Department of Veterans Affairs

Journal of Rehabilitation Research and Development Vol. 38 No. 1, January/February 2001
Pages 13–22

Testing the validity of erythema detection algorithms

Brian Riordan, MEBME; Stephen Sprigle, PhD, PT; Maureen Linden, MSBME
Center for Rehabilitation Technology, Helen Hayes Hospital, West Haverstraw, NY 10993

Abstract—Dermatology has quantified skin color for monitoring progress of treatments. The most common and effective means of erythema detection is visual inspection of the skin. However, for people with darkly pigmented skin, erythema can be masked by melanin. Tissue Reflectance Spectroscopy (TRS) is a noninvasive method of quantifying skin color. Most commonly, TRS quantifies erythema caused by cosmetics, topical ointments, UV light, or other irritants. Recently, TRS has been used to characterize the presence of erythema due to reactive hyperemia or Stage I pressure ulcers. The objective of this study was to compare the reliability and validity of erythema detection algorithms by determining their sensitivity and specificity. Two algorithms, Diffey and Helen Hayes Hospital (HHH), had sensitivity exceeding 85% and specificity exceeding 75%, but most algorithms demonstrated adequate validity across all subjects. The validity of the HHH algorithm did not change with the skin pigmentation of the subject. The results of this comparison will be useful to researchers interested in using TRS to detect erythema in people with different skin pigment levels.

Key words: *algorithm, erythema, melanin, skin pigmentation, tissue reflectance spectroscopy.*

INTRODUCTION

The field of dermatology has long been interested in inspecting skin color and changes in skin color during and after treatment. One common noninvasive measurement is via tissue reflectance spectroscopy (TRS). The TRS uses the characteristic absorption of light by the constituents of the skin to measure the amount of various constituents present. A white light is shone on the skin while detectors measure the returning light. From the spectral characteristics of the light returned, the relative amounts of each skin chromophore can be determined. Most erythema studies using TRS target the effects that cosmetics, topical ointments, UV light, or other irritants on the skin (1,2). Recently, TRS has been used to characterize the presence of erythema due to reactive hyperemia or Stage I pressure ulcers (3).

To detect and quantify the erythema, several algorithms using TRS absorption data have been reported in the literature. These algorithms have been tested primarily with lightly pigmented subjects; hence, their ability to detect erythema in darkly pigmented tissue remains unknown. The objective of this study was to determine the reliability and validity of erythema detection algorithms for all pigment levels. Five methods reported in the literature, and a sixth, developed during ongoing research with diabetic amputees, were investigated. Validity was determined using the sensitivity and specificity of each. In this case, sensitivity is the ability to cor-

This material is based on work supported by the National Institute on Disability and Rehabilitation Research, Award #H133G50018; the National Institutes of Health, Award #HD 31428-03; and the New York State Department of Health.

Address all correspondence and requests for reprints to: Stephen Sprigle, PhD, PT, Center for Rehab Technology, Helen Hayes Hospital, Rt. 9W, West Haverstraw, NY 10993; email: gogators@compuserve.com.

rectly identify tissue with erythema, and specificity is the ability to correctly identify tissue without erythema.

The results of these comparisons will be useful to researchers interested in using TRS to detect erythema in people with different skin pigment levels. One specific benefit of a robust erythema detection algorithm is the development of an instrument for use by health care professionals to detect erythema. This can be useful in monitoring reactive hyperemia or detecting Stage I pressure ulcers in deeply pigmented subjects. Detection of a Stage I ulcer will allow timely intervention to prevent progression of the ulcer.

Tissue Reflectance Theory

The TRS uses the characteristics of light reflected by the constituents of the skin to measure the amounts of various constituents present. A white light is shone on the skin while detectors measure returning light. The light is divided into spectral components by a monochromator and detected by a photomultiplier tube. To eliminate transmission effects of the instrumentation, reflectance from the skin is compared to reflectance from a white standard. The result is a unitless quantity termed relative reflectance. Absorption, or Log Inverse Reflectance (LIR), refers to the amount of light not returned from the skin. This is calculated as the name suggests, with the result expressed in absorption units (au):

$$\text{Absorption} = 1/\log_{10}(\text{relative reflectance}) \qquad [1]$$

The theory of TRS is based upon a simple anatomical model (4). Light passes through the epidermis (melanin layer) and a plexus of blood vessels in the dermis (hemoglobin (Hb) layer) before being reflected off collagen in the lower dermis. This model, shown in **Figure 1**, highlights the human tissue chromophores within the visible spectrum that influence the measurement of erythema: Hb, oxyhemoglobin (oxyHb), and melanin.

