Using hue, saturation and intensity compared with three other pressure ulcer grading systems

Linda Rafter, Tim Reynolds

Abstract

Background: New methods of defining an early grade of pressure ulcer could have great economic benefit to both patient and health service. Objectives: The objective of the study was to find out whether pressure ulcers can be assessed and classified by using computer image analysis of hue, saturation and intensity (HSI) in order to make a prognostic assessment of early skin damage. Aims: To determine whether different grades of pressure ulcers exhibit specific hues. To develop a consensus of opinion from four experts on the grade of pressure ulcers which can be compared with hue. Methods: Data was analysed from 141 patients from a previous randomised controlled trial. All grade 2 pressure ulcers and above were photographed and a computerised image analysis made to measure the hue, saturation and intensity of ulcers and the skin tone of the subject's adjacent healthy skin. This data was analysed against two classifications of pressure ulcers — the National Pressure Ulcer Advisory Panel and Stirling plus digits — and a consensus panel of experts. Results: It was shown that the mean and standard deviations of the recorded HSI values could discriminate between normal skin, grade 2a and grade 2b pressure ulcers. Conclusions: The computer could identify early pressure ulcer formation with reasonable accuracy. A second comparison with the consensus panel demonstrated a positive correlation between the computerised model and the panel's grades of pressure ulcers. Conflict of interest: None.

KEY WORDS

Classification of pressure ulcers
Image analysis
Skin assessment
Screening of pressure ulcers
Hue, saturation and intensity

Classification of pressure ulcers is very complex. In particular the first stage of a pressure ulcer is very difficult to define. The four stages of hyperaemia make it confusing for the practitioner to distinguish the aetiology of pressure ulcers (Collier, 1999). There appears to be consensus on grade 3+ pressure ulcers because the aetiology of these ulcers is clearly defined (European Pressure Advisory Panel, 1998). However, accurate identification of all stages needs to be understood fully in order to enable pressure ulcers to be prevented or treated early.

Early detection of superficial pressure ulcers could help to prevent development into deep ulceration. Evidence is accumulating that hyperaemia and non-blanching hyperaemia are identifiable risk factors in the development of a pressure ulcer and that the subsequent nursing care provided will hopefully be successful at reversing skin damage (Nixon and McGough, 2001).

The problems associated with any classification of pressure ulcers is the sensitivity, specificity and the predictive validity of the tool that is employed for the screening assessment.

General principles of screening

Screening can be defined as the presumptive identification of an unrecognised disease or defect by the application of tests, examinations or other procedures which can be applied rapidly. Thus, screening tests sort out those apparently well people who may have a condition from those who probably do not. Screening tests are not diagnostic and if the screening test gives a positive result, a diagnostic test would be required (Whitby, 1974; Galen and Peters, 1986; Mant and Fowler, 1990).

Screening individuals can be divided into four groups:

True positives (TP): those individuals who have a positive screening test result and who have the condition screened for

True negatives (TN): those individuals who have a negative screening test result and who do not have the condition screened for

False positives (FP): those individuals who have a positive screening test result but do not have the
condition screened for

False negative (FN); those individuals who have a negative screening test result but do have the condition screened for.

These groupings are useful because they allow the calculation of indices which can be used to assess the effectiveness of a screening tool or programme. The derived indices are sensitivity, specificity, predictive value and efficiency.

Sensitivity (equivalent to detection rate)

Sensitivity is defined as the ability of a test to give a positive result in people who have the condition being screened for (i.e. positivity of disease). The sensitivity is the proportion of patients that are correctly identified by the test (Altman et al, 2000).

Sensitivity = \[ \frac{TP}{TP+FN} \]

Specificity (equivalent to false positive rate)

Specificity is defined as the ability of the test to give a negative result in people who do not have the condition being screened for (i.e. negativity in health). The specificity is the proportion of true negatives that are correctly identified by the test (Altman et al, 2000).

