EXPERTS’ ASSESSMENT OF COLOR IN BURN-WOUND PHOTOGRAPHS AS A PREDICTOR OF SKIN GRAFT

A dissertation submitted to the Kent State University College of Nursing in partial fulfillment of the requirements for the degree of Doctor of Philosophy

by

Rose Ann Urdiales Baker

June 2011

© 2011
A dissertation written by

Rose Ann Urdiales Baker

BSN, University of Akron, 1972

MSN, Kent State University, 1981

PhD, Kent State University, 2011

Approved by

Ratchaneewaa Ross, PhD, RN
Co-Chair, Doctoral Dissertation Committee

Diana Biardi, PhD, MSN, RN
Co-Chair, Doctoral Dissertation Committee

Kathleen Tusiae, PhD, MSN, RNCS
The University of Akron

Bilal Bomani, PhD, MBA
Research and Technology Directorate
NASA Glenn Research Center

Accepted by

Karen W. Budd, PhD, RN
Director, Joint, PhD in Nursing Program

Laura Dzurec, PhD, RN
Dean, College of Nursing
Acknowledgements

I would like to acknowledge my spiritual guides, the beloved scholars, St. Thomas the Apostle, St. Thomas á Becket, St. Thomas Aquinas, St. Thomas á Kempis, St. Thomas More; as well as St. Jude the Apostle, Patron of Hopeless Cases, and St. Rita of Casio, Patroness of the Impossible.

I am deeply appreciative of the support I have received from my dissertation committee, Co-Chairs Ratchneewan Ross, PhD and Diana Biordi, PhD, Kathleen Tusaie, PhD, and Bilal Mark McDowell Bomani, PhD, MBA. Their scholarly contribution has helped hone the quality of my dissertation, but equal to their scholarly contribution is the patience and comfort they provided as I developed my research. Ann Jacobson, PhD contributed her expertise as a dissertation committee member early in the development of my research. I greatly appreciated her reviews of my work. I particularly wish to remember Richard A. Zeller, PhD, who served at the beginning of my dissertation. He worked closely with me to design my statistical analysis. His untimely death deprives us all of his genius, but his inspiration and nurturance in the conduct of research will always be with us.

In the effort to render excellent prose, Harold Fry, PhD, has been a knowledgeable and patient reviewer. He very ably has applied his expertise in teaching the English language to the review of my manuscripts. I am very grateful to him.

I especially thank my nurse colleagues, Mary A. Mondoletti, MSN, APRN, BC; Maryke Bard, BSN, RN, Rosalee Unwin, ADN, RN, Lori Treibner, APN and Cindy Bennett, APN. I particularly thank Robert L. Klein, MD, who encouraged me so many times in my research career. They participated in this study by identifying appropriate participants for inclusion in the research and reviewed the photographs to extract the necessary data; without this most fundamental assistance, this study could not have been conducted. Their contribution cannot be overstated.

Most importantly, I acknowledge my husband, Bob, who has sustained me in so many ways. He is patient, understanding and encouraging. He has made many sacrifices to fulfill my dream of completing the dissertation. Without him I could not have finished this work. While words cannot fully convey my gratitude for your presence and your continuing support, at their simplest and yet most profound, I say, “Thank you, Bob.”

My families, my father Anthony in person and mother Anita in spirit, have provided me a basis in life that goes beyond dissertation support, which I gratefully acknowledge. They both lived the lesson of the importance of learning and scholastic achievement in their support, encouragement, and pride in my achievements and sacrifices. I especially acknowledge the inspiration of my maternal grandmother, Rosita. She survived a political oppression that suppressed the education of women. She came to the United States of America, determined to see that her daughters would be taught to read and write. As her daughters have been taught, I have been taught. I am immensely grateful for the privilege of my education.
# TABLE OF CONTENTS

Chapter One .................................................................................................................................. 1
  Burn-Wound Depth, Nomenclature, Definitions, Appearance and Treatment .................................. 3
  Burn-wound depth, nomenclature and definitions ........................................................................... 3
  Appearance .................................................................................................................................. 4
  Treatment .................................................................................................................................... 4
  Burn-wound Assessment: Color and Interrater Reliability .............................................................. 5
  Color as criterion in burn-wound assessment ................................................................................ 5
  Informal clinical accuracy of color assessments ............................................................................ 5
  Formal research: Accuracy of color assessments ........................................................................... 5
  Interrater reliability ....................................................................................................................... 7
  Use of photographs in treatment and research .............................................................................. 7
  Methods in interrater reliability in clinical research ..................................................................... 8
  Implications of photographs in interrater reliability studies in clinical research ........................... 9
  Theory ....................................................................................................................................... 9
  Symptom Management Theory .................................................................................................... 9
  Benner’s Use of the Dreyfus Model of Skill Acquisition Applied to Nursing .............................. 13
  Diagnosis by experts .................................................................................................................... 13
  Burn-wound experts ..................................................................................................................... 13
  Research Questions ....................................................................................................................... 14
  Summary ..................................................................................................................................... 15

Chapter Two Technology in Burn Wounds ...................................................................................... 17
  History of Burn-wound Assessment Using Prevailing Technology .............................................. 17
  Jackson’s Burn-wound Theory ....................................................................................................... 17
  Gold Standard Assessment ........................................................................................................... 19
  Technologies Today in Burn-wound Assessment and Their Effectiveness .................................... 20
  Assessment by Physiological Change ......................................................................................... 21
  Assessment by Color .................................................................................................................... 24
  Potential Machine Vision Assessment Using CMIS ...................................................................... 30
  Summary ..................................................................................................................................... 31

Chapter Three Method .................................................................................................................. 32
  Design and Objectives .................................................................................................................. 32
  Sample ....................................................................................................................................... 32
  Instrument .................................................................................................................................. 33
  Reviewers .................................................................................................................................... 34
  Procedure of the Review ............................................................................................................. 34
  Reliability ................................................................................................................................... 34
  Types of Reliability Measurement ............................................................................................... 35
  Validity ......................................................................................................................................... 37
  Data Analysis ............................................................................................................................... 37
  To examine predictors of skin graft ............................................................................................. 38
  Correlation analysis ..................................................................................................................... 39
  Human rights considerations ......................................................................................................... 39

Chapter Four Results ....................................................................................................................... 41
  Demographics ............................................................................................................................... 41
  Data Analysis for Burn-Wound Experts ...................................................................................... 43
  Burn-Wound Experts Assessment and Reliability ......................................................................... 43
  Correlation Analysis .................................................................................................................... 43
  Logistic Regression Analysis ....................................................................................................... 44
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interclass Correlation Analysis</td>
<td>45</td>
</tr>
<tr>
<td>Chapter Five  Discussion and Implications of the Study</td>
<td>47</td>
</tr>
<tr>
<td>Research Questions</td>
<td>47</td>
</tr>
<tr>
<td>Relationships Among TBSA, LOS, and Skin Graft</td>
<td>48</td>
</tr>
<tr>
<td>Predictors of Skin Graft</td>
<td>49</td>
</tr>
<tr>
<td>Interrater Reliability Using HIM</td>
<td>49</td>
</tr>
<tr>
<td>Lack of Quantification for Color in Burn-wound Assessment</td>
<td>50</td>
</tr>
<tr>
<td>Symptom Management Theory</td>
<td>52</td>
</tr>
<tr>
<td>Benner’s Use of the Dreyfus Model of Skill Acquisition Applied to Nursing</td>
<td>53</td>
</tr>
<tr>
<td>Nurse Expertise in Burn-wound Assessment</td>
<td>53</td>
</tr>
<tr>
<td>Function of Photograph Analysis</td>
<td>56</td>
</tr>
<tr>
<td>Summary</td>
<td>57</td>
</tr>
<tr>
<td>Glossary</td>
<td>59</td>
</tr>
<tr>
<td>References</td>
<td>61</td>
</tr>
<tr>
<td>Appendices</td>
<td>67</td>
</tr>
<tr>
<td>Appendix A  Procedure for the Decision Flow to Apply or Not Apply a Skin Graft</td>
<td>68</td>
</tr>
<tr>
<td>Appendix B  Lund and Browder Charts for Estimation of Burn-Wound Size</td>
<td>69</td>
</tr>
<tr>
<td>Appendix C  Human Intelligence Matrix</td>
<td>71</td>
</tr>
<tr>
<td>Appendix D  Consent for Photographs</td>
<td>72</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1 Procedure for the Decision Flow to Apply or Not Apply a Skin Graft ....................... 68

LIST OF TABLES

Table 1 Frequencies and percentages of Skin Graft Status .......................................................... 42
Table 2 Frequencies of photographs ............................................................................................ 42
Table 3 Correlation Analysis for TBSA, LOS and Skin Graft ......................................................... 43
Table 4 Graft, Zone Percent, and Color Correlation Analysis ....................................................... 44
Table 5 Graft Zone Percent and Color Logistic Regression .......................................................... 45
Table 6 Zone Percent Interclass Correlation ................................................................................ 46
Table 7 Color Interclass Correlation ............................................................................................ 46
This dissertation examined 1) if burn-wound color percent (Red and White), in different burn-wound locations (Midrange and Central), and extent of burn zones (Zone Percent) were predictors of skin graft application; and 2) the interrater reliability of the three burn-wound experts conducting the burn wound assessment using burn-wound photographs. Descriptive variables included cause of burn wound, Total Body Surface Area burned (TBSA), and Length of Stay (LOS).

The sample for this exploratory correlational analysis was 50 color photographs of indeterminate burn-wounds and three burn-wound experts who examined burn-wound color and depth using the Human Intelligence Matrix (HIM), developed by the researcher. Using Pearson’s correlation analysis, weak but significant associations were found between TBSA, LOS and skin graft (r = -.084, p <.05), LOS and skin graft (r=.195, p<.01), and TBSA and LOS (r=.210, p<.01). The fact that skin graft was correlated with TBSA and LOS suggests that TBSA is an indicator of skin graft and skin graft leads to a longer hospital stay. Red and White, not Yellow, Green, Brown or Black were found to be significant colors for interrater reliability in selected locations.
(Midrange and Central) of burn-wounds at selected time points (Baseline, Midtrajectory and Final).

To predict skin graft at each time point, logistic regression was used. Results showed that none of the independent variables (Zone Percent, Red Percent and White Percent) were predictors of skin graft. Burn-wound experts agreed on the assessment of Central Zone and the Red Percent, which suggests that the instrument was helpful in quantifying color (Red) for the burn-wound area (Zone Percent). This dissertation was 1) the first to use the HIM, an instrument in assessing color, 2) a first step toward developing a reliable matrix of colors for future use in an accurate machine vision system, and 3) the first test of interrater reliability involving nurses in burn wound studies. This dissertation illustrates that more reliable burn assessments are critical, particularly in the treatment and comfort of patients with indeterminate burn-wounds.
CHAPTER ONE

Burn wounds in humans can be devastating in their appearance and consequences. The correct diagnosis of the depth of a burn-wound is critical to the choice of care, treatment, and optimal outcomes for the patient (Clarke, 1992). Burn-wound treatment is a choice between surgical grafting of skin to the burn-wound (skin grafting) or spontaneous healing (Baker & Mondozi, 2011), and is highly dependent on the correct diagnosis of the depth and prognosis of a burn wound. Burn wounds can be of any depth, but of interest to this dissertation are those called “indeterminate” where the depth is not easily determined, and thus, the treatment is uncertain. Because the treatment is uncertain, patients and health care professionals often must wait until more sequelae present themselves sufficient to resolve a treatment decision of either skin grafting or spontaneous healing. If the waiting period can be shortened, it is to the benefit and comfort of the patient, which are the ultimate goals of health care professionals.

Children constitute a significant portion of patients with burn wounds. Safe Kids Worldwide (2007) identified the frequency of burn wounds in children as the sixth leading cause of children’s death in 2005. In adults, current war-related burn-wounds are a leading cause of burn-wounds and issues in burn care, as was the case for earlier wars. For other adults, most burns are related to work hazards, such as electrical contact or steam. Elderly adults are more vulnerable to burn wounds. Adults over 64 years of age, who represent 12% of the U.S. population, suffered more than 30% of fire deaths (U.S. Department of Homeland Security, 2008). Children, the war injured, and the elderly continue to be at high risk for burn wounds, and the need is great for more efficient and comfort-related treatment of burns.
Assessments of burn-wound depth by visual assessments of color occur daily in the clinical arena as part of the standard clinical practice of burn-wound experts. However, human visual assessment is open to mistakes, particularly related to color interpretation of burn wounds. A reliable machine vision system which has the advantage of avoiding human error and providing consistent, objective data has already been used in the early recognition of various health problems such as cancer, Darrier Ferrand disease, decubitus, venous ulcers, inguinal verneuil disease, and tuberculosis, (Bon, et al., 2000; Ganster, et al., 2001; Jovanovic & Kihiczak, 2010 McDowell, Gray & Rogers, 2006; Theran & Macq, 1996; Patel & Lopez-Terrada, 2008). Because of the nature and location of burns, a reliable machine system of visualizing burns has not yet been devised (for examples, see Chapter 2). Therefore, the researcher developed the HIM, a paper and pencil instrument based on the Jackson model (Jackson, 1953), including two important aspects: zones of injury and hues, or color types of burn wound. Zones of injury are comprised of: 1) zone of necrosis or coagulation (Central); 2) zone of stasis (Midrange); and 3) zone of hyperemia (Peripheral). Colors of the burn wound include: 1) Red; 2) Yellow; 3) White; 4) Green; 5) Brown; and 6) Black. The HIM allows the burn-wound expert to assess the percentage of each zone and the percentage of each color within a zone. It was the first tool to quantify burn wound color which may help burn wound experts to be more accurate with their assessment of burn-wound depth than without the tool.

Therefore, this dissertation examined two research questions. The first question asked whether burn-wound color (the variable hereafter called Colors) and the depth of the burn wound (the variable hereafter called Zone Percent) are predictors of skin graft application (skin graft vs. Non skin graft--assuming spontaneous healing).

The second research question of this study examined the interrater reliability of the three burn-wound experts conducting the burn wound assessments, using the HIM. This was one of the few studies of interrater reliability of persons using clinical observations, and the first research
study to include burn-nurses as burn-wound expert reviewers. Previous research studies on the assessment of burn-wound depth have relied only on burn surgeons.

