ILKKA KAARTINEN

Assessment of Skin Scars in Clinical Practice and Scientific Studies

Introducing Spectrocutometry,
a new objective method

■

UNIVERSITY OF TAMPERE

To my family

When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind: it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science.

Sir William Thomson (Lord Kelvin),
Popular Lectures and Addresses 1891–1894

Table of Contents

6

Abbreviations

ARD	Automatic relevancy detection
BVD	Blood vessel density
CCLS	Computer controlled lighting system
CI	Confidence interval
DTM	Dermal torque meter
ECC	Estimated concentration change
EGF	Epidermal growth factor
EAI	Equal appearing interval
Hb	Haemoglobin
ICC	Intraclass correlation coefficient
IL-1	Interleukin 1
RS	Reflectance spectrophotometry
LED	Light emitting diode
LSCM	Laser scanning confocal microscopy
MSS	Manchester scar scale
MTG	Mepilex Transfer® Group
PDGF	Platelet derived growth factor
POSAS	Patient and Observer Scar Assessment Scale
RCM	Reflectance confocal microscopy
SDI	Standardized digital imaging
SG	Suprathel® group
SGDS	Split-thickness skin graft donor site
SIAscopy	Spectrophotometric intracutaneous analysis
SpM	Spectral modelling
STSG	Split thickness skin graft
TGFB	Transforming growth factor β
US	Ultrasonography
UV	Ultraviolet
VSS	Vancouver Burn Scar Scale
VAS	Visual analogue scale

VEGF   Vascular endothelial growth factor

List of original communications

The present study is based on the following articles, which have been referred to in the text by their Roman numerals (I-IV)

I. Kaartinen IS, Kuokkanen H. Suprathel causes less bleeding and scarring than Mepilex Transfer in the treatment of donor sites of split thickness skin grafts.
J Plast Surg Hand Surg. In press.

II. Välisuo P, Kaartinen I, Kuokkanen H, Alander J. The colour of blood in skin: a comparison of Allen's test and photonics simulations. Skin Res Technol. 2010 Nov;16(4):390-6.*

III. Kaartinen I, Välisuo P, Alander J, Kuokkanen H. Objective scar assessment – a new method using standardized digital imaging and spectral modelling. Burns. 2011 Feb;37(1):74-81.

IV. Kaartinen I, Välisuo P, Botchko V, Alander J, Kuokkanen H. How to assess scar hypertrophy? A comparison of subjective scales and Spectrocutometry: a new objective method. Wound Repair and Regeneration. 2011, 19(3): 316-23.*

The original publications are reproduced with permission of the copyright owners.
* These articles are also included in the dissertation of MSc Petri Välisuo in the University of Vaasa.

1. Abstract

Scars are formed after surgical procedures and different types of trauma, especially burn injuries. Some scars heal quickly, while others develop into hypertrophic scars or keloids. These problematic scars often need to be treated for months or even years. While there have been new innovations in scar treatment in recent years, effective treatment options are lacking. There is no scientific evidence of the efficacy of the most common scar treatment and prevention methods.

Scar assessment is needed in clinical practice in centres that treat burn victims and other patients with problematic scars. To date, however, it is still uncommon, and existing methods for assessing scars are unknown to many. In scientific scar studies, subjective assessment with scar rating scales, the Vancouver Burn Scar Assessment Scale (VSS) in particular, has been most commonly used. Although it has been reported that subjective assessment is unreliable, it has been considered the gold standard in scar assessment.

Reflectance spectrophotometry has been used in skin colour measurement. Digital imaging has been popularized in dermatology and wound studies. We developed a new method for assessing scar hypertrophy, namely Spectrocutometry, which combines standardized digital imaging, reflectance spectrophotometry and spectral modelling, in addition to producing objective estimations of the concentrations of haemoglobin (Hb) and melanin in skin scars.

In this study, 22 split-thickness skin graft (STSG) donor sites were treated randomly with two dressing options, Suprathel® and Mepilex Transfer®. The pain, bleeding and scarring of the donor sites were recorded and the donor sites were assessed using the Spectrocutometry method. In an experimental study, 100 Allen's tests were performed on 20 healthy subjects, and the skin reflectance spectrum was measured form the skin of the palm. The measured reflectance values were matched with data from a Monte Carlo simulation study to examine the relationship between skin Hb concentration and skin colour. Using the information gained from the Allen's test study, Spectrocutometry was employed to calculate the estimated concentration

change (ECC) of melanin and Hb from the STSG scars, and the values were compared with subjective rating by three observers. Finally, the Spectrocutometry method was tested by measuring 37 scars after melanoma surgery by three observers. The reliability and validity of the Spectrocutometry method was compared with conventional scar assessment using the VSS as well as the Patient and Observer Scar Assessmnet Scale (POSAS).

The STSG donor site model was found feasible in dermal wound studies. Donor site treatment with Suprathel® was associated with significantly less pain, bleeding and scarring when compared to Mepilex Transfer® treatment.

The changes measured during the Allen's test matched those produced by the Monte Carlo simulation. Skin colour was found to be nonlinearly related to skin blood concentration.

The Spectrocutometry method was found to be more sensitive in detecting differences in STSG donor site scars than subjective rating. There was a statistically significant correlation between the ECC of Hb and melanin and the POSAS vascularity and pigmentation subscales (r=0.63 and 0.60, respectively, $p<0.001$). The Intraclass correlation coefficient for the VSS and POSAS scales was lower than acceptable in STSG donor site scar assessment (r=0.26–0.40 and 0.51–0.69, respectively, $p<0.05$).

In melanoma scar assessment, the reliability of Spectrocutometry was high (r= 0.88–0.89, $p<0.005$). The ICC for VSS and POSAS was below acceptable limits (r=0.66 and 0.60, respectively, $p<0.005$). A strong dependency and a positive correlation was found between the scar Hb concentration and pain in Bayesian network analysis.

As a conclusion, objective information on haemoglobin and melanin concentrations in skin scars can be obtained non-invasively using the Spectrocutometry method. This method can be used both in clinical practice and in scientific studies. Many options exist that are proposed to be efficient in the treatment of scar hypertrophy, but evidence of even the most widely used options is weak. It is necessary that these treatments undergo a critical evaluation with reliable objective methods. Future innovations in scar prevention and therapy should be critically analysed and their cost-effectiveness measured before adopting them into wider use, especially considering the increasing costs of public health services.

Spectrocutometry is an accurate and highly reliable method for scar assessment studies and clinical practice. Because it enables the differentiation and measurement of the changes in oxy- and deoxy-haemoglobin, it may prove a useful tool in wound studies as well.

2. Tiivistelmä

Ihon arvet syntyvät yleensä kirurgisten toimenpiteiden tai erilaisten vammojen, erityisesti palovammojen, seurauksena. Arpikudoksen kasvu on yksilöllistä. Usein arvenmuodostus on vähäistä, mutta ääritapauksissa se voi johtaa patologiseen arpeutumiseen eli ns. hypertrofisen arven tai keloidin muodostumiseen. Näiden arpien hoitaminen on vaikeaa ja pitkäkestoista, eikä yksittäisiä tehokkaita hoitomuotoja ole toistaiseksi olemassa. Lisäksi tieteellinen näyttö käytössä olevien, yleisesti hyväksyttyjen hoitojen tehosta on puutteellista.

Arpien systemaattiseen arviointiin on kehitetty useita menetelmiä, joiden käyttöä suositellaan palovamma- ja muita ongelma-arpia hoitavilla klinikoilla. Olemassa olevien menetelmien käyttö on kuitenkin vielä harvinaista. Tieteellisissä arpitutkimuksissa on yleensä suosittu subjektiivisia mittareita, joista tunnetuin ja toistaiseksi eniten käytetty on Vancouver Burn Scar Assessment Scale (VSS). Vaikka VSS:n ja muiden vastaavien asteikkojen luotettavuus on kyseenalaistettu useissa tutkimuksissa, on subjektiivista arviointia edelleen pidetty parhaana vaihtoehtona.

Reflektanssispektrofotometria (RS) on menetelmä, jota on sovellettu ihon ja myös arpien sekä muiden ihomuutosten värin kvantitatiivisessa mittaamisessa. Digitaalinen kuvantaminen on yleistynyt etenkin haavatutkimuksissa ja dermatologian alalla. Spektrocutometria on uusi menetelmä, jossa yhdistyy reflektanssispektrofotometria, standardoitu digitaalinen kuvantaminen ja spektraalinen mallinnus. Menetelmä mahdollistaa ihon ja ihomuutosten, kuten arpien, hemoglobiini- ja melaniinipitoisuuden arvioimisen.

Ensimmäisessä osatutkimuksessa hoidettiin 22:a osaihonsiirteen ottokohdan haavaa käyttäen rinnakkain satunnaistetusti kahta erilaista hoitotuotetta, Suprathelia® ja Mepilex Transferia®. Ottokohdan kipua, vuotoa ja arpeutumista mitattiin kliinisesti, ja lisäksi ottokohdan arvet tutkittiin Spectrocutometrian avulla.

Toisessa tutkimuksessa suoritettiin 100 Allenin testiä 20 terveelle koehenkilölle, ja testin aikana mitattiin ihon reflektanssispektri kämmeneltä. Mitatut reflektanssiarvot

sovitettiin Monte Carlo -simulaatiosta saatuun, ihon väriä eri veripitoisuuksilla mallintavaan aineistoon. Mallinnuksesta saadun tiedon perusteella laskettiin ihonsiirteen ottokohtien arven hemoglobiini- ja melaniinipitoisuuden muutos normaaliin ihoon verrattuna. Laskettuja arvoja verrattiin kolmen arvioitsijan tekemiin subjektiivisiin arvioihin. Lopuksi Spectrocutometrian avulla arvioitiin 37 melanoomaleikkauksen jälkeistä arpea 20 potilaalla. Mittauksen ja arpien subjektiivisen arvioinnin suoritti 3 itsenäistä arvioitsijaa. Spectrocutometrian luotettavuutta ja validiteettia mitattiin vertaamalla tuloksia subjektiivisen arvion tuloksiin.

Ihonsiirteen ottokohdan käyttö dermaalisen defektin mallina osoittautui toimivaksi. Mepilex Transferiin® verrattuna Suprathel® vähensi tilastollisesti merkitsevästi ottokohdan kipua, vuotoa ja arpeutumista.

Allenin testissä mitatut spektrin muutokset vastasivat simulaatiomallin tuloksia. Ihon värin osoitettiin riippuvan epälineaarisesti ihon veripitoisuudesta. Ihonsiirteen ottokohdan arpia tutkittaessa Spektrocutometria osoitti herkemmin ihon arpeutumisen muutoksia. Hemoglobiinin ja melaniinin pitoisuuden muutos korreloi tilastollisesti merkitsevästi POSAS- ja VSS-asteikkojen vaskulaarisuus- ja pigmentaatioarvioiden kanssa (r=0,63 ja 0,60, p<0,001). VSS- ja POSAS-mittareiden luotettavuus oli hyväksyttävää tasoa alempi (ICC=0,26–0,40 ja 0,51–0,69, p<0,005).

Spectrocutometria osoittautui erittäin luotettavaksi menetelmäksi melanooma-arpia mitattaessa (r=0,88–0,89, p<0,005). VSS- ja POSAS-asteikkojen luotettavuus oli hyväksyttävää tasoa heikompi (r=0,66 ja 0,60, p<0,005). Bayesian network -analyysi osoitti, että arven aiheuttama kipu ja hemoglobiinipitoisuus ovat vahvasti riippuvaisia toisistaan.

Lopputuloksena voidaan sanoa, että Spectrocutometrian avulla voidaan saada arven aktiivisuudesta objektiivista tietoa, jota voidaan käyttää sekä kliinisessä työssä että tieteellisissä tutkimuksissa. On tärkeää, että arpien hoitoon käytettävät ja kehitteillä olevat hoitomuodot arvioidaan tieteellisesti luotettavia objektiivisia mittareita käyttäen. Terveydenhuollon kulujen kannalta myös hoitojen kustannustehokkuus tulisi tutkia.

Spectrocutometria on tarkka ja erittäin luotettava arven aktiivisuuden mittausmenetelmä. Koska sen avulla voidaan mitata erikseen hapettuneen

(oxyhemoglobiini) ja hapettoman hemoglobiinin (deoxyhemoglobiini) pitoisuus, se saattaa osoittautua käyttökelpoiseksi myös haavatutkimuksessa.

3. Introduction

Even the smallest surgical operation leaves a scar. Some scars heal well, leading to an unnoticeable fine line, whereas others are prone to develop scar hypertrophy or, in the worst case, transform into keloids. Wounds closed with excess tension are known to produce more scar tissue (Ogawa et al. 2011). In some cases, tension-free closure cannot be achieved, with skin cancer surgery as an example. Scars produced by trauma, especially burn injuries, are more prone to hypertrophy (English and Shenefelt 1999). This poses a difficult problem for the rehabilitation of burn victims (Serghiou et al. 2009).

Hypertrophic scars (HSC) and keloids are often associated with symptoms such as pain and itching (Wolfram et al. 2009). They can create significant functional problems, such as restricted movement, especially when located near joints (Ladak and Tredget 2009). Untreated HSCs develop contractures that can be debilitating, a fact unfortunately often seen in developing countries (Prasanna et al. 2004).

Scars have a significant impact on patients' psychological welfare (Truong et al. 2005). Besides the somatic problems, HSCs affect especially the acceptability to one's self and others as well as social functioning and emotional well-being (Brown et al. 2008a).

The ongoing studies concerning scars have focused on producing treatments to reduce or even prevent scar formation (Occleston et al. 2008). Based on the scar-free foetal wound healing, the newest innovation, transforming growth-factor (TGF) B3, has shown potential in reducing scar formation (Ferguson et al. 2009).

Skin graft donor site wounds pose a problem especially in burn victims. Research aims at finding an optimal donor site treatment that would alleviate pain and reduce scar formation. The skin graft donor site is a wound of standard depth and offers a unique possibility to study dermal wound healing and skin scarring *in vivo*.

There are several ways to treat and prevent scar hypertrophy, but none have been shown to be effective alone (Mustoe et al. 2002). Instead, scars are treated with a combination of methods, including pressure garment therapy, silicone gel sheeting,

cortisone injections and physiotherapy (Ogawa , Karagoz et al. 2009). The cost of these treatments is considerable (Aarabi et al. 2007). Still, there is no evidence of the effect of the most widely used treatments – silicone gel sheeting has been used for over 20 years but has not been proved to affect the outcome of HSCs (O'Brien and Pandit 2006). Pressure garments are thought to relieve scar hypertrophy or at least alleviate the symptoms. However, recent studies, especially a meta-analysis by Anzarut et al., have questioned the need for pressure garment therapy, which yields considerable costs especially in burn centres (Anzarut et al. 2009, Harte et al. 2009). One obvious reason for the poor quality of scar studies is the lack of objective, reliable methods for scar assessment.

There has been a lot of development in the area of subjective scar assessment, and many scales have been introduced (Brusselaers et al. 2010). Established in 1990, the Vancouver Burn Scar Assessment Scale, usually referred to as the Vancouver Scar Scale (VSS), remains the most well-known and widely used scar rating scale, originally designed for rating burn scars (Sullivan et al. 1990, Baryza and Baryza 1995). The Patient and Observer Scar Assessment Scale (POSAS) is a more recent scale that has been suggested especially in the assessment of linear scars (Truong et al. 2005).

Despite the evolution of scar assessment scales, the interrater reliability of subjective colour evaluation has been shown to be poor, and several observers are usually suggested for reliable assessment. In the clinical scenario, the scar is usually assessed by a single observer and often by many different individuals over time, further weakening the accuracy of subjective rating as a tool in clinical practice. The changes in scar dimensions, such as colour, volume and surface roughness, are unevenly distributed, and different colour changes overlap. These obstacles are not easily overcome by the human observer.

As an objective method, spectrophotometry has been used for over 50 years for skin colour measurement (Yun et al. 2010). It has shown good results in the assessment of different skin lesions, both pigmented and vascular (Takiwaki et al. 2002a, Takiwaki et al. 2004). The scar colour has been measured with both a tri-stimulus colorimeter (Minolta Chromameter) and a narrow-band reflectance meter (Dermaspectrometer; (Draaijers et al. 2004b).

Digital images have been used for measuring wound size, blood circulation and wound healing kinetics (Hansen et al. 1996, Bon et al. 2000, Payette et al. 2005, Molnar et al. 2009).

In the present study, a model for dermal wound healing using skin graft donor sites was designed. This model was used in a clinical comparative study, where the postoperative healing of the donor sites using two optional dressings was evaluated to assess the effect of the dressings on donor site pain, bleeding, exudation and scarring. Secondly, a new method combining the standardized digital imaging (SDI) technique with the computer controlled lighting system (CCLS) and reflectance spectrophotometry (RS) for quantitative scar colour assessment is described. During the development of the new method, which was named Spectrocutomery by the author of this dissertation, the relationship between the skin's reflectance spectrum and its haemoglobin and melanin concentrations was examined. The Spectrocutometry method has been used in the assessment of skin graft donor site scars and linear scars after melanoma surgery. Finally, the reliability of the proposed Spectrocutometry method was tested and compared with conventional subjective rating by means of the VSS and POSAS scales.

4. Review of the literature

4.1 Normal wound healing

The wound healing process is traditionally divided into three phases – inflammation, proliferation and remodelling. This division is arbitrary, and the phases overlap. The first reaction to skin trauma is haemostasis, followed by the inflammation phase. This phase lasts for approximately 4–6 days. Neutrophils are first attracted to the wound, and monocytes follow 48–96 hours later. This response is stimulated by interleukin-1 (IL-1), tumour necrosis factor TNF-α and transforming growth factor TGF-β. Angiogenesis and fibroplasia are mediated by macrophages through the vascular endothelial growth factor (VEGF), fibroblast growth factor, TNF- α, epidermal growth factor (EGF), platelet derived growth factor (PDGF) and IL-1. (Broughton et al. 2006)

The second phase, proliferation, lasts from 4 to 14 days, during which time the scar is formed (Shih and Bayat 2010). Fibroblasts migrate into the wound site and start producing collagen. Fibroblasts differentiate into myofibroblasts that are responsible of wound contracture (Shin and Minn 2004). This process is regulated by many different signals, mainly by PDGF, EGF and TGFβ (Broughton et al. 2006).

The remodelling phase starts roughly at day 8 after the injury and lasts for over a year. During this time, the initially formed excess collagen is digested and the type III collagen is replaced with type I collagen. The wound breaking strength increases over the first 3 months to approximately 80% of the original strength (Janis et al. 2010).

4.2 Healing and treatment of the split-thickness skin graft donor site

A split-thickness skin graft (STSG) includes the epidermis and the superficial dermis. Typically the graft is 0.30–0.45 mm thick, and the donor site is left to heal spontaneously. Healing occurs when the keratinocytes deep in the skin appendages, mainly hair follicles and sweat glands, migrate and cover the wound bed. This is usually completed in roughly 12 days, depending on the thickness of the graft. (Paletta 2006). The donor site healing can be delayed by an infection, or the wound can deepen into a full-thickness defect (Monafo 1996). STSG donor site treatment aims at reducing the risk of infection, facilitating epithelialization and alleviating pain (Barnea et al. 2004). Treatment options can be divided into open, semi-open and closed methods. Semi-open and closed methods have the benefit of keeping a moist environment, which enables the keratinocyte migration. They have been found superior to open ones in several studies (Kilinc et al. 2001, Schwarze et al. 2007).

4.3 Scarless foetal healing

The ability of a human foetus to heal wounds without scarring was described by Rowlatt as early as in 1979 (Rowlatt 1979). The notion has later been confirmed in animal and human studies (Larson et al. 2010). In recent years, efforts have been made to further understand the difference between foetal and adult wound healing in order to transform this knowledge into scar prevention measures.

In human foetuses, scarless wound healing generally occurs until approximately the 24th gestational week, but larger wounds may also cause scarring earlier in the pregnancy (Lorenz and Adzick 1993). This means that the foetal skin is able to produce a collagen matrix that is identical to the original one (Longaker et al. 1990). When compared to the postgestational wound healing process, foetal wound healing has several important characteristics. The rates of the signalling molecules differ in

that there is more TGFβ3 as opposed to TGFβ1 and TGFβ2. Especially increased levels of TGFβ1 have been shown to have a paramount effect on scar formation (Krummel et al. 1988). Furthermore, there are more matrix metalloproteinases than inhibitors of matrix metalloproteinases, which diminishes the accumulation of collagen. (Namazi et al. 2011) The inflammatory phase in foetal wound healing is significantly shorter than in postgestational wounds, indicating an important role of inflammation in scarring (Larson et al. 2010). This is also demonstrated in postgestational animal wounds where the absence of inflammation led to reduced scarring (Ashcroft et al. 1999).

There is more type III collagen in foetal wounds, whereas collagen type I is more present in postnatal wounds (Merkel et al. 1988). There is also more hyaluronic acid in foetal skin than in postnatal skin (Longaker et al. 1989).

