OPEM
www.opem.org

Oriental Pharmacy and Experimental Medicine 2006 **6(2)**, 112-120

Objective evaluation of the color of tongue substance using $L^*a^*b^*$ color coordinates

Young-Jae Park[1,*] and Young-Bae Park[2]

[1]*Department of Acupuncture & Moxibustion, College of Oriental Medicine, Se-Myung University, Je Chun, Chungbuk, 390-711, Korea;* [2]*Department of Biofunctional Medicine and Diagnosis, College of Oriental Medicine, Kyung-Hee University, Korea*

SUMMARY

The purpose of this study was to analyze whether quantitative evaluation of the color of the tongue substance using $L^*a^*b^*$ color coordinates system could minimize the problems arising from the different illuminating conditions or not. In controlled 4 different illuminating conditions (by natural light, flashlight, f-number, shutter speed), 12 healthy subjects were photographed of their tongue substance through a digital camera (C-2100uz, Olympus Co.), both on the top surface and on the bottom surface of the tongue substance by two examiners, twice at 3 day intervals. Clinician evaluation was also performed grading the redness of the tongue substance in the form of 5-points scale by 6 clinicians. As a result, there was no significant difference in color differences between the color of the tongue substance and the reference red card in the 4 different illuminating conditions. Intra-rater reliability was satisfied and even though limitedly, inter-rater reliability was satisfied. Color differences were significantly correlated with the results by the clinicians, although they were applicable limitedly to specific illuminating conditions. Our results indicate that the application of the color differences in tongue diagnosis could not only evaluate the color information quantitatively, but also minimize the problems arising from the different illuminating conditions and that there was the significant difference in the visual evaluation of the red color of the tongue substance, both between the clinicians and between the illuminating conditions.

Key words: Tongue diagnosis; Quantitative evaluation; Color differences; $L^*a^*b^*$; Tongue substance; Reference color; Chinese medicine

INTRODUCTION

Tongue diagnosis is a principal inspection process in chinese medicine, which enables the clinicians to seize the patients' condition of Qi, blood, blood stasis, phlegm, indigestion etc.. Inspection of the tongue substance is characterized by its color, luster, shape and movement and that of the tongue coating is characterized by its color and textural properties such as thickness (Chiu, 2000). Traditionally tongue diagnosis has been performed by the clinicians' subjective evaluation through visual inspection on the patient's tongue in natural light. For traditional tongue diagnosis is subjective and qualitative, an objective method to quantify the inspection parameters, especially the color of the tongue substance is needed.

Several techniques such as colorimeter, spectrometer

*Correspondence: Young-Jae Park, Department of Acupuncture & Moxibustion, College of Oriental Medicine, Se-Myung University, Je Chun, Chungbuk, 390-711, Korea. E-mail: bmpomd@paran.com

and glossoscopy have been utilized to evaluate quantitatively the color of the tongue substance and the tongue coating (Zhang *et al.*, 1992; Han *et al.*, 1993; Cui *et al.*, 2001) but these techniques are limited by cost, convenience and illuminating condition. Current quantitative studies of tongue diagnosis include an objective evaluation of the color of the tongue substance using RGB[R: red, G: green, B: blue] and $L^*a^*b^*$[L^*: luminance, a^*: balance between green(-) and red(+), b^*: balance between blue(-) and yellow(+)] color scale. RGB and $L^*a^*b^*$ values can be extracted from the digitalized images of the tongue substance, photographed by digital camera or scanned from slide films. Although some colors of the tongue substance mentioned in traditional medical literature can be expressed quantitatively by means of RGB and $L^*a^*b^*$ values (Han *et al.*, 1993; Cui *et al.*, 2001), there still exist problems arising from different illuminating conditions during taking photographs. Digitalized tongue images from different illuminating conditions results in different RGB or $L^*a^*b^*$ values in the same subject, and makes it difficult to communicate color information.

Rah et al. evaluated quantitatively the treatment effects of portwine stains using $L^*a^*b^*$ color coordinate system in combination with a personal computer (Rah *et al.*, 2001). They reported that problems arising from different illuminating condition could be minimized by $L^*a^*b^*$ color coordinate system.

Although color analysis using color differences can minimize the variation resulted from the different illuminating conditions, color differences between the normal skin and the specific lesion is a relative color information and it is difficult to compare the absolute color among the patients using color differences extracted from the normal skin and the lesion without a neutral reference color. Bengel WM applied a piece of gray card as a neutral reference object to the assessment of therapeutic results after bleaching procedures and indicated that modulating the brightness of the digitalized images based on the gray card photographed simultaneously with the lesion made it possible to compare the absolute color information among the patients' lesion (Bengel, 2003).

