectangle is shown in white; wrinkles within each AOI are identified and highlighted in white.

TABLE 3. Wrinkle parameters for areas in the forehead, under the eye, and upper lip are analyzed using the SWIRL method

Region	Number	Length (mm)	Width (mm)	Area (mm^2)	Depth (a.u.)
Forehead	57	317.3	37.9	96.6	1244
Under eye	15	94.2	17.2	35.6	327
Under lip	10	41.3	11.1	16.5	296

and can be used for further statistical analysis (Table 3).

Discussions

The SWIRL method was developed in response to the increasing demand for objective and quantitative assessment of facial wrinkles before and after treatment from prescription products, cosmetic products or esthetic procedures. Several methods have been employed for facial wrinkle analysis by various groups including ours (13–15). However, this study is the first to fully validate the method clinically. This method utilizes raking light digital photos to analyze facial wrinkles. Raking light photography does not require expensive setup or highly specialized equipment. The analysis method can be automated in a high-throughput fashion with less than a minute per image.

Upon analyzing raking light digital photographs, the SWIRL method reports five different parameters to quantify wrinkle severity. The five parameters, number, length, width, area, and depth, quantitatively describe the conditions of wrinkles at certain AOI. Importantly, through clinical validation study we show that all five parameters correlate well with clinical grading scores (Fig. 3), suggesting that these five parameters capture the essence of clinical grading. Among the five parameters, wrinkle area best correlates with clinical grading score. This is understandable because wrinkle area takes into account wrinkle number, length, and width, which is typically how clinical grading score is determined. Wrinkle number shows the lowest correlation score with clinical grading. This is possibly caused by two factors. One is that investigators tend to give more weight to wrinkle length and depth than wrinkle number when grading. The other is that in addition to major wrinkles the SWIRL method detects smaller and shorter wrinkles which have less weight in clinical grading.

The sensitivity of the SWIRL method is also validated using a clinical study. This method is capable of detecting improvements in wrinkle severity before and after product use, consistent with clinical grading (Table 2). In the example provided in this study, the improvements are reflected in all five parameters. These data provide some insightful information on wrinkle improvements as well as the analysis method. First, reduction in wrinkle number is more significant than in total wrinkle area. In clinical grading, it is observed that the smaller and shallower wrinkles tend to disappear first after treatment. Those wrinkles contribute equally to the wrinkle number as the larger and deeper wrinkles; however, their contribution to wrinkle area is a lot less. Therefore, this quantitative analysis is consistent with clinical observations. Second, the improvement in wrinkle depth is highly significant in this analysis. This is likely due to the loss of smaller and shallower wrinkles as well. However, this result points to the possibility that perhaps the depth parameter used in this study is not in linear relationship with the actual depth therefore the contribution of shallow wrinkles is artificially higher. Third, although good improvements on all five parameters are observed, the statistical significance is lacking for some parameters. This is caused by the variations in the data set, which requires highly reproducible image capturing.

The SWIRL method utilizes raking light photography taken directly of subjects' faces. Raking light photography is a noninvasive and relatively easy procedure (13). Although images used in this study were taken from Stephens photostation, raking light photographs taken from other photostations are also applicable to SWIRL analysis. However, the quality of the images is the key for valid analysis. Specifically, the lighting condition has to be consistent throughout the study to ensure highly reproducible images. The position of the face needs to be consistent to ensure high-throughput analysis.

Comparing to the traditional replica impression method, the SWIRL method offers a number of advantages. First, multiple facial sites can be analyzed from a single photograph and the AOI can be significantly larger than what can be captured in a single replica impression. Second, the digital images can be archived for an indefinite period of time while replicas need to be processed within a short period of time. Third, the SWIRL technology allows precise location of the AOI in each digital photograph using image software tools.