Finally, foetal fibroblasts synthesize more type III and IV collagen and have more receptors for hyaluronic acid than postnatal fibroblasts (Larson et al. 2010). There is also an absence of myofibroblasts, which are considered essential in hypertrophic scar formation, in foetal wounds (Estes et al. 1994).

4.4 Hypertrophic and keloid scarring

4.4.1 Clinical features of pathologic scars

A hypertrophic scar (HSC) is defined as a skin scar that rises above the skin level but stays inside the borders of the original lesion (Kose and Waseem 2008). A keloid scar, on the other hand, is defined as a scar that grows outside the margins of the original wound (Peacock et al. 1970). Although these two types of pathologic skin scars are often confused, they are fundamentally different in their pathogenesis and natural history (Muir 1990). HSCs usually develop during the first three months after the initial injury or operation, whereas keloid scars appear within or over 12 months after the injury and may reappear even several years after treatment

(Bloemen et al. 2009). Although there are fundamental differences between HSCs and keloids, both result in increased fibroblasts and extracellular matrix formation.

Mustoe et al. have divided scars into six categories: mature scars, immature scars (active scars that mature over time), linear HSCs (e.g., surgical or traumatic, Figure 4), widespread HSCs (e.g., burn scars, Figure 3), minor keloids (e.g., earlobe keloids after piercing, Figure 2) and major keloids (Figure 4; (Mustoe et al. 2002). The first four of these scar types represent the spectrum of a normal physiological wound healing process, and the last two constitute a different pathological type of scarring with distinct clinical and histopathological features.

HSCs tend to form during the first three months after the injury. After the formation of the scar, they then undergo a process of maturation, which usually lasts for 12–24 months. This process can be delayed if there is tension or other irritation in the scar tissue. While active, HSCs are often disfiguring and exhibit symptoms such as pain and itching (English and Shenefelt 1999). They can restrict movement when located near joints. Different therapies against HSCs aim at relieving these symptoms and hastening the maturation process. It is thought that these therapies might be able to prevent or at least diminish scar hypertrophy (Ogawa 2010). However, there is still little evidence to back these arguments, and it remains unclear how much can actually be achieved to treat or prevent hypertrophic scarring (O'Brien and Pandit 2006, Anzarut et al. 2009).

Keloid scars are a difficult clinical entity. There is no single effective treatment against keloids, and a combination of therapies is usually commenced, including cortisone injections, pressure garment therapy, silicone gel sheeting and, in severe and recurrent cases, injections of 5-fluorouracil and irradiation therapy have been suggested (Ogawa , Ogawa 2010). If simple excision is attempted without other treatment modalities, the recurrence rate is high, and in some cases the situation can be worsened with surgery.

4.4.2　The process of scar formation

Normal scarring occurs after any type of injury to the skin, with the exception of the most superficial scratches. There is a critical depth in a skin wound after which scar formation begins, and wounds that are more superficial have the potential to heal without scarring. This depth is shown to be approximately one third of the total skin thickness in normal thigh skin (Dunkin et al. 2007). This is demonstrated in dermal burn injuries, where superficial dermal burns heal without scarring, but deeper dermal burns often develop markedly hypertrophic scars, especially when treated conservatively (Dedovic et al. 1999).

When a linear surgical wound is healing, there is usually little need for the scar to contract. In a larger planar wound, such as a burn wound, contraction occurs as a physiological response to decrease the wound surface. In a HSC, myofibroblasts often persist long after wound closure as a result of tension, among other possible causes. When the tension is released, the rate of myofibroblasts decreases rapidly (Junker et al. 2008).

4.4.3　Histopathological features of hypertrophic scars and keloids

Where HSCs inevitably undergo maturation at some point, keloid scars usually grow continually and spread outside the confines of the original wound. One suggestion is that there is an inability to cease the proliferation phase of wound healing (Butler et al. 2008). There is an increased amount of fibroblasts but also large, randomly arranged collagen fibres and proteoglycans in the scar (Figure 1). They have four histological features that are characteristic in keloid scars: keloidal hyalinized collagen, an advancing edge underneath the epidermis and papillary dermis, horizontal fibrous bands in the upper reticular dermis and prominent fascia-like fibrous bands (Lee et al. 2004). Fibroblast density is increased in both keloids and HSCs, but keloids also exhibit increased fibroblast proliferation rates (Al-Attar et al. 2005).

Another distinct feature of HSCs and keloids is the increased blood vessel density when compared to normal skin and normal scars, which is clinically noticed as increased redness (Ehrlich et al. 1994, Li-Tsang et al. 2005). This reflects the hypermetabolic state of active scars. It has been shown that fibroblasts in keloids and HSCs secrete markedly higher levels of VEGF and TGFB-1 than fibroblasts in normal skin and scars reflecting the increased angiogenesis (Beer et al. 1998, Fujiwara et al. 2005). When keloids and HSCs are compared, there is contradictory evidence about the number and distribution of capillaries. Kurokawa et al. and Beer et al. found fewer capillaries in keloids and observed that they appeared flatter in comparison to HSCs, suggesting that tissue hypoxia would be partially responsible for the formation of keloid scars (Beer et al. 1998, Kurokawa et al. 2010). However, in the study by Amadeu et al., there was no difference in blood vessel density (BVD) and the distribution of capillaries between the two scar types (Amadeu et al. 2003). This contradiction is best explained by the maturation process of keloids, during which the amount of capillaries diminishes, reflecting the inactivation of the scar.
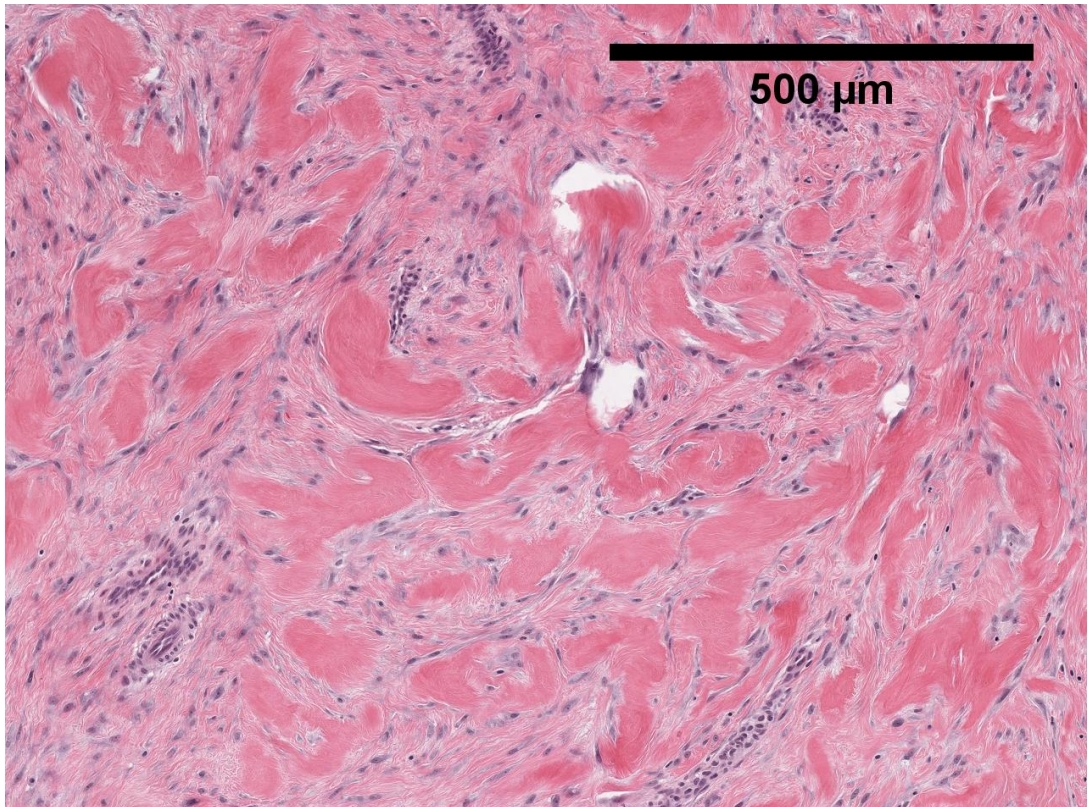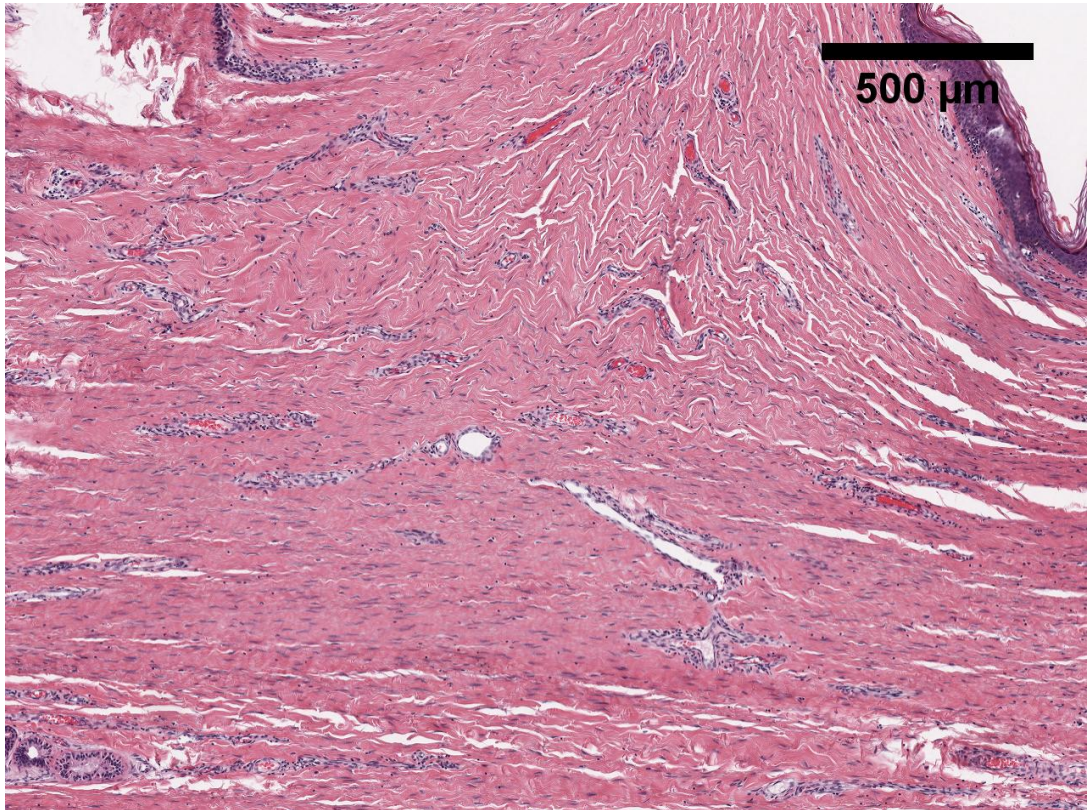
Figure 1. Hematoxylin and eosin staining of a normal scar (above) and a keloid scar (below) from an earlobe keloid.

4.4.4 Genetics of keloid scarring

There are several theories about the pathway of keloid scarring, but the nature of this process is still largely unknown. Current evidence suggests that there is a genetic susceptibility to keloid disease (Brown et al. 2010). There is a predominance of keloid disease in black and hispanic populations as opposed to other ethnic backgrounds (Muir 1990, Bayat et al. 2003a, Kose and Waseem 2008). Patients with a family history of keloid disease have more severe forms of keloids and keloids in multiple sites. There are prevalent anatomical sites in different families with a known tendency to form keloids. (Shih and Bayat 2010) The typical anatomical sites where keloids are formed include the anterior chest, shoulder, scapular and suprapubic regions, which are also subjected to constant mechanical forces due to skin stretching during daily activities (Ogawa 2011). These anatomical sites also have a high sebaceous gland density (Al-Attar et al. 2005). No single gene has been found that would be solely responsible for keloid formation; rather, a pattern where the involvement of different mutations on the same gene or different sets of genes are suspected to be responsible for the process, based on the current clinical evidence. The inheritance pattern is thought to be predominantly autosomal dominant. To date, over 30 different genes have been proposed as possible sources. (Bayat et al. 2004, Brown et al. 2008b, Brown et al. 2008c, Brown et al. 2010, Shih and Bayat 2010)

4.4.5 Incidence and risk factors of hypertrophic scars and keloids

In the developed world, where information on surgical operations is recorded, it is estimated that over 55 million elective surgical operations and 25 million operations due to different types of trauma are performed each year (Bayat et al. 2003b). More than 4 million people acquire scars because of burn injuries each year in the developed world alone, and the number is estimated to be much higher in the

developing world (Bloemen et al. 2009). The rate of hypertrophic scarring in these different situations is unknown. There are contradictory reports of HSCs in different populations. Scar hypertrophy is estimated to occur in 5%–70% of surgical scars after elective surgery and in 30%–67% of burn wounds, with a higher incidence in children (Muir 1990, Dedovic et al. 1999, Mustoe et al. 2002, Bombaro et al. 2003, Li-Tsang et al. 2005). Ethnic background is thought to influence the susceptibility to scarring, with individuals with darker skin more prone to develop HSCs.

Keloids seem to occur less frequently than HSCs. In the year 2000, it was estimated that 11 million people were affected by keloid disease in the developed world alone (Bayat et al. 2003b). Keloid scars tend to occur more during childhood and adolescence and are less frequently seen in older patients (Atiyeh et al. 2005). As opposed to normal skin scarring and HSCs, which appears after sufficiently deep trauma to the skin, keloids are reported to appear after any degree of trauma, including insect bites, vaccinations, tattooing and possibly even with no apparent trauma (Kose and Waseem 2008). The areas most commonly affected are located in the upper parts of the body.

Figure 2. A minor keloid scar.

Figure 3. A widespread hypertrophic scar after skin grafting in a patient with a deep dermal
burn injury.
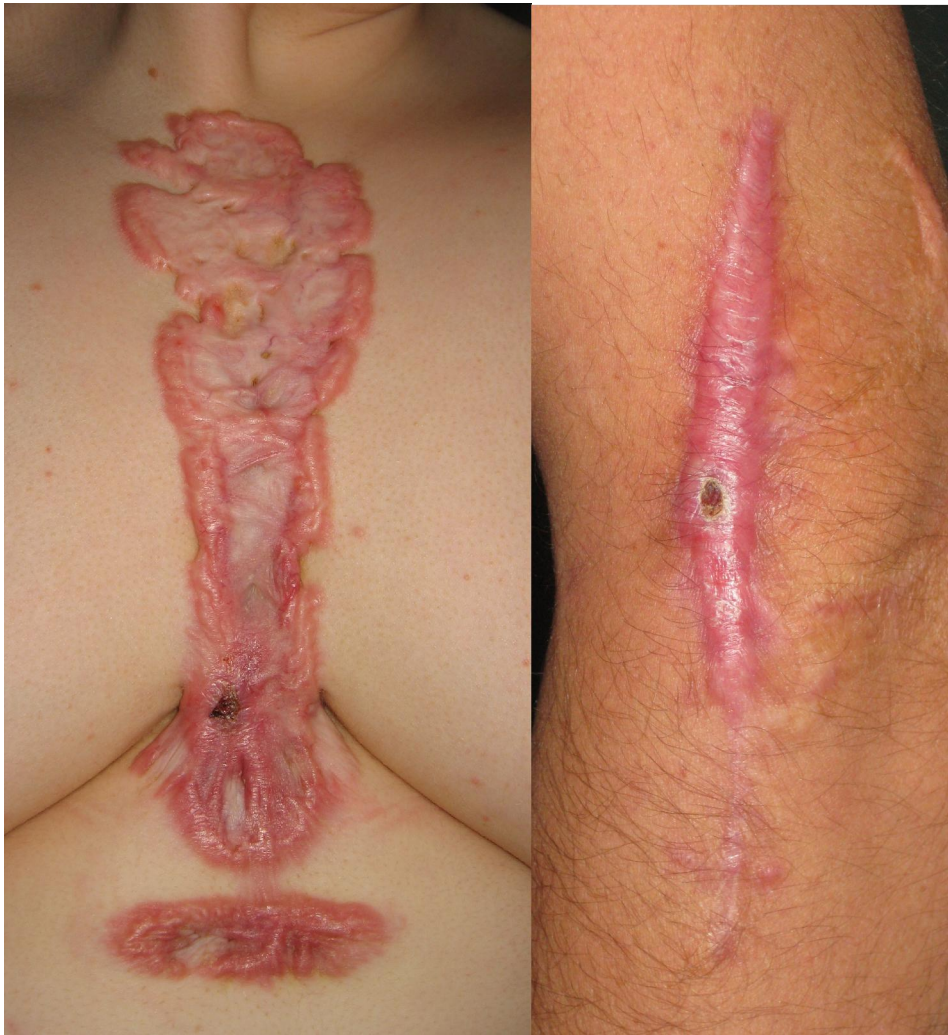
Figure 4. A major keloid scar on the left, a linear hypertrophic scar on the right.

4.5 Requirements of scar assessment methods

An instrument used to measure any health status, in this case the amount of scar hypertrophy or activity, is required to include four (key) features, namely validity, reliability, consistency and feasibility (Duncan et al. 2006).

In scar assessment studies, the validity is assessed by measuring the correlation of a proposed method with either an expert opinion or another known instrument measuring the same parameter. In addition, the validity can be assessed by measuring the correlation of the proposed assessment method with the patient's symptoms (i.e., pain and itching) and the patient's own perception of the scar (Durani et al. 2009).

Any kind of measurement in medicine is subject to some degree of error. In other words, two measurements taken with a certain measurement tool under same circumstances will produce more or less different outcomes. This is naturally reflected in scar assessment methods as well. The ability of an instrument to provide reproducible measurements is expressed as the reliability of the measurement (Fleiss and Shrout 1977). Reliability can be classified into two categories: interrater and intrarater reliability. The former measures the correlation between measurements made by different observers, and the latter measures the correlation between different measurements taken by the same observer. Statistically, reliability can be measured using the intraclass correlation coefficient (ICC). The ICC can be presented for single measures or for average measures, and in scientific publications, both usually are presented. The ICC for single measures refers to the correlation between measurements taken by different observers on one target parameter. The ICC for average measures describes the correlation between the mean of measurements taken by different observers on a group of targets. Mathematically, the ICC of average measures is always higher, and in many studies the reliability is presented as the ICC of average measures. However, there is an important difference between these two indices. For an instrument designed for clinical practice, where only one observer is making the measurements, the ICC of single measures is important, as it describes how reliable these measurements are

when they are repeated by different observers at different points in time. If a measurement tool is shown to be unreliable, it has no value in clinical practice whatsoever. On the other hand, if a method was used in a scientific study by a group of observers to assess a group of targets, the reliability of this method could be described as the ICC for average measures. With more observers, the ICC value rises, but as a drawback, the sensitivity of the method to detect differences between the measured objects deteriorates.

In addition to the ICC value, the confidence intervals should be presented. If a given method has a high ICC, narrow confidence intervals and a statistically significant p-value ($p<0,005$), the method is shown to be reliable, meaning that measurements performed by different observers at the same moment can be expected to be close to similar. In scar assessment studies, a threshold ICC of >0.75 or <0.8 is sometimes used to indicate a reliable method, although this kind of simplification is arbitrary and cannot be used alone without analysing the other results mentioned previously. (Bartko 1966, Fleiss and Shrout 1977, Shrout and Fleiss 1979)

4.6　Current methods for scar assessment

4.6.1　Subjective assessment

4.6.1.1　*Vancouver burn scar assessment scale*

The most widely known and used scar assessment scale is the Vancouver Burn Scar Assessment Scale, generally referred to as the Vancouver Scar Scale (VSS, Table 1). The VSS has four parameters, including vascularity, pigmentation, thickness and pliability, giving a range of 0–14 in the total score. It was originally designed to rate burn scars and was introduced by Sullivan et al. in 1990 (Sullivan et al. 1990). In the original publication, the reliability was poor, but it was concluded that the reliability would improve with time. Baryza et al. proposed the use of an administration tool made of plexiglass that improves the feasibility of the method and increases the reliability (Baryza and Baryza 1995). A modified scale has also been used in several studies. The VSS has been widely used in scar studies, but concerns about its validity and reliability have been raised (Powers et al. 1999). In studies by Nedelec et al., the VSS was used to rate skin graft donor site and burn scars. The reliability of the overall scores and of individual parameters was found to be lower than required for a reliable assessment (Nedelec et al. 2008b, Nedelec et al. 2008a). The VSS was originally designed to assess burn scars but has since been used in clinical studies in rating linear scars as well. Truong et al. tested the scale in the assessment of linear scars after breast cancer surgery. They concluded that the VSS is a valid and reliable scale in the assessment of linear scars, although the level of reliability in their study was lower than what is usually considered acceptable (r=0,64) (Truong et al. 2005).

