Aesthetic Surgery Journal Advance Access published November 24, 2015

Letter to the Editor

The Rainbow Scale: A Simple, Validated Online Method to Score the Outcome of Aesthetic Treatments

Aesthetic Surgery Journal
2015, 1–3

DOI: 10.1093/asj/sjv197
www.aestheticsurgeryjournal.com
OXFORD UNIVERSITY PRESS

Joris A. van Dongen, BSc; Ben M. Eyck, BSc; Berend van der Lei, MD, PhD; and Hieronymus P. Stevens, MD, PhD

Photographs have become an important integral part of documentation in an aesthetic surgery practice, both for medico-legal as well as scientific purposes.[1] Only by evaluating pre- and postoperative photographic documentation are we able to evaluate our success and treatment outcomes in an objective way. Several photographic scales have already been developed for this purpose.[2-4] However, an easy online system to evaluate aesthetic outcomes in a validated, fast, reliable, and standardized way is still lacking.

We have therefore designed what we call the "Rainbow Scale" method based on the Merz Scales (Merz Pharmaceuticals GmbH, Frankfurt, Germany) (Figure 1).[2] In this online system, 6 photographs are presented in a block of two rows of 3 photographs. The photograph of the patient to be evaluated is located in the central position of the lower row. Other photographs show the five grades of severity of the area to be evaluated, with the lowest score left of the photograph to be evaluated, the second score on the upper left side, etc. until the highest score on the lower right side of the photograph. The different scores are thus presented around the patient in a rainbow fashion. In this way, the result of an aesthetic treatment can be scored easily and fast. Both pre- and post-treatment situations can be scored objectively and can be compared with other patients and treatment methods.

We evaluated this system by using it to score the nasolabial fold as a target zone to determine whether our proposed scoring system is easy, reliable, and reproducible.

Using an already validated photographic series of Narins et al[2] with 5 scores of severity of the nasolabial fold, we tested the ease as well as the reliability of our Rainbow Scale in a series of patients with regard to the aesthetic result of the nasolabial fold depth. An online survey was created in SurveyMonkey (SurveyMonkey, Palo Alto, CA). A sample copy of the survey can be viewed as Supplementary Material at www.aestheticsurgeryjournal.com.

The survey contained 15 pre- and postoperative photographs of female patients from the senior authors' database consisting of 5 sets of 3 digital photos per grade, randomly presented throughout the survey in one-page questions. Panelists did not know whether the photographs where pre- or postoperative. Computer randomization was performed with Microsoft Excel Version 14.4.9. (Microsoft, Redmond, WA). Each photograph was edited in the same manner as the reference photographs, in anteroposterior view.

In each question, panelists were asked to rate the photograph in the evaluation spot according to the appropriate corresponding grade on the Rainbow Scale. From May 2015 to June 2015, a group of 5 independent experienced plastic surgeons completed the survey 3 times, but the second and third time the same questions with the same photographs in the survey were randomized in a different order with respect to the previous survey (Table 1). Between each survey there was an interval of at least 2 days. SurveyMonkey tracked scoring time for each plastic surgeon automatically.

Intra- and inter-observer agreements were expressed as a weighted kappa coefficient for all surveys.[5] Statistical analyses were performed using SPSS Statistics version 21.0 (IBM, New York, NY).

Mr van Dongen is a PhD Candidate and Dr van der Lei is a Professor in Cosmetic Surgery, Department of Plastic Surgery, University of Groningen and University Medical Centre Groningen, Groningen, the Netherlands. Mr Eyck is a Student at the Faculty of Medicine, Erasmus University, Rotterdam, the Netherlands. Dr Stevens is a plastic surgeon in private practice, Den Haag, the Netherlands.

Corresponding Author:
Mr Joris A. van Dongen, Boterdiep 94-23, Groningen 9712 LS, The Netherlands.
E-mail: jorisavandongen@gmail.com

Figure 1. The Rainbow Scale based on the Merz Scales (Merz Pharmaceuticals GmbH, Frankfurt, Germany). The photograph of the patient is located in the central position of the lower row. Five grades of the Nasolabial fold Assessment Scale are placed around progressively as a rainbow. Originally published in Narins et al[2] reprinted with permission from Wolters Kluwer Health, Inc.

Table 1. Demographic Information of the Panelists

	Panelists (N = 5)
Male	4
Female	1
Mean age (years)	36 (range, 31-56)
Mean experience time (years)	9 (range, 4.5-20)

Table 2. The Agreement is Expressed as a Weighted Kappa Coefficient

	Intra-Observer Agreement Weighted Kappa Coefficient				
	Plastic Surgeon 1	Plastic Surgeon 2	Plastic Surgeon 3	Plastic Surgeon 4	Plastic Surgeon 5
S1 vs S2	0.79	0.70	0.92	0.88	0.91
S1 vs S3	0.83	0.90	0.88	0.89	0.81
S2 vs S3	0.82	0.92	0.92	0.82	0.90
Mean	0.81	0.84	0.91	0.86	0.87

S1, survey 1; S2, survey 2; S3, survey 3.