In this paper, we developed a system to extract the color differences between the subjects' tongue substance and the reference red card using $L^*a^*b^*$ color coordinate system and analyzed whether the application of color differences using $L^*a^*b^*$ color coordinate system to tongue diagnosis could minimized problem arising from different illuminating condition during taking photographs of the tongue substance or not.

We took photographs of 12 subjects' tongue substance with the reference red color card in controlled 4 illuminating conditions. After taking photographs, we calculated color differences through a personal computer, and analyzed whether there was a significant difference between the color differences extracted from 4 different illuminating conditions or not. We also performed the clinician evaluation of the redness of the tongue substance using Likert-like 5 points scale.

MATERIALS AND METHODS

Subjects

The subjects included 6 healthy male and 6 healthy female volunteers ranging in age from 22 to 32 years (Mean 26.9). None of the subjects was taking any medication and being contracted with any disease resulting in fever, chilling, diarrhea etc., which were recognized to result in changing of the color of the tongue substance in chinese medicine. Informed consent was obtained from all of them.

Taking photographs

The examiners included 2 oriental medical doctors, and the assistants also included 2 another oriental medical doctors. The subjects sit on a chair by the window with a vertical blind and each examiner sit opposite to the subjects. The subjects were requested to expose their tongue substance according to traditional tongue diagnosis process. Two assistants

Table 1. Illuminating conditions by the light sources, the Shutter speed and the F-number of the digital camera*

Illuminating condition	F-number	Shutter speed(s)	Light source
Aic	3.5	1/20	Natural light
Bic	4.5	1/30	Natural light
Cic	6.3	1/400	Flashlight
Dic	8.0	1/800	Flashlight

Aic = Illuminating condition A; Bic = Illuminating condition B; Cic = Illuminating condition C; Dic = Illuminating condition D. *C-2100uz (Olympus Co., Japan).

helped each examiner to take a photograph of the subjects' tongue substance. The assistant A opened or closed the vertical blind to control natural light, and the assistant B put the reference red card parallel to the subjects' exposed tongue substance when each examiner took photographs of the subjects' tongue substance in the 4 illuminating conditions by digital camera. Each illuminating condition was set up changing the light source as well as the shutter speed and the f-number of the digital camera (Table 1). In each illuminating condition, each examiner took 2 photographs on the top surface and the bottom surface of the subjects' tongue substance and totally took 8 photographs on each subject's tongue substance from 4 different illuminating conditions. The digital camera used in this study was 'C-2100 ultra zoom' (Olympus Co., Japan), and the reference red card used was V2 (vivid red) color in 'New Coloring Cards 199a' (Japan Color Enterprise Co. LTD., Japan).

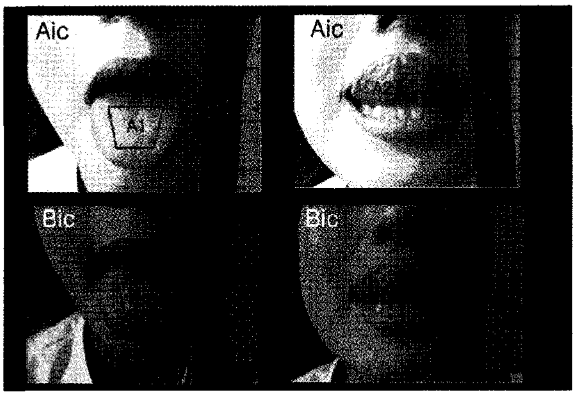

Fig. 1. Digital images photographed in the illuminating condition A and B. Aic = illuminating condition A; Bic = illuminating condition B; A1 = the ROI cropped on the top surface of the tongue substance in the Aic; A2 = the ROI cropped on the bottom surface of the tongue substance in the Aic; B1 = the ROI cropped on the top surface of the tongue substance in the Bic; B2 = the ROI on cropped the bottom surface of the tongue substance in the Bic; R = reference red color.

