Original Article

Comparing Quantitative Measures of Erythema, Pigmentation and Skin Response using Reflectometry

JENNIFER K. WAGNER1, CELINA JOVEL1, HEATHER L. NORTON1, ESTEBAN J. PARRA1,2 and MARK D. SHRIVER1

1Department of Anthropology, The Pennsylvania State University, University Park, PA, USA;
2Department of Anthropology, University of Toronto at Mississauga, Mississauga, Ontario, Canada

*Address reprint requests to Dr Mark D. Shriver, Department of Anthropology, Pennsylvania State University, 409 Carpenter Building, University Park, PA 16802, USA. E-mail: mds17@psu.edu

Received 8 May 2002; in final form 22 July 2002

We measured a number of pigmentation and skin response phenotypes in a sample of volunteers (n = 397) living in State College, PA. The majority of this sample was composed of four groups based on stated ancestry: African-American, European-American, Hispanic and East Asian. Several measures of melanin concentration (L*, melanin index and adjusted melanin index) were estimated by diffuse reflectance spectroscopy and compared. The efficacy of these measures for assessing constitutive pigmentation and melanogenic dose–response was evaluated. Similarly, several measures of erythema (a*, erythema index and adjusted erythema index) were compared and evaluated in their efficacy in measuring erythema and erythemal dose–response. We show a high correspondence among all of the measures for the assessment of constitutive pigmentation and baseline erythema. However, our results demonstrate that evaluating melanogenic dose–response is highly dependent on the summary statistic used: while L* is a valid measure of constitutive pigmentation it is not an effective measure of melanogenic dose–response. Our results also confirm the use of a*, as it is shown to be highly correlated with the adjusted erythema index, a more advanced measure of erythema based on the apparent absorbance. Diffuse reflectance spectroscopy can be used to quantify the constitutive pigmentation, melanogenic dose–response at 7 d and erythemal dose–response at both 24 h and 7 d postexposure.

Keywords: Reflectance spectroscopy, Erythemal dose–response, Melanogenic dose–response

INTRODUCTION

One fundamental property of the skin is its ability to respond to ultraviolet radiation (UVR). In many persons and populations these responses are clearly adaptive where the first response, erythema (redness), is both a signal to the sunburned person to stay inside and also a sign that the immune system is active and the healing process has begun. Neomelanogenesis (tanning) is the second pigmentary response, which is adaptive in facilitating the development of darker skin on exposed regions of the body protecting against future UVR exposures. The degree to which a person or population responds to UVR is highly variable. For example, a person of Northern European ancestry with the classic ‘red hair and freckles’ phenotype experiences a severe burn but obtains only a minimal tan even after repeated exposures; alternatively, another person of the same ancestry may experience no burning reaction and obtain a substantial tan after only one exposure.

Skin response was found to be important clinically when treatment protocols were established in the 1970s for phototherapy regimes for psoriasis and other skin conditions. In 1975, Thomas B. Fitzpatrick and coworkers put forward a classification scheme to assist in determining the initial doses of UVA in PUVA treatments (1). This system involves the classification of a person into one of six...
sun-reactive skin types (I – always burns/never tans; II – usually burns/tans less than average; III – sometimes mildly burns/tans average; IV – rarely burns/tans more than average). This system initially was established for the clinical treatment of white patients primarily of European ancestry, although phototypes for darker skin (i.e. type V – never burns/brown skin; VI – never burns/black skin) were later added (1). This scale, mainly used for convenience alone rather than for its reliability (2), has been criticized previously for various reasons, including its poor predictive value, poor correlation with skin colour and poor correlation with UV sensitivity (2–5). While the Fitzpatrick scale continues to be quite beneficial to clinicians, the widespread availability of quantitative measures that more accurately reflect the chromophore content of the skin has reduced the utility of the Fitzpatrick scale in research on the physiology of skin response. By considering less subjective measures of skin response, we can plan and execute scientific studies of factors that modify skin response, such as particular genes or treatments.

Despite the availability of reflectance spectroscopy since the 1940s, most dermatological research on skin response has not investigated dose–response curves and instead has focused on subjective single-point determinations of either minimal erythema dose (MED) and minimal melanogenic dose (MMD) (4, 6–10). The continuation of this approach is surprising, as many studies (2, 11, 12) have recognized shortcomings of using only one point on the dose–response curve. It is well known that MED is highly dependent on the observer, incident lighting and room temperature. Additionally, MED has been shown not to correlate with constitutive skin pigmentation, with ancestry, or with the stated Fitzpatrick phototype (2). The main reason for the continuation of this approach, as stated by Wee et al. (12), is that despite being a better measure of UVR sensitivity, slopes of dose–response curves are more ‘tedious’ to determine than MEDs.