The Hb absorbs light with a characteristic curve showing broad bands of absorption in the green portion of the spectrum. The oxyHb has two absorption maxima at 542 nm and 574 nm wavelengths. De-oxyhemoglobin (deoxyHb) shows a single maximum at 545 nm. Thus, TRS can theoretically produce information about both the amount of Hb present and its degree of oxygenation. Melanin has a linearly decreasing curve in the spectral range from 500–700 nm. The slope of this curve increases as the melanin content of an individual's skin increases. For people with darkly pigmented skin, the absorption by melanin can be much greater than that of Hb (**Figure 2**). This masking by melanin makes it difficult to detect erythema.

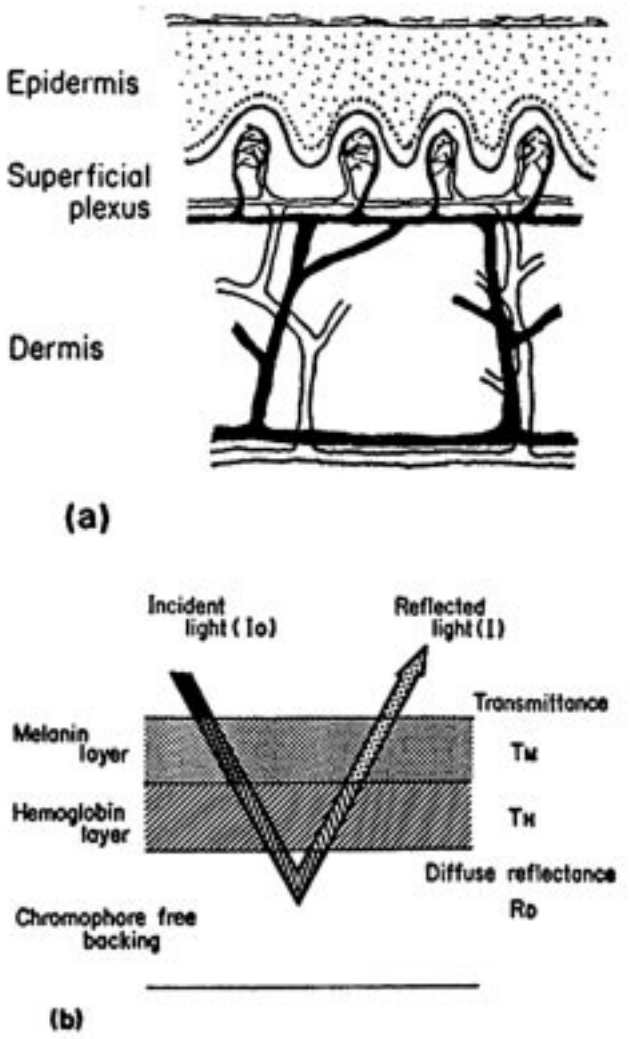

Figure 1.
a) Schematic structure of the skin; b) Optical skin model of three-layered structure with an outer melanin layer, an upper hemoglobin layer, and a backing representing chromophore-free dermis. Reprinted with permission from (Takiwaki S, Serup H. Measurement of erythema and melanin indices. In: Serup H, Jemec GBE (editors). Handbook of noninvasive methods and the skin. Boca Raton: CRC Press; 1955. p. 378, Figure 1 with caption), Copyright CRC Press Boca Raton, Florida.

Figures 3 and **4** illustrate the difference in TRS spectra obtained from lightly and darkly pigmented individuals. The absorption curve of the individual with deep pigmentation has greater overall amplitude (increased light absorption) and the Hb "double hump" is not as easily discerned.

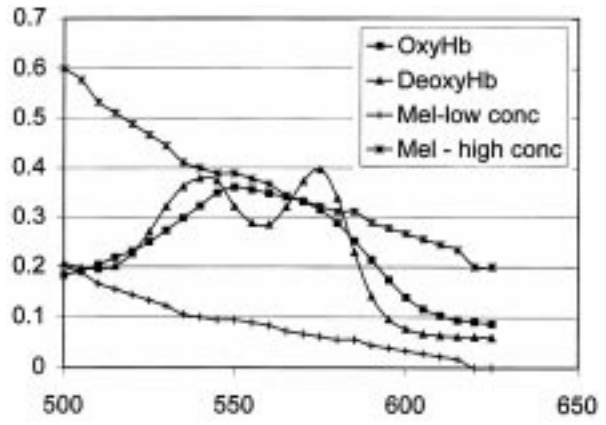

Figure 2.
Characteristic curves of human skin chromophores.