4.6.1.2   The Patient and Observer Scar Assessment Scale

The Patient and Observer Scar Assessment Scale (POSAS) is a scar rating scale that was introduced by Draaijers et al. in 2004. It consists of two separate scales for the observer and the patient (Table 2 and 3). The parameters have a scale of 0–10, and the total score ranges from 0 to 90 points. It achieved a slightly better reliability than the VSS, although neither of the two scales could be considered reliable if used in a clinical setting by one observer. When used in a clinical study with a group of scars by a group of four observers, the reliability could be considered acceptable (Draaijers et al. 2004c). Truong et al. tested the POSAS on a group of linear scars after breast cancer surgery. The reliability of both the VSS and POSAS was tested, and the ICC for both scales was significantly lower than what is considered acceptable (r=0,54 and 0,33, respectively; (Truong et al. 2007).

The POSAS is an important improvement on previous subjective scar assessment methods because it also records the patient's opinion and scar symptoms (Durani et al. 2009). In fact, the scale places more weight on the patient component than on the observer component. The information acquired from the patient component of the scale can be used to test the validity of other scar assessment methods.

4.6.1.3   The Manchester scar scale

The Manchester Scar Scale was introduced by Beausang et al. in 1998. It has four parameters (colour, contour, distorsion and texture) giving a score of 1 to 4, and a visual analogue scale that describes the overall cosmetic appearance of the scars, giving a score of 0–10. Whether a scar is matte or shiny gives a score of 1 or 2, adding up to a total of 4–28. There was a statistically significant correlation between the scores and the histological findings of biopsies taken from the scars (Beausang et al. 1998). Interrater reliability was not measured in the original publication and has not been reported elsewhere. To the author's best knowledge, the scale has not been used in other published clinical studies.

4.6.1.4 *Visual analogue scales*

The use of a Visual analogue scale (VAS) was proposed by Duncan et al. in 2006. Using only one VAS scale (0–10), they found that it was a feasible tool in the visual assessment of skin scars and reflected well the maturation process of a normal skin scar. As can be expected, however, the reliability of this method, measured with ICC for single measures, was low. This means that the VAS scale would not be suitable for scar assessment for clinical purposes. If used by a panel of four observers in a scientific study, the reliability would be acceptable and the method could be used for research purposes. (Duncan et al. 2006)

4.6.1.5 *Other subjective scar assessment scales*

Yeong et al. proposed a scar rating scale in 1997 to improve the reliability of scar rating with the VSS. The system includes four variables – surface roughness, thickness, border height and colour – and the score is given on a 6-step scale for each parameter (-1 to 4). The score is achieved by comparing the scar with the adjacent healthy skin. The interrater reliability was reported to be high, when 8 observers used the scale to rate 10 photographs of scars (Yeong et al. 1997). It is unclear how well the scale would perform in clinical practice. Despite the promising initial results, the scale has not been used subsequently.

In 2005, Masters et al. proposed a scale that was modified from the Yeong scale and combined clinical scar assessment with reference photographs in a manual to increase the reliability. The scale used the same four parameters as the Yeong scale. However, the ICC fell below acceptable limits (r=0.4–0.82 for average measurements for different paramaters, ICC for overall score was not reported; (Masters et al. 2005). It is worth mentioning that the study setting was fairly complicated: the test site was pointed to the raters by the study author, and the group of raters varied between the two stages of the study. To the author's best knowledge, this scale has not been used in (other?) published scar studies.

In 1998, Crowe et al. proposed the so-called Hamilton scale that was designed for rating photographs of hypertrophic scars. The scale had four parameters (surface irregularity, thickness, colour and vascularity). The total score ranged from 0 to 14, and when two observers were used, the scale yielded an acceptable reliability. (Crowe et al. 1998)

Finally, the Stony Brook Scar Evaluation Scale was proposed by Singer et al. in 2007. It includes 5 parameters (width, height, colour, suture marks, overall appearance) and a total score of 0–5 points, where increasing score correlates to scar healing. The reliability in the study was acceptable (Singer et al. 2007).

Variable	Points
Vascularity:	
Normal	0
Pink	1
Red	2
Purple	3
Pigmentation:	
Normal	0
Hypopigmentation	1
Hyperpigmentation	2
Pliability:	
Normal	0
Supple	1
Yielding	2
Firm	3
Banding	4
Contracture	5
Height (mm):	
Normal (flat)	0
> 0 and < 2 mm	1
≥ 2 and < 5	2
> 5	3
Total score	/13

Table 1. The Vancouver Scar Scale (VSS)

	Normal skin							Worst scar imaginable		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Vascularisation	o	o	o	o	O	o	o	o	o	o
Pigmentation	o	o	o	o	O	o	o	o	o	o
Thickness	o	o	o	o	O	o	o	o	o	o
Relief	o	o	o	o	O	o	o	o	o	o
Pliability	o	o	o	o	O	o	o	o	o	o

Table 2. The POSAS scale observer component.

	No, not at all							Yes, very much		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Is the scar painful?	O	o	o	o	o	o	o	o	o	o
Is the scar itching?	O	o	o	o	o	o	o	o	o	o

	No, just like normal skin							Yes, very different		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Is the scar colour different?	O	o	o	o	o	o	o	o	o	o
Is the stiffness of the scar different?	O	o	o	o	o	o	o	o	o	o
Is the thickness of the scar different?	O	o	o	o	o	o	o	o	o	o
Is the scar irregular?	O	o	o	o	o	o	o	o	o	o

Table 3. The POSAS scale, patient component.

4.6.2 Current objective methods

4.6.2.1 *Spectrophotometers*

There are two types of instruments based on reflectance spectrophotometry currently available for the measurement of skin colour: narrow-band reflectance meters and tristimulus colorimeters (Clarys et al. 2000). The Dermaspectrometer (Cortex Technology, Hadsund, Denmark) and the Mexameter MX18 (Courage and Khazada GmbH, Cologne, Germany) are both narrow-band reflectance meters that have been used in scar colour assessment (Nedelec et al. 2008b). They emit light at specific centre wavelengths (568 and 655 nm for the Dermaspectrometer and 568, 660 and 880 nm for the Mexameter) and provide Hb and melanin indexes. The Dermaspectrometer was tested in scar colour measurement and compared with subjective assessment and a tristimulus colorimeter (Minolta Chromameter, Minolta Camera Co, Osaka, Japan) by Draaijers et al. Both the Dermaspectrometer and the Minolta Chromameter were found to be more reliable than subjective assessment, and the Dermaspectrometer was found to be the more feasible of the two objective devices. (Draaijers et al. 2004b) Oliveira et al. found that the Dermaspectrometer was slightly better in erythema measurement than the Minolta Chromameter but concluded that they could not accurately quantify scar pigmentation (Oliveira et al. 2005).

The Mexameter has been used in scientific studies by Nedelec et al. It was found to be acceptably reliable in the measurement of scar colour, although the ICC of single measures was not presented in the study (Nedelec et al. 2008a). The intrarater reliability in another study by the same author was better when compared with subjective assessment, but the instrument was not able to discriminate hypertrophic scars from normal (non-hypertrophic) donor site scars (Nedelec et al. 2008b).

The Mexameter and Dermaspectrometer both have a narrow measuring head aperture (5 and 6 mm in diameter, respectively), which is a possible source of bias; if a larger scar were to be measured, a series of measurements would be needed to gain a mean score for the entire scar.

MicroColor (Dr. Lange GmbH, Germany) is another tristimulus colorimeter that has been used in skin erythema measurement, showing correlation with subjective

assessment and laser Doppler flowmetry. It has not been tested in scar assessment. (Serup and Agner 1990)

LabScan (Hunter Associates Laboratory, Inc., Reston, Va, US) is a full scanning spectrocolorimeter that was tested by Li-Tsang et al. in scar pigmentation assessment. The system uses CIE L*a*b* and CIE L*c*h* systems to describe colour. The reliability of the normal skin pigmentation measurements was good, but for scar pigmentation measurement was lower than acceptable. The vascularity or haemoglobin concentration of the scars was not estimated, although it is a more important predictor of scar activity. (Li-Tsang et al. 2003)

SIAscope (Astron Clinica Ltd., Cambridge, UK) is a newer instrument based on reflectance spectrophotometry which uses known light interaction models to predict the concentrations of Hb, melanin and collagen. It has been used in diagnosing cutaneous melanoma. There is also anecdotal evidence of the usefulness of SIAscopy in burn depth assessment (Tehrani et al. 2008). SIAscopy might be useful in scar assessments studies as well, but information of the method's reliability and feasibility in scar assessment is lacking.

4.6.2.2   Tonometric and suction devices

Attempts have been made to measure objectively the viscoelastic properties of skin scars. Boyce et al. used the Dermal Torque Meter (DTM310; Dia-Stron Ltd., Broomall, Pa.) to assess burn scars resulting from treatment with cultured keratinocytes. The DTM uses a circular measuring head of 2–3 cm with a rotating disc that flexes the scar. The benefit of the DTM in comparison to the other pliability measuring devices is the somewhat larger measuring head. Unfortunately, the reliability of the measurements was not assessed. (Boyce et al. 2000)

A pneumatonometer used by ophtalmologists has been applied to measure burn scars. The results correlated well with the subjective assessment made with the VSS pliability subscale, but there have been no reports of the reliability of the method (Spann et al. 1996, Oliveira et al. 2005).

The Durometer (Model H1000; RexGauge Company, Inc., Buffalo Grove, Ill.) is another instrument for pliability measurement. In the study by Oliveira et al., it showed correlation with the subjective assessment and measurements made with a pneumatonometer in burn scar assessment. However, the authors did not investigate the reliability of the method. (Oliveira et al. 2005)

The concept of tonometry was developed further by the scar-specific modified tissue tonometer (Burns Model; Flinders University, Adelaide, South Australia, Australia). The machine provides a constant force from a 200g weight that pushes the plunger onto the scar surface. The maximum deformation is measured. It has achieved a high interrater reliability rating in studies by Lye et al. and Corica et al. However, it must be noted that in both studies, the measured areas were small and clearly marked on larger scar areas, and one can therefore assume that the reliability would be markedly lower if a mean score of the entire scar had been measured. (Corica et al. 2006, Lye et al. 2006)

The Cutometer (MPA580; Courage and Khazada) is an instrument designed to measure skin compliance. It has a measuring head aperture of 2–6 mm and uses suction to draw in the measured area. This provides information on the viscoelastic properties of the skin. Several authors have tested the method in scar assessment.

Fong et al. first tested the Cutometer in the assessment of postburn hypertrophic scars in 16 patients and demonstrated good interrater reliability (Fong et al. 1997).

Nedelec et al. have used the Cutometer to measure normal skin, skin graft donor site scars and burn scars. The method showed acceptable reliability and correlation with subjective assessment when normal skin and donor sites were measured, but measurements of burn scars were not sufficiently reliable. This is explained by a roofing effect caused by the inability of the small probe to draw in thick scar tissue (Nedelec et al. 2008b, Nedelec et al. 2008a). The same conclusion was made in another study, where the Cutometer showed acceptable reliability when measuring small well-defined areas of normal skin or thin scars, but the reliability was lower than that of subjective rating when hypertrophic scars were measured (Draaijers et al. 2004a).

The Cutometer has been shown to be more sensitive than subjective assessment in detecting small differences in the viscoelastic properties of skin graft donor site scars. This can be useful in clinical studies comparing different treatments for donor sites (Rennekampff et al. 2006). Overall, the Cutometer seems to be a feasible tool

in clinical donor site studies, but does not work well in thick problematic scars. Furthermore, there is no evidence that the method would be reliable in the clinical setting, where quantitative information of the activity of scars is needed to measure the effect of treatment.

The Skin Compliance Device (Sensory Management Services, Baltimore, Md., USA), formerly known as the Derma Durameter, is an electrical instrument designed to quantify the pliability of skin scars in gm/mm². It was tested by Cleary et al. in assessing hand and forearm scars. The reliability of the method was modest despite the fact that only small well-delineated regions of these linear scars were assessed (Cleary C 2007).

4.6.2.3 *Ultrasonography*

Ultrasound has been used for scar assessment in few studies. It has shown potential in scar thickness measurement and is able to measure the subcutaneous parts of the scar. The method entails a certain learning curve, however, and there is no consensus as to how to report the measurements of a scar with a larger surface area (Nguyen et al. 2008, Perry et al. 2010). For larger scars, multiple measurements would presumably be needed to increase the reliability of the measurement. Despite this, it seems that ultrasound is presently the most feasible option when measuring scar thickness, provided that the limitations regarding the reliability can be overcome. (Fong et al. 1997, Van den Kerckhove et al. 2003, Nouveau-Richard et al. 2004, Lau et al. 2005, Nedelec et al. 2008a, Perry et al. 2010)

4.6.2.4   Three-dimensional imaging

Ardehali et al. have introduced three-dimensional imaging (Vectra 3D, Canfield Imaging Systems, Fairfield, N.J. USA) in scar volume measurement (Ardehali et al. 2007). The instrument reliably measures the volumes of manufactured artificial scars. It was able to demonstrate the change in scar volume during serial intralesional cortisone treatment. The interrater reliability was not measured. The system was only tested on small, raised scar types, and is probably less useful in larger scar areas, such as burn scars. However, three-dimensional imaging seems promising in volume measurement of well defined, raised hypertrophic scars and keloids. The method only measures the volume of the part of the scar that is above skin level. This measurement does not include the deeper parts of the scar. This can be a limiting factor when the method's feasibility is considered.

Another three-dimensional measuring device, the PRIMOS (GF Messtechnik GmbH, Teltow, Germany), was tested by Bloemen et al. on burn scars. The device provides arithmetic representations of the scar's surface roughness that correlate with subjective scar assessment. The interobserver reliability of the method was acceptable (Bloemen 2011).

4.6.2.5   Laser Doppler flowmetry

The laser Doppler measures the movement of red blood cells in the skin and superficial subcutaneous tissue. The measured area is as small as 1 mm² and is scanned with the laser beam. Laser Doppler has been shown to distinguish hypertrophic scars from non-hypertrophic scars and normal skin, in addition to correlating with subjective assessment of scar hypertrophy (Leung et al. 1989, Serup and Agner 1990, Oliveira et al. 2005). The instrumentation is costly and the devices available are large and difficult to move.

4.6.2.6   Confocal microscopy

For several years, there has been interest among dermatologists to utilize laser scanning confocal microscopy (LSCM), also referred to as reflectance confocal microscopy (RCM), in clinical practice. RCM uses near-infrared light (830nm) produced by a diode laser. The light penetrates the skin to the depth of 200–500 µm, and cross-sectional images of the epidermis and the superficial dermis can be produced with a resolution of 1–5 µm (Gerger et al. 2009). The method provides images of the skin at different depths. The wavelength is chosen on the basis of the desired penetration depth (Rajadhyaksha 1995). The method has been suggested as an option for conventional punch biopsy and histopathological examination. This would provide instant information non-invasively from dermal structures *in vivo*. RCM could be used clinically to determine the most suitable site for a biopsy. With newer instrumentation, confocal microscopy has become more feasible in the assessment of the skin, with a resolution capable of distinguishing structures at the microanatomical level. (Taylor et al. 2006, Koller et al. 2009)

Most of the interest has been directed at the examination of different skin tumours, including melanoma and non-melanoma skin cancer, and different semi-malignant lesions (Richtig et al. , Sugata et al. 2008b, Ahlgrimm-Siess et al. 2009a, Ahlgrimm-Siess et al. 2009b, Gerger et al. 2009, Lorber et al. 2009, Richtig et al. 2009). In addition, some authors have investigated the findings of RCM in wound healing and scar formation. Sugata et al. examined superficial dermal skin wounds that were created on healthy skin with a suction blister method using the Vivascope 1000 (Lucid Inc., Rochester, NY, USA). They were able to distinguish the structural changes related to wound healing at the dermal and epidermal level (Sugata et al. 2008a). This opens up an interesting possibility for observing wound healing in model wounds such as split thickness skin graft donor sites without the harm of taking a biopsy. Since the technique is completely non-invasive, multiple measurements can easily be performed on a larger wound. The downside is that the section that can be assessed is only less than 1 mm in diameter, giving room for observational bias.

4.7 Skin imaging with reflectance spectrophotometry

4.7.1 Colour chromophores of the skin

Skin colour has been measured for more than 50 years (REF). The two primary colour chromophores in the skin are melanin and haemoglobin (Hb). The colour of the skin is mostly dependant on the concentrations of these two chromophores (Claridge E. 2002, Jung et al. 2004). Hb is deposited inside the red blood cells and thus does not exist outside the vascular network (Latreille et al. 2007). The haemoglobin concentration in the skin is dependent on the skin circulation and can change due to different physiological stimuli.

Melanin, on the other hand, is the product of melanocytes located in the dermal-epidermal border of normal skin. Rapid skin darkening can occur within a few minutes to hours as a consequence of oxidation and polymerization of the existing melanin due to ultraviolet (UV) radiation, especially UVA. This change is small, however, and a true increase in melanin concentration in the skin only occurs several days after UV exposure as a result of the activation of the melanocyte function (Yamaguchi et al. 2007).

When scars are rated subjectively, the colour change is most often divided into two components, pigmentation and vascularity (Durani et al. 2009). Vascularity is a consequence of locally increased circulation and, therefore, increased Hb concentration (Takiwaki et al. 2002a). Pigmentation, on the other hand, is mostly caused by an increased melanin concentration due to a migration of melanocytes to the site of a scar or other pigmented skin lesion (Dressler et al. 2001). Haemociderin is a product of Hb breakdown, and it has previously been thought to be an important cause of local skin pigmentation. However, the concentration of haemociderin is rarely increased, and histological studies have shown an increased amount of melanocytes in these pigmented lesions (Kim and Kang 2002, Caggiati et al. 2008).

4.7.2 Principles of reflectance spectrophotometry in skin colour measurement

Colour change is usually presented in tri-stimulus values, such as the CIE-L*a*b* or RBG colour space. The CIE-L*a*b* is the most well-known and is used in industrial colour measurements. It mimics the human colour perception. Spectrophotometers provide the whole reflectance spectra from at least the range of visible light wavelengths. This information is mathematically converted to tri-stimulus values (Takiwaki et al. 2002c). Therefore, the tri-stimulus colorimeters are designed to obtain only the CIE-L*a*b* values. In skin colour measurement, it has been common to assume that the Hb concentration change can be measured as a*, whereas the melanin would be represented as L*. However, the colour changes caused by alterations in these two chromophores are somewhat overlapping, and hence the use of the CIE-L*a*b* values as an indicator is purely a simplification (Takiwaki et al. 2002b). Despite this, it has been shown that basically any kind of skin disorder can be assessed with spectrophotometry and this assessment can give valuable information of the nature of the disorder (Takiwaki et al. 2004). When compared to the human observer, reflectance spectophotometry is able to detect even very small changes in vascularity/erythema or pigmentation (Latreille et al. 2007).

The most widely used tri-stimulus colorimeters and narrow-band reflectance meters that have been used in skin and scar colour assessment have been presented in section 4.7.2.1. Instead of giving tri-stimulus values, these meters provide reflectance values of certain central wavelengths thought to represent haemoglobin and melanin-related changes. The changes are provided as indexes that are dependent on the intensity of the reflectance spectra (Draaijers et al. 2004b). Although this change is dependent on the change in the concentrations of the skin chromophores, the relationship is not linear, and knowledge of the precise relationship is lacking.

SIAscopy was the first spectrophotometry-based instrument that provided estimations of skin chromophore concentrations (Claridge E. 2002). Moncrieff et al. presented the SIAscopy method in 2002. SIAscopy was used in the assessment of pigmented skin lesions. Simply measuring linear colour changes in the visible light

range was not a sufficient method when calculating the concentrations of skin chromophores (Moncrieff et al. 2002). SIAscopy was originally designed as an alternative to dermoscopy in diagnosing cutaneous melanoma. The wavelengths used in SIAscopy are between 400 and 1000 nm. The algorithms used in SIAscopy have been developed to distinguish features that are prognostic of a melanoma diagnosis (Moncrieff et al. 2002).

4.8   Digital imaging of skin disorders

As high-resolution digital imaging has become mainstream in health care practices, the use of digital imaging in research purposes has become more common. It is commonly used in wound healing follow-up, where digital planimetry can be utilized to record the wound surface area (Papazoglou et al. 2010). Standardized digital imaging has been shown to improve wound treatment results when combined with a wound electronic medical record (Golinko et al. 2009, Rennert et al. 2009).

Digital photography and spectral imaging is common in forensic medicine, and highly sophisticated methods have been developed in crime scene photography for visualization of different substances (Miskelly and Wagner 2005). In plastic surgery, photographs are routinely used for documentation and reporting on surgical results, in addition to telemedicine as well as teaching and research purposes (Nouarei et al. 2005).

There are several obstacles in image-based measurements. Digital imaging is dependent on several factors, including ambient light (whether it is sun light, fluorescent light, electric lamp, camera flash, etc.), the distance and angle to the wound/object, as well as the properties of the imaging device and user-related variations (Välisuo et al. 2010). These factors make conventional digital imaging an unreliable source of measurements.

Provided that these obstacles are be overcome, there are many possibilities in digital imaging.

4.8.1 Colour measurement from digital images

The healing of skin graft donor sites was assessed with digital imaging by Bon et al. A single-lens reflex camera and an image analysing software was applied to assess the healing with the use of two different wound dressings. The analysis of wound healing was based on colour intensity measurement with the image analysing software using an RBG colour space. The authors found the method to be feasible, but did not make any analysis of its reliability (Bon et al. 2000). The imaging of the wounds was not properly standardized.

