

ORIGINAL ARTICLE

Reliability analysis of a novel measurement system for quantifying human skin color

Chien-Chung Chen¹ | Cheng-Yin Chung² | Yi-Wen Chiu^{3,4} | Yu-Hsuan Lin⁵ |
Ling-Shan Tse⁶ | Ching-Ying Wu⁷ | Shang-Jyh Hwang^{3,4} | Ming-Yen Lin^{3,4}  |
for the iH³ research group

¹National Taiwan University Hospital, National Taiwan University, Taipei, Taiwan

²Department of Internal Medicine, Division of Nephrology, Ministry of Health and Welfare, Pingtung Hospital, Pingtung, Taiwan

³Department of Internal Medicine, Division of Nephrology, Kaohsiung Medical University Hospital, Kaohsiung Medical University, Kaohsiung, Taiwan

⁴Department of Renal Care, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan

⁵Taiwan Instrument Research Institute, National Applied Research Laboratories, Hsinchu, Taiwan

⁶Medical Science Liaison, WS Far IR Medical Technology CO., LTD, Taipei, Taiwan

⁷Department of Dermatology, Kaohsiung Municipal Ta-Tung Hospital, Kaohsiung, Taiwan

Correspondence

Ming-Yen Lin, Department of Internal Medicine, Division of Nephrology, Kaohsiung Medical University Hospital, Kaohsiung Medical University, Kaohsiung 807, Taiwan.
Email: mingyenlin3@gmail.com

Funding information

Taiwan Ministry of Science and Technology; WS Far IR Medical Technology CO., LTD, Grant/Award Number: 11016; Taiwan National Health Research Institutes; Taiwan National Health Insurance Administration, Ministry of Health and Welfare, Grant/Award Number: MOHW107-NHI-S-114-000001

Abstract

Background: Precision is crucial in determining the appropriate procedure for implementing further trials. We conducted a study to explore the reliability of a novel measuring system for human skin color.

Methods: The novel skin color measuring system was used to capture the skin color of four volunteers (2 males and 2 females) from the same location on each subject by the same operator. The measurement was repeated for different poses and instrument factors (camera and shooting protocol) in the red, green, and blue (RGB) system. The average color depth in each image was calculated and converted from 0 to 255. The spread of measures and the Bland-Altman plot was displayed to determine each variance source's random error, with the interclass correlation coefficients applied to reflect the reliability.

Result: The RGB color depth in the experiment ranged from 190, 152, and 122 to 208, 170, and 142. The 95% confidential interval of the differences from the means in RGB colors for the different protocols were ± 2.8 , ± 2.6 , and ± 2.1 , respectively. The largest variation in the replicate trials was observed when subjects were in a supine position (standard deviation: 2). The interclass correlation coefficients were greater than 90%, suggesting that the developed system is highly precise.

Conclusion: This study demonstrated that the developed device could stably and reliably detect human skin color across different common sources of variation, and thus could be applied clinically to explore relationships between health/disease and skin color changes.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. Skin Health and Disease published by John Wiley & Sons Ltd on behalf of British Association of Dermatologists.

1 | INTRODUCTION

The skin is the largest organ of the human body composing 18% of body weight which protects the body from external injuries. It also helps regulate body temperature, the pH of bodily fluids, and excrete substances, maintaining a weak acid environment to prevent microbial infection.^{1,2} Several physical features of the skin may reflect specific health problems such as color, temperature, elasticity, dryness, fine wrinkling, atrophy, and laxity. Developing an accurate and precise instrument to quantify physical features of skin could be helpful for health promotion and disease prevention.^{3–5}

Skin color is influenced by several factors such as ethnicity, lifestyle, or even biochemical. It also partly reflects specific social classes that are engaged in more outside activities.⁶ In addition, skin color changes are potentially an indicator of disease. For example, hyperpigmentation (skin color turning to black) is an important differential diagnosis for Addison's disease,⁷ whereas jaundice (skin color turning to yellow) may also be a symptom or sign to identify patients with alcoholic hepatitis.⁸ Determining skin color change by the human eye is usually subjective, ambiguous depends on experience, uncertainty, and not reproducible, therefore an instrument to quantify skin color would be helpful in a range of relevant fields.