Specificity = \[ \frac{TN}{TN+FP} \]

Predictive value

Predictive value gives an idea of how effective a test is in the clinical environment. The predictive value of a positive result is the likelihood that a positive result means that the condition is present and similarly, the predictive value of a negative result gives the likelihood that a negative result indicates the condition is absent. The predictive value is probably the most important test because it varies according to the frequency of the screened condition (Whitby, 1974; Galen and Peters, 1986; Mant and Fowler, 1990). A positive predictive value is the proportion of patients with a positive test result who are correctly diagnosed. A negative predictive value is the proportion of patients with a negative result who are correctly diagnosed (Altman et al, 2000). Predictive value of a positive result = \[ \frac{TP}{TP+FP} \]

Predictive value of a negative result = \[ \frac{TN}{TN+FN} \]

Efficiency

Efficiency relates to the overall performance of screening test.

Efficiency = \[ \frac{TP+TN}{TP+TN+FP+FN} \]

A sensitive test is one which rarely misses disease and a specific test is one which rarely causes false alarms. In general practice the most important feature of a screening test is its specificity because the incidence of any particular disease is generally low. In a hospital setting sensitivity is more important because the population under scrutiny is already selected as being ‘ill’ and the prevalence of disease is thus greater (Mathers and Hodgkinson, 1989). The predictive value of the parameters is affected significantly by the prevalence of the condition being screened for.

Using hue, saturation and intensity to classify pressure ulcers

Very few studies exist on the use of hue, saturation and intensity (HSI) for measurement and classification of pressure ulcers (Russell, 1999). New methods of defining an early stage of pressure ulcer need to be discovered and could have great economic benefit to the NHS and patient care and could result in a decrease in the pain and discomfort caused by pressure ulcers. HSI analysis perceives colour using a similar mechanism to the human eye. Colour is specified using measurements of three primary colours, just as the eye does using nervous impulses from colour-specific rods and cones. Hue measures the wavelength of the main colour; saturation is the amount of white light included within the colour and intensity is a measure of brightness. The HSI components are independent, so any change in brightness or contrast of the original image results in a change of a single HSI component.

Methods

The aims of this study were:

To determine whether different grades of pressure ulcers exhibit specific hues

To develop a consensus of opinion from four experts on the grade of pressure ulcers which can be compared against hue

To determine if 12 images of pressure ulcers from a questionnaire gave measurable hues that enabled identification of the grade of pressure ulcers.
Data was analysed from a randomised controlled clinical trial of Pegasus Cairwave and proactive seating cushion and the Nimbus 3 and Aura seating cushion for the treatment of pressure ulcers (Russell et al, 2000). The trial collected data on 141 patients over a 19-month period. All sacral pressure ulcers that were grade 2 and above on the Torrance classification (1983) were photographed digitally. To ensure consistent treatment of images, a standard ‘macro’ was used. In all images a line corresponding to 4cm was drawn using the scale on the image. Using the mouse the small areas of the ulcer were outlined, then the total ulcer area was outlined. The macro assessed and stored the image data from the outlined areas. Finally normal skin tone was measured using a selected 1cm² area of normal skin in the picture.

Pilot study
The pilot study consisted of 10 grade 2a and 10 Grade 2b pressure ulcers graded using a modified Torrance system (Table 1) taken from the main RCT (Russell and Reynolds, 2000). These were examined to determine whether there were any common factors relating to the HSI and classification of ulcer. Next, using the data from the primary randomised study (Russell et al, 2000), a random selection of patients were chosen. Initially a Microsoft Excel package was used to calculate means and standard deviation of the skin tone plotted against HSI for grade 2a and 2b pressure ulcers (Table 2). Data was then transferred into SPSS for Windows and a box and whisker plot drawn (Figures 1–3). This showed that the hue graduated down, while the saturation and intensity graduated up.

The preliminary findings of this small pilot study (n=30) appeared to suggest that early detection of pressure ulcers was possible and this was deemed worthy of further investigation by the authors’ supervisors at Wolverhampton University and De Montfort University, Leicester.
Follow-on study
To investigate the findings from the pilot further, all the data from the earlier randomised comparison were analysed to see if a larger data set would give a more conclusive answer to whether HSI could detect early pressure ulcer development. It was realised that more investigation was required to check that the images captured were standardised.