**Burn Wound Depth, Nomenclature, Definitions, Appearance, and Treatment**

**Burn wound depth, nomenclature and definitions.** Burn wounds in this study were assessed in three areas: depth (Zone Percent), Total Body Surface Area (TBSA), and Colors. Initial assessments of an ongoing evaluation of the depth of the burn wound have been performed by experienced burn surgeons (Chatterjee, 2006) and are crucial because depth dictates the treatment of choice, and improper treatment can worsen the burn wound by causing additional tissue destruction over time as damaged blood vessels thrombose, ischemia is increased (Hartford, 2002).

Burn wounds are placed in three categories by their depth, 1) the deepest are full-thickness, 2) less deep are partial-thickness, and 3) and indeterminate is indicated when the clinical evaluation cannot definitively ascertain depth (Baker & Mondozi, 2011). Depth is a subjective judgment and has not, to date, been correlated with quantitative measurements.

Based on Jackson (1953), burn-wound depth is classified into three zones of injury: 1) zone of necrosis or coagulation, 2) zone of stasis, and 3) zone of hyperemia. The Jackson criteria of burn-wound depth have been the standard of the field for more than 50 years and are currently used in practice (Jackson, 1953). The zone of necrosis or coagulation defines irreversible injury and is ordinarily most central to the burn wound. The zone of stasis defines the area at risk for becoming necrotic because of ischemic changes in the burn wound. The zone of hyperemia is the area of capillary dilatation and inflammation and is ordinarily most peripheral to the burn wound (Jackson, 1953).

Research has demonstrated that burn wounds are very dynamic in progression of depth in the first 48 hours (Jaskille, Shupp, Jordan, & Jeng, 2009, 2010). The viability of the remaining dermis and its associated vascularity is key to burn-wound diagnostics, color, healing, and associated technologies (Acha, Serrano, Acha, & Roa, 2005).
Appearance. A full-thickness burn wound involves the entire epidermis and dermis, and may extend into subcutaneous tissue, muscle and bone. Full-thickness burns may appear as dry and leathery with the colors of red, white, yellow, green, brown, or black. Partial-thickness burn wounds involve the epidermis and a part of the dermis. These wounds are painful, and may appear, moist, red or white, as well as blistered or edematous. In good quality of care, daily visual assessments of the wound determine whether sufficient dermis is available to support wound healing. If non-viable layers of tissue are present, they are excised. The indeterminate burn-wound has elements of both full-thickness and partial-thickness burn-wound, but are insufficient to make a diagnosis.

Treatment. Burn-wound depth is particularly important to the nature and sequencing of timely burn treatment. If the burn wound is full-thickness, i.e., without dermis, skin grafting is the treatment of choice, and early treatment is even better (Devgan, Bhat, Aylward, & Spence, 2006) (See Figure 1, Appendix A: Procedure for the Decision Flow to Apply or Not to Apply a Skin Graft). Skin grafting requires, in turn, surgical excision of necrotic tissue, as well as topical bacteriostatic ointments and sterile wound covering after the skin graft is applied. If sufficient dermis remains intact in a partial-thickness burn wound, eventually, with the wound covered by sterile dressings and ointment, the burn wound will close with sufficient dermis in what is called spontaneous healing (Baker & Mondonzzi, 2011; Bishop, 2004).

Indeterminate burn-wound depth presents the most difficulty because neither treatment of choice is immediately apparent. In the indeterminate burn-wound, the presence or absence of viable dermis cannot be readily identified. Because of the possibility of burn-wound extension, i.e., deepening and enlarging through reactive physiological processes, the indeterminate burn-wound takes more time to decide treatment. Obviously, the longer the treatment is delayed, the more uncomfortable the patient, and the more worrisome the possible patient outcomes. Thus, the sooner an indeterminate burn-wound can be identified and reclassified into another category, the sooner the best treatment can be instituted.
Burn-Wound Assessment: Color, and Interrater Reliability

Color as a criterion in burn-wounds assessment. Burn-nurses and surgeons use color as the most frequent criterion by which to judge the depth of the burn, and color is therefore, critical to treatment and patient outcomes (Acha et al., 2005; Serrano, Acha, & Acha, 2003; Serrano, Acha, Gomez-Cia, Acha & Roa, 2004). Colors of burn wounds reflect the physiological course of the burn wound. Colors most frequently found in burns are red, white, yellow, green, black, or brown. Red indicates healthy tissue, white refers to ischemia (i.e. diminished vascular flow) or beginning infection, yellow and green indicate infection, while brown or black indicate necrosis (eschar) or coagulation, leading to death of tissue (Baker & Mondonzzi, 2011).

Informal clinical accuracy of color assessments. Burn experts, whether physicians or nurses, teach each other their criteria for assessing burns, and help each other achieve expertise in their clinical assessments (Gordon & Marvin, 2002; Wachtel, Berry, Wachtel, & Frank, 2000). The traditional, widespread method of evaluating wounds, such as burn wounds, is direct clinical visual examination of the injury (Hansen, Sparrow, Kokate, Leland, & Iaizzo, 1997). In burn wound assessment, burn-wound experts visually examine the wounds and teach each other the proper criteria by which to assess burn-wounds and their likely prognosis (Chatterjee, 2006; Gordon & Marvin, 2002).

While they do not assess their clinical interrater reliability formally, the nurses and doctors might be considered to have done so informally by judging the accuracy of their diagnoses and outcomes of patients and informally communicating their judgments to each other (Biordi, 1984). Literature indicates that nurses and doctors teach each other within (Benner, 1984, 2004; Biordi, 1984) and across their disciplines in the clinical arena.

Formal research: Accuracy of color assessments. In formal research, Vermeulen, Ubbink, Schreuder and Lubbers (2007) studied nurses and physicians in their evaluation of surgical wounds’ color in photographs. The subjects used a 3x3 matrix of color (red for granulation tissue, yellow for slough, and black for necrosis); and exudates (wet, moist, and dry)
to mark their evaluation of the condition of open surgical wounds. The researchers calculated the inter- and intra-agreement between categorical assessments in the two groups of color and exudates using the statistic, Kappa.

The matrix for color in Vermeulen et al.’s (2007) study was a simple indication of the presence or absence of one of the three colors, red, yellow, or black without an assignment for the percentage of color for a specific area of the surgical wound. Results indicated a high degree of agreement on color among the nurses and physicians in the Kappa statistic, where $k \geq .75$. Results for exudates ($k = .48$ for physicians; $k = .49$ for nurses) were only fair. No significant difference between nurses and physicians was detected, using the Mann-Whitney test for ordinal data (Vermeulen et al., 2007). Despite the high values of Kappa derived from Vermeulen et al.’s study, it has been cautioned since the 1990s in the area of nursing diagnosis that use of Kappa (proportion of agreement) is not appropriate when a prevalence effect and bias exist (Banerjee & Fielding, 1997).

Mekkes and Westerhof (1995) discussed a problem that remains current, that is, the lack of reliable evaluation in assessing wound healing. Human visional assessment remains hampered by accurate assessment of burn-wound color, especially the brown-black of burn-wound necrosis and the yellow of infection. Necrosis in the burn-wound indicates the need for debridement. Mekkes and Westerhof had designed a digital imaging system (DIA) to measure the shift from black/yellow necrosis to red granulation tissue objectively. Using their DIA, 15 observers analyzed pictures of decubiti and venous ulcers and the authors reported a correlation ($r=.945$) between observers and the DIA system. In spite of the high correlation value in this study, using correlation coefficient has been under criticism because it indicates relative agreement and is insensitive to mean differences as rated by observers (Banerjee & Fielding, 1997).

The accuracy of burn-wound expert assessments has rarely been studied formally. Between 2004 and 2010, three different studies of burn wound assessment of depth by physician burn-wound experts indicated accuracy between 50 and 80% of the time (Bishop, 2004; Devgan
et al., 2006; Hemington-Gorse, 2005; Jaskille et al., 2009, 2010). Only one interrater reliability study exists; it compared 24 burn surgeons, residents, physician assistants and burn nurses on their assessments of two charts indicating burn-wound size but not depth (Wachtel et al., 2000). That is, the study examined percentage of body area and locale of the burn (arm, leg, etc). Findings indicated that there were no statistical differences among physicians and physician assistants, or between nurses and physicians or physicians assistants. Raters with more experience were more accurate (Wachtel et al., 2000). However, the authors based their assessment on scatter plots of means and standard deviations without providing any other statistic while still claiming statistical significance. We can only surmise that statistical significance was found in an unreported analysis. No studies of nurses alone or their assessments of burn-wound depth were otherwise found.

**Interrater reliability.** In an effort to show the outcomes of an inter-professional group of evaluators, which also mimics the clinical world of burn-wound assessments, this dissertation examined the interrater reliability of three burn-wound experts by using color photographs of burns of patient’s wounds which were previously identified as indeterminate. More specifically, this paper examined 1) what burn-wound colors used by burn nurse and burn-surgeon experts to assessment burn wounds as indicated in a defined sample of photographs of burn wounds and 2) the interrater reliability of their choices, which would, in the clinical world, have impact on the accuracy of burn wound estimation and its outcomes in treatment.

**Use of photographs in treatment and research.** The use of photographs in research and clinical practice on color in wounds has had precedent for research studies (Nelson, Boyle, Taggart & Watson, 2006; Jones, Wilson & Andrews, 2003; Vermeulen et al., 2007). For a clinician or a researcher who is anxious to efficiently diagnose burn-wound status for healing, care giving, or research, the use of photographs of the progress of a burn wound have been standard practice in most burn centers. These photographs, because they are taken over a course of treatment and over a specific time frame, are thus a good source of aggregated data available
for specific types of studies or treatment decisions (Jones et al., 2003) and prevent further risk or discomfort for a patient than a real time data collection for every possible burn-wound study that could be concluded (Nelson et al., 2006). The patient with indeterminate burn wounds typically has a prolonged stay in a burn unit. Thus, delays such as can occur in data collection requiring direct, hands-on patient observations, can negatively impact patients with indeterminate burn-wounds.

In clinical practice for both nurses and physicians, photographs of burn wounds are found to be reliable images for teaching burn-wound depth (Baker & Mondonzzi, 2011). For instance, two studies examined the utility and accuracy of images and/or photographs in burn wounds assessments. Hansen and his colleagues (1997) found that computerized imaging could accurately and significantly forecast burn-wound depth and associated histological readings in porcine tissue burns. Jones and his colleagues found that three sizes of digital images were, at each size, reliable in the assessment of burn wounds. They found that the images reliably differentiated among burn-wound edema, erythema, and cellulitis (Jones et al., 2003). This dissertation study used 50 photographs of 22 patients taken at two or three time points along the trajectory of care while in a single burn center.

Methods in interrater reliability studies in clinical research. Interrated reliability studies have been conducted in various clinical phenomena using different methods and instruments. Studied phenomena include skin breakdown, pressure ulcers, moisture lesions, and infection in non-healing ulcers with different methodology using surveys or observations of images, photographs, and clinical procedures. (Beeckman & Schoonhoven, 2007; Kottner & Dassen, 2008; Kottner & Halfens, 2010; Lorentzen & Gottrup, 2006; Suddaby, Barnett & Facteau, 2005). Most of the studies of clinical processes and procedures, which did not use paper and pencil tests, had relatively few observers, from 2 to 7 observers (Lorentzen & Gottrup, 2006; Suddaby et al., 2005). One study (Beeckman & Schoonhoven, 2007) was extraordinary in that it had 1,452 observers examining photographs of pressure ulcers using a paper and pencil
instrument. This dissertation is similar to most studies in that it used the small and reasonable number of three burn-wound experts for the assessment of color in burn wounds.

**Implication of photographs in interrater reliability studies in clinical settings.** The method of examination of burn-wound photographs has implications for disciplines teaching within and across each other in the apprentice-like clinical refinement of skills. Thus, strong interrater reliability can direct a more robust and accurate distribution of knowledge among more experts (Benner, 1984, 2004; Biordi, 1984). This study, while not examining such teaching approaches, nonetheless has implications for future reference in teaching accurate observations. Studies of burn wound photographs have been limited thus far, in that 1) they relied on animal studies (Hansen et al., 1997), or 2) only examined area but not color in human burns (Wachtel et al., 2000), or 3) examined color in only a tiny area of human burn wounds which then were corroborated by painful biopsy (Acha et al., 2005). Due to the three limitations mentioned above, existing research results are limited in their translation to useful clinical practice. Therefore, the present study examined burn-wound experts’ assessments of the color of human-burn wounds using photographs of wounds and color in actual clinical settings.

**Theory**

This study is guided by a combination of two theories and one model. The theories include the Symptom Management Theory (The UCSF School of Nursing Symptom Management Faculty Group, 1994) and the Jackson’s physiological theory (Jackson, 1953) while the model is Benner’s use of the Dreyfus Model of Skill Acquisition Applied to Nursing (Benner, 1984, 2004). In this chapter the Symptom Management Theory and the skills acquisition model are described, whereas the Jackson’s theory is described in relation to the description of burn and the technologies attempted in the assessment of burn-wounds Chapter Two.

**Symptom Management Theory**

The symptom management theory (SMT) defines a symptom as a subjective experience reflecting changes in the biopsychosocial functioning, sensations, or cognition of an individual. A
sign is defined as any abnormality indicative of disease that is detectable by the individual or others (Harver & Mahler, 1990). In 1994, Larson and colleagues published the original SMT, incorporating both signs and symptoms into the assessment of disease to evaluate and verify effectiveness of management strategies (The UCSF School of Nursing Symptom Management Faculty Group, 1994).

The three essential concepts of the theory are symptom experience, symptom management strategies, and symptom status outcomes. Symptom experience is a simultaneous perception, evaluation, and response to change in one’s usual feeling. The experience may include more than one symptom, i.e., a symptom cluster. The patient with a burn wound may experience pain, and increased sensation of heat as symptoms, as well as the sign of tissue destruction. The SMT model illustrates the simultaneous interaction among all three concepts. These interactions may occur in moments, as with a burn wound, or may occur in an iterative process as signs and symptoms resolve and stabilize, as with the ongoing assessment of burn-wound depth (Humphreys et al., 2008).