Several instruments have been recently employed to estimate human skin color but these devices are mainly adopted point imaging systems which are costly and not practical to apply in practice.⁹ We previously reported that the developed skin color system could accurately quantify human skin color,¹⁰ so this study aimed to quantify the reliability of the developed skin color system for catching images from the abdomen which is less influenced by sun exposure.

2 | MATERIALS AND METHODS

2.1 | Imaging system for skin color

The imaging system was co-developed by the Taiwan Instrument Research Institute and Kaohsiung Medical University Hospital. In brief, the imaging system was composed of three major components, a sunlight light-emitting diode ring light, a dark room module, and a digital single-lens reflex camera (Canon DSLR, 77D) with a macro lens (EF-S, 60 mm). The system can catch standardized color through the controlled light source and a precise light censoring module. The newly developed device quantified colors by the Red-Green-Blue (RGB) color system that are easily transformed to *l*, *a*, *b* color system (Figure 1). Previously, we confirmed that the system could detect the color difference that the typical human eye could not identify. To ensure the stability of the color system

What is already known about this topic?

- Several instruments have been recently employed to estimate human skin color for relevant research, however, these devices mainly adopted point imaging systems at extreme costs, making them less practical to apply in practice.

What does this study add?

- This study reports the highly accurate, portable, and low-cost skin color image system testing on four subjects and confirms its reliability regarding the cameras, shot model, and subject posture. The 95% confidential interval of the differences from means in RGB colors for different protocols are ± 2.8 , ± 2.6 , and ± 2.1 , respectively. The interclass correlation coefficients are over 90%, suggesting the developed system is highly precise.

and optimal operational procedure for catching human skin color, we quantified four main variances of colors: cameras, people, subject poses, and acquisition procedures. All of the study procedures were approved by the Institutional Review Board of Kaohsiung Medical University Hospital (KMUHIRBEXEMPT (I)-20170144).

2.2 | Subjects and measurements

The healthy volunteers (2 men and 2 women with an average age of 38 years; range: 26–48) were randomly selected from the Kaohsiung Medical University Hospital located in southern Taiwan (22°N 120°E). Their jobs were mainly indoor so they were not excessively exposed to solar radiation and they reported that they had not been exposed to excessive sunlight 3 months before the experiment. One experienced research assistant performed the experimental procedures in a room without outside interferences in spring 2018. The skin images were all taken from the abdominal skin 10 cm under the central area of the umbilicus (Figure 1). In order to avoid interference from other light sources, the lens gently covers the skin when taking photos. The images were collected from the same body area with three cameras (I, II, III) according to (1) three shooting protocols:

- 10 continuous shots
- 3 shots without picking up the camera
- 3 shots with a camera picked up between each shot;



FIGURE 1 The highly accurate, portable, low-cost skin color image system. The system weighs 1.5 kg and is 30 cm in length and 13 cm in width. It composites three key elements: a sunlight light-emitting diode ring light, a dark room module, and a digital single-lens reflex camera (Canon DSLR, 77D) with a macro lens (EF-S, 60 mm). The plot of the standard color difference was derived from our previous conference paper.¹⁰ The 1d) exhibits that applying the imaging system takes a leaning backward sitting subject's skin color image. (a) The small size of the developed image system (length/width: 30/13 cm). (b) The components of the developed image. (c) Standard color difference of the developed image system. (d) Taking skin color image by leaning backward sitting.

and (2) four postures of observation:

- (a) supine
- (b) leaning backward sitting
- (c) slumped sitting
- (d) standing

2.3 | Data processing and analysis

The instrument RGB system displayed human skin color, which contains three primary colors by 0–255 color depths. We calculated the average color depth ($\overline{CD}$) of each primary color using the following formula:

$$\overline{CD}_{\{R,G,B\}} = \frac{\sum_{c=0}^{255} \text{Color depth} * \text{Pixels}}{\text{Total pixel number in one picture}}$$

The average color depth was obtained for each skin image, and the range of average color depth was described by plots to reflect the maximum changes in skin color measurements in all experimental procedures (human, camera, protocol, and postures). We classified pictures shot in the same camera, posture, and protocols to the same group. The averages and the differences

from the average within protocol and postures were calculated and displayed by Bland-Altman plots.¹¹ The measurements outside the 95% confidence intervals (CIs) of the differences from the average were considered outliers. The number of outliers and dispersions of the differences from the average color depth in each RGB color were evaluated to reflect the consistency between experiments. To understand the stability of our skin color detecting device under different groups, cameras, shooting protocols, and postures, interclass correlation coefficients (ICC) and 95% CIs were displayed. All statistical analyses were performed using SPSS statistical software (version 25; IBM Corp.).