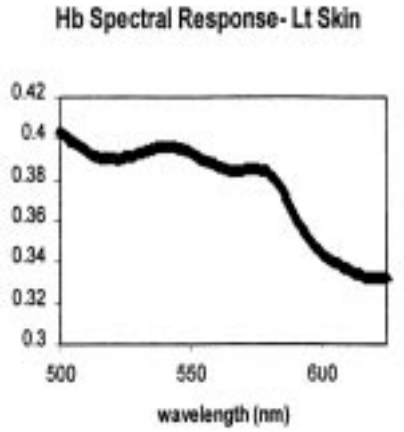

Figure 3.
Absorption spectral response curve for an individual with lightly pigmented skin.

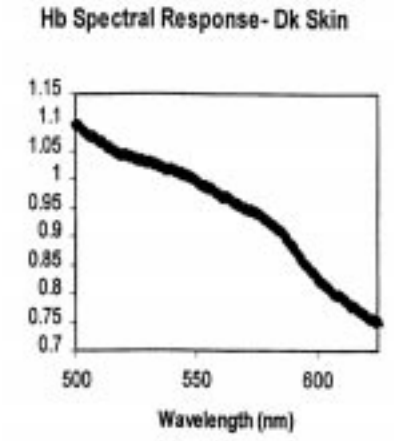

Figure 4.
Absorption spectral response curve for an individual with deeply pigmented skin.

Algorithms Tested

The use of TRS to detect pressure ulcers in people with darkly pigmented skin is dependent on two processes:

1. a means to adjust or correct the signal for melanin

Ideally, melanin compensation would be independent of the amount of blood present and effective for all levels of pigmentation.

2. a robust erythema detection algorithm

Five methods of erythema detection were identified in the literature. A sixth was developed during ongoing research on diabetic amputees and nondisabled controls. Each of these algorithms uses its own method of calculating an index of erythema. Most, but not all, also quantify and correct for melanin. These methods are described briefly below. For detailed descriptions of these algorithms, the reader should consult the cited references.

1. Dawson

Dawson based melanin compensation on the difference of absorption from 650 nm to 700 nm. He averaged absorption values at 645, 650, and 655 nm and subtracted the average of measurements at 695, 700, and 705 nm (5). Because of differences in the response of our instrument, we calculated the index based on the difference between an average of ten wavelengths from 640–650 nm and ten from 690–700 nm, respectively.

$$M_{Daw}=(L_{640\text{-}650}-L_{690\text{-}700})*100 \quad [2]$$

where L_{x-y} is the average absorption in the wavelength range from x to y nm.

Dawson's erythema index, E_{Daw}, is a parameter that is proportional to the area under the Hb absorption curve when an artificial baseline is drawn between 510 nm and 610 nm.

$$E_{Daw}=100\left[A_{560}+1.5*(A_{540}+A_{575})-2(A_{510}+A_{610})\right] \quad [3]$$

where A_n is the absorption of the spectrum at wavelength n (5).

E is corrected for melanin by adding the melanin index scaled by a factor γ(=0.04). This factor was empirically determined based on the assumption that lightly and darkly pigmented subjects have similar blood content.

$$E_C = E_{daw}(1 - \gamma M_{Daw}) \quad [4]$$

2. Ferguson-Pell

Ferguson-Pell defined a melanin index (IMEL) as the slope of the regression line for the bloodless (blanched) absorption spectrum from 500–600 nm (6). After the slope of the bloodless spectrum is subtracted from every spectral curve, the response is attenuated by a multiplication factor, which was a function of IMEL and was empirically determined based on the assumption that lightly and darkly pigmented subjects had similar skin blood content (6).

Ferguson-Pell than quantified erythema by the amount of Hb in the skin surface (IHB) as developed by Feather (7). The IHB is based on the isobestic wavelengths of Hb, or the wavelengths where Hb absorption is not affected by its oxygenation level (absorption of oxyHb=absorption of deoxyHb at these wavelengths). The formula for IHB is shown below (Equation 5), where A_n is the absorption of the spectrum at wavelength n after attenuation compensation for melanin, as described above (7).

$$IHB = 50^* \left[\frac{A_{545} - A_{522}}{23} - \frac{A_{568} - A_{545}}{23} \right] \quad [5]$$

3. Hajizadeh-Saffar

Hajizadeh-Saffar also used the work of Feather to calculate IHB, but first corrected for light backscatter in the epidermis. Melanin compensation is based on the slope of the absorption curve from 650–700 nm. Initial estimates are used to compensate for blood content and oxygenation in the skin (8).

$$M_{675} = \{(A_{700} - A_{650})/50 + (0.060[1\text{-}SaO2c/100] + 0.010)IHBc/80\}^*100 \quad [6]$$

where: A_n is the corrected absorption at wavelength n; SaO_2c is an index of blood oxygenation corrected for backscatter; and, IHBc is the Hb index corrected for backscatter.