Both symptoms and signs are important indicators that call attention to problems detected by patients and clinicians (Dodd et al, 2001). Originally, SMT focused on how a single symptom can be studied. Later, symptom clusters, i.e., two or more symptoms that are related to one another, came to be seen as important expansions of the theory. Symptom clusters may include signs, physiological abnormalities indicative of disease that is detectable by the individual or other, e.g., nausea and poor appetite are symptoms, while vomiting is a sign. Research into symptom clusters may provide new insights into underlying mechanisms of biopsychosocial functioning, sensations or cognition (Miaskowski, 2006). Relevant to this theory is the focus on evaluation, although the evaluation of symptoms and symptom clusters do not address the evaluation of signs, which this dissertation did.
Nursing scholarship embraces the scientific method of discovery to generate new and unique knowledge (Boyer, 1990; American Association of Colleges of Nursing Position Statement on Defining Scholarship for the Discipline of Nursing, 1999). As in this dissertation, nursing scholarship of discovery can take the form of empirical research by investigating burn-wound assessment and using quantitative measures to describe the status of the burn wound. The research is then related to, or develops its own, theory. In this case, investigation of burn-wound assessment relates to the component of SMT, symptom experience, and particularly, the
assessment or evaluation of symptoms and signs. Although SMT does not elaborate on the evaluation of signs, this dissertation regards the evaluation of signs as relevant to the symptom experience for the person, in that the efficient and effective diagnosis of burn-wound depth is a necessary refinement of the burn-wound assessment process. Thus, the dissertation used the scientific method to incrementally develop or build on previous research findings to further develop a theory, or, as in this case, two theories: the combination of physiological theory and STM.

Thus, the concordance between the STM and this dissertation is on evaluation, and eventually from the evaluation, the subsequent management of the patient symptoms. The SMT suggests that symptom evaluation will lead to more effective strategies to minimize or eliminate the symptoms (Humphreys et al., 2008). In this dissertation the management of patient symptoms lies first in the evaluation of the sign (burn depth) and leads to the decision to apply a skin graft. Insofar as evaluation is included in the symptom experience, it is relevant to this study, which focuses primarily on the evaluation of signs. The signs of burn wounds (i.e. burn wound depth) ultimately relate to symptoms experienced by the patient. This dissertation also suggested that the evaluation of symptoms should include, at some relevant physiological stage, the evaluation of signs. While this dissertation chose to focus on signs of burn wound depth, and acknowledges that symptoms may accompany signs, this study focused first, and only, on the reliability of experts in the assessment of a particularly difficult sign, the indeterminate burn wound. The accompaniment of signs to symptoms would be a future research endeavor that would elaborate or challenge current theory.

Symptom management strategies are efforts to avert, delay, or minimize the symptom experience. The SMT contains three ways to effect management: 1) reducing the frequency of the symptom, 2) minimizing the severity of the symptom, and 3) relieving the distress associated with the symptom (Portney et al., 1994). Patients and clinicians may attempt more than one strategy and use a combination of interventions that have a greater effect on the symptom or cluster of
signs and symptoms (Humphreys et al., 2008). As a symptom management strategy, the burn care team would identify the severity of the burn wound by assessing burn-wound depth, a specific sign, to initiate appropriate treatment, to minimize the concurrent symptoms and relieve the distress associated with the symptom.

The management of symptoms and their resulting outcomes often become the responsibility of the patient and family (Dodd et al., 2001). In the acute care setting of the burn center, the management of symptoms is closely related to the signs of the burn wound. This dissertation focuses on the sign of burn-wound depth (using Zone Percent and Color) in order to manage the symptoms of pain, debilitation, compromised movement and infection. The findings which generate new and unique knowledge in burn-wound assessment would expand the knowledge of the symptom experience of a person with a burn wound, and wound be integrated into the holistic model of SMT.

**Benner’s Use of the Dreyfus Model of Skill Acquisition Applied to Nursing**

**Diagnosis by experts.** As in other specialties, nurses and surgeons learn both from scientific sources and from their clinical colleagues. Colleagues may include experts within and across fields (Biordi, 1984). Similarly, burn nurses learn from each other (Gordon & Marvin, 2002) and from physician colleagues (Acha et al., 2005). As they become experts, (Benner, 1984, 2004) nurses develop skills and clinical judgments, while meeting their goals of effective, efficient and reliable treatment of patients. Interestingly, despite burn-wound experts working closely with each other, burn wounds do not seem to solicit the same concordance of interrater reliability as have other modalities and wounds (Vermeulen et al., 2007). What did concern this dissertation is the identification of any areas of agreement in burn-wound assessment that may begin a quantification of assessment features to improve the reliability of burn-wound assessment.

**Burn-wound experts.** For this dissertation, “burn experts” were identified as those who have 1) veto power or, the last word in the diagnosis of the burn, i.e. individuals whose views
will be the final decision in shaping treatments carried out by others, and 2) those whose hours of work on burn units are specifically dedicated to burn wound diagnosis and treatment and which hours are commensurate with formal hours of credentialing in nursing (usually set at 500-1000 hours). Because a work year is set at 2,080 hours, one full year in a burn unit exceeds most post graduate credentialing hours. In the case of this dissertation, burn-nurse practitioners who reviewed photographs had greater than ten years experience in burn-wound nursing care, with greater than five years experience as, specifically, in the more detailed work as a burn-nurse practitioners. In the burn care setting, burn-nurse practitioners assess burn wounds as part of their clinical duties, which requires approximately 25% of their work time. Both burn-nurse practitioners are certified with the American Nurses Credentialing Center, which requires 1,000 hours of clinical practice to achieve certification.

The burn surgeon, also a photograph reviewer, has greater than 30 years experience in burn-wound surgery, and has served as surgical director of the burn center. He completed a burn-care fellowship with the University of California at Davis and is a fellow of the Academy of Surgeons. His competencies and credentials would give him the status of a surgical expert, with a subspecialty in burn wounds. As with the nurses, the burn-surgeon expert who assessed photographs in this dissertation had the prerogative and authority for final diagnostic and treatment decisions. Customarily, these three burn-wound experts assess at what time a burn wound should receive a skin graft or whether the burn-wound should be permitted to heal spontaneously. The treatment decision may be delayed or completed by them at any time throughout the admission of the patient.

**Research Questions**

This study will examine a major criterion, color, for burn-wound assessment and relate the findings of color to the treatment outcome of the burn wound, that is, either skin grafting or spontaneous healing. Further, this dissertation examined the responses of expert burn-wound assessment for interrater reliability. Specifically, the *research questions are:*
a. To what extent was the assessment (Zone percent, Red percent, and White percent) of three burn-wound experts predictive of the skin graft status (skin grafting vs. spontaneous healing) earlier decided by the treating physicians?

b. To what extent do three burn-wound experts, using the HIM in the examination of selected burn wound photographs, agree on percent of burn-wound color according to the zone of burn-wound injury (Central and Midrange)?

**Summary.** Because burn wounds are consequential in their appearance and treatment, the accurate assessment of burn wounds is critical. This dissertation examined two primary research questions mentioned above. Additional descriptive and inferential analyses were conducted on demographics and associated variables, i.e. age, sex, source of burn wound, total body surface area burned (TBSA), and length of stay (LOS).

This chapter discussed burn wound depth, nomenclature, and appearance, i.e., full-thickness or partial thickness and TBSA, as well as treatment. Burn wounds are assessed for depth and TBSA to select the treatment of choice, and to evaluate for any increase in depth or TBSA due to physiological changes. Burn wound colors are described in detail and reflect the physiological course of the burn wound and remain the prevailing criteria for assessment of burn-wound depth. The accuracy of burn-wound expert assessment of color has had little research investigation. This dissertation is one of few to assess inter-rater reliability of expert practitioners in burn assessment and is the first to use burn- nurses as experts in burn-wound assessment.

To study the accuracy of color assessment of burn wounds, the innovative use of photographs in this dissertation was validated in the literature, saving the patient from the risk of additional compromise in burn-wound healing. Burn-wound photographs are reliable images for assessing burn-wound depth, and an excellent source of aggregated data. More accurate knowledge of burn-wound assessment would result from an enhanced understanding of interrater reliability. Symptom management theory and the Dreyfus model provide the theoretical framework for this dissertation.
Chapter Two discusses Jackson’s (1953) physiologically based burn wound theory and the technology of burn wounds diagnosis and treatment, particularly focusing on machine visualization of burn wounds. Method is presented in Chapter Three, Results in Chapter Four, and the study’s Discussion and Implications were presented in Chapter Five.
CHAPTER TWO

Technology in Burn Wounds

History of Burn-wound Assessment Using Prevailing Technology

**Jackson’s Burn-wound Theory.** Jackson’s landmark research described the physical signs important to the diagnosis of burn-wound depth (Jackson, 1953). The clinical motive for his research originated from the empirical discovery that excising necrotic tissue from the burn wound produced better results (Tompkins, Remensnyder, Burke, 1988). Early excision reduced the attendant risks of infection, fibrosis and contracture, i.e., the constriction of the healed burn-wound. The previous procedure had permitted necrotic tissue to spontaneously slough, i.e., fall away from the burn wound, in an effort to preserve any viable tissue that remained. This procedure invited increased risks of infection, fibrosis and contracture.

Jackson established the validity of his theory by first identifying the limitations of previous systems for the diagnosis of burn-wound depth. He reviewed historical German, Latin, and French documents to identify the origins of burn-wound assessment and cites the early 18th century work of Fabry, who developed the first recorded classification of a burn wound (not by burn-wound depth, but by wound appearance). Fabry noted the burns’ 1) redness and blistering of the skin, 2) withering of the skin without charring, and 3) eschar formulation and charring. Later, Kentish and Delpech (Jackson, 1953) used two criteria to assess the burn wound, i.e., inflammation and tissue destruction. Heister and Dupuytren used six levels of burn-wound depth to assess the burn wound. Even later categorizations of burns as first, second and third degree burns may still be found in the literature, but are not recognized any longer as the standard of the field. In fact, today’s classification of the burn wound depth, i.e., partial-thickness and
full-thickness, began in the 1940’s (Jackson, 1953), and is the prevailing classification. Several research studies included in this dissertation cite Jackson’s theory (Hemington-Gorse, 2005; Chatterjee, 2006; Devgan et al., 2006; Jaskille et al, 2009, 2010). Clinical burn-wound literature consistently uses Jackson’s theory in discussing pathogenesis, assessment and treatment (Kagan & Smith, 2000; Bishop, 2004; Singh, Devgan, Bhat & Milner, 2007). Clinical burn-wound literature consistently uses Jackson’s theory in discussing pathogenesis, assessment and treatment (Kagan & Smith, 2000; Bishop, 2004; Singh et al., 2007). A literature review of the past five years found Jackson’s 1953 seminal work cited in 181 articles of burn wound research and treatment.

The appearance of burn wounds were classified according to the day of injury and microscopic examination of tissue, according to a system of classification devised by Jackson, zone of hyperemia, zone of stasis, and zone of coagulation. The descriptive terms for each zone identified the microscopic characteristics that accompanied them. The zone of hyperemia identified a red area that blanched on pressure and became cyanotic if a pneumatic cuff was applied. Jackson interpreted this sign as an indication that circulation of blood and metabolism of tissue was intact. The zone of stasis is initially red, blanching with pressure, and becomes cyanotic with pneumatic cuff application. However, the circulatory responses cease within twenty-four hours indicating that the dermis has become avascular and necrotic, with red cells having hemolyzed. The zone of coagulation contains white coagulated tissue characterized by the destruction of the lumina of blood vessels. However, even microscopic examination may not reveal the extent of necrosis or viability of the dermis, i.e., whether the burn-wound is full-thickness or not.

As burn-wound healing progressed from the day of injury, Jackson observed that the zone of hyperemia healed spontaneously. The zone of stasis healed according to the extent that any vascular supply restored itself by healing. If no vascular supply returned, both the zone of stasis and the zone of coagulation would slough, and not heal. These zones would eventually require
skin grafting. The clear advantage to using Jackson’s zones is to assess burn-wound depth and to determine the optimum time to apply a skin graft, if at all. Jackson’s zones of burn injury still remain the current conceptual understanding of the burn wound (Pham, Gibran & Heimbach, 2007).

Today burn-wound experts continue to use their assessment skills to identify burn-wound depth. However, this procedure retains a level of subjectivity that affects the trustworthiness of the treatment decision. An instrument such as the HIM assists the dependability of the decision to determine the optimal time to apply a skin graft by specifically quantifying Jackson’s zones and the colors found in the burn-wound.

The criteria of the various classifications are related to later technologies which are attempts to simplify and quantify burn assessments. Burn assessment is now aided by current low tech paper charts such as the Lund and Browder charts (see Appendix B) which have replaced older less accurate charts such as the “Rule of Nines”, a body chart of “9 percents” of various body areas. Instead, the Lund and Browder charts, now considered to be the most accurate method of estimation of burn wound size (TBSA), use percentages of burned body surface area for given ages, and is considered to be the most informative diagnostic chart now available. In addition, the chart suggests treatments e.g. electrolyte balance, nutritional support, physical and occupational therapies, etc. (Wachtel et al., 2000).

**Gold standard of assessment.** The assessment of skin loss was first estimated by examining skin circulation with methods such as biopsy and histology (Devgan et al., 2006). Jackson (1953) identified injection of flourescein dye, pinpricking and blanching with pressure. Biopsy of the burn wound with histological examination is considered by some authors to be the gold standard of burn wound assessment (Jackson, 1953; Niazi, Essex, & Papini, 1993; Schiller, Garren, & Curtis, 1997). However, the procedure is considered by most clinicians to be unnecessarily invasive (Hemington-Gorse, 2005), painful (Bishop, 2004), open to sampling error (Devgan et al., 2006), potentially scarring, and expensive (Bishop, 2004; Park et al., 1998). Not
only do biopsy results not necessarily predict burn-wound depth, neither does a biopsy necessarily predict burn-wound healing time (Bishop, 2004). Thus, there is a clear need for accurate burn-wound depth assessment.

