ORIGINAL ARTICLE

Quantitative method for evaluation of aesthetic results after laser treatment for birthmarks

Pawel Szychta · Khalil Al-Nakib · Will Anderson ·
Ken Stewart · Awf Quaba

Abstract Subjective assessment of results after laser treatment for birthmarks does not provide a validated method for clinicians. Previous reports concerning objective evaluation using L*a*b color coordinates were only partially successful due to difficulties in standardizing and comparing colors from pre- and posttreatment photographs. The study aimed to present a reliable and clinically applicable method of aesthetic result assessment after laser treatment for birthmarks. All 48 patients, included in the study, were treated for birthmarks on the face or neck using laser. Each pre- and posttreatment photograph was subjected to triple objective comparative assessment of color fading with use of a computer program, as well as to subjective evaluation by three core physicians and three laypeople. Objective analysis was based on an innovative method using combined L*a*b and hue saturation value color coordinates. Accuracy was higher between objective assessment with the computer program and subjective evaluation by core physicians than by laypeople. Repeatability of results was higher with the use of the computer program than among core physicians or among laypeople. In conclusion, our method may be considered for objective assessment of the results after laser treatment of vascular and pigmented birthmarks.

P. Szychta · K. Al-Nakib · W. Anderson · K. Stewart · A. Quaba
Plastic and Reconstructive Surgery Department,
St John's Hospital, Howden Road West, Livingston, West Lothian
EH54 6PP, UK

P. Szychta (✉)
Department of Oncological Surgery and Breast Diseases, Polish
Mother's Memorial Hospital-Research Institute, ul. Rzgowska,
281/289, Lodz, Poland
e-mail: szychta@yahoo.pl

Introduction

Vascular birthmarks can result in physical and psychological problems and may lead to developmental and behavioral problems [1]. Modern treatment of vascular birthmarks is with 585-nm pulsed dye laser therapy, which can lead to the restoration of normal skin but, in the majority of patients, results in only a partial improvement [2–4].

The evaluation of the treatment outcomes of birthmarks has been derived from subjective and qualitative judgments and does not provide a validated method for clinicians [5]. These subjective observations have relied primarily on visual inspection, which makes treatment effects difficult to evaluate with any precision [6]. The color perception of human eyes varies among individuals; moreover, the problems of illuminant metamerism, different lighting conditions, sensitivity spectrum, and adjustment of the eye to color may be involved in color mismatch and inaccurate comparisons [7].

Apart from lacking standardized objective methods of assessing the results after laser treatment, the grading system of the outcomes differs among institutions, influencing even more the professional communication of results between clinicians. Three-grade classification uses the following grading: excellent for lesion improvement of 75 to 100 %, good for 50 to 75 %, and fair for 25 to 50 % [8]. Four-grade system consists of the following steps: excellent (complete clearance of pigmentary lesion), good (subtotal clearance), fair (moderate clearance), and poor (minimal to no response after treatment) [9]. Perez et al. described the scoring scale 0 to 100 where value 100 stands for total response to

treatment [6]. A universally accepted system would improve the communication among clinicians and scientists.

Previous reports, concerning objective evaluation, included the following: laser Doppler, reflectance spectrophotometry, tristimulus colorimetry, and videomicroscopy [9–12]. The above methods have not been widely incorporated into clinical practice either due to high cost of equipment or time-consuming data processing. Moreover, objective methods of assessment with use of photographs could not standardize colors of the differently shadowed areas of the body existing in the same picture, making the analysis incomprehensive [13, 14].

Therefore, there is a need for a comprehensive, automated, cheap, and clinically applicable objective method to evaluate the outcome and prognosis after treatment. Digital photography with computer analysis appears to be a promising technology for the evaluation of birthmarks. Digital photography and computers are affordable and widely available in clinical practice. Photographs are routine documentation of results after laser treatment. Taking photographs avoids contact with the skin, thereby avoiding the problem of birthmark blanching seen with some techniques. However, most cameras use color presentation, called red green blue. This format is dependent on lightning conditions, device, and focusing; thus, it does not allow accurate comparison of the pre- and postoperative photographs.