Miyamoto et al. used standardized digital imaging to quantify pigmented facial lesions. They used an external lighting device and a special frame for the subject to lean on in order to standardize the distance and angle to the target area. The system was also used in a longitudinal study where pigmented lesions were treated topically, and the effect of treatment could be measured quantitatively with this method (Miyamoto et al. 2002a, Miyamoto et al. 2002b). The setting resembles a professional photography studio, and although the authors of the study achieved a high accuracy in colour quantification, the structure of this kind of system makes it complicated to use.

4.8.2 Digital imaging of chronic wounds

Shai et al. used digital imaging to measure the degree of wound bed preparation. They based their method on the clinical assumption that the colour of the wound bed indicates its clinical state of healing, with red indicating clean wounds, yellow indicating secreting wounds and black indicating necrotic wounds. They created an algorithm to measure a healing index based on expert opinion on different kinds of wounds. This method seemed to correlate well with subjective assessment of the

same group of chronic wounds, indicating that the method would be a feasible option in wounds studies. (Shai et al. 2007)

A similar kind of method was introduced by Bochko et al. to measure chronic wounds. They used a standardized imaging system consisting of a digital camera, a protective dome and a computer-controlled lighting system. The wounds were segmented according to their colour, and the wound and normal skin were segmented with an algorithm. In addition to wound colour, the system can be used to measure wound size accurately, a method referred to as digital planimetry. (Bochko V. 2010)

4.8.3   Digital planimetry

Surface area measurement from digital images is referred to as digital planimetry or computerized planimetry (Oien et al. 2002, Laplaud et al. 2010). The wound surface area can be derived from non-standardized digital imaging, as shown by Papazoglou et al. The wound size was measured with an algorithm using a ruler to calibrate the scale. Although this method gives more accurate information on wound healing than simple measurement of wound diameter, the colour measurement is unreliable because of the effect of the light source. (Papazoglou et al. 2008)

Wound size measurement of burn wounds with standardized digital imaging was presented by Molnar et al. This technique allowed accurate measurement of the burned area in hand burns and showed that the method was feasible in web-based multi-centre trials. (Molnar et al. 2009)

5. Aims of the study

The present study was undertaken in order to:

I   Compare the effects of two adherent wound dressings, Suprathel® and Mepilex Transfer®, on the bleeding, exudation, pain and scarring of split-thickness skin graft donor site wounds

II   Examine the relationship of skin chromophore concentrations and skin colour and to create a model with which the concentrations of these different colour chromophores in the skin can be estimated

III   Test the feasibility and validity of Spectrocutometry in the assessment of skin graft donor site scars in comparison to subjective scar rating scales

IV   Test the validity and reliability of Spectrocutometry in the assessment of linear scars in comparison to subjective scar rating scales

6. Subjects and methods

6.1 Patients and volunteers (Studies I–IV)

For Studies I and III, fourteen patients (nine men and five women, mean[SD] 60[16] years, range 16–78) undergoing split thickness skin grafting (STSG) at the Department of Plastic Surgery at Tampere University Hospital were enrolled, with a total of 22 donor site wounds. All donor sites were located on the thigh. The exclusion criteria were: pregnancy, cortisone treatment, immunosuppression, skin disease, anticoagulation, bleeding disorder or unstable heart disease.

For Study III, the donor site scars were photographed at 14 days, 1 month and 3 months postoperatively, using standardized digital imaging (SDI). From these images, estimated concentration changes (ECC) in Hb and melanin were calculated using Spectral modelling (SpM). The images were also assessed by three observers using the VSS and POSAS scales for colour assessment.

In Study II, 20 healthy volunteers (6 men and 14 women, mean[SD] 31[6] years, range 20-39), with Fitzpatrick skin type I–IV, were enrolled to the experiment. Allen's test was performed on the right upper arm five successive times.

For Study IV, 20 patients (8 men and 12 women, mean[SD] 57[17] years, range 21-82) who had undergone melanoma surgery (excision and a possible sentinel lymph node biopsy or lymph node evacuation) were enrolled. The mean time from melanoma surgery was 17 months (4–30, SD ±7.4). A total of 37 scars were assessed, including 22 excision scars, 12 sentinel lymph node biopsy scars and five lymphadenectomy scars.

6.2 Methods

6.2.1 Skin graft donor site assessment (Study I)

Each of the 22 STSG wounds was divided into proximal and distal halves of equal sizes. They were randomly covered with either Suprathel® or Mepilex Transfer® dressings. The patients stayed on the plastic surgery ward for at least five days after the surgery. The gauze covering the underlying Suprathel® and Mepilex Transfer® dressings was removed on days 1 and 5, and bleeding and exudation was recorded for both dressings individually on a 4-step scale (1=no bleeding, 2=bleeding in the innermost gauzes, 3=bleeding in the outermost gauzes, 4=bleeding throughout the gauzes). The donor site scar was assessed using the Vancouver scar assessment scale (VSS, Table 1) at the outpatient clinic 14 days, 1 month and 3 months after the operation. The cosmetic result was measured using a five-step scale (1=very poor, 2=poor, 3=fair, 4=good, 5=very good). Epithelialization was measured after the removal of the remaining dressings on day 14. Pain related to the donor site wounds was recorded at the time of dressing changes using a visual analogue scale (VAS, 0–10). The donor sites were photographed using the Spectrocutometry method. The digital images were reviewed by three plastic surgeons using the VSS and POSAS colour-related parameters in Study III.

6.2.2 Colour measurement in the Allen's test (Study II)

The reflectance spectrum was measured on 5 specific locations on the palm of each of the 20 volunteers. The measurements were carried out during one session in a quiet, temperature and light controlled room. The Allen's test was performed on all subjects by the author. The spectrum was recorded by Petri Välisuo, MSc, on all occasions.

The spectral measurements were performed during the Allen's test using an HR4000 spectrometer and ISP-REF integrating sphere (Ocean optics Inc., Dunedin, FL,

USA). In the experiment, the measurement was started with both the ulnar and the radial artery occluded manually at wrist level as in a standard Allen's test, followed by release of the circulation and serial measurements every 0.55 seconds until maximal blood concentration in the skin was reached. The model was adapted by adjusting the measured spectral change to the simulation model which was constructed and carried out by Petri Välisuo, MSc, at the University of Vaasa. The results were interpreted against the tri-stimulus values of CIE-l*a*b* and RBG colour spaces in which colour measurements are generally performed. This comparison was used to establish a relationship between skin chromophore concentrations and skin colour that could be used in clinical assessment of skin scars and other skin disorders using the Spectrocutometry method.

6.2.3 Standardized digital imaging used in Spectrocutometry (Studies III and IV)

The instrumentation used for SDI and SpM (Spectrocutometry) in Studies III and IV consists of a digital single-lens reflex camera (DSLR), Fuji IS Pro (Fujifilm Corporation, Tokyo, Japan), and a computer-controlled lighting system build at the University of Vaasa. The instrument is used to capture a standardized, calibrated digital image of the target and to perform spectral modelling in order to calculate ECCs of Hb and melanin in the target area. The lighting unit consists of 120 light emitting diodes (LEDs) with measured central wavelengths of 470 nm, 520 nm and 640 nm, and the _3 dB bandwidths are 19 nm, 25 nm, 15 nm and 50 nm, correspondingly. The imaging software acquires one colour image using red, green and blue LEDs for illumination. The instrument has a rectangular chamber attached to the camera (Figure 3). The LEDs are located at the top of the lighting chamber and illuminate the target. The chamber holds the distance and angle to the target constant, while also blocking out external light that would otherwise interfere with the reflectance measurement.
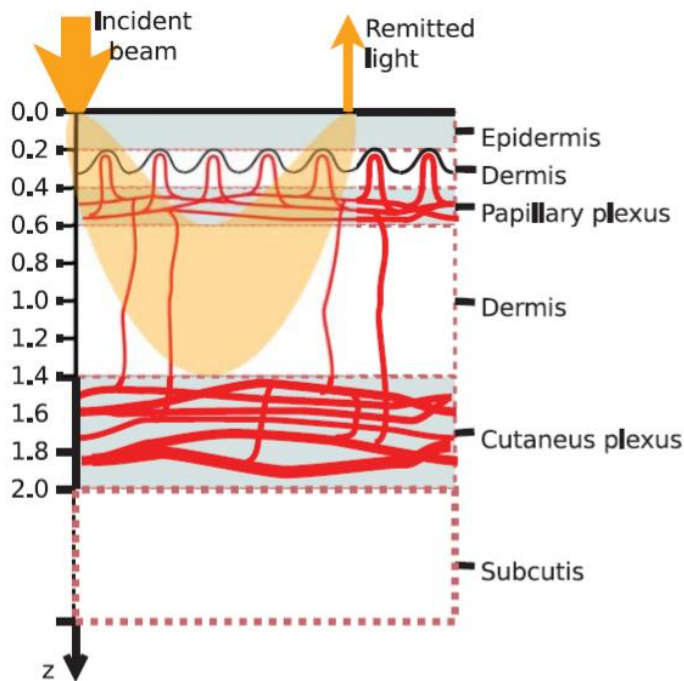
Figure 5. Principles of reflectance spectrophotometry of the skin and the skin model used in the Monte Carlo simulation model. Picture from (Välisuo et al. 2011).

6.2.4 Subjective and objective scar assessment (Studies I, III and IV)

In Study I, the donor site scars were assessed by one observer non-blinded using the Vancouver scar scale.

In Study III, the standardized digital images taken from the scars were viewed and rated by three plastic surgeons, I.K. and two other plastic surgeons serving as independent observers. The rating was accomplished using the colour-related subscales of the VSS and POSAS scales. The rating was done twice: in the first rating, both sides of the donor site scar were viewed at the same time, allowing for comparison between the two sides. In the second rating, the sides were viewed separately to resemble a double blind study setting. This setting was used to assess

to what degree the sensitivity of subjective assessment is affected by inability to compare the test sites.

In Study IV, the VSS and POSAS rating scales were employed. The rating was carried out during a single day in a lighting and temperature controlled room. The rating was done individually by the observers. When Spectrocutometry was used, the observer took a single image of the scar and then viewed it on the computer screen. If the observer approved the image, no more images were taken. There was no training for the use of Spectrocutometry. For the patient component of the POSAS scale, a translated version was used. Otherwise, the original English versions of the scale were used by the observers. The ECCs of Hb and melanin were calculated by Petri Välisuo, MSc, at the University of Vaasa using the algorithm generated in Study II.

6.2.5 Image segmentation and spectral modelling (Studies III and IV)

Image segmentation was used to delineate the scarred area in the images captured with SDI (Figure 6). The surrounding shadowed areas, clothes, naevi etc. were excluded, and the baseline skin in the area surrounding the scar was measured. Because the changes in melanin and Hb concentration overlap, it is necessary to measure the skin baseline from the healthy adjacent skin surrounding the scar. This is a crucial step that allowed for accurate estimation of the chromophore concentration changes.

The reflectance spectrum of the scar and surrounding normal skin was measured in comparison with the reference white as in Study II. The observed reflectance change between the scar and the healthy skin was divided into two main components, which correspond to the concentrations of Hb and melanin. While melanin concentration in the skin changes slowly and remains constant during several measurements

performed the same day, Hb concentration depends on the local and systemic circulation. This is why the concentration was measured as the estimated concentration change between the scar and healthy adjacent skin. The values obtained using Spectrocutometry were the ECC of Hb and melanin.

Figure 6. Image segmentation used to delineate the scar and normal skin.

6.2.6 Statistical analysis (Studies I–IV)

In Study I, the Suprathel® and Mepilex Transfer® treated areas were pooled as two groups, the Suprathel® group (SG) and Mepilex Trasfer Group (MTG). The results of bleeding/exudation measurements, pain measurements and scar assessment were compared between the two groups using the Mann-Whitney U-test. A p-value of <0.05 was considered statistically significant.

In Study III, the reliability of subjective assessment of scar colour using VSS and POSAS colour variables was tested with Intraclass correlation coefficient (ICC). The results were presented for both single and average measurements. The concurrent validity SDI and SpM were assessed by calculating the correlation between the objective and subjective assessments with Spearman's correlation coefficient. To test the sensitivity of these methods, a paired samples Wilcoxon signed-rank test was used to compare the scar assessment results of the two treatment groups, SG and MTG, of Study I.

In Study IV, the reliabilities of the VSS, POSAS and Spectrocutometry in scar assessment were evaluated using the ICC. The results for single and average measurements as well as the confidence intervals were presented. The concurrent validity was assessed by directly comparing the subjective scar ratings with the VSS and the POSAS scales to the results of Spectrocutometry assessments.

A Baeyesian network was constructed using the B-course tool, and the dependencies of the scar assessments and the patient-reported symptoms were analyzed (Pearl 1988, Cooper 1990, Myllymäki P. 2002). Bayesian network analysis is a graphical model, where the probabilistic relationships between the scar symptoms and the clinical findings can be represented. If the B-course tool indicated a dependency between two variables, the correlations between the variables were tested using Pearson's correlation coefficient and Automatic Relevancy Detection (ARD) which can be used to detect nonlinear, non-monotonic dependencies.

7. Results

7.1 Outcome of the skin graft donor site healing comparing Suprathel® and Mepilex Transfer® dressings (Study I)

During the early postoperative period, Suprathel® was significantly less painful than Mepilex Transfer® on days 1 and 5. Suprathel® was associated with significantly less bleeding and exudation on days 1 and 5. (Table 4.) The epithelialization rate did not differ between the two treatment options.

The VSS showed significantly less scarring in the SG at 1 month postoperatively (Table 4). The Spectrocutometry results for the ECC of Hb were higher for MTG than for SG at 14 days and 1 month postoperatively ($p<0.05$). The ECC of melanin was higher in the MTG than the SG at all time points (14 days, 1 month and 3 months postoperatively, $p<0.001$, <0.05 and <0.001, respectively; Table 4). The POSAS rating for vascularity and pigmentation was significantly higher in the MTG than the SG at three months postoperatively ($p<0.05$) and significantly higher pigmentation than in the SG in was recorded in the MTG at 14 days ($p<0,05$) and 1 month ($p<0.001$; Table 4).

	Days from surgery				
	1	5	14	30	90
Pain (VAS) SG / MTG	0.73/1.90 (p<0.05)	0.68/1.95 (p<0.05)			
Bleeding/ exudation, SG / MTG	1.0/3.18 (p<0.001)	1.0/2.55 (p<0.001)			
VSS score SG / MTG			*	2.14/3 (p<0.05)	1.65/2.25 (p<0.05)
ECC of Hb, SG / MTG			0.27/0.31 (p<0.05)	0.20/0.22 (p<0.05)	0.10/0.10
ECC of melanin, SG / MTG			0.067/0.44 (p<0.001)	0.21/0.45 (p<0.05)	0.18/0.41 (p<0.001)
POSAS vascularity, SG / MTG			4.38 / 4.71	3.85/3.93	2.25/2.56 (p<0.05)
POSAS pigmentation, SG / MTG			3.88/4.68 (p<0.05)	4.40/5.20 (p<0.05)	4.43/4.57

Table 4. The results of the pain and bleeding/exudation score, VSS rating, the Spectrocutometry assessment and the POSAS ratings done blindly by three observers using the colour-related parameters. VAS=Visual analogue scale, SG= Suprathel® group, MTG= Mepilex Transfer® Group. P-values for statistically significant differences are shown. *The VSS rating was not done at 14 days postoperatively. Only pain and bleeding/exudation were recorded at 1 and 5 days postoperatively.

7.2 Estimation of skin haemoglobin concentration comparing the results of Allen's test and Monte Carlo simulation (Study II)

A similar skin model as shown in Figure 5 was used for Monte Carlo simulation in Study II. The reflectance values of red, green and blue LEDs are shown in Figures 7 and 8. The measured spectrum correlated with a drop of blood concentration to woughly one third of the original in the simulation model. As shown in Figures 7 and 8, the relationship between the skin colour and skin blood concentration is nonlinear, and when the blood concentration is already high, the increase in blood concentration has less effect on skin colour, which is seen as a plateau of the reflectance values in Figure 8. The colour measurement with the current spectrophotometers is based on a linear measurement of colour intensity. When assessing active scars, the circulation is already above the level of normal skin, and therefore the increase in scar activity and its circulation might be underestimated when mere colour intensity of the scar is measured.
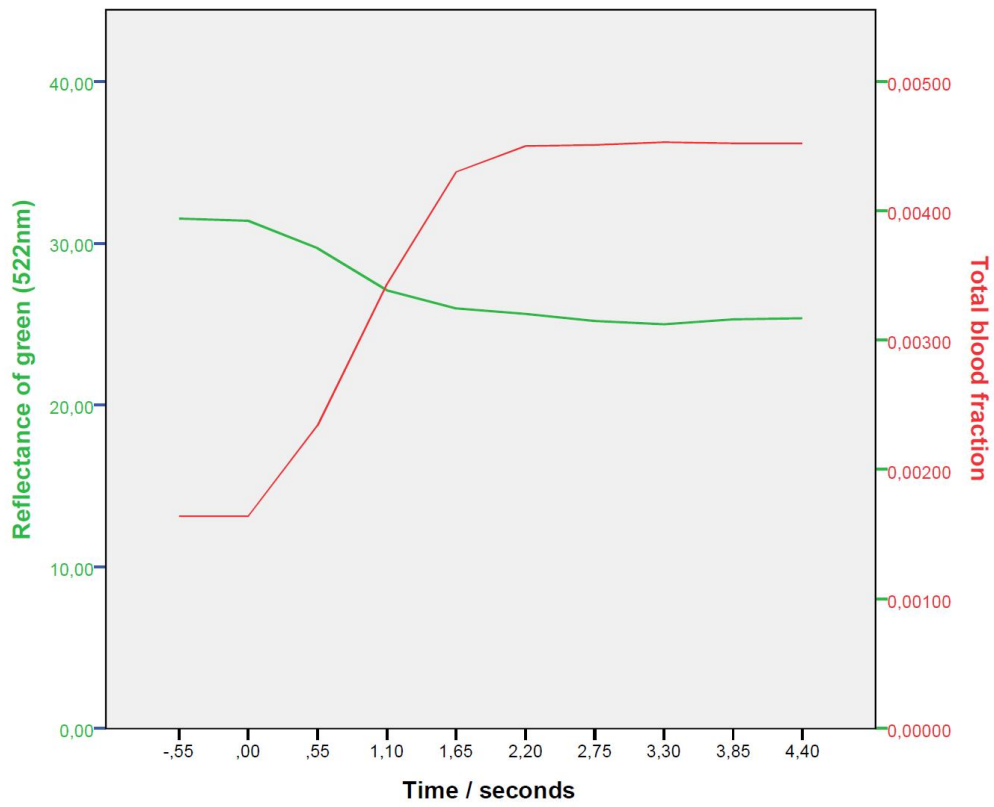
Figure 7. The reflectance of green LED (522 nm) measured from the skin of the palm during the Allen's test and the simulated blood fraction.
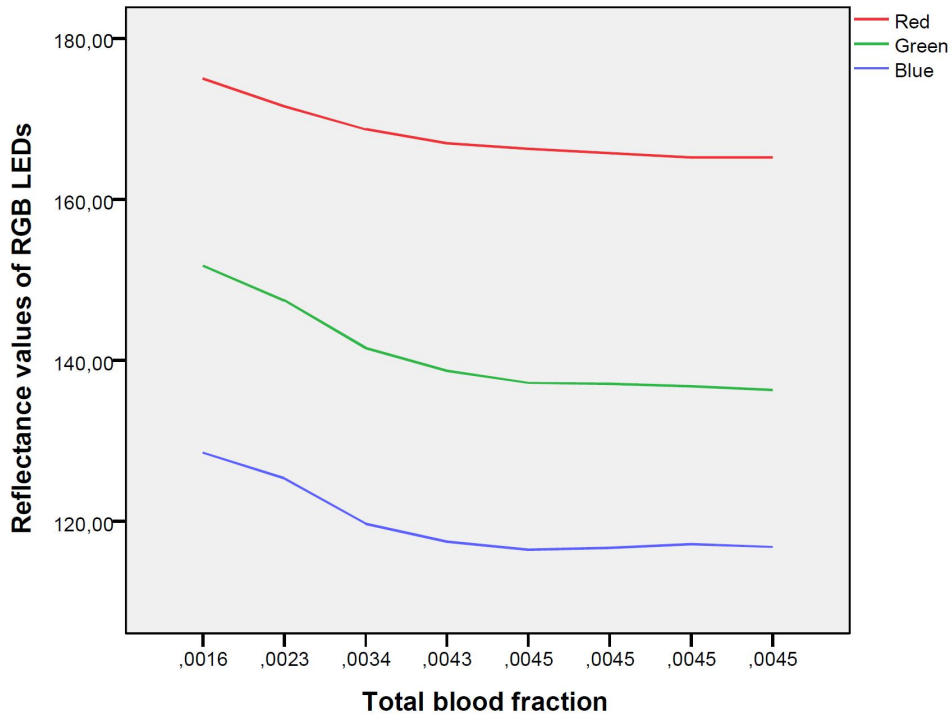
Figure 8. Reflectance values of red, green and blue LEDs measured from the skin of the palm during Allen's test and the simulated total blood fraction.