Computer evaluation

We transported the digital images saved in the digital camera to a personal computer and transformed jpg-format to bmp-format on all the digital images in order to process the digital images using Color Picture Analyzer (Otaka *et al.*, 2002). We cropped region of interest (ROI) for both the tongue substance and the reference red card. As shown in Fig. 1 and Fig. 2, region A1, B1, C1 and D1 are the top surface of the tongue substance and the reference color card and region A2, B2, C2 and D2 are the bottom surface of the tongue substance and region R is the reference red card. On cropping the ROI on the top surface of the tongue substance, we cropped the ROI shaped like a parallelogram and on cropping the ROI on the bottom surface of the tongue substance, we cropped the ROI smaller than that of the top surface to exclude the sublingual veins on it. Color Picture Analyzer could calculate the average value of each R, G and B value dividing the summed RGB value by the number of the pixels in the ROI. Afterwards we transformed R, G and B values to L^*, a^* and b^* values by color picker of Adobe Photoshop (Adobe Systems Incorporated, USA). After extracting 4 sets of L^*, a^* and b^* values from the ROIs on the tongue substance and the reference red card respectively,

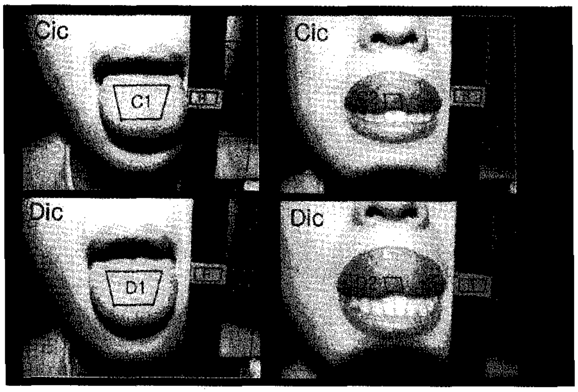

Fig. 2. Digital images photographed in the illuminating condition C and D. Cic = illuminating condition C; Dic = illuminating condition D; C1 = the ROI cropped on the top surface of the tongue substance in the Cic; C2 = the ROI cropped on the bottom surface of the tongue substance in the Cic; D1 = the ROI cropped on the top surface of the tongue substance in the Dic; D2 = the ROI cropped on the bottom surface of the tongue substance in the Dic; R = reference red color.

2 sets of the color differences for the top surface and the bottom surface of the tongue substance were calculated using the coordinate geometry as shown in equation 1.

$$\Delta E = \sqrt{(\Delta L*)^2 + (\Delta a*)^2 + (\Delta b*)^2} \quad \text{Equation 1}$$

(ΔL^*, Δa^* and Δb^* are the color differences between the tongue substance and the reference red card, respectively)

Reliability studies

Intra-rater reliability was assessed by each examiner taking twice photographs of the subjects' tongue substance. The 2nd photographing was taken 3 days after the 1st photographing on the same subjects by the same examiners. We analyzed whether there was a significant difference between the color differences extracted from two digital images photographed at 3 day intervals in each examiner or not. Inter-rater reliability was assessed by two examiners, in which the examiner A took a photograph of one subject's tongue substance and the examiner B repeated it on the same subject immediately. We analyzed whether there was a significant difference between the color differences extracted from two digital images photographed by the examiners or not. Inter-rater reliability was additionally assessed 3 days after the 1st photographing.

Clinician evaluation

Six clinicians were included to evaluate visually the redness of the subjects' tongue substance inspecting the digital images through a computer monitor. Totally 96 digital images of the 12 subjects' tongue substance, including both the top surface and the bottom surface of the tongue substance in 4 illuminating conditions, were presented to each clinician through a blind test. The clinicians were requested to evaluate visually the redness of the top surface of the tongue substance excluding the color information of the tongue coating, and also were requested to evaluate the redness of the bottom surface of the tongue excluding the color of the sublingual veins. By inspecting the digital images presented on the same monitor, each clinician evaluated the redness of the tongue substance using Likert-like 5-points scale: very red (5), above average red (4), average red (3), below average red (2), not at all red (1). Using Likert-like 5-points scale of the redness of the tongue substance, we analyzed whether there was any statistical difference in the redness of the tongue substance between the clinicians and between the 4 illuminating conditions or not. Afterwards we compared the redness of the tongue substance with the color differences.

Statistical analysis

One-way ANOVA test was used to evaluate the statistical difference in color differences extracted from the digital images photographed in the 4 different illuminating conditions. One-way ANOVA test was also used in clinician evaluation in order to analyze the statistical difference in the redness of the tongue substance between the 6 clinicians and between the 4 illuminating conditions. If this analysis showed a significant difference, subsequent comparison was performed using Scheffe's test. Independent-samples *t*-test was used to evaluate the statistical difference in color differences between two examiners, both in the 1st photographing and in the 2nd photographing respectively and paired-samples *t*-test was used to evaluate the statistical difference in color differences between the 1st photographing and the 2nd photographing in each examiner. We compared the color differences for the 4 illuminating conditions with the redness of the tongue substance for the 4 illuminating conditions using Pearson's correlation coefficients. We also compared the redness of the top surface of the tongue with that of the bottom surface of the tongue using Pearson's correlation coefficients. $P < 0.05$ was considered significant.