3 | RESULTS

3.1 | Skin color measurements

Our trial took 768 skin color pictures classified by different cameras, protocols, and postures. The distribution of the RGB color spaces quantified by the developed image system from subjects' skin was similar. The minimum average color depth in the RGB colors of the four subjects was 190, 152, and 122, whereas the maximum was 235, 196, and 165,

respectively (Figure 2). The Bland-Altman plots for the repeated measurements of the different shooting protocols (displayed by a different color) are shown in the supplement. Most average color depths in all Bland-Altman plots were clustered in lower values,

suggesting that one subject's skin color may be lighter than others (Figure S1a–c). The parameters produced from Bland-Altman plot analyses based on the combination of shooting protocols are displayed in Table 1. A similar average color depth in RGB colors (208, 170,

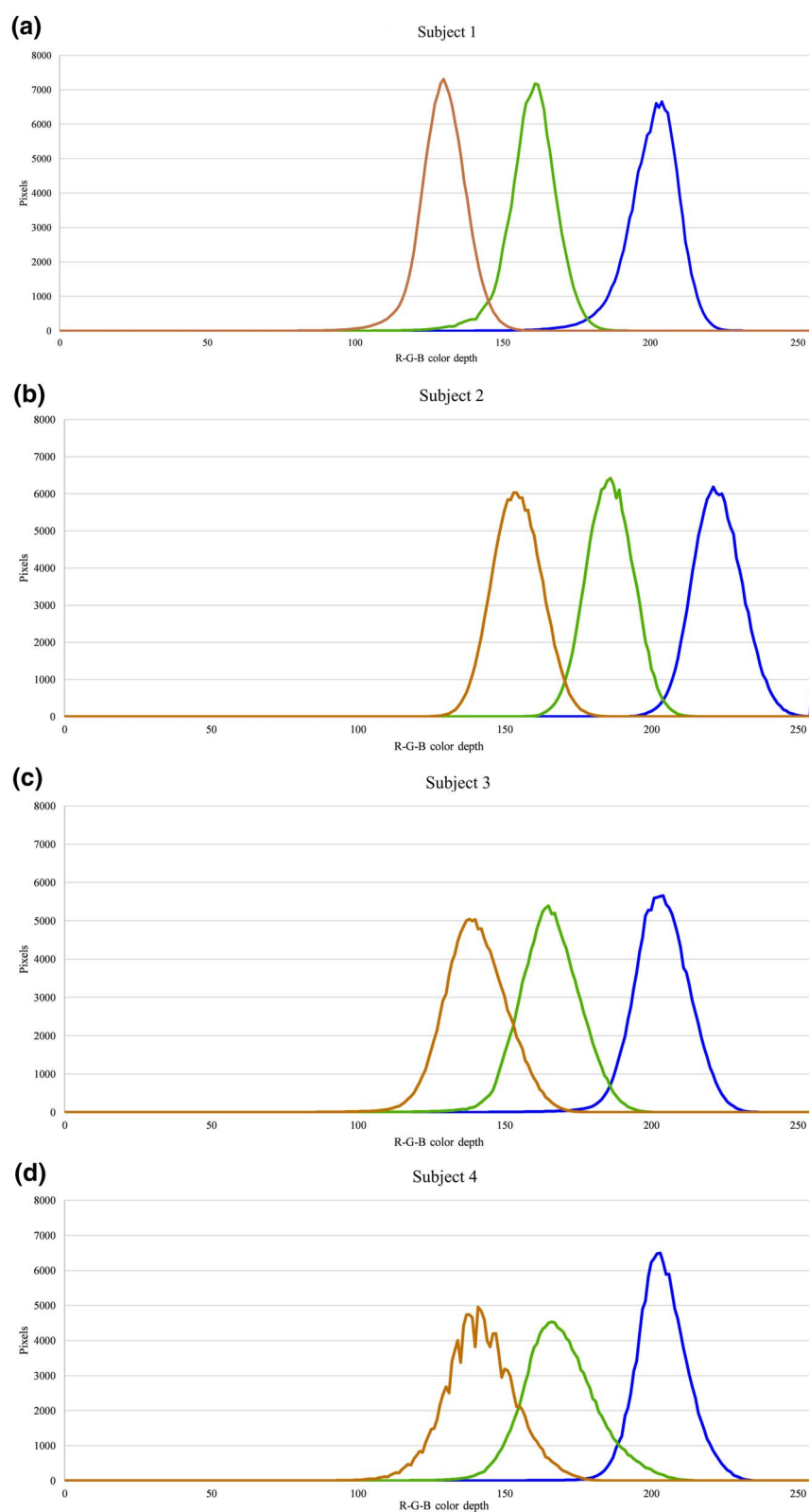


FIGURE 2 The distribution of skin color depth in the red-blue-green colors of (a) subject 1, (b) subject 2, (c) subject 3, and (d) subject 4. Each color depth of the red-blue-green colors was the average of repeated experimental values captured by the developed image system.

TABLE 1 Parameters of the Bland-Altman plot for color depth comparison of the three protocols

Measurements		Difference from average in color depth			Average in color depth		
		Red	Green	Blue	Red	Green	Blue
Protocol A, B, and C ^a (<i>n</i> = 432)	Mean	0.0	0.0	0.0	207.8	169.7	141.0
	Limits of agreement	±2.8	±2.6	±2.1			
	SD	1.4	1.3	1.1			
	No. of outlier (H/L)	12/13	14/10	13/10			
Protocol A and B (<i>n</i> = 288)	Mean	0	0	0	208.3	170.4	141.7
	Limits of agreement	± 2.6	± 2.1	± 1.8			
	SD	1.3	1.1	0.9			
	No. of outlier (H/L)	4/6	7/10	6/8			
By protocol B and C (<i>n</i> = 288)	Mean	0	0	0	207.5	169.2	140.4
	Limits of agreement	±2.8	±2.6	±2.1			
	SD	1.4	1.3	1.1			
	No. of outlier (H/L)	10/11	10/7	9/7			
Protocol A (<i>n</i> = 144)	Mean	0	0	0	208.5	170.7	142.2
	Limits of agreement	± 2.6	± 2.3	± 1.9			