This problem has been previously overcome by converting pixels in the picture into Commission International d'Eclairage (CIE) L*a*b* color scale [15]. CIELab [L*, luminance; a*, balance between green (−) and red (+); and b*, balance between blue (−) and yellow (+)] was intended to provide a standard, approximately uniform color scale which could easily compare color values [16]. CIELab is the most complete color presentation, and it describes the whole variety of colors visible to the human eye. It is claimed to be the reference as the light- and device-independent model. However, the results are still affected by the differences in shadowing between the healthy skin and the birthmark in one picture. Therefore, the studies assessing usefulness of digital two-dimensional color photographs only with CIELab were only partially successful [13, 14].

Innovative additional manipulation of skin shadowing with use of hue saturation value (HSV) color presentation can provide a validated method for objective assessment [17]. HSV is highly compatible with visual perception. Differentiation in color characteristics is useful in image processing. Therefore, HSV can be used to highlight homologously shadowed areas of the lesion and healthy skin. The present study aims to present a reliable method of objective evaluation of results after laser treatment for birthmarks, which can be clinically applicable for everyday surgical practice.

Patients and methods

For the study, we included 48 adults treated for birthmarks on the face or neck using pulsed dye laser in the Plastic and Reconstructive Surgery Department, St John's Hospital, Livingston, UK. The inclusion criteria of the retrospective analysis were the following: birthmarks on the face or neck, patients with follow-up period at least 1 year after final stage of treatment, and pre- and posttreatment digital color photographs taken only by the professional medical photographer at the Medical Photography Department in order to reduce bias to minimum. Laser treatments applied for birthmarks cause posttreatment bruising/contusion that may last for several days; thus, the outcome should be assessed at least 1 month after the treatment. In our study, we used posttreatment photographs taken during the follow-up appointment at least 1 year after the final stage of laser application.

Each pre- and postoperative photograph was subjected to evaluation by three core physicians (consultants in plastic surgery working with lasers on everyday basis) and three laypeople as well as to computerized, triple objective comparative assessment of color shading with use of the program Skin Lesion Color Change (SLCC). In all three instances, the treatment outcome was presented as percentage of success (%), where 0 % corresponded to total failure, and 100 % described total success [6]. Moreover, we performed detailed analysis in relation to location as divided by Leonardo da Vinci: forehead, location 1; middle face, location 2; lower face, location 3; and neck, location 4.

Objective computer assessment

The computer program SLCC has been created by one of the authors (PS) and is copyright protected (Fig. 1). It uses combined color presentation of CIELab with HSV. Algorithm for calculation of treatment results in SLCC consists of two main steps. The first one is conversion of color scale from red green blue (RGB) to the combined scale of CIELab and HSV through transitional standardized RGB (sRGB) and CIE XYZ color presentations (Fig. 2). The above conversion is performed in both, pre- and postoperative photographs. In the second step, the program calculates the degree of color shading between the lesion and healthy skin from the standardized, modified pre- (ΔE) and postoperative ($\Delta E'$) photographs ($\Delta E = \left[\left(\Delta L^* \right)^2 + \left(\Delta a^* \right)^2 + \left(\Delta b^* \right)^2 \right]^{1/2} \rightarrow \text{Result} = (1 - \Delta E'/\Delta E) \times 100\%$). To eliminate any outside variances such as skin tanning/lightening due to seasonal weather changes, as well as to further compensate for any ambient variances from one image to another, the patient's normal skin served as a control/reference when calculating PWS color, as previously described [14].
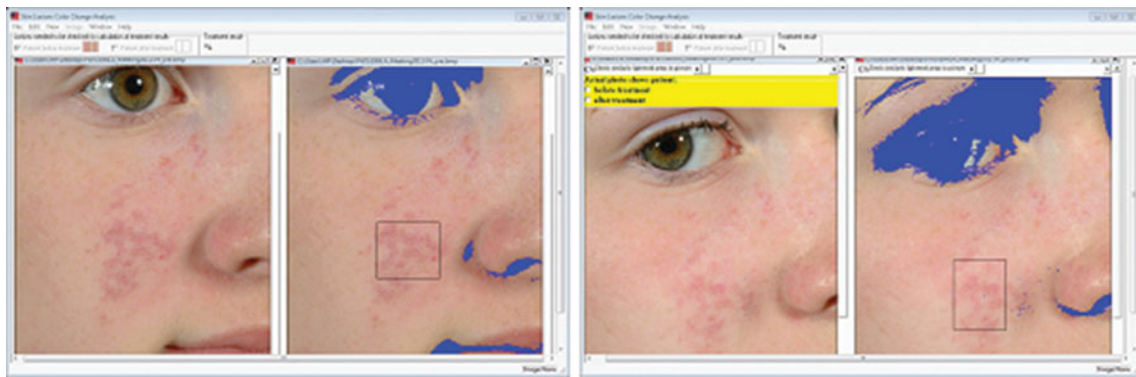
Fig. 1 Main screen of the SLCC program, presenting the process of analyzing the outcome of the laser treatment for birthmarks; **a** preoperative period, **b** state after last stage of treatment