The melanin index can be expressed in terms of the concentration of synthetic melanin required to produce the same response *in vitro*.

$$M_{Haj} = (0.054 - M_{675})^*120.1\ \mu g/cm^2 \quad [7]$$

"True" IHB equals the corrected IHB plus the IMEL expressed in ug/cm^2, scaled by an empirically determined factor (8).

$$IHBt = IHBc + .00047^*M_{Haj} \quad [8]$$

4. Diffey

Diffey noted that Hb had high absorption in the green spectrum and low absorption in the red spectrum, and, further, that erythema caused by vasodilation produced significant increases in green absorption and little change in red absorption. His premise was that differences between red and green absorption were due solely to Hb content and that melanin compensation would be accomplished by comparing the two spectral regions (9).

$$E_{dif} = \log_{10}\left(\frac{REF(635nm)}{REF(565nm)}\right) \quad [9]$$

where REF(x)=Reflectance at wavelength x nm.

5. Tronnier

Tronnier's erythema index is based on the difference between red and green reflectance at a control site and an erythematic site. Melanin compensation is achieved by comparing two sites (10):

$$E_{Tro} = (G - R) - (G_O - R_O) \quad [10]$$

where: G, G_o=reflectance at 545 nm (erythematic, control site) R, R_o=reflectance at 661 nm (erythematic, control site)

One should note that Tronnier subtracted reflectance values at two sites, the first known to have erythema caused by ultraviolet radiation while the other a control site. The control site was measured primarily to evaluate the pigmentation of a subject. This technique produces a single value, characterizing erythema of one site relative to the other.

Because our goal in this analysis is to develop an algorithm useful for determining whether a site has erythema, strict application of Tronnier's algorithm is not

appropriate. Instead, we subtracted green reflection from red reflection using the formula below and quantified the differences between sites with percent differences and Z-score calculations.
where: G=reflectance at 545 nm; R=reflectance at 661 nm.

$$E_{Tro} = (G - R) \quad [11]$$

6. Helen Hayes Hospital (HHH)

A concentration-independent curve for melanin was calculated from *in vitro* data. This slope is scaled for each individual by the difference in absorption from 500 nm to 625 nm. The strengths of this approach include its focus over the area of interest and its ability to accommodate both lightly and darkly pigmented skin. Its weaknesses include an incomplete understanding of how melanin concentration changes over adjacent areas of tissue (i.e., a control site and an adjacent "red" site).

The melanin curve is subtracted from the spectrum. The compensated spectrum is regressed using a standard, concentration-independent absorption curve of Hb (*in vitro* Hb response) as the regressor. The coefficient (beta code this to match the equation and make a note on the hard copy) of this regression is a measure of Hb concentration. If necessary, this model can be expanded to include separate regressors corresponding to oxyHb and deoxyHb.

METHODS

Algorithms were tested by measuring tissue reflectance at an erythematic site and at two adjacent control sites. As mentioned previously, the use of TRS to measure erythema depends upon determining melanin and Hb, so the protocol was designed to permit calculations of both. Each detection algorithm was applied to the same reflectance data. In addition to the six algorithms defined above, a seventh algorithm was tested that applied the HHH melanin correction approach to erythema detection with the Dawson algorithm.

Equipment

A Monolight 6800 series spectrophotometer (Rees Instruments, Smyrna, GA), consisting of scanning monochromator, photodetector, and light source with a fiber optic cable, was used for TRS data collection. A rotating diffraction grating dispersed incident light into component wavelengths that were then converted into electrical signals by the 6117 photodetector. The grating blaze wavelength, 500 nm, yields an optimal operating range from 350–1100 nm, a dispersion of 10 nm/mm, and a resolution of 1 nm in the visible spectrum. A 6162 stabilized tungsten halogen lamp provided the light source. A second order blocking filter (WG345, range 350–640 nm) negated the aliasing effects produced by the rotation of diffraction grating.

The bifurcated fiber optic cable (Fiber Guide Industries, Caldwell, ID) consisted of 2,370 borosilicate fibers. Roughly half the fibers transmitted the incident light, while the remaining fibers carried light back to the photodetector. The end of the fiber optic cable was fitted with a probe with a 3-mm diameter optic aperture. The aperture side of the probe had a slightly convex shape to minimize edge effects during pressure application, and a merosphere was affixed to the top of the probe to interface with the pressure application system.