The literature lacks studies using diffuse reflectance spectroscopy on substantial sample sizes and diverse populations. As it is well known that there is substantial variability in skin response within and between populations, more data are needed on larger samples of multiple populations measured simultaneously with the same equipment to gain a clearer understanding of this variability as well as to help us understand which measures of erythema and neomelanization can best quantify the skin response.

While the primary aim of this study is to investigate measures of skin response, the motivation for this study is noteworthy. There have been only some attempts to map normal pigmentation genes and, as such, not much is known of the inheritance patterns of these genes. Additionally, little is known about which phenotypes and summary statistics best capture general and specific features of skin pigmentation. For these reasons, the goal of gene mapping for pigmentation phenotypes is a major motivation for conducting this study. When working to identify Quantitative Trait Loci it is important that the phenotypes measured are concise, sensitive and physiologically meaningful.

MATERIALS AND METHODS

Samples for Analysis

We have evaluated 397 individuals living in State College, Pennsylvania for constitutive skin colour and a subset of these (n = 246) have been evaluated for solar response. All of the volunteers were adults (i.e. at least 18 yr of age), with the average age of 24.5 yr and a range between 18 and 78 yr. The sample included 244 females and 153 males. The majority of this sample can be divided into four groups, European-American (n = 289), Hispanic (n = 42), African-American (n = 23) and East Asian (n = 29) based on stated ancestry of the persons studied. The Hispanic group includes persons whose ancestry can be traced to Puerto Rico, Colombia, Mexico, Honduras, Panama and Guatemala. Although these individuals represent a large range of cultural and biological diversity and are united primarily by language, the history of Hispanic populations is such that most have significant ancestry from Native Americans, Spaniards and Africans (13). Every individual in this study gave informed consent prior to beginning the procedure, and this research was performed under the approval of the Pennsylvania State University IRB (IRB No. 00M0558-A4) and the General Advisory Committee of the Penn State General Clinical Research Centre.

Exclusion criteria from the skin response protocol included subjects currently taking medication that could possibly alter an individual’s sensitivity to UVR (e.g. antihistamines; antibiotics like tetracycline; sulpha drugs; quinolone derivatives like Cipro or Noroxin; psychotropic drugs; antidiabetic drugs; cardiovascular drugs; oral contraceptives and non-steroidal anti-inflammatory drugs). Additionally, subjects with certain health conditions (e.g. DNA repair enzyme deficiency like Xeroderma Pigmentosum, Blooms Syndrome, Cockayne Syndrome, or Fanconi Syndrome; collagen vascular diseases like rheumatoid arthritis, systemic lupus erythematosis, perarteritis nodosa, scleroderma, dermatomyositis, or polyarteritis nodosa) were also excluded as these conditions may alter an individual’s sensitivity to UVR.

Clinical Methods

The constitutive skin colour was measured on the upper inner side (medial aspect) of both arms on each subject with a DermaSpectrometer (Cortex Technology, Hasund, Denmark) and a Datacolor International Microflash 200D (Lawrenceville, NJ, USA). The DermaSpectrometer is a narrow band spectroscopy instrument with a green diode centred on 568 nm and a red diode centred on 655 nm, while the Microflash 200D is a diffuse reflectance spectrophotometer that uses a prism photodiode to provide information at 10 nm increments along the visual spectrum (i.e. 31 values from 400 to 700 nm) and standard Commission International de l’Eclairage (CIE) tristimulus values (L*, a* and b*). Volunteers also agreeing to participate in the skin response phase of this study had the medial aspect of both arms exposed to six metered doses of UVR. The Solar Simulator model 16S with liquid light guide (Solar Light Co, Philadelphia, PA, USA) was used with no external filters to expose each individual to doses of primarily UVB at 16.8, 21.0, 26.2,
32.8, 41.0 and 51.3 mJ/cm². Persons not responding with visible erythema to the first series of doses were subjected to a second series of doses (41.0, 51.3, 64.0, 80.1, 100.0 and 125 mJ/cm²). Although erythematous responses reach their maximum levels around 10–20 h after exposure (14); for convenience and consistency with the literature, the sites were examined for erythemal responses 24 h after the initial exposure. Neomelanogenic response was examined 7 d after the initial exposure. In addition to measuring exposed sites, an unexposed site on each arm (i.e., a baseline measurement) was measured at both 24 h and 7 d postexposure. After several attempts at measuring erythema with the Dermaspectrometer, it became clear that the measures were inconsistent despite using the utmost care not to press the glass window too hard on the skin. Therefore, the Dermaspectrometer was not used to measure the skin response because the glass window at the end of the machine’s probe must be touching the skin, which can occlude blood from the region being measured. The Microflash includes a positioning template that both facilitates precise probe localization and prevents direct contact between the skin and probe. Therefore, the Microflash was employed to measure the skin response. Additionally, digital photographs were taken to document the responses at both 24 h and 7 d postexposure.

