Received: 28 December 2020 | Revised: 11 February 2021 | Accepted: 24 February 2021

DOI: 10.1111/jocd.14058

JCD Journal of Cosmetic Dermatology WILEY

ORIGINAL CONTRIBUTION

Development and validation of a global photonumeric scale for evaluating skin quality of aged female facial skin

Christine Eiben-Nielson PhD | **Martina Kerscher MD**

Division of Cosmetic Science, Department of Biochemistry and Molecular Biology, University of Hamburg, Hamburg, Germany

Correspondence
Christine Eiben-Nielson and Martina Kerscher, Department of Biochemistry and Molecular Biology, Faculty of Chemistry, Field of Cosmetic Science, University of Hamburg, Papendamm 21, 20146 Hamburg, Germany
Emails: Christine.eiben-nielson@chemie.uni-hamburg.de (C.E.); Martina.kerscher@uni-hamburg.de (M.K.)

Funding information
Universitaet Hamburg

Abstract

Background: Skin quality plays an important role in overall attractiveness. However, so far, no visual grading scales have been published while their development seems to be an essential key step to provide validated grading scales for the evaluation of efficacy of minimally invasive procedures and cosmeceuticals aims on the improvement of skin quality, esthetic research, and clinical application.

Objectives: To develop and validate a visual five-point assessment scale for the evaluation of skin quality of female facial skin.

Methods: The five-point photonumeric Scientific Assessment Scale of Skin Quality is based on six parameters. Fifty standardized photos were rated by 13 experts. This examination was carried out in two cycles with an interval of 4 weeks. The intraclass correlation coefficient contributes to the identification of the inter-rater and intrarater reliability.

Results: Statistical analysis investigated six specific and two general parameters: The results of inter- and intrarater reliability for skin elasticity (ICC 0.816; ICC 0.883), wrinkles (ICC 0.840; 0.885), and age (ICC 0.885; 0.925) were almost perfect. The reliabilities for pigmentation (ICC 0.637; ICC 0.797), erythema (ICC 0.688; ICC 0.797), and overall skin quality (ICC 0.652; ICC 0.756) were substantial and for pore size moderate (ICC 0.405; ICC 0.584). Skin surface roughness (ICC 0.480; ICC 0.645) indicated a substantial intrarater reliability and a moderate interrater reliability. These data revealed good and excellent results.

Conclusions: The Scientific Assessment Scale of Skin Quality represents an innovative universal and reliable measurement instrument for a valid and reproducible evaluation of six parameters of aged female facial skin quality.

KEYWORDS
clinical scale, skin quality, skin quality assessment, visual scale

1 | INTRODUCTION

Skin quality is an omnipresent phenotypical parameter in today's society. New cosmetic procedures and esthetic tools result in an expanding cosmetic industry. The increased life expectancy causes the desire for physical health and attractiveness despite advanced age.[1–6] Perfect skin quality correlates with vitality and youth and is representing the interface to the social surrounding.[7–12] The term skin quality is generally defined as the constitution and quality of the entire characteristics of the human skin. It describes the level

of meeting-identified criteria for specific inherent parameters of the human skin.[13] The condition of the human cutis results from the interaction of multiple components: the skin surface structure, quantity and quality of the hydro-lipid film, the pigmentation, and the tension of the dermis and subcutis. Multiple processes induce effects on the skin. Lines, wrinkles, and loss of firmness and elasticity are typical age signs, as well as pigment disorders (lentigines solares) caused by UV-exposure.[14–16] The evaluation of skin quality requires a paradigmatic definition including the detection of inherent parameters of an aged face. According to extended data revealed in 2019 by a quantitative online survey (so far unpublished results),[13] our study determined six specific parameters for aged skin (Figure 1).