Localized erythema was induced on the shank of subjects using a pneumatic indentor that applied 150 mmHg. The pressure application system consisted of a computer-controlled pneumatically driven piston. The piston is equipped with a load cell (Entran ELF-C1000-10), which provided feedback to the controller and maintained 150-mmHg pressure while accommodating for slight body movements during monitoring. LabTech Notebook Pro, Ver 10 (Laboratory Technologies Corp, Wilmington, MA) was used to control the system. Details of the indentor system have been previously published in this journal (11).

Data Collection

A convenience sample of 20 subjects was recruited from hospital staff. The skin of each subject was classified according to a Munsell color chart (5YR; reference 12). The Munsell chart consists of 33 tiles of color that are scaled on value (2.5–8) and chroma (1–8). The value quantifies the darkness of the skin, with lower numbers indicating darker skin, while the hue quantifies the ruddiness of complexion, with higher numbers indicating reddish complexion. Of the 20 subjects tested, 9 were Caucasian (values from 7–8), 3 were Hispanic (values from 6–7), and 8 were Black (values from 3–6).

A consent form that described the project and risks was provided to each subject. After informed consent was obtained, subjects were positioned in a semi-recumbent position on a mat. The shank of the test leg was exposed

to an area near the tibial flare. This area was tested because it is prone to reactive hyperemia and pressure ulcers in patients with transtibial amputation who ambulate with patellar tendon- bearing prostheses. Test sites that were free of hyper- or hypopigmentation and scarring were identified, cleaned, and shaved, if necessary. Three electrode adhesive rings were affixed to the skin, one marking the test area and the other two at the two adjacent control sites. The electrode adhesive rings were slightly larger than the spectrometer aperture (3-mm diameter), allowing repeated measures to be performed at each test site. The TRS measurements were taken at each control site before inducing erythema, to permit calculation of melanin.

Localized erythema was induced on the shank of subjects using a pneumatic indentor that applied a 150-mmHg pressure for 3 min. Reflectance data from each site was measured with the spectrometer for 30 s at low contact force and 10 s with a contact force sufficient to blanch the skin. The optic head of the spectrophotometer was removed and re-seated on the skin between each measure, to create independent measures at each site.

The protocol was based upon the potential for using TRS as a means to identify sites at risk for pressure ulcer development or any other clinical monitoring. Therefore, TRS measurements were done for short periods of time (40 s) and a researcher positioned and held the optic head against the skin, using a hand-held pressure application system. This pressure application system allowed positioning of the optical head such that light was emitted normal to the skin surface, while ensuring that the pressure applied was less than 40 mmHg. Previous investigation had concluded that, for this optical head and this anatomical location, 40 mmHg does not cause blanching of the skin. The pressure application system also ensured a pressure greater than 150 mmHg during the blanching of the skin required for calculation of IMEL. The minor variations caused by hand-held monitoring assisted in testing the robust nature of each algorithm.

Repeatability was tested by measuring each site twice, in the following sequence:
Control site 1
Control site 1
Control site 2
Control site 2
Erythema site
Erythema site

The control site measures were performed while the test site was being loaded. Because reactive hyperemia is a transient event, the subsequent measures of erythema were not expected to yield the same Hb content; rather they were used to test the detection algorithm's ability to identify the presence of erythema compared to control sites. Reactive hyperemia typically lasts between 50–70 percent as long as the ischemic event (13), in this case 90 s, so the second measure of erythema at the test site falls well within this period.

Data Analysis

This data collection permitted sensitivity to be determined by comparing the erythema test site to the control sites, and allowed specificity to be determined by comparing the control sites. In other words, sensitivity determined whether the test sites were different than the control sites (an erythema site was determined to have erythema or a true positive result) and specificity determined whether control sites were not different (a non-erythema site was determined to be non-erythematic or a true negative response). The melanin correction and erythema detection algorithms were performed as described above.

Within each detection algorithm, four detection criteria were studied. Three Z-scores and a 3-percent change in Hb content were used within each algorithm to determine the optimal criteria.

Z scores were calculated using the formula: $Z=(\mu_{siteA}-\mu_{siteB})/\sigma_{siteB}$.
Percent change was calculated using: $\Delta=(\mu_{siteA}-\mu_{siteB})/\mu_{siteB}$.