	SD	1.3	1.2	1.0			
	No. of outlier (H/L)	10/10	15/14	14/11			
Protocol B (<i>n</i> = 144)	Mean	0	0	0	208.2	170.0	141.3
	Limits of agreement	± 2.4	± 2.1	± 1.6			
	SD	1.2	1.1	0.8			
	No. of outlier (H/L)	4/6	2/6	2/5			
Protocol C (<i>n</i> = 144)	Mean	0	0	0	206.8	168.4	139.4
	Limits of agreement	± 3.1	± 3.1	± 2.5			
	SD	1.6	1.3	1.3			
	No. of outlier (H/L)	8/10	5/3	5/3			

Note: *n* represents the number of combinations of elements (subjects, postures, cameras, and shots). Protocol A: 10 continuous shots; Protocol B: 3 single shots without picking up the camera; Protocol C: 3 single shots with camera picked up between each shot.

Abbreviations: H, high; L, low; SD, standard deviation.

^aWe picked the first 3 times of 10 continuous shots.

142) was found in both shooting protocols A (10 continuous shots) and B (3 shots without picking up the camera) but was smaller for the 2 color depth for shooting protocol C (3 shots with camera picked up between each shot). All standard deviations of the difference from the average of each color space were less than 1.5, and the 95% CI of the differences was less than 3 in line with our previous reported error ($\Delta E < 3$).

3.2 | Variation in skin color measurements between shooting protocols

Regardless of the shooting protocol, the average and variation of the difference from the average showed the

same trend, that is, largest in the R color, followed by G, and B. The number of outliers determined as larger than the upper limit or smaller than the lower limit of the 95% CI of the difference from the average was equally distributed among color spaces in the different protocols. Regarding the occurrence frequency of outliers at each shot in protocol A, the most up and low outliers appeared at the 4th and 8th shots, respectively, with the least occurring at the 6 and 7th shots (Table S1).

3.3 | Variation in skin color measurements between postures

The variations were more obvious between postures than those between shooting protocols (Figure 3).