The effects of dermato-cosmetic tools and treatments should be ratable due to evidence-based criteria. There is a need for a valid measuring instrument.[17] Biophysical measurements, photo-based systems, and clinical scores are established and often used.[18,19] More than 50 former studies and about hundred rating scales exist. This amount of tools shows a heterogeneous variety regarding the development, structure, validation process, statistical analysis, scale level, and overall quality. Due to Dobos et al., one-third of it could be seriously recommended concerning standards based on the Consensus-based Standards for the selection of health status Measurement Instruments (COSMIN)—Checklist.[19,20] This impedes a simultaneous application or comparability of results from different scales. Most scales serve for the evaluation of single criterion.[19] A range of single-factored scales focusing on aging parameter exists[21–25] and also tools abstaining from any visual components are available.[26] A global tool for single or universal evaluation including a standardized construction and validation process is existent.[27,28] But pigment disorders and skin texture parameters so far been ignored. Until now no global five-point photonumeric scale based on a standardized validation process and a statistical analysis is available for assessing skin quality including aging parameters like pigmentation and skin relief. The research focus of this study was to develop a Scientific Assessment Scale of Skin Quality (SASSQ) for aged skin.

FIGURE 1 Parameters for aged skin

2 | MATERIAL AND METHODS

2.1 | Subject selection

The scale development and validation process are based on a photopool of 180 female subjects aged 19–74 years (Ø 37, 40 years) with Fitzpatrick Skin Type (FPST) I-IV. In addition, one picture was selected as a model picture for visualization of the relevant parameters. Exclusion criteria were skin diseases, facial skin lesions, pronounced solar elastosis, tattoos, permanent makeup, treatments with botulinum toxin or fillers (hyaluronic acid or calcium hydroxylapatite) during the last 6 months. Further exclusion criteria were the application of skincare or makeup on the day of the photo documentation. This examination was strictly carried out in accordance with the guidelines of the Good Clinical Practice (GCP) and the Declaration of Helsinki.[29]

2.2 | Photographic method

Photographs were two-dimensional frontal 90° facial portrait images of all qualified subjects. A high-resolution QuantifiCare 2D DermaViz® Camera (QuantifiCare Corporation, USA) was used. A standardized black background and constant light conditions were ensured. The 2D DermaViz® is a digital camera with a light pointer proposing a two-point interaction technique. Its grid-function assures exactly similar distances and angles for a standardized photo documentation.

2.3 | Creation of the photonumeric rating scale

The six scales were developed by a Photo Morphing technology using the software program Adobe® Photoshop® CC (1990–2018, Version 19.1.0) to modify the full-face frontal master picture for each scale progressively. Its stepwise procedure was performed by fixing intensity grades starting with 0 (none) to 4 (very severe). Additionally, real photo sections were identified from the photopool corresponding to the severity grade of each parameter to emphasize the morphed scale examples. The final scale and the summary of the detailed scale description are visualized in Figures 2 and 3.

2.4 | Validation of the rating scales

Validation was based on the assessment of 50 subject cases being evaluated at two different times with an interval of 4 weeks by 13 raters.