RESULTS

L^*, a^* and b^* values of tongue substance extracted from the digital images photographed in the 4 illuminating conditions

In one-way ANOVA test, there was a significant difference in L^* values extracted from the digital images photographed in the 4 illuminating conditions ($F = 12.76$, $P < 0.05$), and subsequent comparison revealed that the L^* value extracted from the illuminating condition B was different from that extracted from the illuminating condition A, C and D (Table 2). There was no significant difference in a^* and b^* values ($F = 0.63$, $P > 0.05$; $F = 2.46$, $P > 0.05$). This meant that L^*, a^* and b^* values extracted from different light sources varied and made it difficult to communicate color information.

Color differences between the bottom surface of the tongue and the reference red card in 4 illuminating conditions

Table 3 presented the mean and standard deviation of color differences between the bottom surface of the tongue substance and the reference red card in 4 illuminating conditions. One-way ANOVA revealed no significant difference in color differences extracted from the 4 illuminating conditions in the examiner A's 1st photographing ($F = 0.40$, $P = 0.75$), 2nd

Table 2. Multiple comparisons of L^* value in the 4 illuminating conditions

Illuminating condition	Aic			Bic		
	Bic	Cic	Dic	Aic	Cic	Dic
Mean difference	26.17*	-5.92	0.00	-26.17*	-32.08*	-26.17*
Illuminating condition	Cic			Dic		
	Aic	Bic	Dic	Aic	Bic	Cic
Mean difference	5.92	32.08*	5.92	0.00	26.17*	-5.92

Aic = Illuminating condition A; Bic = Illuminating condition B; Cic = Illuminating condition C; Dic = Illuminating condition D. *The mean difference is significant at the 0.05 level.

Table 3. Color differences between the bottom surface of the tongue substance and the reference red card in the 4 illuminating conditions

Illuminating condition	Examiner A		Examiner B	
	The 1st photographing	The 2nd photographing	The 1st photographing	The 2nd photographing
Aic	38.06 ± 4.84	36.66 ± 14.68	41.64 ± 6.73	50.37 ± 6.03
Bic	36.97 ± 5.60	44.97 ± 9.94	42.25 ± 6.91	49.97 ± 6.94
Cic	39.51 ± 9.88	40.27 ± 10.58	40.62 ± 11.23	45.57 ± 7.03
Dic	39.64 ± 6.35	42.54 ± 7.27	43.03 ± 6.35	47.97 ± 6.76

Aic = Illuminating condition A; Bic = Illuminating condition B; Cic = Illuminating condition C; Dic = Illuminating condition D.

photographing ($F = 1.25$, $P = 0.30$) and the examiner B's 1st photographing ($F = 0.19$, $P = 0.90$), 2nd photographing ($F = 1.41$, $P = 0.25$) respectively. This meant that application of color differences to tongue diagnosis could minimize problems arising from different illuminating condition during taking photographs of tongue substance.

Reliability studies

Inter- and Intra-rater reliability were assessed on the color differences between the bottom surface of the tongue substance and the reference red card in the illuminating condition A voluntarily. In intra-rater reliability studies, there was no significant difference in color differences for the 1st-and the 2nd-photographing in the examiner A ($P = 0.73 > 0.05$) and in the examiner B ($P = 0.18 > 0.05$). In inter-rater reliability studies, there was no significant difference in color differences between the examiners in the 2nd photographing ($P = 0.99 > 0.05$), but there was a significant difference in color differences between the examiners in the 1st photographing ($P < 0.05$). This meant that inter- and intra-rater reliability were satisfied, although inter-rater reliability test was satisfied limitedly.

Clinician evaluation

In clinician evaluation, the redness of the tongue substance was positively correlated with that of the tongue coating ($r = 0.744$, $P < 0.01$). This meant that there was the significant correlations between the top surface and the bottom surface of the tongue substance in visual inspection by the clinicians and that the color differences between the bottom surface of the tongue substance and the reference red card could be substituted to the color differences between the top surface of the tongue substance and the reference color card in order to exclude the color of the tongue coating.