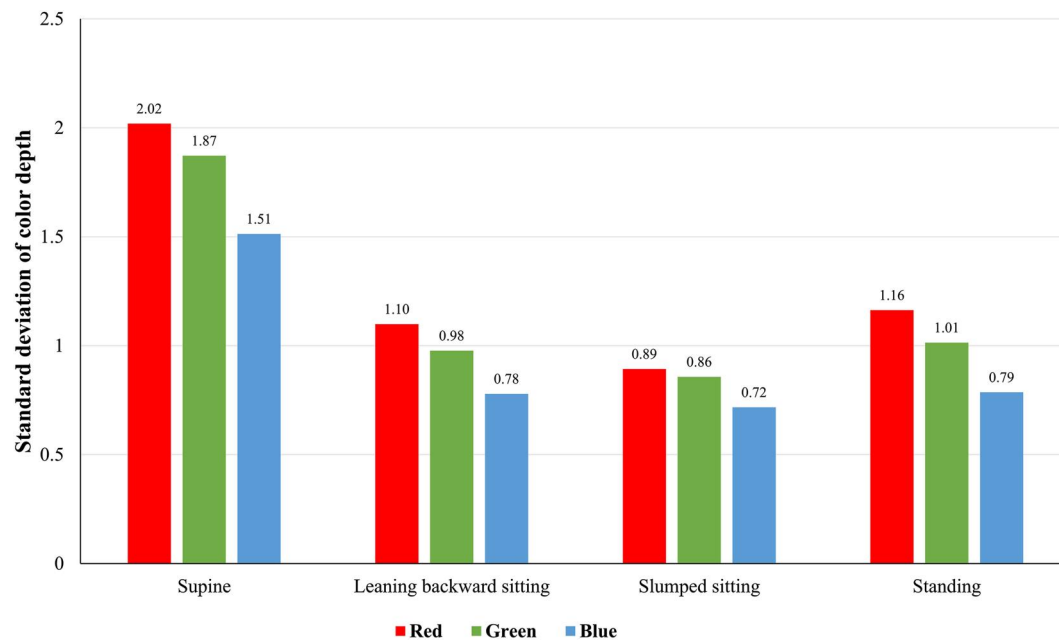


FIGURE 3 Standard deviations of the difference from the average for the red-blue-green colors in different postures

Regardless of posture, blue represents less variation (smallest standard deviation) in color depth compared to the other colors. The standard deviation of differences from the average skin color depth when the subject was in a supine position was nearly two-fold of the other postures.

3.4 | The precision of skin color measurements

Interclass correlation coefficient (ICC) and 95% CI of primary color based on the settings (within-group, camera, protocol, and posture) are shown in Table 2. All replicate skin color measurements were stable under different setting conditions. A perfect ICC (>99%) was observed when looking at measurements at different times (Table 2). The skin color obtained by different cameras reflects lower but remains highly consistent (ICC: 95.3% for red, 92.4% for green, and 93.7% for blue). Large variations may be caused by human factors, with the ICC between postures and protocols reaching 97%.

4 | DISCUSSION

Skin color is considered one of the critical features for identifying specific skin abnormalities, so modern medicine depends on physician subjective eye observation combined with the patient's symptoms to make clinical decisions. Although skin color measurement by an instrument has potential for clinical applications, only

a few studies focus on measuring variations from different sources and consequent implications. The novelty of the current study is the quantification of various sources of uncertainty (subject, protocol, and postures) in human skin color, which should be helpful in future relevant applications.

Skin color changes are mainly detected by the naked eye which lacks objectivity and accuracy. The naked eye can easily discern skin color problems from healthy skin color (either hyperpigmented or hypopigmented), but even experienced physicians find it difficult to grade the severity of skin color disorders. Vitiligo is an autoimmune disorder causing patchy loss of skin pigmentation occurring in 0.76%–1.11% of adults.¹² Several assessment tools incorporating key determinants of skin lesions have been developed to improve vitiligo disorder treatment.¹³ For example, the Vitiligo Extent Tensity Index score mainly depends on cutaneous and hair pigmentation and the proportion of pigmented and depigmented areas in the total body surface.¹⁴ Although the assessment method has largely improved uncertainty caused by subjective clinical judgements, applying these methods to identify changes within limited treated skin lesions remains infeasible. With acceptable costs (~5000 USD), the current device is easy to operate and a non-invasive approach to accurately and reliably quantify human skin color changes. In addition, applying the developed system to explore the relationship between diseases and skin color in the health field could be helpful. More large-scale systematic research is expected to determine meaningful skin color changes in human health/disease conditions.