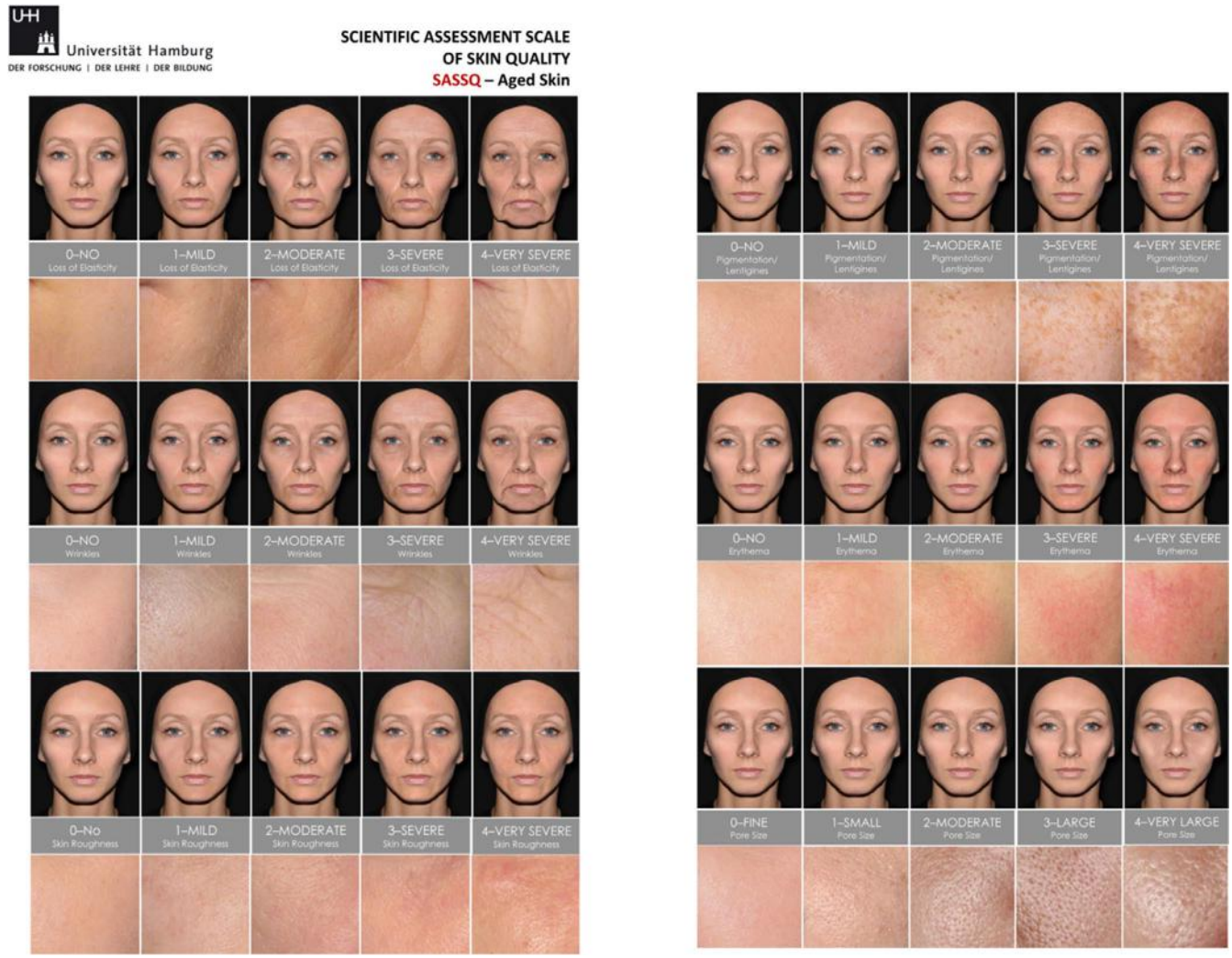

FIGURE 2 SASSQ for aged skin exploring six parameters

Each of the 13 raters got an individual validation set for the first and second cycle (C1; C2). All sets contained a USB stick with a Microsoft® Power-Point-Presentation (PPT) (Version 2011, 14.1, 110310), a corresponding booklet with a rating sheet for each of the 50 subjects, and the SASSQ as a booklet version. The PPT and the rating booklets were randomized individually for each rater and at every rating time. The booklets contained instructions for the validation procedure of the assessments and documented demographic data of the rater (only C1). Each rating page contained a picture of the corresponding subject case, a table for the rating data and it was signed by the rater with date and time information. The rating booklets and the SASSQ consisted of high-quality laser prints. Standardized conditions during both cycles and the assessment of every subject by using all six scales were required. The parameters overall skin quality and the age of the subject were graded as additional general parameters.

To examine the validity of the most relevant parameters of the SASSQ an additional questionnaire was sent with Set 2 (C2). This questionnaire contained six questions about the importance and practical use of the scale, as well as the suitability of the parameters. It was based on a five-point Likert scale.

2.5 | Statistical analysis

Descriptive statistics were carried out by calculating the arithmetic mean (MV), standard deviation (SD), standard error (SE), minimum (MIN), maximum (MAX), and the documentation of missing data at each rating point in time. As index for assessing the reliability of the scales, the intraclass correlation coefficient (ICC) was used for evaluating the reliability between all raters (interrater reliability) and for the reliability between the rating of C1 and C2 of one rater (intrarater reliability). The interrater reliability was identified by ICC for each point in time and is presented as MV of the ratings.