The redness of the tongue substance was significantly different between the 4 illuminating conditions ($F = 8.26$, $P < 0.05$), and subsequent comparison revealed that the redness of the tongue substance in the illuminating condition B was different from that in the illuminating condition A, C and D (Table 4). This meant that the clinicians had a tendency to overestimate the redness of the tongue substance inversely proportion to the luminance of natural light and that visual inspection of tongue substance in diverse illuminating conditions was subjective.

The redness of the tongue substance was significantly different between the clinicians, and subsequent comparison revealed that the redness of the tongue substance evaluated by the clinician B was different from that evaluated by the other clinicians. This meant that there was the significant difference in the visual evaluations of the redness of the tongue substance by the clinicians (Table 5).

Correlations of the redness of the tongue substance with the color differences

The redness of the tongue substance in the illuminating condition A, B and D was negatively

Table 4. Multiple comparisons of the redness of the tongue substance evaluated by 6 clinicians in the 4 illuminating conditions

Illuminating condition	Aic			Bic		
	Bic	Cic	Dic	Aic	Cic	Dic
Mean difference	-0.81*	-0.01	0.14	0.81*	0.79*	0.94*
Illuminating condition	Cic			Dic		
	Aic	Bic	Dic	Aic	Bic	Cic
Mean difference	0.01	-0.79*	0.15	-0.14	-0.94*	-0.15

Aic = Illuminating condition A; Bic = Illuminating condition B; Cic = Illuminating condition C; Dic = Illuminating condition D. *The mean difference is significant at the 0.05 level.

Table 5. Multiple comparisons of the redness of the tongue substance between the clinicians

Clinician		Mean difference	Clinician		Mean difference
A	B	-1.42	D	A	-0.13
	C	-0.63		B	-1.54*
	D	1.25		C	-0.75
	E	-0.48		E	-0.60
	F	-0.29		F	-0.42
B	A	1.42*	E	A	0.48
	C	0.79*		B	-0.94*
	D	1.54*		D	-0.15
	E	0.94*		D	0.60
	F	1.13*		F	0.19
C	A	0.63	F	A	0.29
	B	-0.79*		B	-1.13*
	D	0.75		C	-0.33
	E	0.15		D	0.42
	F	0.33		E	-0.19

*The mean difference is significant at the 0.05 level

correlated with the color differences in the illuminating condition D and positively correlated with those of the illuminating condition A (Table 6). This meant that only specific light source such as flashlight revealed the significant correlations between the redness of the tongue substance and the color differences.

DISCUSSION

Recently quantitative color analysis using RGB and $L^*a^*b^*$ has been applied extensively in chinese medicine as well as ophthalmology, otolaryngology and dentistry (Zhang *et al.*, 1992; Han *et al.*, 1993; David *et al.*, 1998; Cui *et al.*, 2001; Otaka *et al.*, 2002; Ikeda *et al.*, 2003). Although specific color can be expressed quantitatively using RGB and $L^*a^*b^*$, different outer light sources result in the different RGB, $L^*a^*b^*$ value on the same color and this variation of RGB and $L^*a^*b^*$ by different outer light sources makes it difficult to communicate color information. In previous studies of tongue diagnosis using RGB and $L^*a^*b^*$, the variation of outer light source in photographing of tongue substance was not relatively considered weighty and it is not easy practically to take a photograph of a patient's tongue substance using a colorimeter or a standard light source in a darkroom state in most clinics. Our results showed that the application of color differences using $L^*a^*b^*$ color coordinate system

Table 6. Pearson's correlation coefficients between the redness of the tongue substance and the color differences in the 4 illuminating conditions

The redness of the tongue substance	Color differences in 4 illuminating conditions			
	CD in Aic	CD in Bic	CD in Cic	CD in Dic
Aic	0.623*	0.232	-0.469	-0.627*
Bic	0.493	-0.175	-0.197	-0.638*
Cic	-0.438	-0.289	-0.267	-0.481
Dic	0.273	0.337	-0.271	-0.589*

CD = Color differences; Aic = Illuminating condition A; Bic = Illuminating condition B; Cic = Illuminating condition C; Dic = Illuminating condition D. *The mean difference is significant at the 0.05 level.

could not only express the color of the tongue substance quantitatively, but also minimize the variation of color information resulted from outer light source. Color differences satisfied intra-rater reliability and also satisfied limitedly inter-rater reliability. This meant that the color differences could be applied to the tongue diagnosis reproducibly and the color information of the tongue substance could be communicated between the clinicians. It is possible that the significant difference between two examiners in the 1st photographing had resulted from the variation of two examiners' skillfulness on photographing by digital camera.