The clinical usefulness of the method used in the program was assessed with intra- and interexaminer reliability of the results (in other words, repeatability and accuracy, respectively) [18]. Repeatability was analyzed by comparing triple evaluations of the same lesion with the specific method. Accuracy of subjective methods of assessment, by core physicians and by laypeople, was compared to results obtained with SLCC. Both accuracy and repeatability were presented as technical error measurement (TEM) and showed the potential variations in measuring the specific parameter (in units; here, in percentage). Reliability (R) was described using TEM and standard deviation of all single measurements, with scale of values from 0 weakest to 1 strongest.

Statistical tests were used to analyze the results. The normality of the groups was assessed with Shapiro–Wilk test. The parameters were compared with t test for dependent samples, when assessing two groups for accuracy, and with MANOVA, when assessing three groups for repeatability. The difference was claimed as statistically significant when $p < 0.05$.

Results and discussion

In the study, we compared objective assessment with the computer program of vascular birthmarks treated with laser, with the subjective assessment done by core physicians and by laypeople. Our results confirmed the usefulness of the proposed SLCC method. The detailed layout of the results is presented in Table 1.
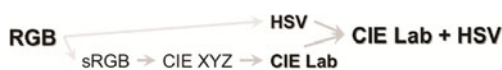

Fig. 2 Algorithm for conversion of color space from RGB to combined coordinates of CIELab and HSV, RGB, sRGB, CIE XYZ, CIELab, and HSV; different color presentations, detailed description in text

The general part of the analysis revealed that the median outcome of the treatment was 61 % when assessed with SLCC; 76.67 % was evaluated by core physicians, and 83.33 %, by laypeople. The variability of the results was similar in analysis done by core physicians and with SLCC (30.29 and 30.38 %, respectively), and was higher than that by laypeople (26.08 %). When assessed with SLCC method, most patients benefited from treatment at the level from 60 to 70 %, whereas in core physicians' opinion, the average level of treatment success was between 70 and 80 %. Distribution of the treatment outcomes assessed with SLCC and by core physicians is presented in Fig. 3. The results were in line with the previous reports, where the most commonly seen outcome after laser treatment of birthmarks was from 60 to 80 % [1, 12, 13]. However, the results of assessment with each of three methods were different with one another; thus, a more detailed analysis of reliability was undertaken.

Validation of usefulness of the three methods was performed in the study with use of parameters TEM and R, characterizing repeatability and accuracy. The repeatability of results was higher with use of computer program in SLCC method ($R = 0.99$) than among core physicians ($R = 0.81$) or among laypeople ($R = 0.63$) (Fig. 4). In another words, the results obtained after multiple objective assessments were almost identical to one another, when assessed with SLCC (TEM = 1.49 %, $p > 0.05$). Additionally, they were less similar among core physicians (TEM = 9.91 %, $p < 0.01$) and differed even more among laypeople (TEM = 13.84 %, $p < 0.0001$). It is difficult to compare our results to previous studies because other authors used only correlation coefficients to describe the interexaminer reliability, which narrowed the possibilities to validate their analysis [13, 14].

The accuracy of the results was also taken into consideration in our study. There was a stronger correlation of the results between objective assessment with SLCC and subjective evaluation performed by core physicians ($R = 0.54$)

Table 1 The detailed layout of analyses of the treatment results, performed by core physicians, laypeople, and with SLCC