**Technologies Today in Burn-wound Assessment and Their Effectiveness**

In an effort to more perfectly visualize and prognosticate burn wounds, certain technologies have been attempted, all of which have significant inaccuracies. The technological assessments include: histological examination, vital dye injection or perfusion fluorometry, ultrasound, light reflectance, laser Doppler with thermography, burn depth indicator (BDI), video camera imaging (where frames can be slowed) (Acha et al., 2005; Bishop, 2004; Devgan et al., 2006; Hansen et al., 1997; Heimbach, Afromovitz, Engrav, Marvin, & Perry, 1984; Jaskille et al., 2009, 2010).

To date, none of the technologies which have been developed meet the immediate and long term needs of the burn patient. The technologies can be divided into two categories:

1) Those which assess physiological changes in the collagen, surface necrosis or surface vascularity of the wound, i.e. histological examination, vital dye injection, or perfusion fluorometry, ultrasound, light reflectance, or laser Doppler, and

2) Those which treat the colors in the burn wound, i.e. tri-stimulus colorimeter, burn-depth indicator, and video camera images and reflective optical multispectral imaging.

These procedures are described in the next sections. Additional technologies studied in animals, but not replicated with humans at this writing, mentioned in the literature include magnetic resonance imaging (MRI) and nuclear imaging (Bishop, 2004), radio-labeled tracers (RLT) for nuclear imaging to map burn depth. RLT introduced a potentially harmful substance and may not be justified for burn-wound assessment. (Sayman, Demir, Cetinkale, Ayan & Onsel, 1999). MRI and RTL will not be discussed.
**Assessment by physiological change.** Physiological changes have been attempted through:

1) histological examination,

2) vital dye injection or perfusion fluorometry,

3) ultra-sound,

4) light reflectance, and

5) laser Doppler with thermography.

described according to their numbered classification above.

Histologically, collagen denaturation suggests a full-thickness burn wound and can only be detected with histological examination (Devgan et al., 2006). However, the procedure is considered by most clinicians to be unnecessarily invasive (Hemington-Gorse, 2005), painful (Bishop, 2004), open to sampling error (Devgan et al., 2006), potentially scarring, and expensive (Bishop, 2004; Park et al., 1998).

Vital dye injection or perfusion fluorometry, e.g., using methylene blue or India ink, can identify surface necrosis, but are generally not able to distinguish between partial-thickness and full-thickness burns any more accurately than visual assessment. Thus, vital dyes have low clinical utility (Bishop, 2004). Indocyanine green (ICG) injection for angiography detects vascularity, but is limited in that topical ointments used in wound coverings interfere with this measurement technique and must be removed from the burn wound at least 10 minutes prior to the ICG. ICG is also somewhat expensive and requires a sophisticated instrumentation for interpretation.

Ultrasound has shown a positive correlation between denatured collagen and burn-wound depth, but the correlation was not substantively better than clinical observation for burn-wound assessment. Noncontact ultrasound used in an animal model (pig) distinguishes the layers of skin to estimate burn-wound depth. However, high frequency ultrasound requires direct contact, which
is also likely to produce pain. While these techniques suggest improved detection, none of them have provided the desired accurate and reliable assessment of burn-wound depth (Bishop, 2004).

Light reflectance using infrared light was significantly more accurate than the burn-wound assessment done by burn-wound experts in predicting burn wounds that would not heal on day three (3). However, infrared light technology is extremely costly (Bishop, 2004). These techniques have not demonstrated acceptable reliability in research studies (Devgan et al., 2006; Park et al., 1998).

The Doppler principle recognizes that moving objects, such as red blood cells, cause a frequency shift of light waves while light reflected by stationary structures remains unchanged in frequency (Kloppenberg, Beethuizen, & ten Duis, 2001). Laser Doppler imaging (LDI) detects blood flow in the small blood vessels surrounding the burn wound, indicating viable dermis in a noninvasive technique that measures the rate and volume of moving blood cells through tissues (Jaskille et al., 2009, 2010).

Consequently, the vascularization of burn wounds was thought to be visualized through laser Doppler techniques. Laser Doppler technology has therefore been found to be most accurate in burns with definitive presence (i.e., greater vascularity) or definitive absence of vascularity (greater depth) burns. Because the indeterminate wound does not have an easily identified vascularity, laser Doppler is not as accurate as clinicians would like. Thus the accurate assessment of the indeterminate burn wound in a timely manner is critical to the treatment of choice and impacts of patient recovery and length of stay (LOS) (Chatterjee, 2006; Jaskille et al., 2009, 2010).

Three reviews of laser Doppler (LD) contained salient insights into the status of laser Doppler imaging (LDI) research (Chatterjee, 2006; Jaskille et al., 2009, 2010). In a critical review of 14 LD studies, Chatterjee (2006) evaluated the clinimetrics of LDI as a method of burn-wound assessment, particularly considering the potential for LD use in clinical practice. Six research studies compared LDI findings with clinical assessment of burn wound healing time. However,
none of these studies demonstrated positive effects of LDI technique over the traditional standard of care.

Specifically, the assignment of burn wound healing time in the literature varied from 12 to 21 days, which prevented any comparison of findings. Five research studies compared LDI findings with histological assessment. These research studies reviewed by Chatterjee (2006) have some limitations. All of these five studies examined only superficial burns and showed no correlation of LDI and clinical assessment of burn wound healing time. One study compared LDI results with clinical assessment in the decision to excise and graft burn wounds in a retrospective review of clinical management. Another study included in the review used a scale for diagnosing of burn-wound depth that was not validated. One study used the LDI to assess the varying effects of burn resuscitation not in humans, but rats. In this review, Chatterjee (2006) concluded that variations in different kinds of LDI equipment do not allow a direct comparison of individual research study results. Although most research studies provided clinical evidence to support the use of LDI assessment, these research studies lack a standardization of equipment and burn-wound assessment conditions such as temperature of the burn wound and scanning distance from the burn wound. LDI should be limited to only research activities because it is costly (a Doppler unit cost $68,000 in 2001) and not standardization (Chatterjee, 2006; Hemington-Gorse, 2005).

This argument is further supported by another critical evaluation of the literature. Jaskille (Jaskille et al., 2009, 2010) evaluated forty-one studies on laser Doppler with the enhancements of dynamic and infrared thermography, and found that different machines, scanning mechanisms and directions, varying skin thicknesses by person and by wound, and physiological issues of shock, moisture, edema and the laser beam angle and penetration did not permit endorsement of laser Doppler in the clinical arena. Several factors impede finding strong positive results: 1) use of different makes of machines; 2) use of different scanning distances; 3) scanning different body areas with different tissue thickness; 4) different skin thickness among individuals; 5) movement of scanned area; 6) burn-wound temperature; 7) burn-wound moisture; 8) incident angle of the
laser; 9) presence or absence of shock; 10) extent of tissue edema; and 11) whether the scan is performed directly on the tissue or through a transparent dressing. A uniform, absolute “scanning unit” representing the measurement of tissue perfusion has not yet been developed (Jaskille et al., 2009 & 2010).

LDI offered the best data-supported estimates of accuracy, i.e. the estimates of burn-wound depth supported by research findings. However, LDI accuracy is contingent on optics, which is affected by the heterogeneity and curvature of burn-wound tissue, topical substances, and ambient light as well as wound infection (Devgan et al., 2006). The single advantage of LDI is an accurate prediction of a small burn wound in a stable patient. However, because the method of LDI relates to the measurement of burn wound vascularity, the reading of LDI is affected by temperature, distance from the wound, wound humidity, angle of recordings, extent of tissue edema, and presence of shock. Such variables make LDI measurements extremely labile. The review of current research studies found that difference in method made broad conclusions difficult. Such methodological differences included a small number of participants in each research study (less than 20 in most studies), different models of LDI technology, different scanning distances and different body areas having been scanned (Jaskille et al., 2009 & 2010).

Finally, the cost of laser Doppler, cited at $68,000 per machine in 2001, further prohibits its use in most clinical practices (Hemington-Gorse, 2005). No more current estimates of cost were identified. Because technology remains inconclusive, human visualization of wounds remains the prevailing clinical diagnostic path.

**Assessment by color.** The tri-stimulus colorimeter assessed burn-wound depth by scanning the burn wound without direct contact thereby eliminating pain and possible contamination of the burn wound. However, only burn wounds that were superficial or deep-partial thickness were assessed. Research results showed that the overall accuracy of partial thickness burn-wound depth assessment was significantly more accurate using a colorimeter than clinical scoring. Specifically, however, the challenge of discriminating between partial-thickness
and full-thickness burn wounds is not addressed using the tri-stimulus colorimeter (Mastronicola, Romanelli & Barchini, 2005).

The Burn Depth Indicator (BDI) device records ratios of red/green, red/infra-red, and green/infra-red reflected light. Because skin is relatively transparent to short-wave infrared light, the measurement of hemoglobin is possible. Reduced hemoglobin absorbs more light than oxygenated hemoglobin. Thus, thrombosed vessels in full-thickness burn wounds are more visible in infrared light than open vessels of partial-thickness burn wounds. Results using the BDI indicated significantly greater accuracy in predicting burn-wounds that would not heal independently within 21 days as compared with the clinical assessment of experienced burn surgeons. There was no significant difference between the BDI and clinical assessment to predict a burn wound that would heal within 21 days. Because clinical assessment of burn wounds is inaccurate except for obvious full-thickness or shallow partial thickness burns, better burn care requires speedier identification of burn-wound depth and treatment (Heimbach et al., 1984).

Afromovitz, Van Liew, and Heimbach (1987) used the BDI to describe the optical properties of burn wounds and identified two salient questions relevant to burn wound assessment: 1) What is the best discriminate line separating the data for healing and nonhealing levels of burned skin? 2) To what extent is the BDI technique independent of the patient or burn-wound site? Although the researchers stated that identified burn-wound depth was correlated with reflectance ratio values, no significance of the correlation was reported. They also reported that the BDI would be able to predict whether or not a burn wound would heal in 21 days. However, no reliability analysis was reported for this pattern.

Video camera imaging as reported by Hansen et al. (1997) detail the limitations of the human observer in analyzing burn wound and pressure ulcer images, particularly the inability to quantify slow or subtle changes in an image. Hansen et al. conclude that color image processing that quantifies epidermal color variations would be of particular use to burn-wound experts. The animal model for a burn wound used young, mixed-breed pigs. Burn wounds were examined for
up to 27 days and analyzed using color imaging obtained with a video camera and processed with computer digitization. The technique has improved the reliability of tracking the visual appearance of the burn wound over time. This study supports the use of color imaging for burn wound assessment. However, more reliable identification of burn-wound depth as measured by color variations of epidermis and dermis are needed using human participants.

Several novel techniques and future directions in burn-depth assessment were identified by Devgan et al. (2006): 1) optical measurement as reflective optical multispectral imaging, fiber-optic confocal imaging, and orthogonal polarization spectral imaging, 2) nuclear measurement, and 3) non-contact and high-frequency ultrasound. Burn-wound depth is related to light reflected from the burn wound. Reflective optical multispectral imaging performs spectral analysis of the reflected light from the burn wound, using the concept that necrotic tissue, scarring and dermal vessel oxygen saturation alters the absorption of light. To assess tissue structure and function in an animal model, optical coherence tomography uses polarity measurements of collagen birefringence amplitude, orientation and di-attenuation. Reduction of collagen birefringence may be related to burn-wound depth.

Fiber-optic confocal imaging in an animal model is microscopy of subsurface living tissue to detect denatured collagen, a characteristic of necrotic tissue. The illumination of tissue with blue light frequencies elicits autofluorescence of necrotic tissue, which is directly proportional to burn-wound depth. Orthogonal polarization spectral imaging (OPSI) illuminates the tissue with polarized light within the hemoglobin spectrum, which produces a non-invasive assessment of microcirculation of the burn wound.

This research (Devgan et al., 2006) referred to partial-thickness burn-wounds as superficial. Superficial burn wounds had small visible dermal capillaries in the field of view. Deep or full-thickness burn wounds showed large thrombosed vessels in a criss-crossed network. While demonstrating a visible difference between superficial and full-thickness burns, the difficulty in clinical observation of a burn wound remains in discriminating between partial-
thickness and full-thickness burn wounds, not superficial and full-thickness burn wounds. OPSI requires direct tissue contact, which is likely to produce pain and possibly contamination of the burn wound. OPSI covers only a small area in a single reading, which requires multiple viewings for a burn wound (Devgan et al., 2006).

Jones et al. (2003) studied the reliability of digital images for burn-wound assessment using a Nikon Coolpix 995 camera. The Nikon Coolpix 995 camera utilized three different file sizes: 2.25, 5.5 and 9 Mbyte (1024x768, 1600x1200 and 2048x1536 pixels respectively). Their interest in digital imaging emerged from the concern that the correct diagnosis of partial thickness burn wounds is one of the most challenging aspects of burn-wound treatment. The differential diagnosis reflects the need for surgical debridement and skin grafting for burn wounds that will not heal spontaneously from the remaining intact dermis. Using 60 participants, they obtained in-person assessment and two digital imaging assessments—with an interval between assessments of four to six weeks. Factors studied were depth of burn (partial or full thickness), the presence of infection, or edema. Digital imaging was assessed for usefulness of location image and quality of all digital images (location and burn wound). Cross-tabulation compared agreement between in-person assessment and digital image assessment. No test of significance in differences between in-person assessment and digital image assessment was conducted. Results indicated no differences between the three file sizes of digital images as tested by the Kappa score.

In general, burn-wound evaluators use color observation as a standard of care to assess burn-wound status, particularly important in the assessment of burn-wound depth. In their argument for the need to automate images of burn-wounds for future reference and treatment in the entire world’s burn centers, Serrano et al., (2003) discussed a mathematical algorithm for segmenting color on an image, which they considered an improvement on their earlier work. They used color and texture information to perform their segmentation with three steps to analyze an image. First, they developed a uniformly diffused image and next they converted the color image into a single grayscale image; then they finally applied an automatic thresholding to the
image in order to segment the image and distinguish burn regions in the image. The researchers used 30 images to test the feasibility of their technique, but noted that the entire process is quite difficult. In a later study, Acha et al. (2005) built on their understanding that burn surgeons use color to determine burn-wound depth, and now examining burn-wound texture, which remains a largely unexplored field. Now they used their color image segmentation to distinguish between burn wounds and healthy skin, for the eventual end of distinguishing between full and partial-thickness burn-wounds. Results were inputted into a software package called Fuzzy-Artmap neural network, which serves as a classifier for three types of burn-wound depth: superficial, deep dermal and full thickness.