It is practically difficult to extract the pure $L^*a^*b^*$ value of the tongue substance from the top surface of the tongue because of the existence of tongue coating on the top surface of the tongue. For the redness for the top surface of the tongue substance significantly correlated with the redness for the bottom surface of the tongue substance ($r = 0.744$, $P < 0.01$), we substituted the color differences on the top surface of the tongue substance to those on the bottom surface of the tongue substance in order to exclude the color of the tongue coating.

Color differences between the bottom surface of the tongue and the reference color card could be also affected by sublingual veins located on the bottom surface of the tongue substance, but it was not difficult to crop the ROI excluding sublingual veins by Color Picture Analyzer(Otaka *et al.*, 2002). The significant difference of the redness evaluated by the clinicians between 4 illuminating conditions meant that there existed the variation of visual inspection resulted from different illuminating conditions and the statistical difference in the redness of the tongue substance between the clinicians and that clinician evaluation was subjective.

Although color analysis of tongue substance using color differences could minimize the problem arising from different illuminating conditions, it satisfied the inter-rater reliability limitedly, and the correlation coefficients between the color differences extracted from color difference geometry and the redness of the tongue substance evaluated by the clinicians were relatively low. For this problem, it is possible that most clinicians evaluated visually the color of the tongue substance unlike the processing of the color analysis through a computer or that there were included the relatively small number of the subjects in our study.

In conclusion, the color differences between the subjects' tongue substance and the reference red card can not only analyze the color of the tongue substance quantitatively, but also minimize the variation of illuminating condition and makes it possible to communicate the color information of the tongue substance in clinical application. In the future, the study of color analysis of the tongue substance using color differences for the large number of subjects is needed.

Limitations and future directions

In this study, we evaluated the color of tongue substance in a small group of young men and women. Future studies should be performed for more normal subjects of all ages to generalize the characteristics of color of tongue substance by color differences methodology. This study was performed by only one model of digital camera and it will also be meaningful to investigate whether color differences methodology will satisfy the inter-camera reliability, when applying different models of digital camera.

REFERENCES

Bengel WM. (2003) Digital photography and the assessment of therapeutic results after bleaching procedures. *J. Esthet. Restor. Dent.* **15**, 21-32.

Chuang-Chien Chiu. (2000) A Novel Approach Based on Computerized Image Analysis for Traditional Chinese Medical Diagnosis of the Tongue. *Comput. Methods Programs Biomed.* **61**, 77-89.

Cui Min-gui, Xu Bao-yu, Huang Shi-jing, Weng Wei-liang. (2001) Quantitative Study on Tongue Diagnosis in Stroke Patients. *Zhongguo Zhong Xi Yi Jie He Za Zhi* **21**, 670-673.

David G. Hanson, Jack Jiang, Wan Chi. (1998) Quantitative

Color Analysis of Laryngeal Erythema in Chronic Posterior Laryngitis. *J. Voice* **12**, 78-83.

Dong Kyun Rah, Soo Chan Kim, Kwang Hoon Lee, Beyoung Yun Park, Deok Won Kim. (2001) Objective Evaluation of Treatment Effects on Portwine Stains Using $L^*a^*b^*$ Color Coordinates. *Plast. Reconstr. Surg.* **108**, 842-847.

Han Xin-min, Yu Yong-min, Sun Tian-ting, Chen Zheng-quan, Lin Min-jia, Wu Xie-wei, Mao Hui-ming, Zhang Fang. (1993) Quantitative Analysis of 200 Human Tongue Pictures. *Zhongguo Zhong Xi Yi Jie He Za Zhi* **13**, 110-111.

Ikeda T, Nakanishi A, Yamamoto T, Sano H. (2003) Color differences and color changes in Vita Shade tooth-colored restorative materials. *Am. J. Dent.* **16**, 381-384.

Isao Otaka, Kenjiro Kumagai, Yoko Inagaki, Masaru Shimoyama, Katsuyuki Saegusa, Takeshi Hara. (2002) Simple and Inexpensive Software Designed for the Evaluation of Color. *Am. J. Ophthalmol.* **133**, 140-142.

Zhang Bo-li, Xu Zong-pei, Liu Hua-yi, Zhang Jin-ying, Gao Xiu-mei, Mao Bing-xie. (1992) Chromometric Study of Tongue Picture. TianJin *J. Traditional Chinese Med.* **4**, 38-40.