TABLE 2 The interclass correlation coefficient of color depth in different variant sources

Variant source	Red			Green			Blue		
	ICC	95% CI		ICC	95% CI		ICC	95% CI	
Inter group	99.2%	98.9%	99.4%	99.3%	99.1%	99.5%	99.3%	99.1%	99.5%
Camera	95.3%	94.2%	96.2%	92.4%	90.6%	93.9%	93.7%	92.3%	95%
Protocol	98.3%	97.8%	98.7%	98.4%	97.9%	98.8%	98.3%	97.7%	98.8%
Posture	97.8%	97.3%	98.3%	97.9%	97.4%	98.3%	97.6%	96.9%	98.1%

Note: Variation sources were from the camera (I, II, III), protocol [(A) 10 continuous shots, (B): single shots without picking up the camera, and (C) 3 single shots with camera picked up between each shot], and posture [(a): supine, (b): leaning backward sitting, (c): slumped sitting, and (d): standing].

Abbreviations: CI, confidence interval; ICC, interclass correlation coefficient.

The variability and accuracy of skin color measurement by instruments and algorithms have been explored by previous studies.^{15,16} Wang et al. analyzed the effects of aperture size, pressure, and measurement distance on the variability of facial color detected by a Konica Minolta CM700d spectrophotometer and a PhotoResearch PR650 telespectroradiometer.¹⁵ Similar to our observation, the differences between measurements for different facial and hand locations were larger than instrument repeatability and the inter-instrument agreement. The information reflects that more studies exploring skin color variabilities in different body locations are needed. Another study different from our intentions emphasized that not only the stability of one image system but also the image format and skin color characterization algorithm play a key role in determining facial color accuracy.¹⁶ It is worth noting of even a common commercial facial color detection device named Visia analysis system (Canfield Scientific Inc.) does not transparently offer whether its measurement is standardized to an international color standard. Therefore, a scientific guideline for making skin color instrument development process transparent should be developed.

The experimental results indicate that the 95% CI of the overall difference in color depth of RGB colors are ± 2.8 , ± 2.6 , and ± 2.1 , which are compatible with the accuracy of this device ($\Delta E = 3$). Although skin color results from the experimental settings were consistently reliable, several novel findings are worthy of note. First, the replicate measurements of 10 continuous shootings were not as stable as those of single shooting approaches, possibly due to the instability of the mechanical camera shutter. In addition, single shooting without picking up the camera provides a smaller variation in measurements than shooting with picking up the camera, reflecting that moving the measurement location may contribute additional variations to repeated skin color measurements. The most inconsistent results were observed in the supine position compared to the other postures. The exact causes are unknown; however, lying down may alleviate abdominal tension, increase diaphragmatic breathing, and add

measurement variability. Finally, the lowest ICC among our experiments was tested between different cameras, which suggested a single camera for one study may be needed if skin color changes are the main research interest.

Some color spaces are commonly applied for detecting color by devices. Our developed devices adopted the RGB rather than HSV or L*A*B color space, which is more instinctive and straightforward for medical professionals to apply. Although HSV and L*A*B color spaces are easily calculated, a recent study showed a weak correlation between the L value and human perception of skin color, which may limit its application in skin color detection.¹⁷ Several potential skincare applications of this device should be emphasized. First, the traditional Fitzprick skin type may lose its accuracy in distinguishing the borderline skin type between adjacent categories. The device undoubtedly could categorize human skin type on a tremendously detailed scale. Then, in a similar previous study, the device could trace the progression and the response of vitiligo and its therapies.¹⁸ Third, it could detect flap failure by measuring the color of the flap to determine whether the flap was successful after plastic surgery. Finally, it could assist a physician in determining the progression of jaundice, so they are no longer dependent on periodic blood tests.¹⁹ More studies exploring its feasibility and accuracy in different healthcare domains are warranted.