Statistical Methods

Apparent absorbance (AA) was determined for all of the reflectance measurements at each 10 nm increment provided by the Microflash (15). The percent reflectance (PR) at a specific wavelength was placed into context by relating it to the reflectance of a ‘blank’ at the equivalent wavelength (i.e., relating the object’s reflectance to the maximum reflectance possible). Fig. 1(A) shows data (i.e. one baseline measurement, one response at 24 h, and one response at 7 d postexposure) from one person displayed as PR, and in Fig. 1(B) that same data is shown transformed into AA. For example, the AA of a measurement at 580 nm is equal to the logarithm of the quotient of the PR of the blank at 580 nm and the PR of the object at 580 nm [i.e., log(PR blank / PR object) = AA]. AA could be determined by using a white tile as the blank or by using each individual’s baseline measurement as the blank. While both methods produced consistent results, the white tile was chosen for this study because it was a more efficient way of determining the AA of a large number of samples and preserves the use of the baseline measurement in the analysis of the response. Linear regression was used to compare the skin response to the UV dose and the slope of the entire dose–response curve was calculated.

Measures of melanin concentration included the melanin (M) index, the CIElab lightness metric (L*) and adjusted melanin (AM) index (slope 650:700 nm). The M index was determined directly from the Dermaspectrometer and also using the Microflash with conversion formulas designed to emulate the DermaSpectrometer. As the DermaSpectrometer’s red diode is centred at 655 nm with a half width of 30 nm (as displayed in Fig. 1A), the M index was calculated from a weighted average of the PR data at 640, 650, 660 and 670 nm using the following formulas:

\[
M \text{ index} = 100 \times \log \left( \frac{1}{\text{PR}_{650\text{ nm}}} + \frac{1}{2} \text{PR}_{660\text{ nm}} + \frac{1}{2} \text{PR}_{670\text{ nm}} \right) / 100
\]

Another commonly used measure of melanin is the lightness (L*) from the Commission International de l’Eclairage L*, a*, b* (CIElab) colour system. This system allows for any colour to be described using three values: L* as the summary on a light–dark scale, a* as the summary on the red–green scale and b* as the summary on the yellow–blue scale [Please refer to (16) for further description of the CIElab colour system]. The L* measure is provided by Microflash (and many other...
commonly used reflectometers) and has been one of the traditional dermatological indices for measuring melanin absorbance. The Microflash calculates this measure from the PR data using the 1976 CIElab conversion standards.

Finally, following the method using spectral trace information proposed by Kollias et al. (17), the AM index was calculated as the slope of AA levels from 650 to 700 nm (See Fig. 1B for illustration). This index was used to measure both melanin and melanogenic dose–response.

Neomelanization was evaluated using the change in the measures of melanin in relation to the change in dose of UVR. For example, the change in L* vs. the change in dose (ΔL*), the change in M vs. the change in dose (ΔM), and the change in AM vs. the change in dose (ΔAM), were all tested as to their appropriateness to measure neomelanization.

Measures of erythema included the erythema (E) index, the CIElab red–green metric (a*) and adjusted erythema (AE) index. Microflash measurements of erythema were also converted for comparison with DermaSpectrometer measurements. The latter instrument’s green diode is centred at 568 nm with a half width of 30 nm (as diagramed in Fig. 1A), so a weighted average of Microflash measurements was used to simulate the DermaSpectrometer. As the objective of the E index is to assess only light absorbed by oxyhaemoglobin and deoxyhaemoglobin, absorbance of melanin at the corresponding wavelengths must be accounted for. For reasons mentioned, the following formulas were used:

\[
E_{\text{index}} = 100 \times \log\left(\frac{1}{\text{eqn 3}} - \log\left(\frac{1}{\text{eqn 1}}\right)\right) \tag{4}
\]

As with the measures of melanin, a widely accepted CIElab colour system index, a*, was used to measure erythema. Again, this is provided directly from the Microflash instrument using the 1976 CIElab conversion standards. In the CIElab colour system, a* is a measure of the contrast between the redness and greenness of an object.