Within the three types of burn depth, the Fuzzy-Artmap neural network also differentiates blisters and gives five examples of different appearances such as rose-brown for blisters, bright red for superficial dermal burns, pink-white for deep dermal burns with a dotted-appearance, yellow-beige for the first appearance of full thickness burns and brown for the second appearance of full-thickness burns. The imaging system was a Canon EOS 300D digital camera with a resolution of 1536x1024 pixels and developed calibration methods based on the Macbeth Color Checker DC chart, which is specifically designed for the calibration of digital images.

Acha et al. (2005) repeated their earlier work on grayscale imaging, and using 16 features to represent a burn wound and developed a neural network technique based on the adaptive resonance theory developed by Grossberg and Carpenter. Based on the neural network analysis, these authors arrive at a metric called a positive predictive value which represents the probability of a good fit with histological examination of burn wound biopsy. A panel of twelve experts in burn diagnostics analyzed 38 photographs of all etiologies of burn-wounds and successfully classified the images in agreement with the imaging system 82% of time. While this research study indicated an improvement in burn-wound depth detection, it only allowed for assessment of a tiny area of the burn wound, and yet also required biopsies, which can be painful and difficult for a burn-wound patient.
In a classic study on digital imaging, Otsu (1979) discussed the calculation of a nonparametric and unsupervised method of automatic threshold selection for picture segmentation to maximize the separability of the resultant classes in gray levels of histograms obtained for a black and white computer image, a threshold selection where only the gray-level histogram suffices without other *a priori* knowledge, using discriminate analysis. The gray-level histogram suffices for threshold selection to automatically select an optimal threshold. Threshold selection is essential for unsupervised decision problems of pattern recognition through discriminant analysis because it directly approaches the feasibility of evaluating the “goodness” of the threshold. Otsu compared the valley sharpening technique, which restricts the histogram to the pixels with large absolute values of derivative (Laplacian or gradient), with the difference histogram method, which selects the thresholds at the gray level with the maximal difference. Using a parametric technique, where the histogram is approximated in the least square sense by a sum of Gaussian distributions and statistical decision procedures applied, calculations are tedious and may be unstable.

Gaussian distributions are meager approximations of real modes. Otsu concludes that threshold selection where only the gray-level histogram suffices is the most simple and standard method for automatic threshold selection characterized by the following advantages: 1) a simple procedure where the zeroth and the first order cumulative moments of the gray-level histogram are utilized; 2) A straightforward extension to multi-thresholding problems is feasible by virtue of the criterion on which the method is based; 3) An optimal threshold (or set of thresholds) is selected automatically and stably, not based on the differentiation (i.e., a global property) of the histogram; 4) further important aspects can also be analyzed (e.g., estimation of class mean levels, evaluation of class separability, etc); and 5) the method is general which covers a wide scope of unsupervised decision procedures.
Potential machine vision assessment using CMIS. Although outside the purview of this dissertation, if the results of this research study would indicate a need for machine vision assessment, then the following would likely be used to assess photographs of burn wounds.

Morphologic feature extraction (MFE) based upon mathematical morphology, i.e., Geodesy, is an important aspect of machine vision. Thiran and Macq (1996) used MFE to recognize cancerous tissue by extracting essential shape characteristics and eliminating irrelevancies. The four basic operations are dilation, erosion, opening and closing, with morphological reconstruction. The technique is used to remove the background noise from the image, operate a segmentation of the nuclei of the cells, and an analysis classification of the image (cancerous or not) is finally produced. Image acquisition used a high quality optical microscope, a high-resolution CCD B/W camera and an acquisition hardware/software system, running on a UNIX workstation. Every image was acquired with 256 gray levels, with a magnifying factor of 60 on the microscope, under immersion. Manual corrections were operated on the contrast and intensity in such a way that the digitized image became visually acceptable for further manual classification by a specialist. No other filtering was operated. The final resolution of the images was $0.16\mu m$ per pixel. In this range of magnification, the shape analysis process is independent of the magnification factor, since all measured values will be normalized according to the size of the objects.

Eighty-three images of biopsies of lung (Set 1 =35; Set 2= 20) and digestive tract (Set 1=28) tissue were analyzed for the four criteria of malignancy: 1) nucleocytoplasmic ratio; 2) anisonucleosis; 3) nuclear deformity; and 4) hyperchromasia by an experienced specialist and the automatic digitized image. Theran and Macq report a 90% accuracy of classification for Set 1 and 85 % for Set 2. They reported a good correlation, but did not report r-values.

The Compact Microscope Imaging System (CMIS) with intelligent controls was developed by NASA for use in space exploration. The CMIS is able to automatically detect and track microscopic changes in cells, cell morphology, and surfaces as well as metric feature extraction
and object identification. Metric feature extraction refers to the ability of the CMIS system to extract scientific data from an interface of interest. Object identification is the ability to discriminate form at the interface of interest. (McDowell, 2004; McDowell, 2006; McDowell & Gray, 2004). The CMIS has been shown to be highly test-retest reliable in research on tuberculosis (McDowell, 2006). The CMIS would potentially measure color differentiating the burn wound from the surrounding normal skin as described above. The CMIS uses the Human Vision System (HVS) to delineate color. Humans make mistakes related to fatigue. Their decisions are based on their expertise, but that may not always produce objective decisions. The CMIS system has the potential to avoid human error, which happens with an observation conducted over any extended period of time.

The best machine visualization utilizes the Human Vision System (HVS) composed of three levels (McDowell, 2006):

- Hue, the color type;
- Saturation, the vibrancy of the color where the lower saturation indicates a greater fading of the color; and
- Value, the brightness of the color.

**Summary.** At present, laser Doppler imaging has emerged as a leading, although not perfect, technology as a clinical aid limited by inaccuracies. As reviewed, other technologies have been researched and tested in humans. The machine visualization technology, found to be reliable to assess burn wounds, holds the promise of supporting and enhancing human vision in the assessment of burn-wound depth without invasive procedures. Detecting burn-wound color via machine vision is the key component to enhancing the assessment of burn-wound depth. Various machine vision systems have demonstrated inroads into a supportive and enhanced technology, but contain considerable limitations. A machine vision assessment using CMIS has potential for development because of its use of the HVS system to detect color so important to accurate burn-wound depth assessment.
CHAPTER THREE

Method

Design and Objectives

The objectives of this exploratory, correlational analysis are to examine: 1) if burn-wound color (Red and White) and Zone Percent are predictors of skin graft application; and 2) the interrater reliability of the three burn-wound experts using the HIM.

Sample. The sample in this study has two sub-samples: Three burn wound experts and 50 color photographs. As established in Chapter One, the use of photographs of the progress of a burn wound, standard practice in most burn centers, is a good source of aggregated data available for specific types of studies or treatment decisions (Hansen et al., 1997; Jones et al., 2003) and augment the use of real time assessment (Nelson et al., 2006). The sample of 50 color photographs for this dissertation is drawn from archived burn-wound photographs of discharged burn patients whose burn-wounds were identified as clinically indeterminate at the time of admission as determined by two burn-wound nurse-experts who are separate from the dissertation color assessment experts and did not review photographs for the assessment of burn-wound colors. These burn-wound nurse experts had three or more years experience in burn-wound care. No partial-thickness or full-thickness burn-wound photographs were included.

A burn-wound was considered indeterminate upon admission when clinical judgment could not definitively assess depth as partial-thickness or full-thickness. The burn-wound must be assessed to identify the stability of burn-wound dermal vascularity when spontaneous healing is expected, or if ischemic changes in the burn-wound continue to destroy dermis, and a skin graft is indicated (Baker & Mondonzzi, 2011). Using this primary data set, inference about burn-wound depth was obtained from burn-wound experts by their assessment of color.
All 50 photographs in this study were designated as indeterminate upon admission by burn-wound nurse experts. However, the reader is reminded that no photographs of the skin graft application were included. Thus, if the skin graft did occur, it became the clinical validation of the burn-wound depth, i.e., full-thickness. Photographs from patients without skin grafts subsequently indicated that burn-wound depth had been clinically determined to be partial-thickness, capable of spontaneous healing, while the selection of photographs from patients with skin grafts subsequently indicated that burn-wound depth was clinically determined to be full-thickness. The variable of skin graft status (skin graft vs. no skin graft), therefore, was used to identify which of this sample of photographs had either full or partial-thickness or burn-wound depth, even though upon admission, burn-wound depth was indeterminate. The date of the skin graft application demonstrated at what time the burn-wound was clinically determined to be a full-thickness burn-wound, (thus requiring a skin graft application).

As previously discussed in Chapters One and Two, the color of the burn-wound is key to the assessment of burn-wound depth. Thus, the sample of photographs was selected so that burn wound thickness was distributed, and thus, the distribution of color would be similarly distributed and not be attributed only to partial-thickness burns or only to full-thickness burns.

**Instrument.** The HIM is a paper and pencil instrument that this researcher devised, which represented or matched the burn-wound assessment conducted in the burn center using the Jackson model (Jackson, 1953). This researcher formulated a table which included Jackson’s three zones of injury: 1) zone of necrosis or coagulation (Central); 2) zone of stasis (Midrange); and 3) zone of hyperemia (Peripheral), as well as the hues or color types of the burn wound: 1) Red; 2) Yellow; 3) White; 4) Green; 5) Brown; and 6) Black. Thus, the table would allow the burn-wound expert to assess the percentage of each zone and the percentage of each color within a zone. (See Appendix C).

The HIM was used to describe burn-wound color, and is valid in construct. Construct validity refers to the extent to which the empirical measurement of color measures what it is
supposed to measure (wound depth) (Carmines & Zeller, 1979). The HIM is a simple listing of possible colors of the burn wound, i.e., red, yellow, white, green, brown, and black, and does not require a narrative description of the burn-wound or an independent listing of color by the burn-wound expert. Until this instrument was devised, no research identified had so precisely assessed color of the burn wound and its area using Jackson’s model for burn-wound assessment, or assigned percentages of color to a wound.

Face validity was used after the HIM was developed. A panel of three burn-wound nurse experts, none of whom functioned as reviewers of photographs, examined the HIM in terms of its appearance (Polit & Beck, 2008). Because the HIM was developed based on a well-accepted theory—the Jackson’s theory, including colors of burn wound as a significant indicator of wound depth, the HIM which includes all of the colors identified by Jackson, is considered valid in its construct (Carmine & Zeller, 1979).

Reviewers. Using the HIM, three burn-wound experts, i.e., two burn nurse-practitioners and a burn surgeon, reviewed photographs to assess color, the principle indicator for depth of the burn wound. Burn-wound experts evaluated photographs using case numbers assigned to each photograph and were unaware of patient identity in the photographs. Burn-wound experts used the HIM instrument to evaluate each photograph’s colors by assigning percentage to each color.

Procedure of the review. In this study, each burn-wound expert assessed 50 burn-wound photographs, taken at three time points: (1) Baseline upon admission, (2) Midtrajectory in treatment, and (3) the Final photograph time, which occurs either before the skin graft or upon discharge. These photographs were stored on a computer compact disc (CD) located in each expert’s office, simulating the experience of reviewing burn-wound photographs in the course of treatment. Using HIM, each expert assigned percentages to burn-wound colors (red, yellow, white, green, brown, and black) in each photograph. The sum of each color is 100%. This researcher offered to transcribe the entries of color selection and percentage as verbally reported by each burn-wound expert reviewer. The transcription procedure provided by this researcher
served to simulate the standard procedure of dictating findings in the clinical review of burn-wound photographs. The burn surgeon and both burn-wound nurse practitioner accepted this researchers offer to transcribe.

**Reliability.** Concept information obtained in research is reliable when it is accurate and consistent. Statistical reliability (now referred to as reliability) refers to the probability that information obtained in a different sampling or by different observers will give the same results (Polit & Beck, 2008).

Measurement of reliability involves a basic formulation where the observed score (X) equals the true score (t) plus random measurement error (e), i.e., \( X = t + e \). The true score is an unobservable quantity that cannot be directly measured. The true score is identified as the average score that would be obtained if the variable were measured an infinite number of times. True scores are hypothetical. However, a true score is a central factor in understanding reliability measurement. Random measurement error represents the random disturbance that affects the true score in measurement. As a quality of randomness, observed scores would be distributed in a random fashion, thus “positive” errors would be as likely to occur as “negative” errors. Therefore, reliability can be expressed as error variance. Reliability is the ratio of the true to the observed score. (Carmine & Zeller, 1979).

**Types of reliability measurements.** Definitive indices of interrater reliability are not universally agreed upon by authorities (Orwin, 1994). Four indices are suggested as foundational in the measurement of reliability: 1) agreement rate expressed as a percentage, 2) kappa and weighted kappa as a family of indices adapted to various circumstances, 3) intercoder or interrater correlation using the Pearson Product Moment correlation (\( r \)), and 4) interclass correlation (\( r_i \)) computed as a ratio of the variance divided by the variance plus error.

Agreement expressed as a percentage requires the formula of the number of observations agreed upon divided by the total number of observations. With nominal data, a percentage of agreement can be computed, but chance or “guessed” agreement cannot be controlled (Orwin,
Interval data used in this dissertation suggests more advanced analytical measures.

The Kappa statistic measures agreement for categorical data (Orwin, 1994) when the number of evaluators is two or greater, and the categories for evaluation is also two or greater (Green, 2010). Each category uses the same category value. Thus, the smallest variable set that may be used with Kappa is a 3 x 3 matrix, with a category value of two for each category, e.g., yes or no. Kappa values ≥ .75 indicate a high agreement, and values ≤ .40 indicate a low agreement. Values of >.40 and < .75 indicate a fair to good level of agreement (Green, 2010). Because Kappa analysis is designed for use with categorical data, it will not permit the most exhaustive inference in the analysis of interval or continuous data. Thus, Kappa was not an appropriate analysis for this study.

Interclass correlation is used with categorical and interval data and permits the researcher to choose not to use the between-rater variance in the error term according to form of $r_I$ to be applied. Consequently, results require a substantial between-items variance to produce a significant indication of agreement (Orwin, 1994).