Method of assessment	Parameter	Percentage of success (%)							Repeatability (R) Accuracy (A), compared to SLCC				R/A
		Mean	Median	Min	Max	QR/2	SD	v (%)	TEM	R	F (R) t (A)	p	
Core physician	Respondent 1	73.54	85.00	0.00	100.00	12.50	25.56	34.75					
	Respondent 2	69.38	75.00	10.00	95.00	13.75	20.85	30.05	9.91	0.81	5.74	<0.01	R
	Respondent 3	67.19	70.00	10.00	100.00	17.50	21.01	31.27					
Core physicians	Overall	70.03	76.67	6.67	98.33	12.92	21.21	30.29	13.81	0.54	4.32	<0.0001	A
Lay person	Respondent 1	75.21	85.00	0.00	100.00	15.00	25.35	33.70					
	Respondent 2	81.88	90.00	30.00	100.00	15.00	18.30	22.35	13.84	0.63	13.34	<0.0001	R
	Respondent 3	70.00	80.00	20.00	100.00	17.50	22.60	32.28					
Laypeople	Overall	75.69	83.33	30.00	100.00	13.33	19.74	26.08	17.29	0.29	5.98	<0.0001	A
SLCC	Measurement 1	59.69	61.00	20.00	91.00	12.50	18.49	30.98					
	Measurement 2	59.46	61.00	20.00	90.00	12.25	18.20	30.60	1.49	0.99	0.40	>0.05	R
	Measurement 3	59.73	61.00	22.00	93.00	11.75	17.77	29.75					
SLCC	Overall	59.63	61.00	21.00	91.33	12.25	18.11	30.38	Results with SLCC used as reference when compared to other methods				
Core physicians	Location 1	71.67	70.00	51.67	95.00	18.33	21.60	30.14	7.39	0.79	0.18	>0.05	A
	Location 2	66.16	71.67	16.67	98.33	15.83	20.41	30.85	15.34	0.45	3.11	<0.01	A
	Location 3	70.33	82.50	6.67	91.67	5.83	28.50	40.52	13.86	0.66	1.89	>0.05	A
	Location 4	76.39	78.33	61.67	90.00	8.33	10.56	13.83	9.90	0.42	2.31	>0.05	A
Laypeople	Location 1	75.83	78.33	60.00	86.67	10.83	13.16	17.35	3.51	0.91	1.37	>0.05	A
	Location 2	71.88	80.00	30.00	100.00	16.67	22.58	31.41	20.10	0.22	3.74	<0.005	A
	Location 3	78.00	85.00	33.33	96.67	5.00	21.09	27.04	15.84	0.45	4.11	<0.005	A
	Location 4	81.67	83.33	63.33	96.67	15.00	14.26	17.46	13.61	0.25	2.95	<0.05	A
SLCC overall	Location 1	72.75	73.00	58.00	87.00	9.25	12.28	16.89	–				
	Location 2	54.17	57.00	20.00	87.00	15.00	19.58	36.15					
	Location 3	59.90	64.00	23.00	84.00	11.00	18.36	30.66					
	Location 4	66.33	69.50	40.00	78.00	6.50	14.21	21.42					

than between objective assessment with SLCC and subjective evaluation done by laypeople ($R=0.29$) (Fig. 5). It means that the results calculated with the computer program and SLCC were less different to the assessment done by core physicians (TEM=13.81 %, $p<0.0001$) than by laypeople (TEM=17.29 %, $p<0.0001$). Both results obtained with SLCC and by core physicians were characterized by similar distribution of treatment outcomes (Fig. 3).