Ethical approval was given by South Staffordshire Local Research Ethics Committee (LREC) for the collection of data that was used in this study. Further approval was not required as data has already been collected from previous studies and would not involve further patient interventions.

Consensus panel
A consensus panel of four experts discussed the HSI box plots and the images used to derive them when compared with the European Pressure Ulcer Advisory Panel classification (EPUAP, 1998). The panel agreed that HSI analysis lined up well with the classification system. The panel were then asked to classify digital images of the 33 patients whose HSI had been measured and compared them against the Stirling plus digits (Reid and Morison, 1994) and the European Pressure Ulcer Advisory Panel grading systems. Twelve photographs were then randomly selected for a nurse reliability postal questionnaire (n=200) on the classification of pressure ulcers using Stirling plus digits classification and the European Pressure Ulcer Advisory Panel system (Russell and Reynolds, 2001).

Precision and accuracy of the camera and lighting
In order to check the precision, consistency and accuracy of the camera and lighting, 10 digital images were taken from four volunteers’ arms and legs. The digital images were taken behind curtains on the same ward that was used for the trial, in a middle bed space next to the window at the same time of day that the photographers had taken photographs of the patients in the pilot study. These images were then analysed using the special macro using the rectangle box normally used for skin tone as the wound outline sequences. The data was then processed in the same method as the pilot study and transferred to Microsoft Excel.

The sample population consisted of 226 images taken from four healthy volunteers, 495 images of healthy skin from patients over the age of 75 years, 510 with grade 2a ulcers, 492 images of grade 2b ulcers and 53 images of grade 3 ulcers. All images were analysed using image analysis data reported in the study.
as an 8-bit number meaning that it has a range of 00000000 to 11111111 in binary which is equivalent to 0–255. These values represent the colour of the skin on a colour wheel as three separate variables: hue, saturation and intensity. The actual value appertaining to any individual hue is simply the number assigned by the camera chip to a colour of that hue (or saturation, or intensity). The specification of an 8-bit number is important because it indicates that modulus arithmetic applies. In modulus arithmetic numbers form a constrained range and when they reach the top of that range they immediately return to zero. Thus, for example, the hue could change from a high value (e.g. 245) to a low value (e.g. 5) as a result of a combination of small negative changes in value (-15 units) and a large positive change (+240 units). Therefore traditional mean and standard deviation statistics did not function here. To calculate mean and ±sd for modular data it was first necessary to examine the data to determine whether the mean value was on the 0–255 boundary or in the centre of the distribution. If the distribution crossed the 255–0 boundary, the data was recentred to 128 by modulus addition, and the means and ±sd recalculated and the resultant mean was again recentred by modulus subtraction of 128 (Figure 4).

To demonstrate the changes in skin HSI measurements of the skin with progression of pressure ulcer grade, ellipses showing the distribution of hue and saturation for normal skin, grade 2a and grade 2b pressure ulcers were drawn using MicroSoft Excel.

Histograms were plotted to demonstrate the distribution of the continuous variables of HSI of normal skin of young healthy people, and patients with normal skin, grade 2a, 2b and grade 3 ulcers using the figures in Table 3.

A further stage of analysis used the formula for an ellipse using...
MicroSoft Excel to draw the hue and saturation on the normal skin, grade 2a and grade 2b pressure ulcers to demonstrate 1 and 2 standard deviations and demonstrate the variability of each group. The differences between 2a and 2b ulcers were tested by comparing standard errors of the mean. Finally the data from the questionnaire photographs using the consensus panel answers of the European Pressure Ulcer Advisory Panel was plotted against the computerised analysis.