Different models of the ICC exist. For this study, the Shrout-Fleiss estimation was used under the assumption that the same raters (being evidently a random subset of all possible raters) would have rated a constant group of subjects at two points in time. The two-way-random-model and the subsumption "absolute agreement" were selected.[30] The interpretation of the data was obtained due to Landis and Koch (Table 1).[31] The 95% confidence interval (95% CI) was assessed for the intrarater reliability. The test-retest reliability was analyzed by the Pearson Correlation Coefficient (PKK) for

PARAMETER	INTENSITY SCALE				
	0=NONE	1=MILD	2=MODERATE	3=SEVERE	4=VERY SEVERE
Elasticity	No Loss of Elasticity	Slight Loss of Elasticity	Moderate Loss of Elasticity	Severe Loss of Elasticity	Very Severe Loss of Elasticity
Wrinkles	No Wrinkles	Slight Wrinkles	Moderate Wrinkles	Severe Wrinkles	Very Severe Wrinkles
SKIN SURFACE ROUGHNESS	No Skin Surface Roughness	Slight Skin Surface Roughness	Moderate Skin Surface Roughness	Severe Skin Surface Roughness	Very Severe Skin Surface Roughness
Lentigines / Pigmentation	No Uneven Pigmentation/ Lentigines	Slight Uneven Pigmentation/ Lentigines	Moderate Uneven Pigmentation/ Lentigines	Severe Uneven Pigmentation/ Lentigines	Very Severe Uneven Pigmentation/ Lentigines
Erythema	No Erythema	Slight Erythema	Moderate Erythema	Severe Erythema	Very Severe Erythema
Blemishes	No Blemishes -No Comedones -No Papules Per Facial Half	Slight Blemishes -1-3 Comedones -1-3 Papules Per Facial Half	Moderate Blemishes -4-6 Comedones -4-6 Papules Per Facial Half	Severe Blemishes -7-10 Comedones -7-10 Papules Per Facial Half	Very Severe Belmishes -≤ 11 Comedones -≤ 11 Papules Per Facial Half
Pore Size	0=FINE	1=SMALL	2=MODERATE	3=LARGE	4=VERY LARGE
	Fine Pore Size	Small Pore Size	Moderate Pore Size	Large Pore Size	Very Large Pore Size

FIGURE 3 SASSQ: Scale description

TABLE 1 ICC Interpretation due to Landis and Koch

Kappa statistik	Strength of agreement
<0.00	Poor
0.00–0.20	Slight
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Substantial
0.81–1.00	Almost perfect

TABLE 2 PKK interpretation by Bravais and Pearson

R	Strength of relationship
$0.00 \le r < 0.10$	Non
$0.10 \le r < 0.30$	Slight
$0.30 \le r < 0.50$	Moderate
$0.50 \le r < 0.70$	High
$0.70 \le r < 1.00$	Very high

the intrarater reliability and in accordance with Bravais and Pearson (Table 2).

For each scale, 650 rating combinations were possible for each point in time (50 subjects x 13 experts) for the six scales and for the two general parameters overall skin quality and age. In summary, both cycles show 10.400 possible rating combinations.

3 | RESULTS

3.1 | Subject characteristics

The mean age of the 50 female subjects of the rating PPT, building up the defined population for rating skin quality, was 42–56 years (range 22–74).

3.2 | Expert characteristics

Of the 13 expert raters (11 female, 2 male), seven experts were dermatologists specialized in esthetic dermatology. All dermatologists (six from Germany, 1 from the UK) were well-established experts with professional working expertise between 10–20 years. Six raters were cosmetic scientists. Four were graduated as Ph.D. doctors, two had M. Ed. degrees. They all had professional working expertise of 5 years and more as cosmetic consultants, lecturer in the field of cosmetic science, study coordinator in esthetic science, and product developer.

3.3 | Scale validation

The statistical results show comparably homogeneous values at C1 and C2 for the MV (and SD) as for the median values (Table 3). The highest MV was analyzed for the wrinkle scale (MV 1.69 C1; 1,70 C2 (wrinkles scale)) and for the assessment of the overall skin quality (MV 2.05 C1; 2.03 C2). The MV of the Shrout Fleiss estimates for the interrater reliability of C1 and C2 are visualized in Table 4. Almost perfect ICCs were analyzed for the elasticity and wrinkles scales with an ICC of 0.816 (elasticity scale) and 0.840 (wrinkle scale) and for the age assessment (ICC 0.885). ICCs with a substantial strength of agreement were calculated for the pigmentation scale (ICC 0.637) and for the erythema scale (ICC 0.688). Also, the additional general parameter overall skin quality was assessed with a substantial ICC of 0.652. ICCs of 0.480 and 0.405 were analyzed for the skin surface roughness scale and for the pore size scale. These data are moderate results.