7.3   Donor site scar assessment with Spectrocutometry and subjective rating scales (Study III)

The results of the subjective ratings and Spectrocutometry measurements are shown in Table 4. In the first assessment, the VSS scale detected statistically significant differences between SG and MTG in vascularity assessment at 14 and 30 days and in pigmentation at all time points ($p < 0.05$). The POSAS scale detected less vascularity and pigmentation in SG at all time points ($p < 0.05$).

In the second assessment, the VSS detected less vascularity in SG at 30 days only ($p < 0.05$). The POSAS detected lower vascularity in SG at 90 days and lower pigmentation in DG (SG?) at 14 and 30 days ($p < 0.05$). The Spectrocutometry results

showed a higher ECC of Hb in MTG at 14 and 30 days and a higher ECC of melanin at all time points ($p<0.05$).

In the pooled data, the POSAS showed a statistically significant reduction in vascularity between 14 and 30 days and between 30 and 90 days in both assessments ($p<0.05$). This was also demostrated in the Spectrocutometry measurements ($p<0.05$). The ECC of melanin was higher at 30 days than at 14 days ($p<0.05$). This was also evident in the second assessment with the POSAS scale ($p<0.05$).

The concurrent validity between the subjective ratings and Spectrocutometry results was calculated using Spearman's correlation coefficient (Table 5). There was a statistically significant correlation between the ECC of Hb and the POSAS and VSS vascularity subscales ($r= 0.63$ and 0.74, respectively, $p<0.001$). Similarly, there was a statistically significant correlation between the ECC of melanin and the POSAS and VSS pigmentation subscales ($r=0.6$ and 0.53, respectively, $p<0.001$).

	Correlation coefficient	p-value
Skin graft donor site scars (study III)		
VSS vascularity vs. ECC of Hb	0.74	<0.001
POSAS vascularity vs. ECC of Hb	0.63	<0.001
VSS pigmentation vs. ECC of melanin	0.53	<0.001
POSAS pigmentation vs. ECC of melanin	0.6	<0.001
Melanoma scars (Study IV)		
VSS vascularity vs. ECC of Hb	0.4	0.015
POSAS vascularity vs. ECC of Hb	0.5	<0.001
VSS pigmentation vs. ECC of melanin	0.6	0.002
POSAS pigmentation vs. ECC of melanin	0.67	<0.001

Table 5. Correlations between the subjective scar ratings and spectrocutometry results in Studies III and IV. Spearman's correlation coefficient used in Study III and Pearson's correlation coefficient in Study IV.

7.4 Reliability and concurrent validity of subjective assessment and Spectrocutometry in scar assessment (Studies III and IV)

The reliability of the subjective ratings fell below acceptable limits in both Studies III and IV (Table 6). In Study III, the VSS showed unacceptably low reliability levels in both assessments. At the time point of 1 month postoperatively, there was no statistically significant correlation between the three observers when the VSS was used to rate scar pigmentation. For this reason, the VSS pigmentation was not used in the second assessment. The vascularity rating using the VSS yielded slightly better reliability (r= 0.2–0.45 for single measures, p<0.01), but again could not be considered reliable enough. In the second assessment, the VSS had slightly better reliability for vascularity rating, but it did not reach an acceptable level.

The POSAS scale achieved a better outcome than the VSS in Study III. In the first assessment, the POSAS vascularity rating ranged between 0.25 and 0.51 for single measures (p<0.001), and in the second assessment, between 0.32–0.43 (p<0.001). The POSAS pigmentation rating seemed to be the most reliable parameter in Study III and the only one that achieved nearly acceptably reliable results in both assessments (r=0.73–0.76 and 0.68–0.82 in the first and the second assessments, respectively, p<0.001). The results remained close to each other between the two assessments.

In Study IV, the VSS and POSAS scales and the Spectrocutometry method were used independently by three observers to assess linear scars after melanoma surgery. The ICCs of these ratings are shown in Table 3. The reliability of Spectrocutometry was superior to the subjective scales and achieved acceptable levels (r= 0.88–0.89 and 0.96 for single and average measures, respectively, p<0.005). The VSS gained a somewhat higher reliability rating than the POSAS in this study, demonstrating an acceptable reliability when used by three observers, but not for single measures (r= 0.86 and 0.66 for average and single measures, respectively, p<0.005). The POSAS resulted in slightly lower reliability levels, but again, it showed an acceptable reliability when three observers were used (r= 0.60 and 0.82 for single and average measures, respectively, p<0.005).

The correlations of subjective and objective assessment with the Spectrocutometry method were measured using Pearson's correlation coefficient and are shown in Table 6. There was a statistically significant correlation between the vascularity and pigmentation assessments in Studies III and IV.

A Bayesian network analysis of the results in Study IV was undertaken applying the B-course tool. The dependencies, the Pearson and Spearman estimates and the Automatic relevancy detection (ARD) results found in the Bayesian network analysis between the patient-reported scar symptoms and the measured parameters are shown in Table 7. The strongest dependency was between the scar pain and ECC of Hb. There was also a positive correlation between these two parameters in regard to the Pearson's correlation coefficient and ARD. This fits the hypothesis that metabolically active scars with increased circulation are prone to cause symptoms. Interestingly, there was no correlation between the pigmentation measurements and scar symptoms.

Figure 9. Three images taken from the same skin scar individually by three independent observers, showing a similar distance and angle to the wound.

	ICC of single measures	ICC of average measures	Confidence intervals (average measures)	p-value
Skin graft donor site scars				
VSS vascularity, 1st rating	0.4	0.67	0.66–0.82	<0.001
VSS vascularity, 2nd rating	0.32	0.58	0.37–0.72	<0,001
VSS pigmentation 1st rating *	0.26	0.52	0.54–0.69	<0,001
POSAS vascularity, 1st rating	0.51	0.76	0.67–0.82	<0,001
POSAS vascularity, 2nd rating	0.56	0.79	0.72–0.85	<0,001
POSAS pigmentation, 1st rating	0.75	0.9	0.87–0.93	<0,001
POSAS pigmentation, 2nd rating	0.69	0.89	0.83–0.90	>0,001
Melanoma scars				
Subjective assessment				
POSAS	0.6	0.82	0.69–0.90	<0.001
VSS	0.66	0.86	0.75–0.92	<0,001
Objective assessment				
ECC of haemoglobin	0.88	0.96	0.92–0.98	<0,001
ECC of melanin	0.89	0.96	0.93–0.98	<0,01

Table 6. Intraclass correlation coefficients of subjective ratings and Spectrocutometry assessments in studies III and IV including confidence intervals of average measures.

	r	p	α
POSAS pliability and scar pain	0.49	0.25	1.2
ECC of Hb and Scar pain	0.64	0.36	1.5
ECC of HB and Scar colour	0.48	0.48	

Table 7. Dependencies found in the Bayesian network analysis between the patient-reported scar symptoms and scar assessment results. r=Pearson's correlation coefficient, p=Spearman's correlation coefficient, α=ARD.

8. Discussion

8.1 Treatment of the split-thickness skin graft donor sites

The number of options for skin graft donor site treatment is vast, and new innovations are frequently introduced (Voineskos et al. 2009). However, precise knowledge of their effect on donor site wound healing is lacking. One reason for this is the fact that wound studies are extremely complicated to arrange; the wound healing ability is different in every patient and also depends on the general condition of the patient (Paletta 2006).

The choice of the wound dressing has an important impact on wound healing characteristics (Paddle-Ledinek et al. 2006). When studying dermal wounds, the split-thickness skin graft donor site seems to be the closest to what one can achieve when a standardized defect is needed (Barnea et al. 2004, Argirova et al. 2007). Because the skin harvesting is accomplished using a dermatome, the depth of the injury is known (Malpass et al. 2003). The remaining problem is the individual characteristics of the patient and his/her ability to heal. The rate of scarring after dermal injury is subject to inter-individual differences.

The aim of Study I was to compare the effects of two modern, semi-occlusive dressings on the donor site healing using a randomized, prospective study setting. The method used in Study I was designed to minimize the effects of inter-individual variation in wound healing ability. The treatment options were compared on the same donor site, and the position of the dressings was randomized. This setting uses the patient as his/her own control. The setting is not a true randomized controlled trial, but rather a randomized case-control trial. A similar kind of method was used by Schwarze et al. (Schwarze et al. 2007). This setting offers a unique opportunity to test the effect of a treatment method on dermal wound healing *in vivo*. In order to achieve clinically solid information, a novel treatment option should be compared

71

against the best known alternative, and not against a method that is already known to be poor.

In Study I, the dressings that were compared remained on the wound until day 14, which rendered direct comparisons of the rate of epithelialization between the study groups impossible. In this setting, a punch biopsy taken from the wound at 7 days postoperatively would be beneficial (Kilinc et al. 2001). In our study, only speculations of the epithelialization rate can be made based on the measured parameters of pain, bleeding and the resulting scar. However, there are ethical issues involved in the biopsy which obviously causes pain and discomfort to the patient and, theoretically, might delay the wound healing process. In a study that uses dressings that are supposed to be changed periodically, it is possible to measure epithelialization directly (Uysal et al. 2006).

Suprathel® has been used in recent years especially in the treatment of dermal burn injuries and skin graft donor sites and has also been reported in one case of toxic epidermal necrolysis (Schwarze et al. 2007, Pfurtscheller et al. 2008). It has been reported to alleviate pain in conservatively treated burns (Uhlig et al. 2007, Schwarze et al. 2008). Silver-coated dressings have been used in several studies and are reported to have beneficial effects on wound healing and to reduce the risk of infection (Innes et al. 2001). Suprathel® has a lower pH than other dressings, and this seems to create an antiseptic effect comparable or even superior to topical silver (Ryssel et al. 2010, Ryssel et al. 2011). The use of topical silver is also reported to have cytotoxic effects on keratinocytes and fibroblasts, which raises questions regarding its usability in donor site wounds (Paddle-Ledinek et al. 2006, Burd et al. 2007).

A benefit shown in our study but not in the previously published studies is the haemostatic effect of Suprathel® on skin graft donor site wounds. This is a considerable benefit that can be advantageous when large wound surfaces are created with the skin harvest or when skin grafts are harvested from a patient that has a bleeding disorder.

8.2 Spectrocutometry in measuring the concentrations of melanin and haemoglobin in skin scars

The skin is the largest organ in humans. It is relatively transparent to light, which allows measurements to be taken noninvasively. The reflectance spectrum of the skin chromophores, Hb and melanin, are well known (Dawson et al. 1980). Based on this knowledge, indexes that reflect their concentrations have been used in studies involving skin scars and other skin disorders (Takiwaki et al. 2004). Spectrophotometry is a widely-known method for quantitative colour measurement (Gonzalez et al. 2010). The devices used in the medical field are presented in chapter 4.7.2.1. Although they do offer the possibility of quantitative colour measurement, they do not accurately reflect the concentration changes in Hb and melanin (Claridge et al. 2002, Välisuo et al. 2011).

The aim of Study II was to analyze the relationship of the measured spectral change and chromophore concentrations of the skin *in vivo*. Since actual measurement of haemoglobin concentrations in living skin is not possible, a modified Monte Carlo simulation was used instead. Monte Carlo simulation is a deterministic model that was used in this case to calculate the probabilities of individual photons to be reflected or absorbed in the skin at different blood concentrations (Nishidate et al. 2011). The model was adjusted by matching the results to the measured reflectance values *in vivo*.

The relationship between the haemoglobin concentration and skin colour was found to be non-linear. The current scales for scar assessment, however, are equal-appearing interval (EAI) scales that assume that the colour change would follow a linear path (Brandt et al. 2009). The spectrophotometers used previously in scar assessment also measure the linear change in colour intensity, which means that the change in scar activity is often underestimated, whether using a subjective scale or a commercial spectrophotometer.

As a conclusion, if scar circulation is to be estimated indirectly by measuring the colour change, one has to consider the non-linear relationship of haemoglobin concentration and colour change. Spectrocutometry is a method that can produce estimations of the ECC of Hb and melanin to a satisfying degree and can therefore

be considered a more accurate method in scar measurement than the conventional spectrophotometers.

8.3 The reliability and validity of subjective scar assessment with VSS and POSAS

Attempts have been made by many to create an ideal method for subjective scar rating (Powers et al. 1999, Teot 2002, van Zuijlen et al. 2002, Durani et al. 2009, Idriss and Maibach 2009, Vercelli et al. 2009). This approach is, in many ways, problematic and demands several simplifications. There are different types of skin scars, and it would be unrealistic to assume that completely different scar types could be measured with a single scale (Mustoe et al. 2002). On the other hand, the clinical importance of a scar often depends more on the location than the type or size of the scar (Duncan et al. 2006). Scars that are easily hidden are less disadvantageous than ones that are difficult to cover.

Subjective assessment with different rating scales has long been the mainstay of scar assessment studies. As mentioned earlier in chapter 4.4.1, the VSS and POSAS scales are the most widely used and have been validated. For this reason, we wanted to test these scales in Studies III and IV. The results of our studies are quite similar as those published earlier, and they demonstrate that neither of the two scales can be considered a reliable tool in clinical practice (Powers et al. 1999, Draaijers et al. 2004c, Truong et al. 2005, Truong et al. 2007, Nedelec et al. 2008a). This is due to the limitations of a human as an observer. As shown in Study III, human observers are good in detecting differences when they are able to make direct comparisons. However, they lack the ability to objectively determine a mean value of a measured parameter, i.e., colour change or a mean thickness or pliability estimation.

Based on the findings in Studies III and IV, one cannot draw direct conclusions as to which of the two scales used in the studies would be more recommendable. The POSAS scale has 11 parameters, 6 of which are patient-reported parameters. We found no correlation between the patient and observer components of the scale, nor

was there a correlation between the overall patient score and the Spectrocutometry results. The correlation between the observer and patient component was not reported by Draaijers et al. who introduced the scale (Draaijers et al. 2004c). Although the patient component does not add to the reliability of the scale, it does bring an important documentation of the patient's perception of the scar (Truong et al. 2007). This information is particularly useful when testing the validity of objective measurement methods. The patient component can be considered additional information, and the observer component can be used as the principal outcome measure in clinical practice. Overall, of the existing scales, the POSAS seems to be the most comprehensive one (Vercelli et al. 2009).

It is noteworthy that in the POSAS scale, all the parameters have equal weight. In our study, it was shown that pliability is the most important parameter clinically, but in the scale it is only responsible for one eleventh of the total score. In the Bayesian network analysis, there was no dependency or correlation between the overall POSAS score and patient-reported symptoms. The only dependency found was between scar pain and the individual POSAS parameter of pliability.

Unlike the POSAS, which is based on a visual analogue scale, the VSS is a parametric scale and might therefore be easier to use for someone with little experience. The pigmentation subscale has only three steps and describes the mere pattern of the pigmentation and not its intensity. In the modified version of the scale, there is another step, "mixed pigmentation," but this brings little additional value to the scale (Nedelec et al. 2008b). The VSS is designed to be used in rating burn scars (Sullivan et al. 1990). Studies III and IV involved STSG donor site scars and linear scars, which might have affected the reliability results. However, the VSS has also been used in the assessment of linear scars by others (Truong et al. 2005).

Overall, in Studies III and IV, the subjective rating scales were found to be unreliable when used by a single observer. When used by a group of three observers in a study setting, the reliability was acceptable. Others have found the POSAS scale to be more reliable than the VSS (Draaijers et al. 2004c). Van de Kar et al. have also found the POSAS to be reliable in rating linear scars. It is debatable whether the reliability could be increased with more experience and standardized training (van de Kar et al. 2005).

8.4 Objective methods in scar assessment

There are numerous objective methods suggested for scar assessment in previous studies (Verhaegen et al. 2011). They are mostly designed to measure the same parameters as the subjective scales – colour, pliability and thickness in particular (Perry et al. 2010). In several publications, they are found to be preferable to the subjective rating scales (Oliveira et al. 2005). However, no instrument to date has been widely accepted in clinical practice. Although labelled as objective methods, many of the measurement devices introduced so far have a certain amount of subjectivity involved (Perry et al. 2010, Verhaegen et al. 2011). If one considers a scar with a large surface area and uneven colour and thickness, it is obvious that the observer can affect the outcome by choosing a certain spot for the measurement when the measuring head diameter of the instrument is small. Multiple measurements would be required to arrive at a more reliable outcome, or, preferably, a method that produces a mean value of the whole scar should be applied. Of the objective methods previously described, laser Doppler flowmetry as well as some three-dimensional imaging systems offer this possibility. This was also the aim when designing the suggested Spectrocutometry method.

We tested the validity and reliability of Spectrocutometry in the assessment of SGDS scars and linear scars in Studies III and IV. The aim of the studies was to compare the suggested method to the most widely known and commonly used rating scales, the VSS and POSAS.

Spectrocutometry was found to be a reliable tool in the assessment of skin scars. It has many advantages compared to some other objective methods. First of all, it captures a standardized digital image of the scar that can be used as documentation or for subjective evaluation by a panel of observers in the case of a clinical randomized trial. It is superior to conventional digital photography, which is affected by ambient light and the experience of the photographer (Välisuo et al. 2010). The external lighting chamber holds the distance and the angle to the scar constant, resulting in similar images despite different users, as shown in Figure 4.

Spectrocutometry achieved high interrater reliability in Study IV when compared with previously published data (Draaijers et al. 2004b, Nedelec et al. 2008a). It must be noted that the study setting we have used was not altered in any way to facilitate

the assessment, but it was comparable to a clinical setting where the observers make an independent assessment. The Bayesian network analysis in Study IV revealed interesting details. The most important finding was that there was a strong dependency between the scar pain and its ECC of Hb. This was also demonstrated in the positive correlation showed in the Pearson's correlation coefficient and the ARD results. This indicates that scars which have an increased circulation and are clinically active are most often symptomatic. Hence, these scars are usually clinically important (Li-Tsang et al. 2003). As presented in section 4.1, HSCs and keloids have a higher BVD than normal skin and skin scars (Ehrlich et al. 1994, Amadeu et al. 2003). The BVD decreases as the scar maturates, reflecting the inactivation of the scarring process (Beer et al. 1998). This process is clinically important, because it indicates the need for further scar treatment. Scar treatment generally aims at gradual inactivation and maturation of the scars (Ogawa 2010). Since Hb in scars is located only inside the blood vessel, it can be used as an indicator of scar BVD and activity. By measuring the decreased vascularity of the scar, one can effectively monitor the efficacy of the treatment. Spectrocutometry provides essential information on the maturation process of the scar.

It can be stated that there are other important dimensions in scars, such as volume, thickness and pliability, that cannot be measured by means of Spectrocutometry. However, they can be considered as secondary changes with regard to scar activity (Broughton et al. 2006). As shown in Study IV, the symptoms of the scar are less dependent on the other modalities. Of the POSAS subscales, scar pain was dependent only on pliability.

For a thorough objective assessment of a hypertrophic or keloid scar, several objective methods would be needed. Based on the findings in Study IV, one should be able to measure scar thickness, pliability and circulation. For thickness measurement, US seems to be the most feasible option (Nouveau-Richard et al. 2004, Lau et al. 2005). For pliability measurement, the Cutometer is the most frequently applied option, but it is not reliable in thick scars (Nedelec et al. 2008a). At the moment, there is no ideal method for the assessment of scar pliability (Verhaegen et al. 2011).

RCM offers a unique opportunity to obtain information of microscopic accuracy and seems to be a promising tool in scientific studies of skin scars (Koller et al. 2009).

For colour-based measurements, several commercially available instruments exist (Perry et al. 2010). However, with the exception of SIAscopy, they do not yield information on the physiological changes in scar circulation (Tehrani et al. 2008). SIAscopy has not been tested for scar assessment, however. For scar assessment, Spectrocutometry appears to be a feasible, valid and reliable method that can be recommended for use in clinical practice and scientific studies.

9. Conclusions

I   Suprathel® was superior to Mepilex Transfer® in the treatment of split-thickness skin graft donor sites with regard to pain, bleeding and scar formation.

II   There is a nonlinear relationship between skin colour and the concentration of melanin and Hb. Reflectance measurements of the skin using Spectrocutometry can be used to accurately estimate the concentration changes in the skin chromophores.

III   Spectrocutometry can be used for objective assessment of skin graft donor sites. Subjective scar assessment of donor site scars is unreliable.

IV   Spectrocutometry is a valid and feasible method and more reliable than subjective rating in the assessment of linear skin scars.

10. Future perspectives

In the future, the concept of Spectrocutometry is planned to be used and tested in both scar and wound studies. There is a limited amount of good-quality studies on recent scar prevention and treatment methods, and we believe that Spectrocutometry can be used in comparative studies of different scar therapies as an objective means to assess scar activity. In addition, the findings of the previous studies have enabled us to make reliable measurements of oxygen saturation by gauging the concentration of oxyhaemoglobin from the skin. This may open interesting opportunities especially in wound studies, as it is known that wound oxygenation is an important factor in wound healing. We wish to proceed with our studies in assessing chronic wounds of different backgrounds, with particular focus on diabetic foot ulcers. The Spectrocutometry method can be used to make precise measurements of wound size (digital planimetry) and to estimate wound oxygenation, and these measurements may yield interesting and valuable data when different wound therapies are compared.