**Results**

The normal skin of young healthy people (not patients) showed that the hue is distributed in the centre of the graph while the saturation and intensity are at the lower end (Figure 5). Patients with normal skin show the hue has moved to the end of the bar graph but the saturation and intensity remain at the lower end of the graph (Figure 6). The figures show quite a different picture with 2a ulcers the hue had a nearly normal distribution curve. The saturation was still at the top end of the graph, but the intensity was spread evenly across the graph (Figure 7). The graph for the 2b ulcers demonstrated another distribution of hues in the middle with saturation next to the hues and the intensity at the extreme of the graph (Figure 8). The graph for grade 3 ulcers demonstrated similarities between the 2b ulcer hues, and saturation were in similar places with higher frequencies and not so normal distribution curves (Figure 9). The intensity was in the same place with smaller frequencies. Figure 9 clearly shows a distribution which lies on the 0–255 boundary indicating the need for modulus correction.

### Modular statistics

Table 3 shows the mean and standard deviation (±sd) for hue, saturation and intensity for each category. The mean ±sd was also calculated using modular statistics; this demonstrated a more definitive split between the normal skin of the healthy volunteers and that of patients. The modular mean for the 2a ulcers were at 248, ±sd 75, whereas the 2b ulcers were 233 ±sd 39 and the grade 3 ulcers 237 ±sd 46 which is very close.

### Table 3

<table>
<thead>
<tr>
<th>Normal skin (healthy subjects)</th>
<th>Modular calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hue</td>
<td>Saturation</td>
</tr>
<tr>
<td>max</td>
<td>201.3181</td>
</tr>
<tr>
<td>min</td>
<td>116.2239</td>
</tr>
<tr>
<td>mean</td>
<td>152.1538</td>
</tr>
<tr>
<td>sd</td>
<td>21.85559</td>
</tr>
<tr>
<td>N</td>
<td>266</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Normal skin (patients)</th>
<th>Modular calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>max</td>
<td>227.1515</td>
</tr>
<tr>
<td>min</td>
<td>92.06829</td>
</tr>
<tr>
<td>mean</td>
<td>180.6765</td>
</tr>
<tr>
<td>sd</td>
<td>20.78225</td>
</tr>
<tr>
<td>N</td>
<td>495</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grade 2a pressure ulcers</th>
<th>Modular calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>max</td>
<td>204.2815</td>
</tr>
<tr>
<td>min</td>
<td>61.00513</td>
</tr>
<tr>
<td>mean</td>
<td>128.9006</td>
</tr>
<tr>
<td>sd</td>
<td>19.77104</td>
</tr>
<tr>
<td>N</td>
<td>510</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grade 2b pressure ulcers</th>
<th>Modular calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>max</td>
<td>216.9531</td>
</tr>
<tr>
<td>min</td>
<td>54.04493</td>
</tr>
<tr>
<td>mean</td>
<td>115.2024</td>
</tr>
<tr>
<td>sd</td>
<td>25.61451</td>
</tr>
<tr>
<td>N</td>
<td>492</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grade 3 pressure ulcers</th>
<th>Modular calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>max</td>
<td>181</td>
</tr>
<tr>
<td>min</td>
<td>76.02</td>
</tr>
<tr>
<td>mean</td>
<td>113.6735</td>
</tr>
<tr>
<td>sd</td>
<td>18.37002</td>
</tr>
<tr>
<td>N</td>
<td>53</td>
</tr>
</tbody>
</table>

### Table 4

<table>
<thead>
<tr>
<th>Specificity and sensitivity of HSI screening</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test positive (i.e. in 2b ring)</td>
</tr>
<tr>
<td>Test negative (in normal skin ring)</td>
</tr>
</tbody>
</table>

**Ellipse**

Figure 10 shows that the hue and saturation distributions for grade 2a and 2b ulcers overlapped significantly. Statistical tests showed no significant
Figure 10. Hue and saturation distributions for grade 2a and grade 2b pressure ulcers.

Figure 11. Hue and saturation distributions for grade 2a and grade 2b pressure ulcers plotted against actual measurements of pressure ulcers graded by the consensus panel.

difference but for clarity both ellipses are shown in Figure 10. This demonstrates that the majority of the data plots in the expected ellipses: 10 of the 12 normal skin tone results are in the normal ellipse; all of the pressure ulcers are within the red 2b ellipse, but the purple 2a ulcer ellipse also encloses one normal skin tone. The effectiveness of the HSI method was evaluated using standard parametric statistical tests for sensitivity and specificity (Table 4) which shows that the HSI test scored 100% in sensitivity and 83.3% in specificity.