Objective assessment of clinical studies is highly desired. Quantitation of parameters can provide many detailed information that are previously combined into a single clinical grading score. We achieved both with SWIRL analysis of wrinkles. Through this study, we demonstrated that SWIRL analysis can potentially shed light on how the wrinkles are changing and how the changes relate to product action. It is taking us a step forward toward better understanding of the actions and changes produced by prescription and cosmetic wrinkle treatment products and by medical procedures.

Acknowledgments

We thank Dr. LeeAnn Boerma for her comments on the manuscript and Mr. Paul Kavanaugh and Mr. Matthew Batchelor for their initial development on image analysis macro. We are extremely grateful to all our employees at Stephens & Associates for their excellent team work and dedication.

References

1. Fisher GJ, Kang S, Varani J, Bata-Csorgo Z, Wan Y, Datta S, Voorhees JJ. Mechanisms of photoaging and chronological skin aging. Arch Dermatol 2002; 138: 1462–1470.
2. Kang S, Fisher GJ, Voorhees JJ. Photoaging: pathogenesis, prevention, and treatment. Clin Geriatr Med 2001; 17: 643–659.
3. Yaar M, Gilchrest BA. Photoageing: mechanism, prevention and therapy. Br J Dermatol 2007; 157: 874–887.
4. Rabe JH, Mamelak AJ, McElgunn PJ, Morison WL, Sauder DN. Photoaging: mechanisms and repair. J Am Acad Dermatol 2006; 55: 1–19.
5. Leyden J, Stephens TJ, Herndon JH Jr. Multicenter clinical trial of a home-use nonablative fractional laser device for wrinkle reduction. J Am Acad Dermatol 2012; 67: 975–984.
6. McCall-Perez F, Stephens TJ, Herndon JH Jr. Efficacy and tolerability of a facial serum for fine lines, wrinkles, and photodamaged skin. J Clin Aesthet Dermatol 2011; 4: 51–54.
7. Carruthers A, Carruthers J, Hardas B et al. A validated grading scale for crow's feet. Dermatol Surg 2008; 34 (Suppl. 2): S173–S178.
8. Day DJ, Littler CM, Swift RW, Gottlieb S. The wrinkle severity rating scale: a validation study. Am J Clin Dermatol 2004; 5: 49–52.
9. Fitzpatrick RE, Goldman MP, Satur NM, Tope WD. Pulsed carbon dioxide laser resurfacing of photo-aged facial skin. Arch Dermatol 1996; 132: 395–402.
10. Griffiths CEM, Wang TS, Hamilton TA, Voorhees JJ, Ellis CN. A photonumeric scale for the assessment of cutaneous photodamage. Arch Dermatol 1992; 128: 347–351.
11. Kappes UP. Skin ageing and wrinkles: clinical and photographic scoring. J Cosmet Dermatol 2004; 3: 23–25.
12. JCSS-Guideline. Guidelines for evaluation of anti-wrinkle products. J Japnaese Cosmet Sci Soc 2007; 31: S411–S431.
13. Stephens TJ, Oresajo C, Goodman G, Yatskayer M, Kavanaugh P. Chapter 6. Novel, compelling non-invasive techniques for evaluating cosmetic products. In: Draelos ZD, ed. Cosmetic dermatology: products and procedures. West Sussex, UK: Wiley-Blackwell; 2010: 47–54.
14. Fujimura T, Haketa K, Hotta M, Kitahara T. Global and systematic demonstration for the practical usage of a direct in vivo measurement system to evaluate wrinkles. Int J Cosmet Sci 2007; 29: 423–436.
15. Hillebrand GG, Miyamoto K, Schnell B, Ichihashi M, Shinkura R, Akiba S. Quantitative evaluation of skin condition in an epidemiological survey of females living in northern versus southern Japan. J Dermatol Sci 2001; 27 (Suppl. 1): S42–S52.

Address:
Lily I. Jiang
Thomas J. Stephens & Associates Inc.
3310 Keller Springs Road Suite 130
Carrollton
TX 75006
USA
Tel: + 972-392-1529
Fax: + 972-392-2347
e-mail: ljiang@stephens-associates.com