This study has some limitations. First, the study only included Asian subjects located in a subtropic area, so the results may not be generalizable to other uncollected variances such as environmental factors. Although the reliability of the system was not tested in extreme environmental conditions (temperature, humidity..., et al.), it should not affect the standard applications in most usual conditions. All study subjects were healthy and had little hair, so the reliability of our device is unknown when applying it to participants with much hair and intense skin color such as jaundice. In addition, there are numerous skin color-changing diseases caused by different pigmentation disorders. Sensitivity of the device in determining each

pigmentation disease needed to be further tested. Finally, the interoperator variances in skin color measurements were not evaluated as previously,²⁰ however, this variance is usually controlled by an operator training programme before initiating a formal experiment.

5 | CONCLUSION

In conclusion, this study demonstrated that the developed device could stably and reliably detect human skin color across different sources of variation including operating protocols, and respondents' postures. Furthermore, this device has the potential to be used to explore the relationships between health/disease and skin color changes.

ACKNOWLEDGEMENTS

We thank the Taiwan Instrument Research Institute and National Applied Research Laboratories for providing technical development and support for the imaging system.

The color imaging system has been registered by patent number I683287 to the Intellectual Property Office, Ministry of Economic Affairs, Taiwan, and US20190371004A1 to the United States Patent and Trademark Office. The authors would like to sincerely thank the Taiwan Ministry of Science and Technology (grant numbers: MOST: 106-2314-B-002-253-, MOST:107-2314-B-037-122-, MOST: 108-2314-B-037-110-, and MOST 109-2314-B-037-107-MY2), the Taiwan National Health Research Institutes (grant numbers: NHRI-EX108-10505PI and NHRI-EX110-10717PI), the Taiwan National Health Insurance Administration, Ministry of Health and Welfare (grant numbers: MOHW107-NHI-S-114-000001) and WS Far IR Medical Technology CO., LTD (grant number: 11016) for supporting this study.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

AUTHOR CONTRIBUTIONS

Chien-Chung Chen: Data curation (equal); Formal analysis (equal); Methodology (equal); Visualization (lead); Writing – original draft (lead). **Cheng-Yin Chung:** Writing – original draft (equal). **Yi-Wen Chiu:** Conceptualization (equal); Methodology (lead); Supervision (equal); Writing – review & editing (equal). **Yu-Hsuan Lin:** Funding acquisition (equal); Methodology (equal); Writing – review & editing (equal). **Ling-Shan Tse:** Resources (equal); Writing – review & editing (equal). **Ching-Ying Wu:** Writing – original draft (equal); Writing – review & editing (equal). **Shang-Jyh**

Hwang: Resources (equal); Writing – review & editing; (equal). **Ming-Yen Lin:** Data curation (lead); Formal analysis (lead); Methodology (lead); Writing – original draft (equal); Writing – review & editing (lead).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ETHICS STATEMENT

All of the study procedures were approved by the Institutional Review Board of Kaohsiung Medical University Hospital (KMUHIRBEXEMPT(I)-20170144).