The final summary statistic used to measure erythema was AE determined according to Kollias et al. (17). Fig. 1B diagrams how both AM and AE are calculated. First, the slope of the dose–response curve at AA_{680} to AA_{700} nm was determined (i.e. AM). Using this slope and the equation of a line (i.e. \( Y = mX + b \)), the line was extended back to 580 nm (point Y). This shows the absorbance of the skin due to melanin. Accordingly, the magnitude of the absorption by haemoglobin was extracted using the simple formula:

\[
AE = AA_{580\text{nm}} - Y_{580\text{nm}} \tag{5}
\]

Erythemal dose–response was measured by using the same logic used to measure neomelanization. The slopes of the dose–response curves were calculated. Accordingly, the change in E vs. the change in dose (ΔE), the change in a* vs. the change in dose (Δa*), and the change in AE vs. the change in dose (ΔAE) were used to measure the burning response.

As the DermaSpectrometer was not used to measure skin response measurements due to clinical complications discussed above, neither the DermaSpectrometer E index nor the M index was used for the comparative analyses presented below. Instead, the M and E indices calculated from the Microflash were used. Prior to making this decision, linear regression was performed on the data from both machines. On measurements of unexposed skin both the M and E index from the Microflash were highly correlated with those reported by the DermaSpectrometer (\( R^2 = 0.851, P < 0.0001 \) and \( R^2 = 0.816, P < 0.0001 \), respectively).

**RESULTS**

**Measures of Erythema and Erythemal Dose–Response**

In order to compare the various measures of erythema and erythemal dose–response, we calculated the correlation coefficients (\( R^2 \)) comparing three measures of erythema: a*, E index and AE to each other (Fig. 2). These correlation coefficients are provided in Table 1, and all are significant (\( P < 0.0001 \)). AE is highly correlated with both a* and E (\( R^2 = 0.854 \) and 0.989, respectively), and the E index and a* also are highly correlated (\( R^2 = 0.834 \)).

The dose–response slopes of the summary statistics used to measure erythema at 24 h postexposure were also compared to evaluate their efficacy in measuring erythemal dose–response. As shown in Fig. 3, ΔAE and Δa* are highly correlated (\( R^2 = 0.970 \)) at 24 h postexposure; however, the correlation between ΔAE and ΔE index or between Δa* and ΔE index are considerably lower (\( R^2 = 0.612 \) and 0.585, respectively).

![Fig. 2. Comparing measures of erythema. Shown are the relationships between AE and a* (open circles) and AE and E (filled circles) for unexposed skin.](image)

**Table 1. Comparisons of measures of erythema and their use in measuring erythemal dose–response at 24 h and 7 d postexposure**

<table>
<thead>
<tr>
<th></th>
<th>Erythema</th>
<th>EDR at 24 h</th>
<th>EDR at 7 d</th>
</tr>
</thead>
<tbody>
<tr>
<td>AE vs. a*</td>
<td>( R^2 = 0.854 )</td>
<td>( R^2 = 0.973 )</td>
<td>( R^2 = 0.974 )</td>
</tr>
<tr>
<td>AE vs. E</td>
<td>( R^2 = 0.989 )</td>
<td>( R^2 = 0.612 )</td>
<td>( R^2 = 0.662 )</td>
</tr>
<tr>
<td>a* vs. E</td>
<td>( R^2 = 0.834 )</td>
<td>( R^2 = 0.585 )</td>
<td>( R^2 = 0.635 )</td>
</tr>
</tbody>
</table>

\( ^a \) All relationships were significant (\( P < 0.0001 \)); \( ^b \) EDR uses ΔAE, ΔE and Δa*.
Furthermore, the efficacy of these three indices was examined at 7 d postexposure. The relationships are similar to those at 24 h postexposure, as ΔAE and Δa* are more strongly correlated than ΔE index with either ΔAE or Δa* (See Table 1 for details).

### Measures of Melanin and Melanogenic Dose–Response

Table 2 lists the correlation values for the relationships between the measures of pigmentation and melanogenic dose–response, and all results are significant (P < 0.0001). Fig. 4 displays the relationships between the AM index and L* and between the AM index and M index. All of the measures of melanin are highly correlated, with L* and AM having the highest correlation ($R^2 = 0.905$).

The dose–response slopes of the summary statistics used to measure pigmentation were evaluated at 7 d postexposure for their ability to measure melanogenic dose–response (Fig. 5). The correlation between ΔAM and ΔM index is moderate ($R^2 = 0.535$). More importantly, ΔL* is narrowly correlated with ΔAM ($R^2 = 0.075$) or ΔM index ($R^2 = 0.145$) at this time. These results suggest that ΔAM and ΔM index are better able to measure melanogenic dose–response than L*.