The stability of the ratings was demonstrated by a bivariate scatter plot (bubble plot). Figure 4 illustrates the rating of all experts at C1 and C2 for all parameters.

The intrarater reliability for the elasticity and wrinkle scales and for the assessment of age showed almost perfect MV of the ICCs

TABLE 3 Descriptive statistics of the interrater reliability

Inter-rater reliability				
	Rating 1		Rating 2	
Score-parameter	MV + SD	Median (range)	MV + SD	Median (range)
Elasticity	1.61 ± 1.06	2.0 (0–4)	1.67 ± 1.06	2.0 (0–4)
Wrinkles	1.69 ± 1.07	2.0 (0–4)	1.70 ± 1.05	2.0 (0–4)
SSR	1.60 ± 0.68	2.0 (0–4)	1.59 ± 0.65	2.0 (0–4)
Pigmentation	1.42 ± 0.75	1.0 (0–4)	1.41 ± 0.72	1.0 (0–4)
Erythema	1.39 ± 0.89	1.0 (0–4)	1.33 ± 0.84	1.0 (0–4)
Pore size	1.33 ± 0.53	1.0 (0–4)	1.39 ± 0.51	1.0 (0–4)
Overall SQ	2.05 ± 0.71	2.0 (0–4)	2.03 ± 0.70	2.0 (0–4)
Age	41.48 ± 13.44	40.0 (17–80)	42.62 ± 13.67	41.0 (18–80)

TABLE 4 Shrout-Fleiss estimates for the inter-rater reliability

Inter-rater reliability				
		95% CI		
Score-parameter	MV_Z1_Z2 ICC Shrout-Fleiss	Min	Max	Strength of agreement
Elasticity	0.816	0.739	0.880	Almost perfect
Wrinkles	0.840	0.774	0.895	Almost perfect
SSR	0.480	0.359	0.611	Moderate
Pigmentation	0.637	0.525	0.745	Substantial
Erythema	0.688	0.586	0.783	Substantial
Pore size	0.405	0.279	0.545	Moderate
Overall SQ	0.652	0.546	0.755	Substantial
Age	0.885	0.836	0.925	Almost perfect

for all raters (elasticity ICC 0.883; wrinkles ICC 0.885; age ICC MV 0.925). Substantial strength of agreement was calculated for the skin surface roughness scale with an ICC of 0.645, the pigmentation scale with an ICC of 0.797, the erythema scale (ICC 0.797), and the ratings for the overall skin quality (ICC 0.756). The PKK values demonstrated a very high strength of relationship for four scales and the two additional parameters. The values of a high strength of relationship were shown for two scales (see Table 5).

Regarding the questionnaire, the data of 12 experts could be analyzed. All experts verified the importance of skin quality in daily practice. 10 experts agreed that objective methods for the evaluation of skin quality are currently missing. That the parameters of the SASSQ are suitable for evaluation skin quality was verified by all 12 experts as well as the fact, that the assessment is easy and comprehensive. All experts affirm the practical application of the scale in daily clinical practice or science.

4 | DISCUSSION

The analysis of the validation process resulted in moderate, substantial, and almost perfect ICCs. The highest intra- and interrater reliabilities with a very high strength of relationship were detected for elasticity, wrinkles, and age. Pigmentation, erythema, and overall skin quality showed a very high strength of relationship. The PKK for the intrarater reliability was high for the skin surface roughness and pore size scale.

In fact, for these parameters, the lowest ICC values with moderate results were detected. For the explanation of the presented data, the interpretation of the ICC values was based on Landis and Koch[31] as it was used for the discussion of the majority of existing instruments. Due to the interpretation of Koo et al.,[30] the data were even lower and would be showing poor reliability. This might be attributed to the assessment of subject images instead of a live evaluation. The method of photo evaluations itself has in evidence its limits, because of the photo quality and an overall high rating complexity. Especially for skin surface roughness scale and pore size scale a validation of smaller skin areas (optionally cheek skin areas) would be an improvement. Additively the technique of live rating of all scales (or just of skin roughness and pore size) might be a preferable option for revalidation.