Z scores have the benefit of using both the mean and standard deviations and represent the normalized value from a distribution. Percent change uses only the mean at both sites, but is more intuitive for most persons.

Within each detection algorithm and detection criteria, the data were categorized into a 2×2 logic table divided into True-positive, True-negative, False-positive, and False-negative designations. Tabulated results were used to calculate:
Sensitivity=True-positive/(True-positive+False-negative)
Specificity=True-negative/(True-negative+False-positive)
where: True-positive=erythema sites identified as erythema sites
True-negative=non-erythema sites identified as non-erythema sites
False-positive=non-erythema sites identified as erythema
False-negative=erythema sites identified as non-erythema

Reliability was determined by analyzing the two independent measures at each control site. Because neither site was erythematic, the repeated measures should

be consistent quantitatively and should not differ by the threshold criteria. In other words, a site should not be deemed different from itself after repeated measures. Therefore, repeatability was calculated in 2 manners: 1) Intraclass correlation coefficient (ICC) to measure the quantitative consistency of the erythema measurement and specificity of the control sites [true-negative/(false-positive+true-negative)] to determine how often repeated measures of the same site differed by the threshold (false-positive result).

RESULTS

The optimal Z-score and percent difference for each of the algorithms and the resulting validity measures are contained in **Table 1**. The "optimal" criteria were determined using four conditions: 1) sensitivity≥0.80 for the three groups (All Subjects, Dark Skin Subjects, and Light Skin Subjects); 2) sum of sensitivity+specificity value; 3) sensitivity≥specificity; and, 4) dark skin total sum. Appendix A lists the sensitivity and specificity of the six criteria used for each algorithm. **Table 2** lists the reliability results, including ICC and the specificity of the control sites for each optimal criterion.

The HHH was the only algorithm to exceed 80 percent sensitivity and specificity levels for both the Z-score and percent difference thresholds. The Diffey algorithm, using Z=2.5 threshold, approached these sensitivity and specificity levels with values at 88 percent and 78 percent, respectively.

The reliability of all algorithms was high if using ICC as the measure, but only Dawson, IHB, and HHH

Table 1.
Algorithm validity

Algorithm	Z Score Criteria	All Subjects		Dark-Skin Subjects		Light-Skin Subjects	
		Sensitivity	*Specificity*	*Sensitivity*	*Specificity*	*Sensitivity*	*Specificity*
Dawson	3	0.809	0.765	0.815	0.718	0.804	0.81
Dawson-HHH	2	0.827	0.758	0.815	0.667	0.839	0.845
Diffey	2.5	0.88	0.784	0.852	0.744	0.907	0.821
Hajizadeh	1	0.691	0.883	0.667	0.897	0.714	0.869
HHH	4	0.9	0.846	0.944	0.859	0.857	0.833
IHB	1	0.609	0.901	0.611	0.91	0.607	0.893
Tronnier	3	0.818	0.79	0.796	0.718	0.839	0.857
Algorithm	**% Difference Criteria**	**All Subjects**		**Dark-Skin Subjects**		**Light-Skin Subjects**	
		Sensitivity	*Specificity*	*Sensitivity*	*Specificity*	*Sensitivity*	*Specificity*
Dawson	10	0.818	0.769	0.741	0.782	0.893	0.81
Dawson-HHH	5	0.836	0.747	0.815	0.744	0.857	0.75
Diffey	20	0.845	0.741	0.796	0.769	0.893	0.714
Hajizadeh	20	0.832	0.68	0.87	0.636	0.8	0.714
HHH	40	0.873	0.887	0.907	0.91	0.833	0.861
IHB	20	0.84	0.703	0.842	0.704	0.839	0.702
Tronnier	20	0.845	0.698	0.796	0.705	0.893	0.69

Table 2.
Algorithm reliability

Algorithm	ICC	Reliability	
		Using Z Scores	*Using % Difference*
Dawson	.795	.63	.75
Dawson-HHH	.838	.53	.52
Diffey	.891	.65	.63
Hajizadeh	.655	.37	1.0
HHH	.812	.75	.88
IHB	.998	.72	.43
Tronnier	.921	.60	.55

demonstrated reliability exceeding 70 percent when calculated with respect to the optimal threshold. This difference can be attributed to the meanings of each respective reliability measure. The ICC reflects the variance of values within each repeated measure compared to the variance across subjects. Conversely, repeatability with respect to difference thresholds is obviously dependent on the threshold magnitudes and how they relate to the variation of same-site, repeated measures.