11. Acknowledgements

This dissertation is a result of the work conducted in Tampere University Hospital during the years 2008 to 2011. During this time there was close collaboration with the Department of Electrical Engineering and Automation of the University of Vaasa and the instrument presented in this study, the Spectrocutometer, is a result of this collaboration.

First and foremost, I want to express my gratitude to my supervisor and mentor Docent Hannu Kuokkanen. Hannu has guided and encouraged me in my scientific work. At the same time, throughout my residency in plastic surgery, he has been an outstanding teacher and has provided a personal example of the skills and the mindset of a great surgeon.

The co-operation with Dr Tech Petri Välisuo from the University of Vaasa has been seamless throughout this study. It is my great pleasure to thank Petri for his brilliant thinking and dedication to this project. It is fair to say that his input has directed my own scientific interests into an exciting yet unpredictable direction. I would also like to acknowledge Professor Jarmo Alander from the University of Vaasa for his contribution in this study, and his thorough revision of the original publications.

Docent Virve Koljonen and Docent Patrik Lassus from Helsinki University Hospital have made a significant contribution by revising this dissertation. I want to express my deepest appreciation for their in-detail corrections and comments that have improved the final result tremendously.

Undoubtedly, this scientific work has sometimes occupied my time, adding to the workload of my dear colleagues in the Department of Plastic Surgery. I want to thank them for their patience. Also, I wish to express my appreciation for my colleagues MD Minna Kääriäinen, MD Kimmo Tanttula and MD Kai Saariniemi, who contributed as volunteer observers in studies III and IV.

12. References

Aarabi S, Longaker MT, Gurtner GC. Hypertrophic scar formation following burns and trauma: new approaches to treatment. PLoS Med 4:e234,2007.

Ahlgrimm-Siess V, Horn M, Koller S, Ludwig R, Gerger A, Hofmann-Wellenhof R. Monitoring efficacy of cryotherapy for superficial basal cell carcinomas with in vivo reflectance confocal microscopy: a preliminary study. J Dermatol Sci 53:60-64,2009a.

Ahlgrimm-Siess V, Massone C, Scope A, Fink-Puches R, Richtig E, Wolf IH, Koller S, Gerger A, Smolle J, Hofmann-Wellenhof R. Reflectance confocal microscopy of facial lentigo maligna and lentigo maligna melanoma: a preliminary study. Br J Dermatol 161:1307-1316,2009b.

Al-Attar A, Mess S, Thomassen J, Kauffman L, Davison S. Keloid Pathogenesis and treatment. Plast Reconstr Surg 117:286-300,2005.

Amadeu T, Braune A, Mandarim-de-Lacerda C, Porto LC, Desmouliere A, Costa A. Vascularization pattern in hypertrophic scars and keloids: a stereological analysis. Pathol Res Pract 199:469-473,2003.

Anzarut A, Olson J, Singh P, Rowe BH, Tredget EE. The effectiveness of pressure garment therapy for the prevention of abnormal scarring after burn injury: a meta-analysis. J Plast Reconstr Aesthet Surg 62:77-84,2009.

Ardehali B, Nouraei SA, Van Dam H, Dex E, Wood S, Nduka C. Objective assessment of keloid scars with three-dimensional imaging: quantifying response to intralesional steroid therapy. Plast Reconstr Surg 119:556-561,2007.

Argirova M, Hadjinski O, Victorova A. Acticoat versus Allevyn as a split-thickness skin graft donor-site dressing. A prospective comparative study. Annals of Plastic Surgery 59:415-422,2007.

Ashcroft GS, Yang X, Glick AB, Weinstein M, Letterio JL, Mizel DE, Anzano M, Greenwell-Wild T, Wahl SM, Deng C, Roberts AB. Mice lacking Smad3 show

accelerated wound healing and an impaired local inflammatory response. Nat Cell Biol 1:260-266,1999.

Atiyeh BS, Costagliola M, Hayek SN. Keloid or hypertrophic scar: the controversy: review of the literature. Ann Plast Surg 54:676-680,2005.

Barnea Y, Amir A, Leshem D, Zaretski A, Weiss J, Shafir R, Gur E. Clinical comparative study of aquacel and paraffin gauze dressing for split-skin donor site treatment. Ann Plast Surg 53:132-136,2004.

Bartko JJ. The intraclass correlation coefficient as a measure of reliability. Psychol Rep 19:3-11,1966.

Baryza MJ, Baryza GA. The Vancouver scar scale. An administration tool and its interrater reliability. Journal of Burn Care and Rehabilitation 16:535-538,1995.

Bayat A, Bock O, Mrowietz U, Ollier WE, Ferguson MW. Genetic susceptibility to keloid disease and hypertrophic scarring: transforming growth factor beta1 common polymorphisms and plasma levels. Plast Reconstr Surg 111:535-543; discussion 544-536,2003a.

Bayat A, Bock O, Mrowietz U, Ollier WE, Ferguson MW. Genetic susceptibility to keloid disease: transforming growth factor beta receptor gene polymorphisms are not associated with keloid disease. Exp Dermatol 13:120-124,2004.

Bayat A, McGrouther DA, Ferguson MW. Skin scarring. BMJ 326:88-92,2003b.

Beausang E, Floyd H, Dunn KW, Orton CI, Ferguson MW. A new quantitative scale for clinical scar assessment. Plast Reconstr Surg 102:1954-1961,1998.

Beer TW, Baldwin HC, Goddard JR, Gallagher PJ, Wright DH. Angiogenesis in pathological and surgical scars. Hum Pathol 29:1273-1278,1998.

Bloemen MC. An objective device for measuring surface roughness of skin and scars. J Am Acad Dermatol Epub:2011.

Bloemen MC, van der Veer WM, Ulrich MM, van Zuijlen PP, Niessen FB, Middelkoop E. Prevention and curative management of hypertrophic scar formation. Burns 35:463-475,2009.

Bochko V. VP, Harju T., Alander J. Lower extremity ulcer image segmentation of visual and near-infrared imagery. Skin Res Technol 16:190-197,2010.

Bombaro KM, Engrav LH, Carrougher GJ, Wiechman SA, Faucher L, Costa BA, Heimbach DM, Rivara FP, Honari S. What is the prevalence of hypertrophic scarring following burns? Burns 29:299-302,2003.

Bon FX, Briand E, Guichard S, Couturaud B, Revol M, Servant JM, Dubertret L. Quantitative and kinetic evolution of wound healing through image analysis. IEEE Trans Med Imaging 19:767-772,2000.

Boyce ST, Supp AP, Wickett RR, Hoath SB, Warden GD. Assessment with the dermal torque meter of skin pliability after treatment of burns with cultured skin substitutes. J Burn Care Rehabil 21:55-63,2000.

Brandt MG, Moore CC, Micomonaco D, Fung K, Franklin JH, Yoo J, Doyle PC. A Prospective Randomized Evaluation of Scar Assessment Measures. The Laryngoscope 119:841-845,2009.

Broughton G, 2nd, Janis JE, Attinger CE. The basic science of wound healing. Plast Reconstr Surg 117:12S-34S,2006.

Brown BC, McKenna SP, Siddhi K, McGrouther DA, Bayat A. The hidden cost of skin scars: quality of life after skin scarring. J Plast Reconstr Aesthet Surg 61:1049-1058,2008a.

Brown JJ, Ollier W, Arscott G, Ke X, Lamb J, Day P, Bayat A. Genetic susceptibility to keloid scarring: SMAD gene SNP frequencies in Afro-Caribbeans. Exp Dermatol 17:610-613,2008b.

Brown JJ, Ollier WE, Arscott G, Bayat A. Association of HLA-DRB1* and keloid disease in an Afro-Caribbean population. Clin Exp Dermatol 35:305-310,2010.

Brown JJ, Ollier WE, Thomson W, Bayat A. Positive association of HLA-DRB1*15 with keloid disease in Caucasians. Int J Immunogenet 35:303-307,2008c.

Brusselaers M, Pirayesh A, Hoeksema P, Verbelen J, Blot S, Monstray S. Burn scar assessment: A systematic review of different scar scales. Journal of surgical research 164:115-123,2010.

Burd A, Kwok CH, Hung SC, Chan HS, Gu H, Lam WK, Huang L. A comparative study of the toxicity of silver-based dressings in monolayer cell, tissue explant, and animal models. Wound Repair Regen 15:94-104,2007.

Butler PD, Longaker MT, Yang GP. Current progress in keloid research and treatment. J Am Coll Surg 206:731-741,2008.

Caggiati A, Rosi C, Franceschini M, Innocenzi D. The nature of skin pigmentations in chronic venous insufficiency: a preliminary report. Eur J Vasc Endovasc Surg 35:111-118,2008.

Claridge E, Cotton S, Hall P, Moncrieff M. From colour to tissue histology: physics based interpretation of images of pigmented skin lesions. MICCAI 1:730-738,2002.

Claridge E. CS, Hall P., Moncrieff M. From colour to tissue histology: physics based interpretation of images of pigmented skin lesions. MICCAI 1:730-738,2002.

Clarys P, Alewaeters K, Lambrecht R, Barel AO. Skin color measurements: comparison between three instruments: the Chromameter(R), the DermaSpectrometer(R) and the Mexameter(R). Skin Res Technol 6:230-238,2000.

Cleary C SA, Nick G. Reliability of the Skin Compliance Device in the Assessment of Scar Pliability. Journal of Hand Therapy 20:232-238,2007.

Cooper GF. The computational complexity of probabilistic inference using Bayesian belief networks. Artificial intelligence 42:393-405,1990.

Corica GF, Wigger NC, Edgar DW, Wood FM, Carroll S. Objective measurement of scarring by multiple assessors: is the tissue tonometer a reliable option? J Burn Care Res 27:520-523,2006.

Crowe JM, Simpson K, Johnson W, Allen J. Reliability of photographic analysis in determining change in scar appearance. J Burn Care Rehabil 19:183-186,1998.

Dawson JB, Barker DJ, Ellis DJ, Grassam E, Cotterill JA, Fisher GW, Feather JW. A theoretical and experimental study of light absorption and scattering by in vivo skin. Phys Med Biol 25:695-709,1980.

Dedovic Z, Koupilova I, Brychta P. Time trends in incidence of hypertrophic scarring in children treated for burns. Acta Chir Plast 41:87-90,1999.

Draaijers LJ, Botman YA, Tempelman FR, Kreis RW, Middelkoop E, van Zuijlen PP. Skin elasticity meter or subjective evaluation in scars: a reliability assessment. Burns 30:109-114,2004a.

Draaijers LJ, Tempelman FR, Botman YA, Kreis RW, Middelkoop E, van Zuijlen PP. Colour evaluation in scars: tristimulus colorimeter, narrow-band simple reflectance meter or subjective evaluation? Burns 30:103-107,2004b.

Draaijers LJ, Tempelman FR, Botman YA, Tuinebreijer WE, Middelkoop E, Kreis RW, van Zuijlen PP. The patient and observer scar assessment scale: a reliable and feasible tool for scar evaluation. Plast Reconstr Surg 113:1960-1965; discussion 1966-1967,2004c.

Dressler J, Busuttil A, Koch R, Harrison DJ. Sequence of melanocyte migration into human scar tissue. Int J Legal Med 115:61-63,2001.

Duncan JA, Bond JS, Mason T, Ludlow A, Cridland P, O'Kane S, Ferguson MW. Visual analogue scale scoring and ranking: a suitable and sensitive method for assessing scar quality? Plast Reconstr Surg 118:909-918,2006.

Dunkin CS, Pleat JM, Gillespie PH, Tyler MP, Roberts AH, McGrouther DA. Scarring occurs at a critical depth of skin injury: precise measurement in a graduated dermal scratch in human volunteers. Plast Reconstr Surg 119:1722-1732; discussion 1733-1724,2007.

Durani P, McGrouther DA, Ferguson MW. Current scales for assessing human scarring: a review. J Plast Reconstr Aesthet Surg 62:713-720,2009.

Ehrlich HP, Desmouliere A, Diegelmann RF, Cohen IK, Compton CC, Garner WL, Kapanci Y, Gabbiani G. Morphological and immunochemical differences between keloid and hypertrophic scar. Am J Pathol 145:105-113,1994.

English RS, Shenefelt PD. Keloids and hypertrophic scars. Dermatol Surg 25:631-638,1999.

Estes JM, Vande Berg JS, Adzick NS, MacGillivray TE, Desmouliere A, Gabbiani G. Phenotypic and functional features of myofibroblasts in sheep fetal wounds. Differentiation 56:173-181,1994.

Ferguson MW, Duncan J, Bond J, Bush J, Durani P, So K, Taylor L, Chantrey J, Mason T, James G, Laverty H, Occleston NL, Sattar A, Ludlow A, O'Kane S. Prophylactic administration of avotermin for improvement of skin scarring: three double-blind, placebo-controlled, phase I/II studies. Lancet 373:1264-1274,2009.

Fleiss JL, Shrout PE. The effects of measurement errors on some multivariate procedures. Am J Public Health 67:1188-1191,1977.

Fong SS, Hung LK, Cheng JC. The cutometer and ultrasonography in the assessment of postburn hypertrophic scar--a preliminary study. Burns 23 Suppl 1:S12-18,1997.

Fujiwara M, Muragaki Y, Ooshima A. Upregulation of transforming growth factor-beta1 and vascular endothelial growth factor in cultured keloid fibroblasts: relevance to angiogenic activity. Arch Dermatol Res 297:161-169,2005.

Gerger A, Hofmann-Wellenhof R, Samonigg H, Smolle J. In vivo confocal laser scanning microscopy in the diagnosis of melanocytic skin tumours. Br J Dermatol 160:475-481,2009.

Golinko M, Margolis D, Tal A, Hoffstad O, Boulton AJ, Brem H. Preliminary development of a diabetic foot ulcer database from a wound electronic medical record: A tool to decrease limb amputations. Wound Repair Regen 17:657-665,2009.

Gonzalez FJ, Martinez-Escaname M, Munoz RI, Torres-Alvarez B, Moncada B. Diffuse reflectance spectrophotometry for skin phototype determination. Skin Res Technol 16:397-400,2010.

Hansen GL, Sparrow EM, Kommamuri N, Iaizzo PA. Assessing wound severity with color and infrared imaging of reactive hyperemia. Wound Repair Regen 4:386-392,1996.

Harte D, Gordon J, Shaw M, Stinson M, Porter-Armstrong A. The use of pressure and silicone in hypertrophic scar management in burns patients: a pilot randomized controlled trial. J Burn Care Res 30:632-642,2009.

Idriss N, Maibach HI. Scar assessment scales: a dermatologic overview. Skin Res Technol 15:1-5,2009.

Innes ME, Umraw N, Fish JS, Gomez M, Cartotto RC. The use of silver coated dressings on the donor site wounds: a prospective, controlled matched pair study. Burns 27:621-627,2001.

Janis JE, Kwon RK, Lalonde DH. A practical guide to wound healing. Plast Reconstr Surg 125:230e-244e,2010.

Jung B, Choi B, Durkin AJ, Kelly KM, Nelson JS. Characterization of port wine stain skin erythema and melanin content using cross-polarized diffuse reflectance imaging. Lasers Surg Med 34:174-181,2004.

Junker JP, Kratz C, Tollback A, Kratz G. Mechanical tension stimulates the transdifferentiation of fibroblasts into myofibroblasts in human burn scars. Burns 34:942-946,2008.

Karagoz H, Yuksel F, Ulkur E, Evinc R. Comparison of efficacy of silicone gel, silicone gel sheeting, and topical onion extract including heparin and allantoin for the treatment of postburn hypertrophic scars. Burns 2009.

Kilinc H, Sensoz O, Ozdemir R, Unlu RE, Baran C. Which dressing for split-thickness skin graft donor sites? Ann Plast Surg 46:409-414,2001.

Kim W, Kang WH. Role of dermal melanocytes in cutaneous pigmentation of stasis dermatitis: a histopathological study of 20 cases. Journal of Korean Medical Science 17:648-654,2002.

Koller S, Gerger A, Ahlgrimm-Siess V, Weger W, Smolle J, Hofmann-Wellenhof R. In vivo reflectance confocal microscopy of erythematosquamous skin diseases. Exp Dermatol 18:536-540,2009.

Kose O, Waseem A. Keloids and hypertrophic scars: are they two different sides of the same coin? Dermatol Surg 34:336-346,2008.

Krummel TM, Michna BA, Thomas BL, Sporn MB, Nelson JM, Salzberg AM, Cohen IK, Diegelmann RF. Transforming growth factor beta (TGF-beta) induces fibrosis in a fetal wound model. J Pediatr Surg 23:647-652,1988.

Kurokawa N, Ueda K, Tsuji M. Study of microvascular structure in keloid and hypertrophic scars: density of microvessels and the efficacy of three-dimensional vascular imaging. J Plast Surg Hand Surg 44:272-277,2010.

Ladak A, Tredget EE. Pathophysiology and management of burn scars. Clinics in Plastic Surgery 36:661-674,2009.

Laplaud A, Blaizot X, Gaillard C, Morice A, Lebreuilly I, Clement C, Parienti J, Dompmartin A. Wound debridement: Comparative reliability of three methods for measuring fibrin percentage in chronic wounds. Wound Repair Regen 18:13-20,2010.

Larson BJ, Longaker MT, Lorenz HP. Scarless fetal wound healing: a basic science review. Plast Reconstr Surg 126:1172-1180,2010.

Latreille J, Gardinier S, Ambroisine L, Mauger E, Tenenhaus M, Guehenneux S, Morizot F, Tschachler E, Guinot C. Influence of skin colour on the detection of cutaneous erythema and tanning phenomena using reflectance spectrophotometry. Skin Res Technol 13:236-241,2007.

Lau JC, Li-Tsang CW, Zheng YP. Application of tissue ultrasound palpation system (TUPS) in objective scar evaluation. Burns 31:445-452,2005.

Lee JY, Yang CC, Chao SC, Wong TW. Histopathological differential diagnosis of keloid and hypertrophic scar. Am J Dermatopathol 26:379-384,2004.

Leung KS, Sher A, Clark JA, Cheng JC, Leung PC. Microcirculation in hypertrophic scars after burn injury. J Burn Care Rehabil 10:436-444,1989.

Li-Tsang CW, Lau JC, Chan CC. Prevalence of hypertrophic scar formation and its characteristics among the Chinese population. Burns 31:610-616,2005.

Li-Tsang CW, Lau JC, Liu SK. Validation of an objective scar pigmentation measurement by using a spectrocolorimeter. Burns 29:779-784,2003.

Longaker MT, Chiu ES, Harrison MR, Crombleholme TM, Langer JC, Duncan BW, Adzick NS, Verrier ED, Stern R. Studies in fetal wound healing. IV. Hyaluronic acid-stimulating activity distinguishes fetal wound fluid from adult wound fluid. Ann Surg 210:667-672,1989.

Longaker MT, Whitby DJ, Adzick NS, Crombleholme TM, Langer JC, Duncan BW, Bradley SM, Stern R, Ferguson MW, Harrison MR. Studies in fetal wound healing, VI. Second and early third trimester fetal wounds demonstrate rapid collagen deposition without scar formation. J Pediatr Surg 25:63-68; discussion 68-69,1990.

Lorber A, Wiltgen M, Hofmann-Wellenhof R, Koller S, Weger W, Ahlgrimm-Siess V, Smolle J, Gerger A. Correlation of image analysis features and visual morphology in melanocytic skin tumours using in vivo confocal laser scanning microscopy. Skin Res Technol 15:237-241,2009.

Lorenz HP, Adzick NS. Scarless skin wound repair in the fetus. West J Med 159:350-355,1993.

Lye I, Edgar DW, Wood FM, Carroll S. Tissue tonometry is a simple, objective measure for pliability of burn scar: is it reliable? J Burn Care Res 27:82-85,2006.

Malpass, Snelling, Tron. Comparison of donor-site healing under Xeroform and Jelonet dressings: Unexpected findings. Plastic and Reconstructive Surgery 112:430-439,2003.

Masters M, McMahon M, Svens B. Reliability testing of a new scar assessment tool, Matching Assessment of Scars and Photographs (MAPS). J Burn Care Rehabil 26:273-284,2005.

Merkel JR, DiPaolo BR, Hallock GG, Rice DC. Type I and type III collagen content of healing wounds in fetal and adult rats. Proc Soc Exp Biol Med 187:493-497,1988.

Miskelly G, Wagner J. Using spectral information in forensic imaging. Forensic Sci Int 155:112-118,2005.

Miyamoto K, Takiwaki H, Hillebrand GG, Arase S. Development of a digital imaging system for objective measurement of hyperpigmented spots on the face. Skin Res Technol 8:227-235,2002a.

Miyamoto K, Takiwaki H, Hillebrand GG, Arase S. Utilization of a high-resolution digital imaging system for the objective and quantitative assessment of hyperpigmented spots on the face. Skin Res Technol 8:73-77,2002b.