However, photo rating seems to be a more practical solution for use in daily practice or for clinical trials aiming to objectively changes in skin quality. Further investigations should aim at future

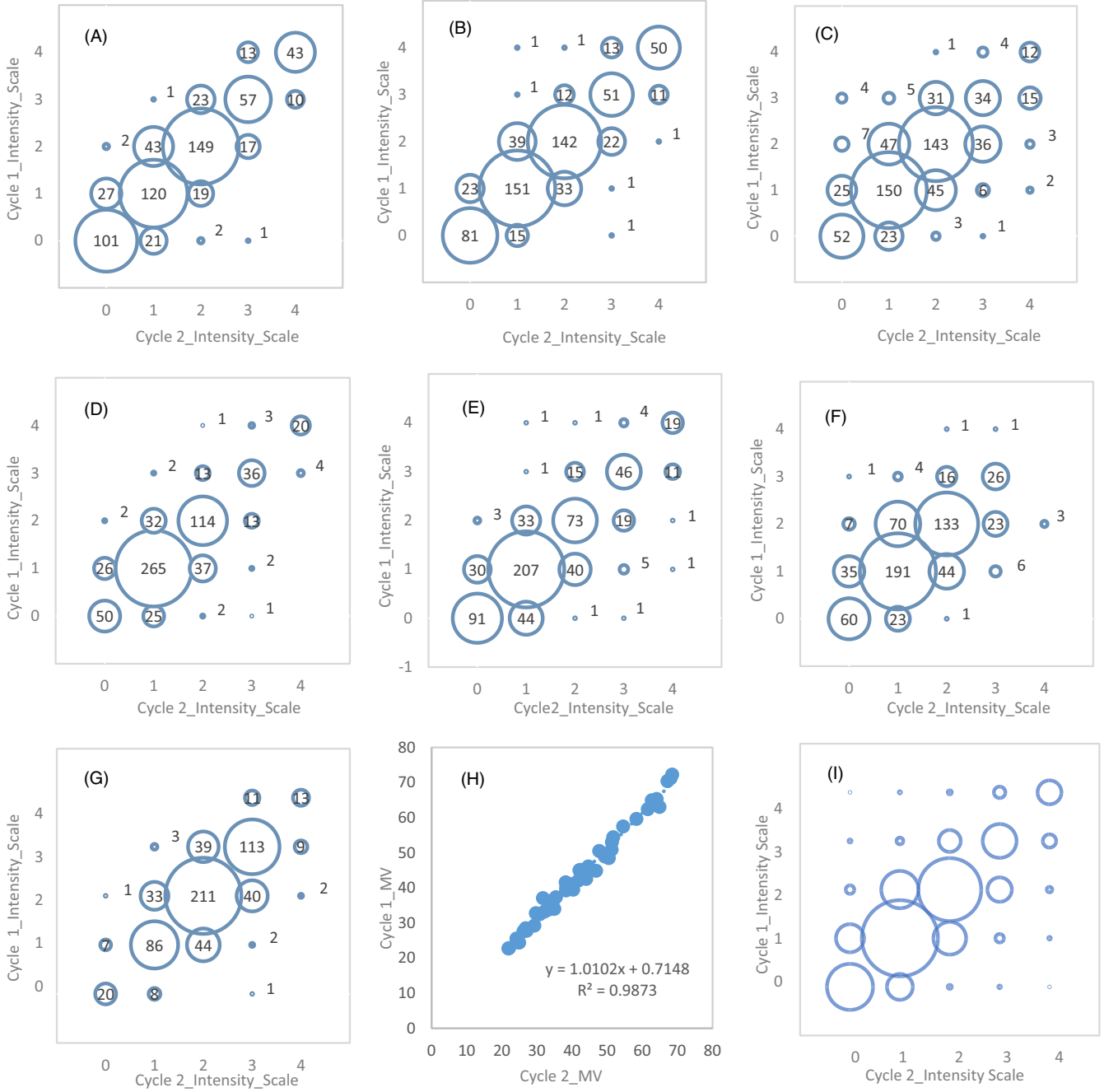

FIGURE 4 Bubble plots of the rating combinations indicating intrarater reliability. Bubble plots illustrate the relationship of variables of a scatter plot. Bubble size is proportional to the frequency of rating combinations between rating point 1 (Cycle 1) and rating point 2 (Cycle 2). The reliability is high if the bubbles are located along the diagonal and low if the bubbles are scattered randomly in the plot. Both axes (Cycle 1 + Cycle 2) are labeled with the scale severity grades between 0–4. Short forms: (A) elasticity scale), (B) wrinkle scale, (C) skin surface roughness scale, (D) pigmentation scale, (E) erythema scale, (F) pore size scale, (G), overall skin quality, (H) age, (I) SASSQ—aged skin unit

experiences in using of the scales in esthetic practice and science to test their sensitivity. The assessments of real skin in 3D-modus could further result in even more severe intense grades than the assessments of 2-D pictures. A useful addition could be the analysis of male skin and skin deriving from different ethnic groups to elucidate differences in aging.