Fig. 3 Distribution of treatment outcomes among patients, assessed with SLCC and by core physicians
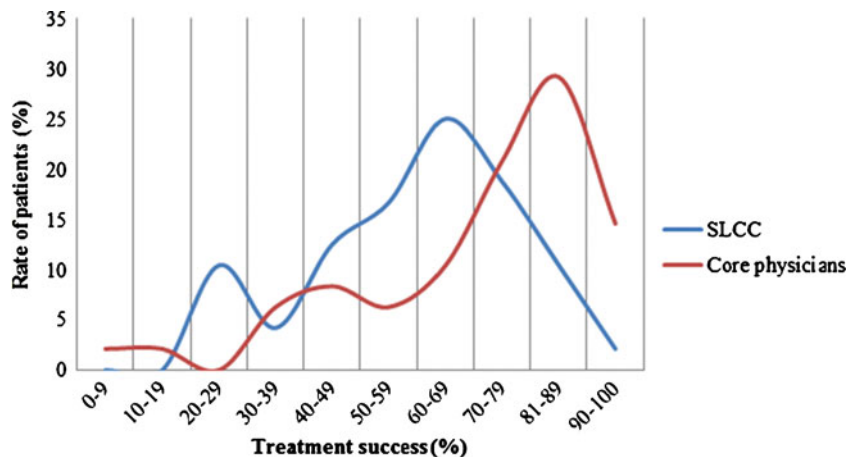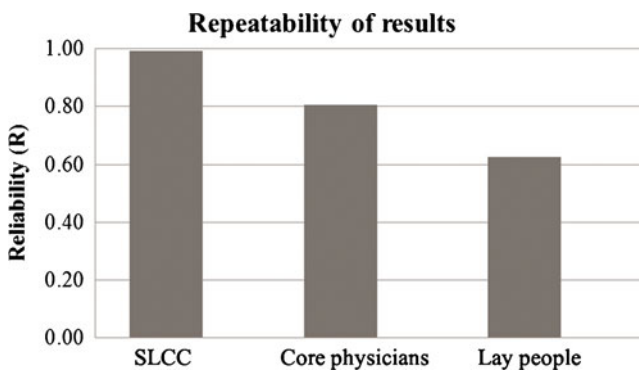
Fig. 4 Repeatability [presented as reliability (R)] of the results obtained by core physicians, laypeople, and with SLCC
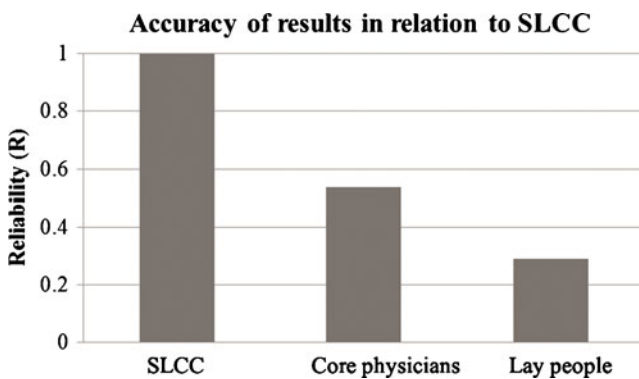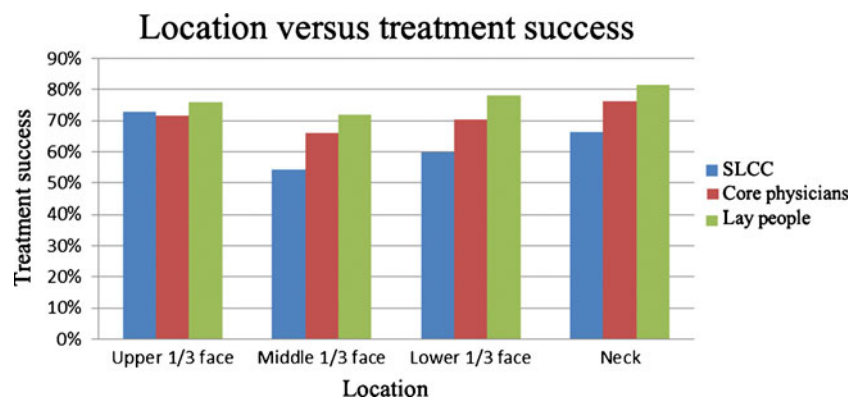

Fig. 5 Accuracy (presented as reliability) of the results obtained by core physicians, laypeople, and with SLCC

However, it should be noted that the core physicians gave slightly more favorable opinion concerning the treatment results than SLCC. Similarly, the previously described studies concerning usefulness of CIELab in assessing laser treatment could not be taken into comparison because they used only correlation coefficient [13, 14].

The detailed analysis in relation to the location revealed that core physicians, laypeople, and SLCC differed in results in location 2 (71.67, 80.00, and 57.00 %, respectively; $p<0.05$), location 3 (82.50, 85.00, and 64.00 %, respectively; $p<0.05$), and location 4 (78.33, 83.33, and 69.50 %, respectively; $p<0.05$), whereas the outcomes were assessed similarly in the case of lesions in location 1 (70.00, 78.33, and 73.00 %, respectively; $p>0.05$) (Table 1). Generally, core physicians gave higher scores in most areas, and laypeople were even more optimistic. This corresponds well to the previous reports [13, 14].