In the review of interrater reliability research, some authors stated that their research identified interrater reliability with the report of confidence intervals for variables of biopsy and visual assessment (Hansen et al., 1997) and plotting of means and standard deviations for comparison of two assessment charts for size of the burn wound (Wachtel et al., 2000). Confidence intervals, i.e., the association of probability values with interval estimates, means and standard deviations as measures of central tendency, are all measures of association, not reliability. Confidence intervals are, at best, an indirect method of assessment of interrater reliability (Orwin, 1994; Loether & McTavish, 1980).

This research study focused on burn-wound colors shown in past photographs, which signal burn-wound depth. Thus, burn-wound color is the observable response, while the underlying unobservable response is burn-wound depth. The measuring of the observable empirical
observation of color links then to the abstract concept of burn-wound depth. It is not possible for human vision to clearly see the depth of a burn wound, assessing the actual level of tissue destruction and the actual level of ischemia, but healthy human vision can assess and process color.

Validity. The degree to which an instrument measures what it is supposed to measure is validity (Polit & Beck, 2008). Validity is the extent to which an instrument provides data compatible with other relevant evidence (Diers, 1979). Construct validity refers to the extent to which the empirical measurement of color measure what it is supposed to measure (burn-wound depth) (Carmine & Zeller, 1979). The reliability and validity of an instrument are interdependent qualities. An instrument must be reliable to be valid, i.e., stable, consistent and accurate. An unreliable instrument may be measuring factors associated with random error and thus not a valid indicator of the target variable (Polit & Beck, 2008).

Data Analysis

Frequencies, percentages, means, and standard deviations were used for descriptive analysis. To examine inter-rater reliability, the intraclass correlation (ICC) was used in this study, yielding reliability coefficients as the magnitudes of relative agreement among the three raters on color percentages (Hoyt, 2010). These coefficients are ratios of between rating variance to total variance, comparing covariance of ratings with total variance. There are three types of ICC, based on the rating design: 1) one-way ANOVA for raters nested in targets (different sets of raters assessed same sets of photographs); 2) two-way fixed design for raters crossed with targets (different sets of raters assessed different photographs); and 3) two-way fixed model for fixed raters crossed with targets (one set of raters assessed same photographs) (Hoyt, 2010; Shrout & Fleiss, 1979). The appropriate model for this dissertation is the third model, two-way fixed, because the reviewers are a fixed effect and the variables of the photograph assessment are also fixed.
To examine predictors of skin graft. Logistic regression analysis was used because the outcome of this analysis is dichotomous (skin graft vs. spontaneous healing) (Foster, Barkus, & Yakovsky, 2006). The independent variables were Zone Percents of the burn wound (Central and Midrange separately), as well as the Colors Red and White. These Zones Percents represented the zones of injury more likely to contain an increased burn-wound depth, which would be more closely assessed to determine the need for a skin graft.

In wound assessment, several colors typically indicate the depth of injury in the burn wound and the presence of infection. However, in this study, only red and white will be used as independent variables to predict skin graft status because these two colors are the most significant colors generally used by burn wound experts in clinical practice to determine skin graft status. While red indicates the presence of vascularity, white represents the absence of vascularity, thus, indicating the viability of tissue, i.e. dermal destruction.

The direct method of logistic regression is used to place all independent variables into the regression equation because there is no specific order or importance for the independent variables. Multivariate outliers were examined using the standardized residuals. No outliers were found, based on the cutoff, absolute value of 3.0 (Polit, 2010). To test the assumption of multicollinearity, the Pearson Product Moment Correlation matrix was produced among the independent variables. Since no correlation exceeded .80, multicollinearity was considered non-problematic (Polit, 2010). Also, using multiple regression technique, values of Tolerance test and Variance Inflation Index (VIF) were produced, yielding acceptable Tolerance values of over .10 and values of VIF of less than 5, indicating no multivariate multicollinearity (O’Connell & Amico, 2010; Tabachnick & Fidell, 2007) All independent variables, therefore, were included in the analysis. The assumptions of linearity, normal distribution, and homogeneity of the independent variables were not tested since they are not required in logistic regression.

The omnibus test identified the significance of the overall model. The Chi-square identified the association of the independent variables to the dependent variable, i.e. whether or
not the model fits the data. The Nagelkerke $R^2$ indicates the percentage of the dependent variable predicted by the independent variables (Foster et al., 2006).

Logistic regression calculates an odds ratio, i.e., a measure of association between two variables, while controlling for other variables. The estimated odds ratio for a dependent variable in the regression equation is the exponential beta ($\text{ExpB}$), which is the value of beta exponentiated. $\text{ExpB}$ is calculated using the constant $e = 2.71828$ raised to the power of the $B$ ($e^B$). The regression equation predicts the odds of the independent variables (zone percent of burn wound, as well as the colors Red and White) being in one of the two groups on the dependent variable (skin graft or spontaneous healing). In this study skin graft = 1, while spontaneous healing = 0. (Foster et al., 2006).

**Correlational analysis.** As an additional analysis, the Pearson Product Moment Correlation for the independent variables of length of stay (LOS) and total body surface area burned (TBSA) were individually analyzed with the dependent variable, or skin graft status to identify any association. Also analyzed were predictors (means of Zone Percent and the Colors Red and White) of skin graft by Photograph Time (Baseline, Midtrajectory and Final) each for the Central and Midrange Zones.

**Human Rights Considerations**

Akron Children’s Hospital Institutional Review Board granted expedited approval, with the current renewal date of March 22, 2012. Kent State University Institutional Review Board granted a Level II expedited approval, acknowledging Akron Children’s Hospital Institutional Review Board as the primary Institutional Review Board for this protocol. As indicated in the discussion of the research design, the sample consisted of archived photographs of burns with known subsequent treatment outcomes from already discharged burn patients, adults and children who have been treated at the hospital. Each patient or patient’s family provided permission to take the photographs and to use them for educational and therapeutic purposes. The hospital’s burn center maintains an historical library of photographs of burns for such educational and
therapeutic purposes. This library has been assembled for approximately the last 50 years and is available to the educators in nursing and medicine of the hospital. Thus, the hospital has agreed to allow this researcher access to the photographs of patients with burn wounds of indeterminate depth upon admission. Although the photographs were found per patient number, they were anonymous in that the reviewers did not know patient names. Patient permission had been gained by the hospital prior to taking the photographs, and permission had been granted by patients for the photographs’ use in research and teaching. (See Appendix D).
CHAPTER FOUR

Results

Demographics

A non-random sample of 50 retrospective photographs was selected from 22 admitted patients of the past two years (2008 and 2009). Patients’ age ranged from one year to 95 years, with a mean of 34 years (S.D. = 25.1). Both males (50%) and females (50%) were included. Total body surface area burned (TBSA) ranged from 1.5 % to 65%, with a mean of 15% (S.D. = .149). Thirteen (59%) had a TBSA of <10%, six (27%) had a TBSA between 10% and 20 %, with three (14%) having a TBSA of > 35%. A flame (40%) was most frequently the source of the burn wound, followed by steam (31.8%), contact (18.2%) and other (9.1%). Length of stay ranged from six to 53 days, with a mean of 20 days.

More patients received a skin graft (64%) than did not. The mean post-burn wound day for skin grafting was 10. Three patients (14%) received a skin graft within a week of the burn wound. After two weeks an additional five patients (23%) had received a skin graft, with six patients (27%) waiting from 16 to 32 days to receive their skin graft. Table 1 represents skin graft frequencies.

Upon admission, all 22 patients had clinically indeterminate acute burn wounds. Burn-wound depth is indeterminate when clinical assessment cannot definitively assess depth as partial-thickness or full-thickness. An indeterminate burn-wound by definition must be further assessed to identify the stability of burn-wound dermal vascularity, or whether, if unstable this leads to dermal destruction (Baker & Mondozi, 2011).
Fourteen (68%) of the 22 patients had a skin graft in the course of their treatment; the other eight did not. For each patient, photographs of their burns were included in the data.

Table 1.

*Frequencies and percentages of Skin Graft Status (n = 22)*

<table>
<thead>
<tr>
<th>Skin graft status</th>
<th>N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Skin graft</td>
<td>8 (32%)</td>
</tr>
<tr>
<td>Skin graft</td>
<td>14 (68%)</td>
</tr>
<tr>
<td>Within 7 days</td>
<td>3</td>
</tr>
<tr>
<td>After 14 days</td>
<td>5</td>
</tr>
<tr>
<td>In 16 to 32 days</td>
<td>6</td>
</tr>
</tbody>
</table>

For each patient, photographs of their burns were included in the data reflecting these times: (1) Baseline upon admission, (2) Midtrajectory in treatment, and (3) Final photograph times, which occurs either before the skin graft or upon discharge. Because photographs were obtained retrospectively from the library of burn-wound photographs, not all participants had photographs for each of the times listed. Nineteen participants had Baseline photographs, 18 participants had Midtrajectory photographs, with 2 participants having two Midtrajectory photographs. Only eleven participants had a Final photograph. Thus the photograph is the unit of analysis with a total of 50 photographs. Table 2 represents photograph frequency.

Table 2.

*Frequencies of photographs*

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Mid-trajectory</th>
<th>Final</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photographs</td>
<td>19</td>
<td>20</td>
<td>11</td>
<td>50</td>
</tr>
</tbody>
</table>
Data Analysis for Burn-Wound Experts

**Burn-wound experts’ assessment.** Three burn-wound experts used the Human Intelligence Matrix (HIM) to evaluate each photograph. Colors assessed included red, yellow, white, green, brown, and black. All colors were used in the photograph assessment by at least one burn-wound expert. The sum of the colors was 100% in each photograph. The photographs were labeled in order of post-burn-wound day: Baseline, Midtrajectory, and Final. Burn-wound experts responded by case number and were unaware of patient identity.

**Correlation analysis.** The Pearson Product Moment Correlation for the independent variables of length of stay (LOS) and total body surface area burned (TBSA) were individually analyzed with the dependent variable of graft or no graft (spontaneous healing) to identify any association. Between TBSA and skin graft the relationship was negative and weak but statistically significant, i.e., as the TBSA increased, the frequency of skin graft decreased, (r = -.084, p <.05). A positive and weak but significant correlation was detected between LOS and skin graft (r=.195, p<.01). A positive and weak but significant correlation was detected between TBSA and LOS (r=.210, p<.01). Table 3 represents the correlation analysis.

Table 3.

<table>
<thead>
<tr>
<th>Correlation Analysis for TBSA, LOS and Skin Graft</th>
<th>Pearson</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Body Surface Area (TBSA) by Skin Graft</td>
<td>-0.084</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Length of Stay (LOS) by Skin Graft</td>
<td>0.195</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Total Body Surface Area (TBSA) by Length of Stay (LOS)</td>
<td>0.210</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>

Also analyzed were the variables of skin graft vs. spontaneous healing, and the means of Zone Percent, as well as the Colors Red and White by Photograph Time (Baseline, Midtrajectory and Final) each for the Central and Midrange Zones. For the Central Zone, at the Midtrajectory Photograph Time, a negative and weak but significant correlation was found for Zone Percent (r
= -.254; p < .05), with positive and weak but significant correlations for Red (r = .267; p < .05) and White (r = .268; p < .05).

For the Midrange Zone, at the Baseline Photograph Time, negative and weak but significant correlations were found for Red (r = -.265; p < .05) and White (r = -.264; p < .05). At the Midtrajectory Photograph Time, a negative and weak but significant correlation was found for Zone Percent (r = -.268; p < .05), with positive and weak but significant correlations for Red (r = .293; p < .05) and White (r = .294; p < .05). For the Final Photograph, all means for Zone Percent, Red and White coincidently were the same for each burn-wound expert, representing a constant value. The correlation could not be computed. Table 4 represents the correlation analysis.

Table 4.

Graft, Zone Percent and Color Correlation Analysis

<table>
<thead>
<tr>
<th>Photograph Time</th>
<th>Mean Zone Percent</th>
<th>Mean Color Red</th>
<th>Mean Color White</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>.054</td>
<td>-.241</td>
<td>-.240</td>
</tr>
<tr>
<td>Midtrajectory</td>
<td>-.254*</td>
<td>.267*</td>
<td>.268*</td>
</tr>
<tr>
<td>Final</td>
<td>.a</td>
<td>.a</td>
<td>.a</td>
</tr>
<tr>
<td>Midrange Zone</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>.039</td>
<td>-.265*</td>
<td>-.264*</td>
</tr>
<tr>
<td>Midtrajectory</td>
<td>-.268*</td>
<td>.293*</td>
<td>.294*</td>
</tr>
<tr>
<td>Final</td>
<td>.a</td>
<td>.a</td>
<td>.a</td>
</tr>
</tbody>
</table>

Note. *Correlation is significant at the 0.05 level.
.a Cannot be computed because at least one variable is constant.

Logistic regression analysis. Logistic regression analysis identified the relationship among Zone Percent (percent of the Zone estimated by each burn-wound expert, and the colors Red and White with the dependent variable of graft (application of graft or spontaneous healing).
for all three photograph times. The omnibus tests of model coefficients found no statistical significance. The Nagelkerke $R^2$ (.031) indicated that 3.1% of variance is accounted for by the model. No variables demonstrated significance for the Wald test. The ExpB values could not indicate any significant odds ratios. Table 5 represents the logistic regression analysis.

### Table 5.

**Graft Zone Percent and Color Logistic Regression**

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>p value</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1 Zone</td>
<td>.296</td>
<td>.546</td>
<td>.293</td>
<td>.588</td>
<td>1.344</td>
</tr>
<tr>
<td>Zone Percent</td>
<td>.951</td>
<td>1.087</td>
<td>.764</td>
<td>.382</td>
<td>2.587</td>
</tr>
<tr>
<td>Red</td>
<td>.606</td>
<td>1.266</td>
<td>.229</td>
<td>.633</td>
<td>1.832</td>
</tr>
<tr>
<td>White</td>
<td>-.241</td>
<td>.308</td>
<td>.613</td>
<td>.434</td>
<td>.786</td>
</tr>
</tbody>
</table>

**Interclass correlation analysis.** The two-way fixed model for interclass correlation analysis (ICC) is the appropriate model for this dissertation because the burn-wound experts are a fixed effect and the variables of the photograph assessment are the random effects (Hoyt, 2010; Shrout & Fleiss, 1979). With the ICC values for Central Zone at each photograph time, the three burn-wound experts showed that they fairly agreed on Baseline photograph time (ICC = .383), and moderately agreed on Midtrajectory photograph time (ICC = .557). The Final photograph time poorly agreed (ICC = .271). With the ICC values for the Midrange zone showed that the three burn-wound experts poorly agreed on all three time points (Baseline ICC = .019; Midtrajectory ICC = .273; and Final ICC = -.422) (LeBreton & Senter, 2008). Table 6 represents the Zone Percent ICC analysis.