Further work with more data is indicated to determine the optimum screening cut-offs. This demonstrates that the standard normal, grade 2a and 2b have definite parameters. Thus the computerised analysis of HSI parameters could be used to identify early pressure ulcer formation. A second ellipse using the information from the consensus panel demonstrated a positive correlation between the computerised model and the panel’s pressure ulcer grading (Figure 11).

Discussion

Strengths of the study

The original population from which ellipses were derived included a large sample of 141 patients with sacral pressure ulcers, which has demonstrated that there are distinct differences to be seen between normal skin tone and grade 2a ulcers and further differences with grade 2b ulcers when the skin has been broken. There is not such a distinct difference between normal skin and grade 3 ulcers. This means that the model does detect when there is a break in the epidermis. Lyder (1991) outlines a grade 1 pressure ulcer as a skin area that ranges from pale pink to bright red in colour, that is non-blanchable, warm to touch, that has erythema that does not resolve within two hours, skin oedema and an intact dermis. It has been demonstrated that hue, saturation and intensity can be used to identify early grades of pressure ulcers and breaks in the epidermis.

Skin assessment plays a key role in determining a patient’s risk of developing a pressure ulcer. Observation of erythema should direct the nurse to consider if skin changes are normal physiological responses or abnormal and the clinical significance of the assessment (Nixon and McGough, 2001). The subsequent interventions and care provided will hopefully reverse the skin damage if undertaken promptly. There is a lack of clarity on whether redness should be persistent or transient. Furthermore, transient may be described as anything from 30 minutes to 48 hours (Versluysen, 1986; Yarkony et al, 1990). Nixon and McGough (2001) state that pathological and clinical symptoms of trauma, non-blanching erythema with swelling and induration can resolve without any superficial skin loss (Brooks and Ducan, 1940; Lowthian, 1994). This still requires further research.
In this study the consensus panel provided the model ‘definitively graded’ pressure ulcer images that were used as a gold standard against which to compare the ellipse data derived from another study. Further development of the HSI ellipse concept would allow production of an instrument that would be the equivalent of having the expert at the bedside to classify the skin. This would be very useful not just in hospitals but also in primary care.

The argument surrounding the classification of pressure ulcers for research and auditing is plagued with dilemmas in achieving good inter-rater reliability and validity of definition (Lyder, 1991; Healey, 1995). The results of this study could put an end to the issue of adoption of a national grading system as the statistical results of using HSI demonstrate good reliability. A common classification would enable improved communication and standardised documentation.

Weaknesses of the study
A limitation of the study was that the sample size for the grade 3 ulcers was insufficient. The sample size was dictated by the number of patients recruited to the randomised controlled trial and therefore further work needs to be undertaken with a greater sample size. The complexity of analysing data meant it was very time consuming and this may prohibit use in day-to-day nursing. This model has only analysed sacral pressure ulcers and not heels as they are a particularly difficult surface to analyse because their convex shape means that an image taken by a camera does not give equal weighting to all areas and because fluctuations in light intensity may be more pronounced.

Conclusion
There is very little research to investigate the reliability and validity of classification systems. A computerised system removes the element of human error; classifications of pressure ulcers are increasingly being used to predict patient outcomes of care and therefore accuracy of classification will assist in this and is clearly essential. Conversely, inaccuracy of grading could render league table comparisons unfeasible. Incidence and prevalence data also uses classification of pressure ulcers and this can lead to more consistent collection of information. Ultimately this could result in a higher standard of patient care and prevent pressure ulcers developing by instigating prompt intervention.

Key Points
- New methods of defining an early grade of pressure ulcer could have great economic benefit to both patient and health service.
- This study has found that different grades of pressure ulcers exhibit specific hues.
- Computerised analysis of hue, saturation and intensity measurements can identify early pressure ulcer formation with reasonable accuracy.
- Comparison with a consensus panel demonstrated a positive correlation between the computerised model and the panel’s grades of pressure ulcers.

References