In summary, our data present statistically significant good or very good results to use the SASSQ as tool for skin quality "photo" and live assessments, whose easy use and adequate parameters were confirmed by all experts. The two rating cycles took place with time lag between two and 4 weeks to prevent memory effects. This time distance corresponds to the time period of common follow-up appointments for treatment evaluations and fresh-ups in esthetic practice as well as the validation process itself of other established instruments,[27,28] which are used for live as well as photo assessments.

TABLE 5 Shrout-Fleiss estimates; test-retest reliability for the intrarater reliability

Intrarater reliability							
		95% CI (MV)					
Parameter	ICC MV Shrout-Fleiss	Min	Max	Strength of agreement	PCC (*r*)	*p*	Strength of relationship
Elasticity	0.883	0.796	0.932	Almost perfect	0.889	0.000	Very high
Wrinkles	0.885	0.804	0.933	Almost perfect	0.890	0.000	Very high
SSR	0.645	0.469	0.773	Substantial	0.657	0.026	High
Pigmentation	0.797	0.673	0.878	Substantial	0.801	0.000	Very high
Erythema	0.797	0.660	0.881	Substantial	0.812	0.000	Very high
Pore Size	0.584	0.369	0.738	Moderate	0.605	0.001	High
Age	0.925	0.865	0.957	Almost perfect	0.916	0.000	Very high
Overall SQ	0.756	0.610	0.853	Substantial	0.767	0.000	Very high

In conclusion, the developed SASSQ with the parameters elasticity, wrinkles, skin surface roughness, pigmentation, erythema, and pore size represents a valid, objective, and global photonumeric 5-point scale for the evaluation of aged female facial skin quality for use in science and in practice.

ACKNOWLEDGMENTS

We thank the following collaborators which were involved in the validation process of the scale: Braun K MSC, Buntrock H PHD, Gout U MD, Kleine-Börger L MSC, Mader-Brodmeier J DSC, Ogilvie A MD, Pavicic T MD, Reuther T MD, Sattler G MD, Sattler S MD, Steinkraus S MD, Streker M PHD, Weis D PHD. All collaborators agreed in the publication of the script. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

Both authors are employees of the University of Hamburg. Apart from this, the authors declare no conflicts of interest that are directly relevant for the content of this publication. There was no external sponsor or funding involved in the project.

SUBMISSION STATEMENT

This paper is exclusively submitted to the *Journal of Cosmetic Dermatology* and has not been and will not be published elsewhere.