Analysis for the Color Red found that the three burn-wound experts poorly agreed for Baseline photograph time (ICC = .285), and fairly agreed for Midtrajectory photograph time (ICC
= .395), and Final photograph time (ICC = .340). The three burn-wound experts agreed poorly on the Color White at all three photograph times. Table 7 represents the Color ICC analysis.

Table 6.

**Zone Percent Interclass Correlation**

<table>
<thead>
<tr>
<th>Zone</th>
<th>Baseline</th>
<th>Midtrajectory</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central Zone</td>
<td>ICC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICC</td>
<td>.383</td>
<td>.557</td>
<td>.271</td>
</tr>
<tr>
<td>Midrange Zone</td>
<td>ICC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICC</td>
<td>-.019</td>
<td>.273</td>
<td>-.422</td>
</tr>
</tbody>
</table>

Table 7.

**Color Interclass Correlation**

<table>
<thead>
<tr>
<th>Color</th>
<th>Baseline</th>
<th>Midtrajectory</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>.285</td>
<td>.395</td>
<td>.340</td>
</tr>
<tr>
<td>White</td>
<td>.002</td>
<td>.055</td>
<td>-.086</td>
</tr>
</tbody>
</table>
CHAPTER FIVE

Discussion and Implications

This dissertation has three major implications. First, it was the first study that included burn-nurse experts as reviewers in burn-wound assessment. Second, it was also the first to quantify color, a significant factor in burn wound healing, in burn wound assessment. Third, this study used photograph analysis, one of the most frequently used methods to clinically assess burn wounds. This chapter discusses research results and their implications, using research questions as guidelines.

Research Questions

The research questions for this dissertation are:

a. To what extent was the assessment of three burn-wound experts predictive of the outcome earlier decided by the treating physicians (skin grafting or spontaneous healing)?

b. To what extent do three burn-wound experts, using their human vision in the examination of selected burn wound photographs, agree on percent of burn-wound color according to the zone of burn-wound injury (Inter-rater reliability)?

The clinical significance for accurately identifying specific predictors for the decision to apply a skin graft (as discussed in Chapter One) is to apply the skin graft in a timely manner when properly indicated, and to avoid an application when spontaneous healing would give the best result. The findings, discussion and implications of the analysis develop an underpinning for the development of improved burn-wound assessment using machine technology.
Relationships among TBSA, LOS, and Skin Graft

The correlation analysis described the association among the sample of photographs using the variables of total body surface area burned (TBSA), length of hospital stay (LOS), and skin graft application. Findings indicated that more patients received a skin graft (64%) than did not. The mean post-burn-wound day for skin grafting was 10. Although LOS had an upper limit value of 32 days, the decision to perform a skin graft was made relatively early, within 10 days on average.

The decision to perform a skin graft has clinical risks. The clinical implications with the use of a skin graft include: (1) increased pain and discomfort with the surgical procedure of skin removal from the donor site and suturing of skin to the burn wound; (2) infection; (3) non-healing of the burn wound even with the application of a skin graft; and (4) increased cost with a longer burn center stay compared to spontaneous healing. The two-tailed correlation analysis revealed only weak correlations between TBSA, LOS and skin graft. Between TBSA and skin graft the relationship was inverse, i.e., as the TBSA increased, the frequency of skin graft decreased. In this particular sample of burn-wound photographs, the larger TBSA burn-wounds were only partial-thickness burn wounds, which did not require a skin graft. (See Table 4) Because of the weak correlations, although significant, the correlation analysis suggests that other variables impacted the decision to perform a skin graft. Very likely one of those variables is burn-wound depth, assessed by color (Acha et al., 2005).

The correlation analysis also described the association of values when a graft was applied, and the respective weak correlations of the means for Zone Percent, and the Colors Red and White. Since the Midtrajectory photograph time had weak but significant correlations for all three means, this correlational pattern suggests that, retrospectively, burn-wound expert reviewers somewhat agreed at this photograph time (that is midway through the patient’s stay) with the clinical decision-makers for burn-wound depth.
Predictors of Skin Graft

Logistic regression analysis identified the relationship between the independent variables Zone (Central, Midrange, and Peripheral) and Zone Percent estimated by each burn-expert reviewer, as well as the photograph Colors Red and White with the dependent variable of graft (application of graft or spontaneous healing). No statistical significance was identified in the logistic regression analysis, which very likely would be attributed to sample size. Only 3.1% of the variance in the variables was accounted for in the regression line. Thus, contrary to what is found in the literature where burn surgeons use color as the main determinant in assessing the burn wound (Acha et al., 2005), findings here did not identify either the Color Red or White as significant variables. This finding, accompanied by the earlier finding of low correlations between skin grafting and LOS suggests that other factors were part of the decision to graft. This makes sense when we recognize that indeterminate burn wounds present a clinical challenge to assess burn-wound depth and make the decision to apply a skin graft. An increased sample size may improve the level of explained variance.

Interrater Reliability Using HIM

With the ICC values for Central Zone at each photograph time, the three burn-expert reviewers fairly agreed at Baseline photograph time, moderately agreed at Midtrajectory photograph time and poorly agreed at the Final photograph time. With the ICC values for the Midrange Zone, the three burn-wound experts poorly agreed on all Photograph Times. That is, burn-wound experts agreed fairly or moderately only at the Baseline and Midtrajectory Photograph Times for the Central Zone, indicating the difficulty in obtaining agreement for assessing burn-wound depth and to make the decision to apply a skin graft (LeBreton & Senter, 2008) (See Table 6).

Analysis for the Color Red found that the three burn-expert reviewers poorly agreed for Baseline, and fairly agreed for Midtrajectory as well as Final. For the Color White, all three burn-wound experts poorly agreed at all three photograph times. Burn-wound experts found the best
agreement on for the Color Red for Midtrajectory and Final photograph times (See Table 7). (See discussion of the frequency and importance of the Color Red and White.

Although agreement is not found consistently for all photograph times, the Color Red did have poor (Baseline) to fair (Midtrajectory and Final) agreement. Burn-wound experts had some success in agreement on color. The identified agreement suggests that the HIM did assist the burn-wound expert reviewers to quantify burn color (Red) according to time (Baseline, Midtrajectory and Final). Thus, the HIM could be a teaching device to help observers learn where to focus their attention on indeterminate burn-wounds. Moreover, the colors red & white should be explained as to why they are important.

Burn-wound depth assessment would benefit from the agreement, i.e. consistency and reliability associated with machine technology. Research studies that investigate the reliability of an interval measurement of burn-wound color are very few. Acknowledged to be a difficult research design to construct and analyze, an interval level of assessment is essential in the development of a precise measurement instrument for technological instrumentation. Such instrumentation was discussed in Chapter Two.

**Lack of Quantification for Color in Burn-wound Assessment**

This dissertation resonated consistently with previous research studies, in which burn-wound experts demonstrate much difficulty in assessing the color of a burn wound (Bishop, 2004; Hemington-Gorse, 2005; Devgan et al., 2006; Jaskille et al., 2009 & 2010). The literature indicated that burn-wound experts lack a precise assessment of color. The need to address the consistent and reliable assessment of color was not documented in the literature.

Burn-wound experts lack agreement in most of the assessment of burn-wound color even though the assessment of burn-wound color is widely accepted as the most important variable in the assessment of burn-wound depth. However, only one research study, Acha et al. (2005), documented that aspect of assessment using burn surgeons. Agreement on color assessment is based upon two factors. The first factor is the ability to perceive color accurately. Knowledge of
burn-wound pathophysiology and experience in assessing burn wounds educates the burn-wound expert to detect the nuances of color present in the burn wound. The second factor is the ability to link color perception to the progression of burn-wound ischemia that produces additional tissue damage. The conclusion reached on the actual color of the burn wound influences the decision: 1) to apply a skin graft to avoid extending the length of stay; or 2) to avoid the surgical procedure of a skin graft by waiting for spontaneous healing. The indeterminate burn-wound offers the toughest challenge to burn-wound assessment because of its very nature. Daily burn-wound assessment is often inconclusive, and the patient must await additional assessment to come to a conclusion on burn-wound depth.

As indicated previously, individual clinical assessment remains the most widespread method of determining burn-wound depth. With some reports of only a chance percentage (50%) of agreement on burn-wound depth (Bishop, 2004; Devgan et al., 2006; Hemington-Gorse, 2005; Jaskille et al., 2009, 2010), it is not surprising that the burn-wound experts in this dissertation could reliably agree only narrowly on the assessment of burn-wound color, the principle indicator of burn-wound depth. The results of the burn-wound experts in this dissertation corroborate the findings in burn-wound research literature, which indicated that burn-wound experts do not produce a robust assessment for the color of a burn wound.

The specialty of burn-wound care, reflected in the assessment practice of burn-wound experts, lacks a reliable method of assessing burn-wound depth, the critical variable in the decision to graft or not to graft the burn wound. The application of a skin graft may be unnecessarily delayed, as the determination of burn-wound depth is confirmed to require the replacement of tissue using a skin graft. When the LOS is unnecessarily extended, the patient is not enrolled in a timely fashion for out-patient burn-wound care while the burn-wound heals to ultimate closure. The skin graft may be unnecessarily applied, as the determination of burn-wound depth is assessed prematurely and the burn-wound is not permitted to heal spontaneously. This decision then presents with unnecessary surgery, which includes the risks and expenses of:
1) a surgical procedure including anesthesia; 2) donor-site wound pain and additional pain control measures; 3) donor-site wound care; and 4) the probability of a closure result less satisfactory than spontaneous healing (Baker & Mondoizzi, 2011). Findings in this dissertation did not identify a significant broad series of predictors for the outcome of skin graft or spontaneous healing. A future research study would need to include a larger sample size.

Both the variables of Central Zone and the Color Red demonstrated some burn-wound assessment agreement. In this dissertation the HIM provided a very detailed assessment instrument to quantify burn-wound color and area. Burn-wound experts were able to agree on Central Zone and the Color Red. Burn-wound experts demonstrated an agreement on assessment before the Final photograph time, when the burn care team presumably reached the clinical decision to apply a skin graft or allow for spontaneous healing.

The HIM may account for the reliable results for these photograph times. Better quantification of color may better inform decisions about the status of the burn-wound and appropriate treatment. How can better quantification of burn-wound color and thus burn-wound depth be achieved? One possibility to aid Burn-wound experts may be assisted by machine technology that would enhance the ability to perceive the color of a burn wound. Just as machine technology has allowed the assessment of internal organs with radiographic and magnetic imaging, a technically verifiable and accurate machine vision system may reveal a reliable and valid matrix of colors to more precisely assess the burn wound. A machine vision system may perceive color accurately and link color perception to the progression of burn-wound ischemia that produces additional dermal destruction.

**Symptom Management Theory.** This dissertation is not meant to test or examine the Symptom Management Theory (SMT). However, supportive to the theoretical framework of this dissertation, SMT includes the concept of symptom cluster, which may include signs, i.e., physiological abnormalities indicative of disease that is detectable by the individual. The assessment of a burn wound is clinically addressed as a symptom cluster. Important signs in the
burn wound are Color and Zone Percent studied as indicators of burn-wound depth. This dissertation is the first to study the signs of burn wound by quantitatively measuring the assessment of color and zone.

At most, burn wound experts agreed fairly on the color Red as assessed in the photographs, and moderately on Zone Percent. The lack of agreement in assessing these signs of burn-wound depth suggests a need to expand the links between signs and burn-expert reviewer competence to assess burn-wound depth. In the clinical setting, links between signs and symptoms, i.e., pain, debilitation, compromised movement and infection may contribute to increased competence in burn-wound depth assessment. Because of the demonstrated strength in variability, burn-wound experts must develop keener assessment capabilities.

**Benner’s Use of the Dreyfus Model for Skill Acquisition Applied to Nursing**

This dissertation examined the extent to which burn-wound experts agreed in the zone percent and color of burn wounds. Of interest to note is that despite the close working relationship of these burn-wound experts, agreement in interrater reliability was limited. While this is an interesting phenomenon in and of itself, it is not the subject of this interrater reliability analysis. How burn-wound experts learn and how they determine and fix accuracy is a subject of future research.

**Nurse Expertise in Burn-wound Assessment**

This dissertation, which examined the use of nurse-practitioners as burn-wound experts for burn-wound depth assessment is the first study identified in the literature using nurses as burn-wound experts. The only identified research study that included nurses as evaluators investigated burn-wound size, not burn-wound depth (Wachtel et al., 2000). Nurses are essential for the day to day treatment of the burn-wound patient, particularly to assess the burn wound with every burn-wound treatment. Nurses understand that burn-wound depth is a significant determinant of treatment, comfort, and mortality and burn-wound depth assessment is highly valid only for very superficial and for very deep burn wounds. Burn-wound assessment is
markedly less accurate for intermediate burn-wound depth, the toughest challenge to burn-wound assessment, which has been borne out in this study. Burn-wound depth is indeterminate when clinical judgment cannot definitively assess depth as partial-thickness or full-thickness. Nurses must observe the burn wound to identify status, i.e., whether burn-wound vascularity in the dermis is stable, or if ischemic changes in the burn-wound continue to destroy dermis. As discussed in the literature, clinical burn-wound depth of any sort estimate is not accurate in about half to three-fourths of assessments, where the most frequent error is depth overestimation (Bishop, 2004; Devgan et al, 2006; Hemington-Gorse, 2005; Jaskille et al., 2009 & 2010).