Molnar JA, Lew WK, Rapp DA, Gordon ES, Voignier D, Rushing S, Willner W. Use of Standardized, Quantitative Digital Photography in a Multicenter Web-based Study. Eplasty 9:e4,2009.

Monafo W, Bessey, P. Wound Care. In Herndon DN, ed. Total Burn Care. W.B. Saunders, 1996: 88-98.

Moncrieff M, Cotton S, Claridge E, Hall P. Spectrophotometric intracutaneous analysis: a new technique for imaging pigmented skin lesions. Br J Dermatol 146:448-457,2002.

Muir IF. On the nature of keloid and hypertrophic scars. Br J Plast Surg 43:61-69,1990.

Mustoe TA, Cooter RD, Gold MH, Hobbs FD, Ramelet AA, Shakespeare PG, Stella M, Teot L, Wood FM, Ziegler UE. International clinical recommendations on scar management. Plast Reconstr Surg 110:560-571,2002.

Myllymäki P. ST, Tirri H., Uronen P. B-course: a web-based tool for Bayesian and causal data analysis. J Artificial Int Tools 11:369-388,2002.

Namazi MR, Fallahzadeh MK, Schwartz RA. Strategies for prevention of scars: what can we learn from fetal skin? Int J Dermatol 50:85-93,2011.

Nedelec B, Correa JA, Rachelska G, Armour A, LaSalle L. Quantitative measurement of hypertrophic scar: interrater reliability and concurrent validity. J Burn Care Res 29:501-511,2008a.

Nedelec B, Correa JA, Rachelska G, Armour A, LaSalle L. Quantitative measurement of hypertrophic scar: intrarater reliability, sensitivity, and specificity. J Burn Care Res 29:489-500,2008b.

Nguyen DQ, Potokar T, Price P. A review of current objective and subjective scar assessment tools. J Wound Care 17:101-102, 104-106,2008.

Nishidate I, Wiswadarma A, Hase Y, Tanaka N, Maeda T, Niizeki K, Aizu Y. Noninvasive spectral imaging of skin chromophores based on multiple regression analysis aided by Monte Carlo simulation. Opt Lett 36:3239-3241,2011.

Nouarei S, Frame J, Nduka C. Uses and abuses of digital imaging in plastic surgery. International Journal of Surgery 3:254-257,2005.

Nouveau-Richard S, Monot M, Bastien P, de Lacharriere O. In vivo epidermal thickness measurement: ultrasound vs. confocal imaging. Skin Res Technol 10:136-140,2004.

O'Brien L, Pandit A. Silicon gel sheeting for preventing and treating hypertrophic and keloid scars. Cochrane Database Syst Rev CD003826,2006.

Occleston NL, O'Kane S, Goldspink G, Ferguson M. New therapeutics for the prevention and reduction of scarring. Drug Discovery Today 13:973-981,2008.

Ogawa R. The most current algorithms for the treatment and prevention of hypertrophic scars and keloids. Plast Reconstr Surg 125:557-568,

Ogawa R. The most current algorithms for the treatment and prevention of hypertrophic scars and keloids. Plast Reconstr Surg 125:557-568,2010.

Ogawa R. Mechanobiology of scarring. Wound Repair Regen 19 Suppl 1:s2-9,2011.

Ogawa R, Akaishi S, Huang C, Dohi T, Aoki M, Omori Y, Koike S, Kobe K, Akimoto M, Hyakusoku H. Clinical applications of basic research that shows reducing skin tension could prevent and treat abnormal scarring: the importance of fascial/subcutaneous tensile reduction sutures and flap surgery for keloid and hypertrophic scar reconstruction. J Nippon Med Sch 78:68-76,2011.

Oien R, Håkansson A, Hansen B, Bjellerup M. Measuring the size of ulcers by planimetry: a useful method in the clinical setting. J Wound Care 11:165-168,2002.

Oliveira GV, Chinkes D, Mitchell C, Oliveras G, Hawkins HK, Herndon DN. Objective assessment of burn scar vascularity, erythema, pliability, thickness, and planimetry. Dermatol Surg 31:48-58,2005.

Paddle-Ledinek JE, Nasa Z, Cleland HJ. Effect of different wound dressings on cell viability and proliferation. Plast Reconstr Surg 117:110S-118S; discussion 119S-120S,2006.

Paletta C, Pokorny, J., Rumbolo, P. Skin grafts. In Mathes S, ed. Plastic Surgery. Philadelphia: Saunders Elsevier, 2006: 293-316.

Papazoglou E, Zubkov L, Mao X, Neidrauer M, Rannou N, Weingarten MS. Image analysis of chronic wounds for determining the surface area. Wound Repair Regen 18:349-358,2010.

Papazoglou ES, Weingarten MS, Zubkov L, Neidrauer M, Zhu L, Tyagi S, Pourrezaei K. Changes in optical properties of tissue during acute wound healing in an animal model. J Biomed Opt 13:044005,2008.

Payette JR, Kohlenberg E, Leonardi L, Pabbies A, Kerr P, Liu KZ, Sowa MG. Assessment of skin flaps using optically based methods for measuring blood flow and oxygenation. Plast Reconstr Surg 115:539-546,2005.

Peacock E, Madden J, Trier W. Biologic basis for the treatment of keloids and hypertrophic scars. South Med J 63:755-760,1970.

Pearl J. Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann 1988.

Perry DM, McGrouther DA, Bayat A. Current tools for noninvasive objective assessment of skin scars. Plast Reconstr Surg 126:912-923,2010.

Pfurtscheller K, Zobel G, Roedl S, Trop M. Use of Suprathel dressing in a young infant with TEN. Pediatr Dermatol 25:541-543,2008.

Powers PS, Sarkar S, Goldgof DB, Cruse CW, Tsap LV. Scar assessment: current problems and future solutions. J Burn Care Rehabil 20:54-60; discussion 53,1999.

Prasanna M, Mishra P, Thomas C. Delayed primary closure of the burn wounds. Burns 30:169-175,2004.

Rennekampff HO, Rabbels J, Reinhard V, Becker ST, Schaller HE. Comparing the Vancouver Scar Scale with the cutometer in the assessment of donor site wounds treated with various dressings in a randomized trial. J Burn Care Res 27:345-351,2006.

Rennert R, Golinko M, Kaplan D, Flattau A, Brem H. Standardization of wound photography using the wound electronic medical record. Adv Skin Wound Care 22:32-38,2009.

Richtig E, Ahlgrimm-Siess V, Koller S, Gerger A, Horn M, Smolle J, Hofmann-Wellenhof R. Follow-up of actinic keratoses after shave biopsy by in-vivo reflectance confocal microscopy--a pilot study. J Eur Acad Dermatol Venereol 24:293-298,

Richtig E, Gerger A, El-Shabrawi-Caelen L, Szkandera J, Hofmann-Wellenhof R. Reflectance confocal microscopy in early diagnosis of cutaneous metastases of breast cancer. Gynecol Oncol 115:510-511,2009.

Rowlatt U. Intrauterine wound healing in a 20 week human fetus. Virchows Arch 381:353-361,1979.

Ryssel H, Andreas Radu C, Germann G, Kloeters O, Riedel K, Otte M, Kremer T. Suprathel-antiseptic matrix: in vitro model for local antiseptic treatment? Adv Skin Wound Care 24:64-67,2011.

Ryssel H, Germann G, Riedel K, Reichenberger M, Hellmich S, Kloeters O. Suprathel-acetic acid matrix versus acticoat and aquacel as an antiseptic dressing: an in vitro study. Ann Plast Surg 65:391-395,2010.

Schwarze H, Kuntscher M, Uhlig C, Hierlemann H, Prantl L, Noack N, Hartmann B. Suprathel, a new skin substitute, in the management of donor sites of split-thickness skin grafts: results of a clinical study. Burns 33:850-854,2007.

Schwarze H, Kuntscher M, Uhlig C, Hierlemann H, Prantl L, Ottomann C, Hartmann B. Suprathel, a new skin substitute, in the management of partial-thickness burn wounds: results of a clinical study. Ann Plast Surg 60:181-185,2008.

Serghiou M, Cowan A, Whitehead C. Rehabilitation after a burn injury. Clin Plast Surg 36:675-686,2009.

Serup J, Agner T. Colorimetric quantification of erythema--a comparison of two colorimeters (Lange Micro Color and Minolta Chroma Meter CR-200) with a clinical scoring scheme and laser-Doppler flowmetry. Clin Exp Dermatol 15:267-272,1990.

Shai A, Avrin V, Zvulunov A, Dinstein I. Scaled Healing Assessment Index: a novel method for measuring the degree of wound bed preparation. Skin Res Technol 13:227-235,2007.

Shih B, Bayat A. Genetics of keloid scarring. Arch Dermatol Res 302:319-339,2010.

Shin D, Minn KW. The effect of myofibroblast on contracture of hypertrophic scar. Plast Reconstr Surg 113:633-640,2004.

Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 86:420-428,1979.

Singer A, Arora B, Dagum A, Valentine S, Hollander J. Developement and validation of a novel scar evaluation scale. Plast Reconstr Surg 120:1982-1987,2007.

Spann K, Mileski WJ, Atiles L, Purdue G, Hunt J. The 1996 Clinical Research award. Use of a pneumatonometer in burn scar assessment. J Burn Care Rehabil 17:515-517,1996.

Sugata K, Kitahara T, Takema Y. Changes of human skin in subepidermal wound healing process. Skin Res Technol 14:436-439,2008a.

Sugata K, Nishijima T, Kitahara T, Takema Y. Confocal laser microscopic imaging of conspicuous facial pores in vivo: relation between the appearance and the internal structure of skin. Skin Res Technol 14:208-212,2008b.

Sullivan T, Smith J, Kermode J, McIver E, Courtemanche DJ. Rating the burn scar. J Burn Care Rehabil 11:256-260,1990.

Takiwaki H, Miyaoka Y, Arase S. Analysis of the absorbance spectra of skin lesions as a helpful tool for detection of major pathophysiological changes. Skin Research and Technology 10:130-135,2004.

Takiwaki H, Miyaoka Y, Kohno H, Arase S. Graphic analysis of the relationship between skin colour change and variations in the amounts of melanin and haemoglobin. Skin Research and Technology 8:78-83,2002a.

Takiwaki H, Miyaoka Y, Kohno H, Arase S. Graphic analysis of the relationship between skin colour change and variations in the amounts of melanin and haemoglobin. Skin Res Technol 8:78-83,2002b.

Takiwaki H, Miyaoka Y, Skrebova N, Kohno H, Arase S. Skin reflectance-spectra and colour-value dependency on measuring-head aperture area in ordinary reflectance spectrophotometry and tristimulus colourimetry. Skin Res Technol 8:94-97,2002c.

Taylor S, Westerhof W, Im S, Lim J. Noninvasive techniques for the evaluation of skin color. J Am Acad Dermatol 54:S282-290,2006.

Tehrani H, Moncrieff M, Philp B, Dziewulski P. Spectrophotometric intracutaneous analysis: a novel imaging technique in the assessment of acute burn depth. Ann Plast Surg 61:437-440,2008.

Teot L. Clinical evaluation of scars. Wound Repair Regen 10:93-97,2002.

Truong PT, Abnousi F, Yong CM, Hayashi A, Runkel JA, Phillips T, Olivotto IA. Standardized assessment of breast cancer surgical scars integrating the Vancouver Scar Scale, Short-Form McGill Pain Questionnaire, and patients' perspectives. Plast Reconstr Surg 116:1291-1299,2005.

Truong PT, Lee JC, Soer B, Gaul CA, Olivotto IA. Reliability and validity testing of the Patient and Observer Scar Assessment Scale in evaluating linear scars after breast cancer surgery. Plast Reconstr Surg 119:487-494,2007.

Uhlig C, Rapp M, Hartmann B, Hierlemann H, Planck H, Dittel KK. Suprathel-an innovative, resorbable skin substitute for the treatment of burn victims. Burns 33:221-229,2007.

Uysal AC, Alagoz MS, Orbay H, Sensoz O. An alternative dressing material for the split-thickness skin graft donor site: oxidized regenerated cellulose. Ann Plast Surg 57:60-64,2006.

van de Kar AL, Corion LU, Smeulders MJ, Draaijers LJ, van der Horst CM, van Zuijlen PP. Reliable and feasible evaluation of linear scars by the Patient and Observer Scar Assessment Scale. Plast Reconstr Surg 116:514-522,2005.

Van den Kerckhove E, Staes F, Flour M, Stappaerts K, Boeckx W. Reproducibility of repeated measurements on post-burn scars with Dermascan C. Skin Res Technol 9:81-84,2003.

van Zuijlen PP, Angeles AP, Kreis RW, Bos KE, Middelkoop E. Scar assessment tools: implications for current research. Plast Reconstr Surg 109:1108-1122,2002.

Vercelli S, Ferriero G, Sartorio F, Stissi V, Franchignoni F. How to assess postsurgical scars: a review of outcome measures. Disabil Rehabil 31:2055-2063,2009.

Verhaegen PD, van der Wal MB, Middelkoop E, van Zuijlen PP. Objective scar assessment tools: a clinimetric appraisal. Plast Reconstr Surg 127:1561-1570,2011.

Voineskos SH, Ayeni OA, McKnight L, Thoma A. Systematic review of skin graft donor-site dressings. Plast Reconstr Surg 124:298-306,2009.

Wolfram D, Tzankov A, Pulzl P, Piza-Katzer H. Hypertrophic scars and keloids--a review of their pathophysiology, risk factors, and therapeutic management. Dermatol Surg 35:171-181,2009.

Välisuo P, Harju T, Alander J. Reflectance measurement using digital camera and a protecting dome with built in light source. Journal of Biophotonics Accepted for publication.:2010.

Välisuo P, Kaartinen I, Tuchin V, Alander J. New closed-form approximation for skin chromophore mapping. J Biomed Opt 16:046012,2011.

Yamaguchi Y, Brenner M, Hearing V. The regulation of skin pigmentation. The Journal of Biological Chemistry 282:27557-27561,2007.

Yeong EK, Mann R, Engrav LH, Goldberg M, Cain V, Costa B, Moore M, Nakamura D, Lee J. Improved burn scar assessment with use of a new scar-rating scale. J Burn Care Rehabil 18:353-355; discussion 352,1997.

Yun I, Lee W, Rah D, Kim Y, Park B. Skin color analysis using a spectrophotometer in Asians. Skin Res Technol 16:311-315,2010.

The colour of blood in skin: a comparison of Allen's test and photonics simulations

Petri Välisuo[1], Ilkka Kaartinen[2], Hannu Kuokkanen[2] and Jarmo Alander[1]

[1]*Department of Electrical Engineering and Automation, University of Vaasa, Wolffintie, Vaasa, Finland and* [2]*Department of Plastic Surgery, Tampere University Hospital, Tampere, Finland*

Background: The colour of the skin reflects many physiological and pathological states of an individual. Usually, the skin colour is examined by the bare eye alone. Several scaling systems have been developed to quantify the sensory evaluation of skin colour. In this work, the reflectance of the skin is measured directly using an objective instrument. Haemoglobin inside the dermal circulation is one of the key factors of skin colour and it also has a major role in the appearance of many skin lesions and scars. To quantitatively measure and analyse such conditions, the relation between the skin colour and the haemoglobin concentration in the skin needs to be resolved.
Methods: To examine the effect of blood concentration on the skin colour, five Allen's tests were performed on 20 persons. The skin colour change was measured using a spectrophotometer by changing the blood concentration by the Allen's test. Light interaction with the skin was simulated with a Monte Carlo model, tuning the blood concentration parameter until the simulated and the measured spectra

matched, yielding the relationship between the skin colour and the blood concentration.
Results: The simulation produced spectra similar to those measured. The change in the blood concentration in the simulation model and in the skin produced changes similar to the spectra. The reflectance of the skin was found to be a nonlinear function of the blood concentration.
Conclusion: The relationship found between skin colour and blood concentration makes it possible to quantify those skin conditions expressed by blood volume better than plain colour.

Key words: spectrophotometry – medical imaging – colour vision – haemodynamics – skin imaging – dermatology – Monte Carlo simulation

© 2010 John Wiley & Sons A/S
Accepted for publication 23 January 2010

THE SKIN is the largest organ of a human being. Still, it is only several millimetres thick and its surface is easily accessible everywhere. The combination of thinness and low optical absorption and sufficient scattering allows visible light to penetrate deep into the skin and scatter back to its surface. The skin reacts to many kinds of local and global stimuli by adjusting the perfusion and thus its blood volume. Injuries and other disorders may also permanently change the physical structures of the skin and bring additional chromophores into the injured area. These changes affect both the scattering coefficient and the absorption spectra of the skin, which can be seen as a difference in colour on the surface. The human colour vision may have been especially adapted to observe changes in skin colour (1). Still, the interpretation of the colour sensation is always subjective and it is better in providing qualitative than quantitative measures (2). Therefore, ma-

chine vision provides essential improvements to colour-based skin analysis. Noninvasive optical measurements are a convenient method to obtain information on the state of the human skin. It may help in diagnosing and following up various skin disorders and other diseases that may manifest as changes in the appearance of the skin (3).

The reason why optical measurements are not used more is that the relation between the skin structure and chromophore concentrations is extremely complicated. The optics of the skin have been studied perhaps for a century. A recently published review (4) covers many areas of the topic and so does Baranoski and Krishnaswamy in (5). Although the colours of the skin chromophores are well known (6–8), the light interaction with skin is so complicated that it cannot be completely solved mathematically, but merely approximated with suitable simplifications. Several simplifications have been used

to model light behaviour in skin, Kubelka–Munk theory, diffusion theory and Monte Carlo multi-layer (MCML) simulation being the most regularly used. The MCML simulation is the most accurate, but it may often be inconvenient and slow.

Tsumura et al. (9, 10) and Claridge et al. (11) suggest that the possible normal skin colour spans a two-dimensional space in the three-dimensional colour space. Therefore, the skin colour can be mostly described by the absorbance of melanin and the absorbance of haemoglobin alone. This model may be accurate when producing different skin colours but these axes are not directly relational to the pigment concentrations, even in the logarithmic space, due to the non-linear nature of skin. Takiwaki et al. (3) measured the spectra of many skin disorders and compared the spectra with normal skin colour. They concluded that the spectral difference in absorbance (SDA) yields useful information about skin disorders. They were able to show that the levels and the shapes of the SDA spectra correlate with the melanin and haemoglobin levels. Yet, the exact shapes and values of the SDA profiles due to haemoglobin- and melanin-level changes are still unknown. Further research is needed to gain a better understanding of the relationships of the reflectance and the skin structure and chromophore concentrations.

There are several sensors that are suitable for spectrophotometric measurements and imaging. Imaging sensors that are based on tri-stimulus or monochromatic sensors are often more affordable and easier to use than actual spectrophotometric cameras. The estimation of skin chromophores does not require excessive number of wavelengths, as each chromophore has a smooth spectrum in the VIS-NIR range (6–8). The sensitivity curves of the sensors and the geometries of the light source and the detector influence the measurements (12). These factors need to be carefully examined to provide more accurate measurements.

In this article, the relationships between skin chromophore concentrations and skin colour are examined in spectral and in three tri-stimulus colour domains. The examination is based on a simulation model similar to the one used in (13, 14). The results of the simulations are compared with spectrophotometric measurements during five Allen's tests performed on 20 persons.

Materials and Methods

Skin reflectance measurement
To measure the reflectance of the skin, it is illuminated with a light source, whose intensity at each wavelength, λ, is $E(\lambda)$. The reflectance of the skin $R(\lambda)$ describes the skin property to remit back the incident light. The remitted light is measured with a sensor, whose sensitivity is $S(\lambda)$. The response, v, of the measurement system is therefore

$$v = \int S(\lambda)E(\lambda)R(\lambda)d\lambda \qquad (1)$$

The bandwidth of the sensor $S(\lambda)$ and the illumination spectra $E(\lambda)$ determine the characteristics of the measurement.

The spectral measurement system measures several narrow band channels. The bandwidths of these channels use either narrowband light sources or narrowband sensors. To record n pure reflectances, R_n, the characteristics of the channel are compensated by calibrating the values with the values of the known reflectance R_w, usually a reference white:

$$R_n = \frac{\int S_n(\lambda)E_n(\lambda)R(\lambda)}{\int S_n(\lambda)E_n(\lambda)R_w} R_w \qquad (2)$$

The bandwidth of the light source and the detector together defines the bandwidths of each channel. The measured reflectance of the channel is the weighted average of the reflectance of the sample within the channel bandwidth. The observation of narrow reflectance peaks or notches requires many narrow channels, whereas fewer and wider bands are sufficient if the spectrum is smoother, which is the case in this study.

Tri-stimulus values
The spectral sensors, which mimic human colour vision, are common and easily affordable. Even though the channel bandwidths of these tri-stimulus sensors are wide, they are in many cases satisfactory for skin analysis, as the chromophores in skin have quite smooth spectra.

The human eye has receptors, called cone cells, for short, middle and long wavelengths. Thus, in principle, three parameters are enough to approximate a colour sensation. The tri-stimulus values of a colour are the amounts of three primary colours in a three-component additive colour model needed to match that test colour.