To investigate whether residual burning affects the measurements of melanogenic dose–response, ΔL* and ΔAM were compared with Δa* at 7 d postexposure. While ΔAM shows only a minimal correlation with Δa* ($R^2 = 0.021$, $P = 0.031$), ΔL* has a high correlation with Δa* ($R^2 = 0.771$).

Table 2. Comparisons of measures of melanin and their use in measuring melanogenic dose–response at 7 d postexposure

<table>
<thead>
<tr>
<th>Pigmentation</th>
<th>MDR$^{a,b}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AM vs. L*</td>
<td>$R^2 = 0.905$</td>
</tr>
<tr>
<td>AM vs. M</td>
<td>$R^2 = 0.771$</td>
</tr>
<tr>
<td>L* vs. M</td>
<td>$R^2 = 0.891$</td>
</tr>
</tbody>
</table>

$^a$ All relationships were significant (P < 0.0001); $^b$ MDR uses ΔL*, ΔM and ΔAM.

DISCUSSION

While all of the summary statistics of erythema investigated appear to be measuring erythema in unexposed skin similarly, there are some major differences at 24 h postexposure. Δa* appears to be equally effective in measuring the burn response as ΔAE; however, the ΔE index does not perform as well as the other two measures. Although many of the ΔE data points are highly correlated with other measures, there is a substantial subset of measures that are very far from the bulk, and it is these measures that are diminishing the usefulness of ΔE. It is unknown as to why this index would be ineffective. At 7 d postexposure, the situation was similar for all measures. Both ΔAE and Δa* are more highly correlated with each other than either is with ΔE, perhaps because they allow for a more precise assessment of the absorbance due to

Fig. 3. Measures of erythemal dose–response at 24 h postexposure. Shown are the relationships between ΔAE (AE slope) and the Δa* (a* slope; open circles) and ΔAE (AE slope) and the ΔE (E slope; filled circles).

Fig. 4. Comparing measures of pigmentation. Shown are the relationships between AM and the L* (open circles) and AM and the ΔE (filled circles) for unexposed skin.

Fig. 5. Measures of melanogenic dose–response at 7 d postexposure. Shown are the relationships between ΔAM (AM slope) and the ΔL* (seven L* slope; open circles) and ΔAM (AM slope) and the ΔM (seven M slope; filled circles).

0.7514, $P < 0.0001$). Thus, it is likely that the main reason why ΔL* shows a much lower correlation with ΔM or ΔAM is because redness remaining at 7 d is influencing the measurement.
erythemal response isolated from other chromophores like melanin or bilirubin. The strong correlation between $\Delta a^*$ and $\Delta AE$ is reassuring, because it indicates that the more commonly used CIElab measure, $a^*$, is a valid substitute for the more elaborately computed but physiologically informed measure, $\Delta AE$.

Given the high correlations, all three measures of melanin ($L^*$, $M$ index and $AM$) are suitable for measuring constitutive skin colour; however, the dose–response slopes ($\Delta L^*$, $\Delta M$ index and $\Delta AM$) are not as strongly related at 7 d postexposure. In particular, $\Delta L^*$ shows a much lower correlation when compared with $\Delta M$ index and $\Delta AM$ than $\Delta M$ and $\Delta AM$ when compared with each other. As $L^*$ is calculated as a composite value of reflectance levels at all wavelengths, residual erythema absorbing at wavelengths less than 650 nm may be confounding its ability to measure neomelanogenesis by making the skin appear darker than it really is.

However, $\Delta AM$ is a more accurate measure of melanogenic dose–response than other traditional indices, as it allows us to measure neomelanization without the complications of residual erythema. Furthermore, as $\Delta AM$ is only measuring wavelengths at the red end of the visual spectrum, where haemoglobin does not absorb, we can assume that the only chromophore being measured is melanin.

The findings of this research are vitally important to the success of studying variation within and between populations and are crucial when using these phenotypes to find the genes responsible for normal variation in skin pigmentation and sensitivity to UVR. Both $\Delta AE$ and $\Delta a^*$ are valid measures of erythema and erythemal dose–response. While $L^*$ is recognized as an effective measure of constitutive pigmentation, it would not be recommended to use it in measuring melanogenic dose–response as it is affected considerably by residual erythema. Instead, $\Delta AM$ is the recommended index when measuring melanogenic dose–response.

Acknowledgements – We would like to thank the volunteers who gave their time, skin and DNA for this research. We also acknowledge the support of the Pennsylvania State University General Clinical Research Centre (M01RR10732). This work was supported in part by a grant from NIH/NHGRI (HG002154) to MDS. Communication with N. Kollias was also valuable.

REFERENCES