In general, nurses are very skilled in holistically managing the signs and symptoms of the patient’s burn wound by assessing the burn-wound depth, eliciting the patient’s related experience, and incorporating relevant information as provided by family and friends. Even though not the focus of this study, it is important to note that other factors affecting burn-wound assessment and wound healing may include patients’ age, race, nutritional status, mobility and psychological disturbances (e.g., anxiety and depression) as well as health professionals’ consensus and spontaneous extraneous circumstances. Therefore, nurses should assess burn-wound depth and healing based on such factors and discuss issues with the patients and families. Thus, nurses consider the strengths and limitations imposed upon patients within the person themselves, their significant others, the environment, and the dimensions of the burn wound experience. Adequately addressing signs and symptoms improve the quality of life and the functional status achieved by patients (Smith & Liehr, 2008).

While these factors are part of the holistic approach that nurses consider in their day to day practice, and are considered as part of the theory development, they fall outside the parameters of this dissertation. Nevertheless, future studies of burn treatment may want to consider these and similar variables as they assess the patient experiences and management of burns.
Important factors in the dependability of burn-wound assessment are the experience of the clinician and the extent of dermal destruction immediately and visually apparent (Baker & Mondozi, 2011). In burn-wound patients, the principle nursing goal is to provide rapid, efficient and effective treatment which assures patient comfort, and the best possible prognosis and outcome. Burn-wound depth is particularly important to the nature and sequencing of timely treatment. If the layer of dermis is insufficient to support healing, skin grafting is required. Like other nursing specialties, burn nursing has its own particular standards of care and conventional understanding of the pathways and decision points to care. Nurses develop a keen skill in assessing and recording the status of the burn wound to detect infection or non-healing of the burn wound.

Burn “wound assessment and care is a learned skill that develops over time”, in which the expert burn nurse teaches “skills to new burn nurses in the… operating room and at the bedside.” Nurses are responsible for teaching other nurses the assessment skills required for burn wounds (Gordon & Marvin, 2002). However, having a reliable, steady, and efficient method of demonstrating color would not only aid in the teaching of others, but also improve burn-wound treatment more rapidly. This is a critical issue that until recently, has had little science or technology to support such important clinical decisions. This dissertation is a step toward the development of a precise science of color on which clinical assessments, prognoses, and treatments can be based.

Without an accurate skill in the assessment of burn color, burn-wound experts attempt to describe and show to each other the important colors, and the important variations of hues and saturations, as evidence that leads to various decision points in the path of treatment. A valid, reliable, steady, and efficient method of demonstrating color would not only aid the teaching of others, but obviously, move a burn treatment in the desired direction more rapidly.
Function of Photograph Analysis

The diagnosis of burn-wound depth based on digital photographs does have reliable agreement with human vision assessment (Devgan et al., 2006). The relationship and efficiency of human vision and machine vision is integral to identifying a more efficient and effective method of assessing a burn wound. If burn-wound color is more accurately assessed, whether by humans or machine, burn-wound depth, and subsequent prognosis and treatment may more effectively be determined. If burn-wound color can be more accurately assessed and developed into a technology the use of this technology in burn-wound treatment has a significant implication in both treatment and the further development of more sophisticated technology.

One of the first steps in building this technology is the development of a reliable matrix that addresses the variations of burn-wound color. Apart from documenting the continued difficulty with the assessment of burn-wound depth using a description of color, the value of this dissertation research is having a documented analysis of a sample of photographs to assess with the CMIS, i.e., the machine vision system. An analysis of these same photographs using the CMIS, without the review of burn-wound experts, would have left an important question unanswered—is it possible that burn-wound experts would have agreed reliably with the clinical assessments of color for these particular photographs? In this dissertation, burn-wound experts only agreed on the variables of assessment (Central Zone and the Color Red). However, in another study, digital images of burn wounds were found to be reliably assessed when compared to the actual physical assessment of the burn-wound (Jones et al, 2003). More reliable assessment is needed particularly in burn wounds of indeterminate depth, the most difficult to assess correctly. Burn-wound depth assessment of obvious partial-thickness and full thickness burn wounds is accurate (Devgan et al., 2006; Heimbach et al., 1984). Improvement in burn-wound depth detection will result in an improved burn-wound treatment and recovery prognosis.

Since burn-wound depth assessment is accurate for the extremes, obvious partial-thickness and full-thickness burn-wounds, those burns identified as indeterminate in depth
provide a distinct challenge to human vision to obtain an accurate burn-wound depth assessment. If burn wounds could be assessed with machine vision, burn-wound expert assessment: 1) could greatly enhance the results of human-vision assessment; and 2) extend the capacity of burn-wound expert assessment with the use of telecommunications often referred to as telemedicine, i.e., the use of telecommunications to provide, enhance or expedite healthcare services by accessing off-site databases, linking clinics or physicians’ offices to central hospitals, or transmitting diagnostic images for examination at another site (American Telemedicine Association, 2010). The photograph of the burn-wound would then become more than an archived image for reference in reviewing the progress of burn-wound healing. Instead, the burn-wound photograph would become an important immediate instrument for differential burn-wound diagnosis and treatment. The decision for burn-wound treatment, particularly of an indeterminate burn-wound, would be enhanced in treatment areas without a burn center. Burn-wound expert assessment could be extended to clinical areas offsite of the burn center. With a documented analysis of burn-wound depth, burn-wound treatment and recovery would be meaningfully improved.

Summary. In summary, this dissertation is a step toward that vision of machine technology, indicating as it does the well meaning but continuing disagreement among burn-wound experts on color in indeterminate burn wounds. Thus, this important documentation of burn–wound expert assessment to assess colors of burn-wounds in photographs is a basis for eventual machine color technology and telemedicine. As a first in-depth study of nurses in burn-wound analysis this study also indicates the need to not only train nurses in burn-wound assessment, but help them reconcile the likely result that indeterminate burn-wound assessment is challenged to predict skin grafting.

Findings in this dissertation indicated that:

1. Weak correlations found between LOS and skin graft, as well as TBSA and skin graft, suggests that other variables determined the need to apply a skin graft or to
allow for spontaneous healing. Small sample size very likely accounts for the lack of findings with logistic regression.

2. The burn-expert reviewers agreed on the assessment of Central Zone and the color Red, which suggests that the HIM was helpful in quantifying color (Red) for the burn-wound according to the area (Zone).

Neither correlation nor logistic regression analysis identified strong significant association in the assessment of color as the measure of burn-wound depth. However, clinical report in the literature matches these findings, because burn-wound experts are reported to have little agreement on the assessment of burn-wound depth. Burn-wound experts agreed that the HIM is a valid instrument, reflecting the expected clinical assessment of a burn wound. Initial reliability analysis for the HIM suggests that this quantitative instrument at minimum calls for additional analysis with other burn-wound experts, and suggests that the assessment of the Color Red is the most reliable of colors for burn-wound assessment.

With a reliable color or matrix of colors for burn-wound assessment, the future use of a machine to gauge burn-wound color will have a first step toward development. The future goal is to devise a machine that is more accurate than burn-wound expert visual assessment alone to increase the accuracy of human vision. With the development of a technically verifiable and accurate machine vision system, a stable and reliable standard for the assessment of burn-wound color can be used for clinical application. Such clinical application would support the differential diagnosis of burn-wound depth, enhancing human vision, and thus defining the best treatment choice for burn wounds. Since burn-wound experts are assessing the color characteristics of the burn wound, a machine vision system which could more reliably describe the color of a burn wound would assist burn-wound expert assessment. As indicated in this dissertation, more reliable assessments are critical, particularly in the program and treatment of the indeterminate burn wound.
GLOSSARY

Angiography, the radiographic visualization of blood vessels following introduction of contrast material used as a diagnostic aid in such conditions as stroke syndrome and myocardial infarction.

Birefringence, The resolution of splitting of a light wave into two unequally reflected or transmitted waves by an optically anisotropic medium such as calcite or quartz. Also called double refraction.

Clinimetrics, the practice of assessing or describing symptoms, signs, and laboratory findings by means of scales, indices, and other quantitative instruments.

Confocal, imaging used in a conventional epifluorescence microscope, where short wavelength light (e.g., blue light when fluorescein is being used as the fluorophore) is reflected by a chromatic reflector through the objective and bathes the whole of the specimen in a fairly uniform illumination.

Dermis, the layer of the skin deep to the epidermis, consisting of a dense bed of vascular connective tissue; it is divided into a papillary layer and a reticular layer.

Edema, the presence of abnormally large amounts of fluid in the intercellular tissue spaces of the body, usually referring to subcutaneous tissues.

Epidermis, the outermost, nonvascular layer of the skin.

Eschar, slough produced by a thermal burn, by a corrosive application, or by gangrene.

Fluid resuscitation, fluid replacement strategies to forestall hypovolemic shock in a burn patient.

Ischemia, deficiency of blood in a part, usually due to functional constriction or actual obstruction of a blood vessel.

Necrosis, the sum of the morphological changes indicative of cell death and caused by the progressive degradative action of enzymes; it may affect groups of cells or part of a structure or an organ.

Slough, 1. necrotic tissue in the process of separating from viable portions of the body. 2. to shed or cast off.

Thrombus, a stationary blood clot along the wall of a blood vessel, frequently causing vascular obstruction.

Vital dye, one that penetrates living cells and colors certain structures without serious injury to the cells.
§Zones of burn wound, 1. zone of hyperemia, the outer periphery of the burn wound containing viable cells and vasodilatation influenced by local inflammatory mediators. 2. zone of stasis, containing both viable and nonviable cells, capillary vasoconstriction and ischemia; the zone at risk for conversion to necrosis with hypoperfusion, desiccation, edema and infection. 3. zone of coagulation, where no viable cells are present.


http://education.yahoo.com/reference/dictionary/entry/birefringence


http://www.loci.wisc.edu/optical-sectioning/confocal-imaging


REFERENCES


APPENDICES A, B, C, and D
Appendix A

Figure 1. Procedure for the Decision Flow to Apply or Not Apply a Skin Graft.

1. **Burn-Wound Color indicates injured tissue**
2. **Burn-Wound Experts assess burn-wound depth**
   - **Burn Wound will heal spontaneously**
     - **No Graft**
   - **Burn Wound will not heal spontaneously**
     - **Graft**
3. **The CMIS, a machine vision system to assess burn-wound color**
Appendix B

Lund and Browder Charts for Estimation of Burn-Wound Size

<table>
<thead>
<tr>
<th>AREA</th>
<th>inf.</th>
<th>1-4</th>
<th>5-9</th>
<th>10-14</th>
<th>15</th>
<th>adult</th>
<th>part.</th>
<th>full</th>
<th>total</th>
<th>Donor areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>19</td>
<td>17</td>
<td>13</td>
<td>11</td>
<td>9</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neck</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ant. Trunk</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post Trunk</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R Buttock</td>
<td>2½</td>
<td>2½</td>
<td>2½</td>
<td>2½</td>
<td>2½</td>
<td>2½</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Buttock</td>
<td>2½</td>
<td>2½</td>
<td>2½</td>
<td>2½</td>
<td>2½</td>
<td>2½</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genitalia</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R U Arm</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L U Arm</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R L Arm</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L L Arm</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R I Hand</td>
<td>2½</td>
<td>2½</td>
<td>2½</td>
<td>2½</td>
<td>2½</td>
<td>2½</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Hand</td>
<td>2½</td>
<td>2½</td>
<td>2½</td>
<td>2½</td>
<td>2½</td>
<td>2½</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R Thigh</td>
<td>5½</td>
<td>6½</td>
<td>8</td>
<td>8½</td>
<td>9</td>
<td>9½</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Thigh</td>
<td>5½</td>
<td>6½</td>
<td>8</td>
<td>8½</td>
<td>9</td>
<td>9½</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R Leg</td>
<td>5</td>
<td>5</td>
<td>5½</td>
<td>6</td>
<td>6½</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Leg</td>
<td>5</td>
<td>5</td>
<td>5½</td>
<td>6</td>
<td>6½</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R Foot</td>
<td>3½</td>
<td>3½</td>
<td>3½</td>
<td>3½</td>
<td>3½</td>
<td>3½</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Foot</td>
<td>3½</td>
<td>3½</td>
<td>3½</td>
<td>3½</td>
<td>3½</td>
<td>3½</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TOTAL
1. Contaminated or infected at entry
2. Time elapsed until treated
3. Pain
   a. Amount at entry
   b. How much relieved by local treatment
   c. Amount of pain during hospital course
4. Date when ready for grafting
5. Date when healed
6. Size of area unhealed at discharge (also indicate on chart)
7. Date of discharge
8. Contracture - type of healed lesion
Appendix C

Human Intelligence Matrix

Human Intelligence Matrix for Burn Wound Color

To score, indicate percentage of attribute (0-100%) according to area-zone.

<table>
<thead>
<tr>
<th>Hue</th>
<th>Area/Zone</th>
<th>Central % of burn</th>
<th>Midrange % of burn</th>
<th>Peripheral % of burn</th>
<th>Non-burn Skin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yellow</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Green</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brown</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Definition for scoring attribute of color.

Hue is the color type, i.e., red, blue or yellow.
Appendix D

Consent for Photographs

MEDICAL AUDIO VISUAL DEPARTMENT REQUISITION

REQUESTING PHYSICIAN OR DEPARTMENT HEAD

PATIENT:  
NAME:  
AGE:  
SEX:  
ROOM NO.:  
DATE:  

ANTERIOR  
POSTERIOR  
RIGHT LATERAL  
LEFT LATERAL

ANTERIOR  
POSTERIOR  
RIGHT LATERAL  
LEFT LATERAL

DORSAL  
PALMER  
DORSAL  
PLANTAR

DORSAL  
PALMER  
DORSAL  
PLANTAR

CIRCLE AND DESCRIBE EACH AREA TO BE PHOTOGRAPHED

OD  
OS

NUMBER OF SLIDES TAKEN:

ORIGINAL SLIDES TO:
1. ❑ AV SLIDE FILES
2. 
3. 
4. 
5. 
6. 

DUPLICATE SLIDES TO:
1. 
2. 
3. 
4. 
5. 
6. 

PRINTS TO:
1. 
2. 
3. 
4. 
5. 
6. 

PHOTOGRAPHIC RELEASE

I WE CONSENT TO THE TAKING AND PUBLICATION OF ANY PHOTOGRAPHS, MOTION PICTURES, AND VISUAL AIDS FOR THE PURPOSE OF ADVANCING MEDICAL EDUCATION AND/OR FOR HOSPITAL PUBLICITY PURPOSES.

(PARENT / GUARDIAN / SELF)  
(DATE)

(WITNESS)  
(DATE)