DISCUSSION AND CONCLUSIONS

The results indicate that certain algorithms are able to detect erythema in skin of varying pigmentation. To date, few erythema research studies addressed skin pigmentation. This conclusion may be used in the development of clinical tools to detect erythema.

Often, sensitivity and specificity are inversely related in clinical tests that measure a continuous variable (e.g., blood sugar, blood pressure). Since erythema is a continuous variable, changing the threshold criteria used to indicate erythema can alter the sensitivity and specificity. This project tested six thresholds and designated "optimal" thresholds for each algorithm according to the criteria listed above. Sensitivity was deemed more important than specificity but the sum total of both were also included.

Calculating both sensitivity and specificity is important in the study of detection algorithms because these parameters affect Type I and Type II errors. Balancing these errors depends on many factors such as costs of the test and costs associated with false positives. In certain instances, one may need to maximize sensitivity in order to ensure that sites with erythema are identified with minimal error. Alternatively, maximizing specificity would ensure that skin free of erythema is identified accurately.

If a detection algorithm was used clinically to identify tissue at risk for ulcer formation, the risk of incorrectly identifying erythema when it was not present (false-positive) would probably lead to increased monitoring of that site. While this error might lead to additional staff time, it should not have any adverse effects on the patient. Conversely, if an erythema site was incorrectly identified as not erythematic (false-negative), the patient is at risk. Therefore, in judging algorithms, sensitivity was deemed more important than specificity because the higher the sensitivity the lower the risk of a false-negative finding, and thus a lower risk to the patient.

Two algorithms, HHH and Diffey, exceeded 85 percent sensitivity and the other 5 algorithms had criteria that led to a sensitivity exceeding 80 percent. These are acceptable values for clinical tests, especially given the concomitant specificity values of most of the algorithms (>75 percent).

The Z=1 criteria of IHB and Hajizadeh are suspect as viable thresholds, as are the 10 percent and 5 percent difference criteria for the two Dawson algorithms. The inherent variability of TRS, as with many physiological variables, leads to questioning the use of very low detection thresholds. In other words, a low detection threshold might lack robustness in a clinical setting as opposed to a controlled research setting. This intuition, however, was not completely corroborated by the reliability measures. While the Hajizadeh algorithm returned relatively high false-positive results using a Z=1 threshold, IHB exhibited a specificity of 72 percent. Reliability of the Dawson-HHH algorithm (52 percent) was affected by the low percent-difference threshold, but the Dawson algorithm exhibited an acceptable reliability (75 percent).

The limitations of the study include both risks to external and internal validity. In this study, erythema was induced via a reactive hyperemic response to localized ischemia. While this is consistent with pressure ulcer formation, these results might not generalize to erythema induced by other means (e.g., topical ointments, UV, full limb ischemia). Similarly, the small number of subjects might also limit external validity, although this study used a variety of skin pigmentations to test the validity of each algorithm. An internal validity risk might be due to instrumentation bias. The influence on these algorithms is unclear. Theoretically, the algorithms should act independently of the spectrometer used, as long as its accuracy, resolution, and precision are adequate. The spectrometer used in this study had adequate spectral range and a 1-nm resolution.

APPENDIX

Raw data

	Criteria	Results					
Algorithm	*Z Score*	*Overall*		*Dark*		*Light*	
		Sens	Spec	Sens	Spec	Sens	Spec
HHH	5	0.855	0.870	0.907	0.897	0.804	0.845
	4	0.900	0.846	0.944	0.859	0.857	0.833
	3.5	0.918	0.778	0.963	0.756	0.875	0.798
IHB	2	0.373	0.988	0.315	1.000	0.429	0.976
	1.5	0.500	0.963	0.481	0.974	0.518	0.952
	1	0.609	0.901	0.611	0.910	0.607	0.893
Dawson	5	0.752	0.877	0.759	0.810	0.745	0.940
	4	0.782	0.815	0.796	0.744	0.768	0.881
	3	0.809	0.765	0.815	0.718	0.804	0.810
Dawson-HHH	4	0.727	0.870	0.685	0.821	0.768	0.917
	3	0.782	0.809	0.741	0.744	0.821	0.869
	2	0.827	0.759	0.815	0.667	0.839	0.845
Hajizadeh	2	0.482	0.969	0.444	1.000	0.518	0.940
	1.5	0.573	0.944	0.519	0.974	0.625	0.917
	1	0.691	0.883	0.667	0.897	0.714	0.869
Diffey	3	0.845	0.790	0.833	0.744	0.857	0.833
	2.5	0.880	0.784	0.852	0.744	0.907	0.821
	2	0.880	0.741	0.852	0.718	0.907	0.762
Tronnier	4	0.791	0.815	0.778	0.769	0.804	0.857
	3	0.818	0.790	0.796	0.718	0.839	0.857
	2.5	0.827	0.722	0.815	0.667	0.839	0.774