ORCID

Ming-Yen Lin  <https://orcid.org/0000-0002-8194-4833>

REFERENCES

1. Montagna W. The structure and function of skin. Elsevier; 2012.
2. Chuong C-M, Nickoloff B, Elias P, Goldsmith L, Macher E, Maderson P, et al. What is the 'true' function of skin? *Exp Dermatol*. 2002;11(2):159–87.
3. Lahiri B, Bagavathiappan S, Jayakumar T, Philip J. Medical applications of infrared thermography: a review. *Infrared Phys Technol*. 2012;55(4):221–35. <https://doi.org/10.1016/j.infrared.2012.03.007>
4. Kwon K, Wang H, Lim J, San Chun K, Jang H, Yoo I, et al. Wireless, soft electronics for rapid, multisensor measurements of hydration levels in healthy and diseased skin. *Proc Natl Acad Sci USA*. 2021;118(5). <https://doi.org/10.1073/pnas.2020398118>
5. Lutgers HL, Graaff R, Links TP, Ubink-Veltmaat LJ, Bilo HJ, Gans RO, et al. Skin auto fluorescence as a noninvasive marker of vascular damage in patients with type 2 diabetes. *Diabetes Care*. 2006;29(12):2654–9. <https://doi.org/10.2337/dc05-2173>
6. Goldsmith AH, Hamilton D, Darity W, Jr. Shades of discrimination: skin tone and wages. *Am Econ Rev*. 2006;96(2):242–5. <https://doi.org/10.1257/000282806777212152>
7. Neto RAB, de Carvalho JF. Diagnosis and classification of Addison's disease (autoimmune adrenalitis). *Autoimmun Rev*. 2014;13(4-5):408–11. <https://doi.org/10.1016/j.autrev.2014.01.025>
8. Singal AK, Louvet A, Shah VH, Kamath PS. Grand rounds: alcoholic hepatitis. *J Hepatol*. 2018;69(2):534–43. <https://doi.org/10.1016/j.jhep.2018.05.001>
9. Wang M, Xiao K, Wuergler S, Cheung V, Luo MR, ed. Measuring human skin colour. In: *Color and Imaging Conference*. Society for Imaging Science and Technology; 2015.
10. Li C-R, Tsai H-Y, Yang C-C, Lin M-Y, Huang K-C, Lin Y-H, ed. Quantifying the color changes in bruised skin using a color-calibrated imaging system. In: *2020 IEEE international symposium on medical measurements and applications (MeMeA)*. IEEE; 2020.
11. Giavarina D. Understanding bland altman analysis. *Biochem Med*. 2015;25(2):141–51. <https://doi.org/10.11613/bm.2015.015>
12. Gandhi K, Ezzedine K, Anastassopoulos KP, Patel R, Sikirica V, Daniel SR, et al. Prevalence of vitiligo among adults in the United States. *JAMA Dermatol*. 2022;158(1):43–50. <https://doi.org/10.1001/jamadermatol.2021.4724>
13. Peralta-Pedrero ML, Morales-Sánchez MA, Jurado-Santa Cruz F, De la Torre-García ME, Cruz-Peralta ES, Olguín-García MG. Systematic review of clinimetric instruments to determine the

- severity of non-segmental vitiligo. *Australas J Dermatol*. 2019;60(3):e178–e85.
14. Feily A. Vitiligo Extent Tensity Index (VETI) score: a new definition, assessment and treatment evaluation criteria in vitiligo. *Dermatol Pract Concept*. 2014;4(4):81. <https://doi.org/10.5826/dpc.0404a18>
 15. Wang M, Xiao K, Luo MR, Pointer M, Cheung V, Wuerger S. An investigation into the variability of skin colour measurements. *Color Res Appl*. 2018;43(4):458–70. <https://doi.org/10.1002/col.22230>
 16. He R, Xiao K, Pointer M, Bressler Y, Liu Z, Lu Y. Development of an image-based measurement system for human facial skin colour. *Color Res Appl*. 2022;47(2):288–300. <https://doi.org/10.1002/col.22737>
 17. Leem S, Kim SJ, Kim Y, Shin JG, Song HJ, Lee SG, et al. Comparative analysis of skin characteristics evaluation by a dermatologist and the Janus-III measurement system. *Skin Res Technol*. 2021;27(1):86–92. <https://doi.org/10.1111/srt.12915>
 18. Hayashi M, Okamura K, Araki Y, Suzuki M, Tanaka T, Abe Y, et al. Spectrophotometer is useful for assessing vitiligo and chemical leukoderma severity by quantifying color difference with surrounding normally pigmented skin. *Skin Res Technol*. 2018;24(2):175–9. <https://doi.org/10.1111/srt.12410>
 19. Maya-Enero S, Candel-Pau J, Garcia-Garcia J, Duran-Jorda X, Lopez-Vilchez MA. Reliability of transcutaneous bilirubin determination based on skin color determined by a neonatal skin color scale of our own. *Eur J Pediatr*. 2021;180(2):607–16. <https://doi.org/10.1007/s00431-020-03885-0>
 20. Sommers M, Beacham B, Baker R, Fargo J. Intra-and inter-rater reliability of digital image analysis for skin color measurement. *Skin Res Technol*. 2013;19(4):484–91. <https://doi.org/10.1111/srt.12072>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Chen C-C, Chung C-Y, Chiu Y-W, Lin Y-H, Tse L-S, Wu C-Y, et al. Reliability analysis of a novel measurement system for quantifying human skin color. *Skin Health Dis*. 2023;3(1):e182. <https://doi.org/10.1002/ski2.182>