Table 3. The Agreement is Expressed as a Weighted Kappa Coefficient

	Inter-Observer Agreement Weighted Kappa Coefficient				
	Plastic Surgeon 1	Plastic Surgeon 2	Plastic Surgeon 3	Plastic Surgeon 4	Plastic Surgeon 5
Plastic surgeon 1		S1 = 0.63	S1 = 0.68	S1 = 0.81	S1 = 0.86
		S2 = 0.84	S2 = 0.78	S2 = 0.78	S2 = 0.82
		S3 = 0.68	S3 = 0.73	S3 = 0.88	S3 = 0.86
Plastic surgeon 2	S1 = 0.63		S1 = 0.76	S1 = 0.80	S1 = 0.73
	S2 = 0.84		S2 = 0.80	S2 = 0.73	S2 = 0.81
	S3 = 0.68		S3 = 0.86	S3 = 0.81	S3 = 0.69
Plastic surgeon 3	S1 = 0.68	S1 = 0.76		S1 = 0.80	S1 = 0.75
	S2 = 0.78	S2 = 0.80		S2 = 0.84	S2 = 0.93
	S3 = 0.73	S3 = 0.86		S3 = 0.68	S3 = 0.78
Plastic surgeon 4	S1 = 0.81	S1 = 0.80	S1 = 0.80		S1 = 0.89
	S2 = 0.78	S2 = 0.73	S2 = 0.84		S2 = 0.73
	S3 = 0.88	S3 = 0.81	S3 = 0.68		S3 = 0.83
Plastic surgeon 5	S1 = 0.86	S1 = 0.73	S1 = 0.75	S1 = 0.89	
	S2 = 0.82	S2 = 0.81	S2 = 0.93	S2 = 0.73	
	S3 = 0.86	S3 = 0.69	S3 = 0.78	S3 = 0.83	

S1, survey 1; S2, survey 2; S3, survey 3.

Table 4. Time to Take One Survey

	Survey 1	Survey 2	Survey 3
Plastic surgeon 1	163 s	162 s	143 s
Plastic surgeon 2	108 s	130 s	331 s
Plastic surgeon 3	339 s	1505 s	429 s
Plastic surgeon 4	112 s	112 s	102 s
Plastic surgeon 5	297 s	147 s	129 s

The mean intra-observer agreement was 0.86 (Table 2). Mean inter-observer agreements within pairs of plastic surgeons were 0.77 for survey 1, 0.81 for survey 2, and 0.78 for survey 3 (Table 3). The average time to take a survey was 195 seconds (range, 102-429 seconds). The average time to rate one photograph was 13 seconds. This relatively short time is potentially due to the ease with which each patient could be evaluated by comparison to the 5 scales presented around in a rainbow fashion (Table 4). One scoring time of 1505 seconds was excluded, because the surgeon took a break during the survey.

Based upon our experience with the Rainbow Scale, and in the past having used a scoring system on paper, we found that this scoring system can be accessed faster and more easily. Although no statistical differences in scoring time between our online method and other methods described in the literature are presented. The arc design of the Rainbow Scale also gives the ability to judge photos better. Moreover, with this Rainbow Scale, the data can be directly translated into programs such as SPSS, thereby reducing the risk of loss

of information. The data can then directly be analyzed statistically. Validation of this scale allows for the creation of an online standardized setup that can be used in other aesthetic outcome measurements such as evaluating other facial wrinkles or breast ptosis. Last but not least, many potential panelists can be reached online by offering a link to the online questionnaire in an email or on a website.

Descriptive information could be added to further improve the validation of the scale. Reliability between 2 observers improves by using a combination of photographic grading and descriptive information.[4]

In conclusion, the Rainbow Scale is a reliable and objective online questionnaire for the assessment of any aesthetic feature. It proved to be a reliable measurement with high intra- and inter-observer agreements that can be performed in an easy online fashion.

Supplementary Material

This article contains supplementary material located online at www.aestheticsurgeryjournal.com.

Disclosures

The authors declared no potential conflicts of interest with respect to the research, authorship, and publication of this article.

Funding

The authors received no financial support for the research, authorship, and publication of this article.

REFERENCES

1. Galdino GM, Swier P, Manson PN, Vander Kolk CA. Converting to digital photography: a model for a large group or academic practice. *Plast Reconstr Surg*. 2000;106(1):119-124.
2. Narins RS, Carruthers J, Flynn TC, et al. Validated assessment scales for the lower face. *Dermatol Surg*. 2012;38(2 Spec No.):333-342.
3. Buchner L, Vamvakias G, Rom D. Validation of a photonumeric wrinkle assessment scale for assessing nasolabial fold wrinkles. *Plast Reconstr Surg*. 2010;126(2):596-601.
4. Shoshani D, Markovitz E, Monstrey SJ, Narins DJ. The modified Fitzpatrick Wrinkle Scale: a clinical validated measurement tool for nasolabial wrinkle severity assessment. *Dermatol Surg*. 2008;34(Suppl 1):S85-S91; discussion S91.
5. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas*. 1973;33(3):613-619.