In relation to lesion location, analysis of accuracy of results obtained with SLCC objective method in comparison to the subjective methods revealed that birthmarks in location 1 were analyzed by core physicians ($R=0.79$, TEM=7.39 %, $p>0.05$) and by laypeople ($R=0.91$, TEM=3.51 %, $p>0.05$) similarly to SLCC method; the results did not differ significantly from analysis with the computer method. In contrast, in the case of lesions present in locations 2, 3, and 4, the accuracy of subjective methods was significantly lower than SLCC ($p<0.05$) (Table 1). It is noticeable that the results of treatment of lesions localized on the forehead were almost identical in three groups (Fig. 6). We suggest that the accurate subjective assessment of lesions localized on the forehead was possible because of the usually good light exposure of this area of the face with almost absent skin shadowing. In conclusion, the SLCC method may be considered for objective assessment of the results after laser treatment for vascular and pigmented birthmarks.

Fig. 6 Comparative analysis of the treatment outcomes in relation to the location of the birthmark

Conflict of interest   None

References

1. Lande RG, Crawford PM, Ramsey BJ (2001) Psychosocial impact of vascular birthmarks. Facial Plast Surg Clin North Am 9:561–567
2. Chen JK, Ghasri P, Aguilar G, van Drooge AM, Wolkerstorfer A, Kelly KM, Heger M (2012) An overview of clinical and experimental treatment modalities for port wine stains. J Am Acad Dermatol 67(2):289–304
3. Rubin IK, Farinelli WA, Doukas A, Anderson RR (2012) Optimal wavelengths for vein-selective photothermolysis. Lasers Surg Med 44:152–157
4. Ting PT, Rao J (2011) Vascular lesions. Curr Probl Dermatol 42:67–80
5. Faurschou A, Olesen AB, Leonardi-Bee J, Haedersdal M (2011) Lasers or light sources for treating port-wine stains. Cochrane Database Syst Rev 11:CD007152
6. Perez B, Abraira V, Nunez M, Boixeda P, Perez Corral F, Ledo A (1997) Evaluation of agreement among dermatologists in the assessment of the color of port wine stains and their clearance after treatment with the flashlamp-pumped dye laser. Dermatology 194:127–130
7. Sicotte M, Ledoux M, Zunzunegui M, Aboubacrine S, Nguyen A, ATARAO group (2010) Reliability of anthropometric measures in a longitudinal cohort of patients initiating ART in West Africa. BMC Med Res Methodol 10:102–111
8. Goldman MP, Fitzpatrick RE, Ruiz-Esparza J (1993) Treatment of port-wine stains (capillary malformation) with the flashlamp-pumped pulsed dye laser. J Pediatr 122:71–77
9. Troilius AM, Ljunggren B (1995) Reflectance spectrometry in the objective assessment of dye laser-treated port-wine stains. Br J Dermatol 132:245–250
10. Nelson JS, Kelly KM, Zhao Y, Chen Z (2001) Imaging blood flow in human port-wine stain in situ and in real time using optical Doppler tomography. Arch Dermatol 137:741–744
11. Pickering JW, Mordon SR, Brunetaud JM (1992) The objective reporting of laser treatment of port wine stains. Laser Med Sci 7:415–421
12. Le KV, Shahidullah H, Frieden IJ (1999) Review of modern techniques in detecting port-wine stain response to laser therapy. Dermatol Surg 25:127–132
13. Fairchild MD (1998) Color appearance models. Addison Wesley, Reading, pp 63–96
14. Rah DK, Kim SC, Lee KH, Park BY, Kim DW (2001) Objective evaluation of treatment effects on port-wine stains using L*a*b* color coordinates. Plast Reconstr Surg 108:842–847
15. Liu A, Moy RL, Victor Ross E, Hamzavi I, Ozog DM (2012) Pulsed dye laser and pulsed dye laser-mediated photodynamic therapy in the treatment of dermatologic disorders. Dermatol Surg 38:351–366
16. Naran S, Gilmore J, Deleyiannis FW (2008) The assessment of port wine stains in children following multiple pulsed-dye laser treatment. Ann Plast Surg 60:426–430
17. Cantrell K, Erenas MM, de Orbe-Payá I, Capitán-Vallvey LF (2010) Use of the hue parameter of the hue, saturation, value color space as a quantitative analytical parameter for bitonal optical sensors. Anal Chem 82:531–542
18. Lahti A, Kopola H, Harila A, Myllyla R, Hannuksela M (1993) Assessment of skin erythema by eye, laser Doppler flowmeter, spectroradiometer, two-channel erythema meter and Minolta chroma meter. Arch Dermatol Res 285:278–282