	Criteria	Results					
	Percent	Overall		Dark		Light	
Algorithm	Difference	Sens	Spec	Sens	Spec	Sens	Spec
HHH	40	0.873	0.887	0.907	0.910	0.833	0.861
	30	0.901	0.813	0.926	0.833	0.872	0.792
	20	0.945	0.722	1.000	0.744	0.893	0.702
IHB	30	0.809	0.732	0.789	0.611	0.821	0.810
	20	0.840	0.703	0.842	0.704	0.839	0.702
	15	0.830	0.638	0.842	0.593	0.821	0.667
Dawson	20	0.591	0.932	0.481	0.936	0.696	0.929
	10	0.818	0.796	0.741	0.782	0.893	0.810
	5	0.882	0.660	0.833	0.654	0.929	0.667
Dawson-HHB	15	0.664	0.932	0.574	0.923	0.750	0.940
	10	0.755	0.840	0.685	0.821	0.821	0.857
	5	0.836	0.747	0.815	0.744	0.857	0.750
Hajizadeh	30	0.784	0.773	0.826	0.712	0.750	0.821
	20	0.832	0.680	0.870	0.636	0.800	0.714
	10	0.847	0.600	0.783	0.591	0.904	0.607
Diffey	20	0.845	0.741	0.796	0.769	0.893	0.714
	15	0.873	0.679	0.833	0.705	0.911	0.655
	10	0.891	0.636	0.870	0.641	0.911	0.631
Tronnier	30	0.736	0.759	0.611	0.782	0.857	0.738
	20	0.845	0.698	0.796	0.705	0.893	0.690
	15	0.861	0.648	0.827	0.679	0.893	0.619

REFERENCES

1. Feather JW, Ryatt KS, Dawson JB. Reflectance spectrophotometric quantification of skin colour changes induced by topical corticosteroid preparations. Br J Dermatol 1982;106:436–40.
2. Wan S, Parrish JA, Jaenicke KF. Quantitative evaluation of ultraviolet induced erythema. Photochem Photobio 1983;37:643–8.
3. Hagisawa S, Ferguson-Pell M, Cardi M, Miller SD. Assessment of skin blood content and oxygenation in spinal cord injured subjects during reactive hyperemia. J Rehabil Res Dev 1994;31(1):1–14.
4. Takiwaki H, Serup, J. Measurement of erythema and melanin indices. In: Serup J, Jemee GBE (editors). Handbook of non-invasive methods and the skin. Boca Raton, FL: CRC Press; 1995. p 377–97.
5. Dawson JB, Barker DJ, Ellis DJ, Grassam E, Cotterill JA, Fisher GW, Feather JW. A theoretical and experimental study of light absorption and scattering by *in vivo* skin. Phys Med Biol 1980;25(4)695–709.
6. Ferguson-Pell M, Hagisawa S. An empirical technique to compensate for melanin when monitoring skin microcirculation using reflectance spectrophotometry. Med Eng Physics 1995;17(2):104–10.
7. Feather JW, Hajizadeh-Saffar M, Leslie G, Dawson JB. A portable scanning reflectance spectrophotometer using visible wavelengths for the rapid measurement of skin pigments. Phys Med Biol 1989;34(7):807–20.
8. Hajizadeh-Saffar M, Feather JW, Dawson JB. An investigation of factors affecting the accuracy of *in vivo* measurement of skin pigments by reflectance spectrophotometry. Phys Med Biol 1990;35(9):1301–15.
9. Diffey BL, Farr PM. Quantitative aspects of ultraviolet erythema. Clin Phys Physiol Meas 1991;12(4):311–25.
10. Tronnier H. Evaluation and measurement of ultraviolet erythema. In: F Urbach (editor). Biologic effects of ultraviolet radiation. Oxford: Pergaman Press; 1969. p. 255–66.
11. Ferguson-Pell MW, Hagisawa S, Masiello RD. A skin indentation system using a pneumatic bellows. J Rehabil Res Dev 1994;31(1):15–9.
12. Young KA. A study of the accuracy of pulse oximetry among varying skin pigment populations [master's thesis]. Long Beach (CA): California State Univ.; May 1994.
13. Landis EM. Micro-injection studies of capillary blood pressure in human skin. Heart 1930;15:209–28.

Submitted for publication March 3, 2000. Accepted in revised form June 22, 2000.