DATA AVAILABILITY STATEMENT

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

ORCID

Christine Eiben-Nielson https://orcid.org/0000-0002-9949-6215

REFERENCES

1. Cotterill J. What is really true about the cosmetic industry? *Int J Dermatol*. 1988;27(10):682-683.
2. Moeran B. The portrayal of beauty in women's fashion magazines. *Fashion Theory*. 2010;14(4):491-510.
3. Gold M. The future of non-invasive rejuvenation technology: devices. *JDD*. 2017;16(6):104-107.
4. Honigman R, Castle DJ. Aging and cosmetic enhancement. *Clin Interv Aging*. 2006;1(2):115-119.
5. Baumann L. How to use oral and topical cosmeceuticals to prevent and treat skin aging. *Facial Plast Surg Clin North Am*. 2018;26(4):407-413.
6. Puizina-Ivić N. Skin aging. *Acta Dermatoven APA*. 2008;17(2):47-54.
7. Morris D. *The Naked Ape: A Zoologist's Study of the Human Animal. Third Impression Edition*. New York, NY: McGraw-Hill; 1967.
8. Jackson LA. *Physical Appearance and Gender: Sociobiological and Sociocultural Perspectives*. Albany/USA: SUNY Press; 1992.
9. Barber N. The evolutionary psychology of physical attractiveness: sexual selection and human morphology. *Ethol Sociobiol*. 1995;16:395-424.
10. Mueser KT, Grau BW, Sussman S, Rosen AJ. You're only as pretty as you feel: facial expression as a determinant of physical attractiveness. *J Pers Soc Psychol*. 1984;46:469-478.
11. Lutz J, Kemper CJ, Beierlein C, Margraf-Stiksrud J, Rammstedt B. Development and validation of a single-item scale for the relative assessment of physical attractiveness. The Attractiveness Rating 1 (AR1). *mda*. 2013;7(2):209-232.
12. Fink B, Penton-Voak I. Evolutionary psychology of facial attractiveness. *Curr Dir Psychol Sci*. 2001;11:154-158.
13. Eiben-Nielson C. *Scientific Assessment Scale of Skin Quality (SASSQ): Estabilishing of Relevant Conditions and Parameters with Evaluation within an Interdisciplinary Research Approach*. Dissertation. Hamburg: University of Hamburg; 2019.
14. Krutmann J, Gilchrest BA. *Skin Aging*. Berlin/Heidelberg: Springer; 2006.
15. Rhie G, Shin MH, Seo JY, et al. Aging- and photo-aging-dependent changes of enzymatic and nonenzymatic antioxidants in the epidermis and dermis of human skin in vivo. *J Invest Dermatol*. 2002;118:1212-1217.
16. Farage MA, Miller KW, Maibach HI. Degenerative Changes in Aging Skin. In: Farage MA, Miller KW, Maibach HI, eds. *Textbook of Aging Skin*, 2nd edn. Berlin/Heidelberg: Spinger Verlag; 2017:15-30.
17. Carruthers A, Carruthers J. A validated facial grading scale: the future of facial ageing measurement tools? *J Cosmet Laser Ther*. 2012;12:235-241.
18. Holle R. *Methods for the Construction and Evaluation of clinical Scores*. Habilitation thesis. Heidelberg: Ruprecht-Karls-University; 1995.
19. Dobos G, Lichterfeld A, Blume-Peytavi U, Kottner J. Evaluation of skin ageing: a systematic review of clinical scales. *Br J Dermatol*. 2015;172(5):1249-1261.

20. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010;19(4):539-549.
21. Tsukahara K, Takema Y, Kazama H, et al. A photographic scale for the assessment of human facial wrinkles. *J Cosmet Sci*. 2000;51:127-139.
22. Leal Silva HG. Facial laxity rating scale validation study. *Dermatol Surg*. 2016;42:370-379.
23. Carruthers A, Donofrio L, Hardas B, et al. Development and validation of a photonumeric scale for evaluation of facial fine lines. *Dermatol Surg*. 2016;42:227-234.
24. Honeck P, Weiss C, Sterry W, Rzany B, Gladys Study Group. Reproducibility of a four-point clinical severity score for glabellar frown lines. *Br J Dermatol*. 2003;149:306-310.
25. Lorenc ZP, Bank D, Kane M, Lin X, Smith S. Validation of a four-point photographic scale for the assessment of midface volume loss and/or contour deficiency. *Plast Reconstr Surg*. 2012;130(6):1330-1336.
26. Allerhand M, Ting Ooi E, Starr RJ, et al. Skin aging and oxidative stress in a narrow-age cohort of older adults. *Eur Geriatr Med*. 2011;2(3):140-144.
27. Merz Pharma UK Ltd (Hrsg.). *The Merz Aesthetics Scales™*. Hertfordshire: Merz Pharma UK Ltd.; 2018. https://www.merzaesthetics.me/scales/.
28. Rzany B, Carruthers A, Carruthers J, et al. Validated composite assessment scales for the global face. *Dermatol Surg*. 2012;38:294-308.
29. WMA Declaration of Helsinki. Ethical Principles for Medical Research involving Human Subjects. Helsinki, Finland, 1964, 64th WMA General Assembly, Brazil 2013.
30. Koo TK, Li MY. A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155-163.
31. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174.

How to cite this article: Eiben-Nielson C, Kerscher M. Development and validation of a global photonumeric scale for evaluating skin quality of aged female facial skin. *J Cosmet Dermatol*. 2021;20:4032–4039. https://doi.org/10.1111/jocd.14058