The tri-stimulus values are most often given in the CIE 1931 colour space, in which they are denoted by X, Y and Z (15).

Any specific method for associating tri-stimulus values with each colour is called a colour space. CIE XYZ, one of many such spaces, is special because it is based on direct measurements of human visual perception, and serves as the basis from which many other colour spaces are defined.

The responses of the channels of the tri-stimulus sensors mimic human colour vision and thus approximate the three CIE standard colourimetric observer (SCO) curves.

$$X = \int_{380}^{680} x(\lambda)E(\lambda)R(\lambda)d\lambda \qquad (3)$$

$$Y = \int_{380}^{680} y(\lambda)E(\lambda)R(\lambda)d\lambda \qquad (4)$$

$$Z = \int_{380}^{680} z(\lambda)E(\lambda)R(\lambda)d\lambda \qquad (5)$$

where $x(\lambda)$, $y(\lambda)$ and $z(\lambda)$ are the CIE SCO curves (15), and X, Y and Z are the resulting colour values in the XYZ space.

The XYZ colour values can be further transformed to other tri-stimulus colour spaces, such as RGB and CIE-$L^*a^*b^*$, using matrix transformations and additional nonlinear gamma correction for RGB.

The skin model
The model used in Välisuo and Alander (13) was used in this study too. It is shown in Fig. 1.

The model consists of the model structure and its model parameters. The parameters of each layer are the absorption coefficient, the scattering coefficient and the anisotropy. To be able to estimate the absorption coefficients, the skin chromophores and their concentrations are needed. The chromophores in the skin are either deposited in skin cells or are carried with blood. The concentrations of chromophores in blood vary with the blood fraction, while the concentration of the chromophores deposited in skin cells remains rather constant in time. The most important chromophores of the skin are haemoglobin and melanin. Bilirubin and β-carotene may have a significant effect too. Water has significant effect for wavelengths longer than 800 nm, but it does not influence the skin colour.
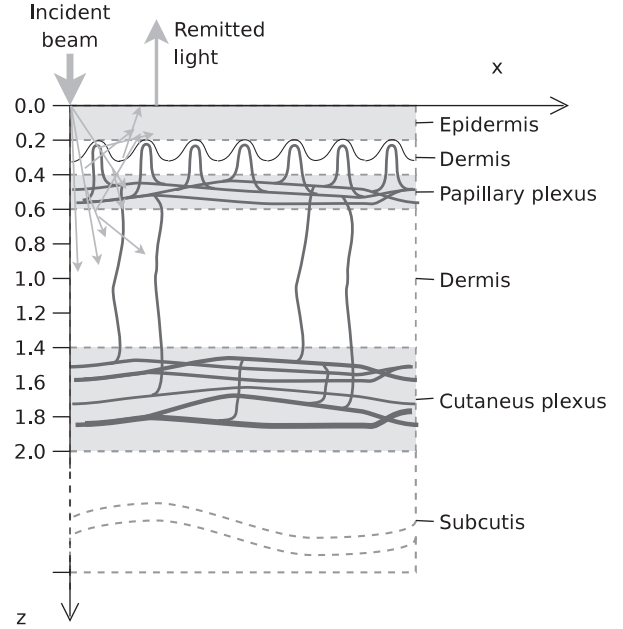
Fig. 1. The model of the skin. The names of the layers are on the right-hand side. The left-hand side displays depths, z, from the skin surface in mm.

TABLE 1. Blood fractions

Layer	Layer name	Nominal blood fraction, f_i
1	Epidermis	0
2	Dermis	0
3	Papillary plexus	0.0556
4	Dermis	0.0417
5	Cutaneous plexus	0.2037
6	Subcutis	0.0417

The blood fractions in the skin layers when the total blood fraction $f_0 = 0.5$.

The absorption of the plain skin, without haemoglobin, melanin and other skin chromophores, is called the skin baseline. The baseline measured by Saidi was used here (16).

The absorption spectra of haemoglobin are described by Horecker in (8) and tabulated by Prahl in (7). Both oxygenated haemoglobin and deoxygenated haemoglobin spectra were used. An oxygen saturation level of 0.70 and a nominal haemoglobin concentration value of 150 g/L or 2.33 mmol/L were assumed. The blood distribution in different skin layers, described by Reuss (17), was used. It is presented in Table 1. The absolute blood concentration in each layer is achieved by multiplying f_i with the blood fraction in the total tissue volume f_0. Reuss reported f_0 values from 0.015 to 0.05.

A formula for reproducing the absorption spectra of melanin is suggested by Jacques in

(18). This formula was also used here as well as Prahl's formula for approximating Mie and Rayleigh scattering in skin.

The spectra of bilirubin and β-carotene are included in the Photochemcad software (19). The normal concentration of total bilirubin in blood is between 5 and 25 µmol/L.

Allen's test

The simulations were validated by measuring the *in vivo* skin spectra from the palm of the hand during Allen's test (20). The HR4000 spectrometer and the ISP-REF integrating sphere (Ocean Optics Inc., Dunedin, FL, USA) were used in the measurements.

The palm of the hand has a rich vascular network and a very low concentration of melanin. Allen's test is normally used to test the patency of ulnar circulation, when harvesting of the radial artery is planned. In this experiment, we used Allen's test to alter the skin's blood volume and to measure the vascularity-related changes in skin spectra. For this purpose, 20 healthy volunteers (ages 20–39), six males and 14 females, with Fitzpatrick skin types I–IV, were enrolled in the study. The experiment was started by performing Allen's test in which both ulnar and radial arteries were occluded at the wrist level by manual compression, and the blood was pumped out from the palm by clenching and unclenching the hand several times. Then, with minimal blood concentration in the skin, the spectra were measured, and while releasing compression of both arteries leading to rapid inflow of blood to the hand, the skin spectra were recorded approximately every 0.55 s until the maximum blood concentration in the skin was reached. This resulted in 10 spectral images from the beginning to the end of each measurement with different degrees of haemoglobin concentrations. The measurements were taken from five different reference points of the right palm, shown in Fig. 2, of each volunteer, adding up to 100 measurements altogether. The experiment was performed in normal room temperature and humidity.

Results

The concentrations of the skin chromophores in the simulation and their reference values are given in Table 2.
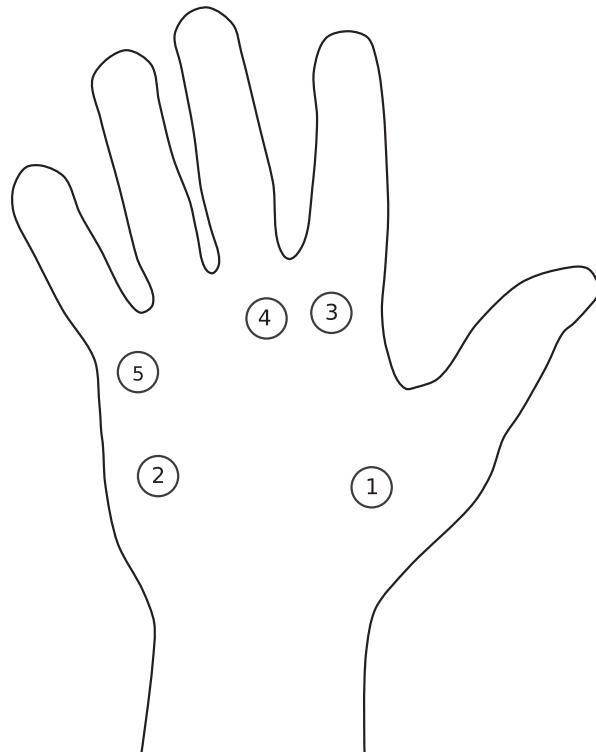
Fig. 2. The measurement locations in the right palm.

TABLE 2. *Chromophore concentrations*

Parameter	Value	Reference value
Melanin	$f_M = 0.012$	[0, 0.05]
Total blood	$f_0 = [0.0016, 0.0045]$	[0.015, 0.05]
Oxysat	0.70	[0.2, 1]
Haemoglobin	$C_{HB} = 2.33$ mmol/L	2.33 mmol/L
Bilirubin	$C_{Br} = 10$ µmol/L	<[5, 25] µmol/L
β-carotene	$C_{BC} = 2$ µmol/L	[1, 4] µmol/L

Skin chromophore concentrations in the simulation model.

The absorbance of the skin in green ($\lambda = 522$ nm) during Allen's test, as a function of time, is shown in Fig. 3. The reflectance of green in the skin decreased to approximately one-third during Allen's test. Then the normal blood volume was retained within 1 s. The curve in the middle represents the median values of all 100 measurements. The topmost line is the 75% percentile and the lowest curve is the 25% percentile.

The simulated and measured reflectance spectra of skin during the Allen's test are shown in Fig. 4. The topmost line shows the reflectance in the beginning, when the blood volume is the smallest. The other curves from the top to bottom correspond to measured and simulated spectra after 1.1, 1.65 and 4.95 s after artery release.
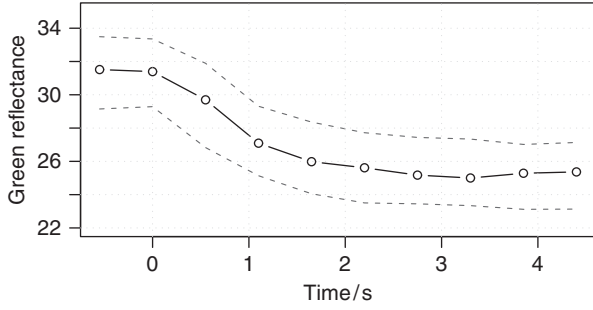
Fig. 3. The median reflectance of skin in green ($\lambda = 522\,nm$) during Allen's test is shown as a continuous line. Dashed lines show the 25% and 75% percentiles.
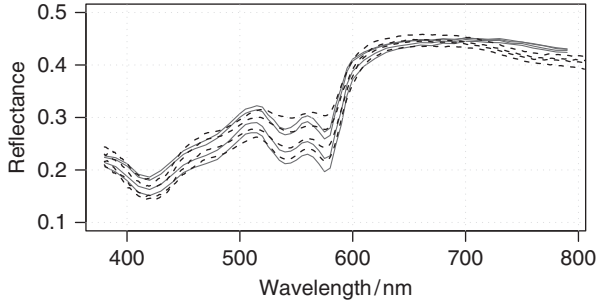
Fig. 4. The measured (continuous line) and simulated (dashed line) skin reflectance spectra during Allen's test, after approximately 0, 0.5, 1 and 5 s after the release of the arteries.
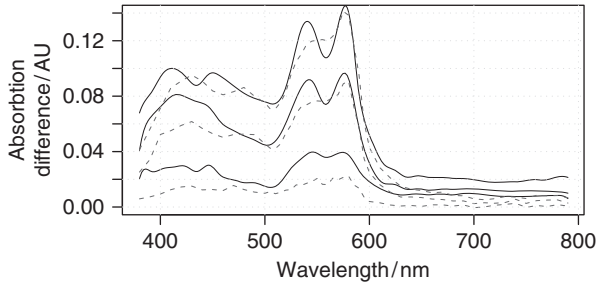
Fig. 5. The measured (continuous line) and simulated (dashed line) difference of the skin absorption spectra with different total blood fractions, f_0.

Fig. 6. The root mean square difference between the simulated reflectance spectra and the observed reflectance spectra when the observed spectra are measured after $\circ = 0$, $\Delta = 0.5$, $X = 1$ and $+ = 5\,s$ after artery release.
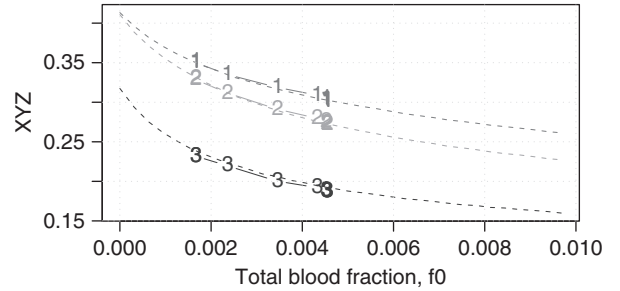
Fig. 7. XYZ reflectance values (1 = X, 2 = Y and 3 = Z) as a function of total blood fraction (dashed line) vs. measurement results (continuous line).
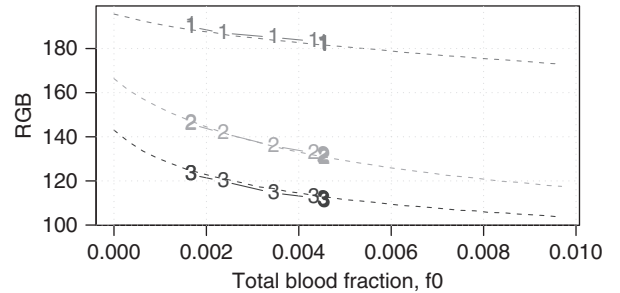
Fig. 8. The RGB values (1 = red, 2 = green and 3 = blue) of skin reflectance with different total blood fractions (dashed line) vs. measurement results (continuous line).

The SDA spectra shown in Fig. 5 were obtained by subtracting the spectra measured in the beginning of Allen's test from the spectra measured 1.1, 1.65 and 4.95 s after arterial release. The SDA spectra calculated from the corresponding simulated spectra are also shown.

The root mean square (RMS) differences of the simulated spectra, corresponding to the observed spectra at different times during Allen's test, can be mostly explained by different blood concentrations, as shown in Fig. 6. The total blood fractions that make the simulated spectra match

closest to the observed spectra are: $f_0 = \{0.16\%, 0.23\%, 0.034\%, 0.45\%\}$.

Figures 7–9 show the different tri-stimulus values as a function of blood volume. The blood fractions for measured values are those minimizing the RMS differences, as reported above. The XYZ values are calculated from the measured and the simulated reflectance spectra using Eqs (3)–(5). The sRGB values are obtained from the XYZ values using a linear transformation and by applying a γ correction. The curves in Fig. 9 are obtained by using a narrow band, red, green and blue LEDs instead of CIE SCO curves as sensor channels.
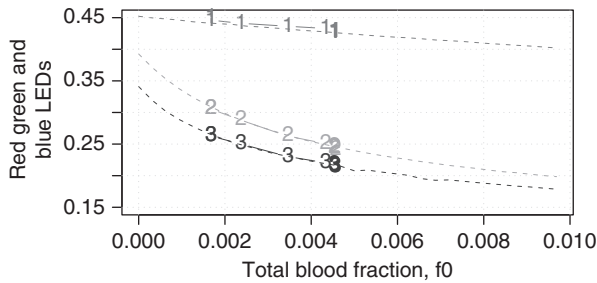
Fig. 9. The reflectance values of (1 = red), (2 = green) and (3 = blue) LEDs with different total blood fractions (dashed line) vs. measurement results (continuous line).

Conclusion and Discussion

The exact knowledge of the relationship between skin blood concentration and its spectra is important when quantifying the intensity of different skin lesions. An increase in vascularity occurs in erythema and also scars. The subjective colour assessment scales and the commercial equipment used in colour measurement assume that the colour change related to increase of haemoglobin in the skin follows a linear model (21). The measurements and simulations in this work show the dependence between the skin reflectance spectra and the blood concentration. The RGB or other tri-stimulus values can be calculated from the reflectance spectra to show how the skin colour depends on the blood concentration. As seen in the results of Allen's test and the simulation model, the skin colour does not linearly follow the blood concentration changes. When the blood concentration is high, further increase has less effect on the colour intensity than when the concentration is low. Based on this knowledge, the intensity of different hyperaemic conditions, especially scars, should be expressed in terms of the relative haemoglobin concentration in the skin rather than the mere intensity of colour.

The simulated skin spectra approximate the measured spectra in the whole visual range, $\lambda \in [380,780]$ nm. The simulation model can be tuned to match to all the spectra measured in Allen's test by tuning only the blood concentration in the simulation model. The RMS error between the simulated and the measured spectra is below 0.013 in all cases. Potentially, the simulation model also holds outside the validated region.

Takiwaki and colleagues measured SDA spectra from different skin lesions and assumed that the main difference between certain lesions is the different blood concentration. The SDA spectra shown in Fig. 5 are similar, but in this work, it is known that the only difference in the skin was the blood concentration. The SDA of the simulated spectra also agrees satisfactorily with the measured SDA, after a suitable amount of bilirubin was added to the model. The SDA spectra of Allen's test have potential in investigating whether certain chromophore is mixed in the blood or deposited in the skin, because a change in the total blood fraction also changes the concentrations of all the pigments in blood. This change is seen as a peak in the SDA spectra. If the pigment is deposited in the skin, the SDA spectra are flat.

The parameters of the model correspond mainly to those mentioned in the previous work. The absolute value of the blood concentration in skin is not very well known and it varies widely between persons, skin regions and time. The total blood fraction, f_0, was from 0.16% to 0.45%, while it was from 1.5% to 5% in the previous work (17). The nominal concentrations of bilirubin and β-carotene are so small that they did not seem to have any measurable effect on the skin colour if they are only in blood. Therefore, it was assumed that the concentration of these pigments in the skin is equal to the concentration in blood. This assumption helped to construct the simulation model to match with the measured spectra using bilirubin and β-carotene concentrations, which are in the middle of their reference minimal and maximal values. The spectra of bilirubin and β-carotene are so close to each other that it is difficult to separate them from each other.

More research is needed to tune the model to match with the measurements even further. More work is also needed to validate the model in the larger range.

Acknowledgements

References

1. Changizi M, Zhang Q, Shimojo S. Bare skin, blood and the evolution of primate colour vision. Biol Lett 2006; 2: 217–221.

2. Nedelec B, Correa JA, Rachelska G, Armour A, LaSalle L. Quantitative measurement of hypertrophic scar: intrarater reliability, sensitivity, and specificity. J Burn Care Res 2008; 29: 489–500.

3. Takiwaki H, Miyaoka Y, Arase S. Analysis of the absorbance spectra of skin lesions as a helpful tool for detection of major pathophysiological changes. Skin Res Technol 2004; 10: 130–135.

4. Igarashi T, Nishino K, Nayar S. The appearance of human skin: a survey, volume 3 of foundations and trends in computer graphics and vision. Now Publishers Inc., 2007.

5. Baranoski GVG, Krishnaswamy A. Light interaction with human skin: from believable images to predictable models. In: SIGGRAPH Asia '08: ACM SIGGRAPH ASIA 2008 courses. New York, NY, USA: ACM, 2008: 1–80.

6. Zijlstra WG, Buursma A, van der Meeuwsen R. Absorbtion spectra of human fetal and adult oxyhemoglobin, de-oxyhemoglobin, carboxyhemoglobin and methemoglobin. Clin Chem 1991; 37: 1633–1638.

7. Prahl S. Optical absorption of hemoglobin. Oregon: Oregon Medical Laser Center, 1999. Available at http://omlc.ogi.edu/spectra/hemoglobin/index.html (accessed 31 January 2008)

8. Horecker BL. The absorbtion spectra of hemoglobin and its derivatives in the visible and near infra-red regions. J Biol Chem 1942; 148: 173–183.

9. Tsumura N, Haneishi H, Miyake Y. Independent-component analysis of skin color image. J Opt Soc 1999; 16: 2169–2176.

10. Tsumura N, Kawabuchi M, Haneishi H. Mapping pigmentation in human skin by multi-visible-spectral imaging by inverse optical scattering technique. In: 8th IS&T/SID Color Imaging Conference. Scottsdale, AZ, USA, 2000.

11. Claridge E, Cotton S, Hall P, Moncrieff M. From colour to tissue histology: physics based interpretation of images of pigmented skin lesions. MICCAI 2002; 1: 730–738.

12. Takiwaki H. Instrumental quantification and its clinical application in the future. Skin Res Technol 2002; 8: 94–97.

13. Välisuo P, Alander J. The effect of the shape and location of the light source in diffuse reflectance measurements. 21st IEEE Int Symp Comput Based Med Syst Jyväskylä, Finland, 2008: 81–86.

14. Välisuo P, Mantere T, Alander J. Solving optical skin simulation model parameters using genetic algorithm. The 2nd International Conference on BioMedical Engineering and Informatics, Tianjin, China 2009.

15. CIE. Commission internationale de l'Eclairage proceedings. Cambridge: Cambridge University Press, 1931.

16. Saidi IS. Transcutaneous optical measurement of hyperbilirubinemia in neonates. PhD thesis, Rice University, Houston, USA, 1992.

17. Reuss JL. Multilayer modeling of reflectance pulse oximetry. IEEE Trans Biomed Eng 2005; 52: 153–159.

18. Jacques SL, McAuliffe DJ. The melanosome: threshold temperature for explosive vaporization and internal absorption coefficient during pulsed laser irradiation. Photochem Photobiol 1991; 53: 769–775.

19. Du H, Fuh RA, Li J, Corkan AL, Lindsey JS. PhotochemCAD: a computer-aided design and research tool in photochemistry. Photochem Photobiol 1998; 68: 141–142. Available at http://dx.doi.org/10.1111/j.1751-1097.1998.tb02480.x.

20. Asif M, Sarkar PK. Three-digit Allen's test. Ann Thorac Surg 2007; 84: 686–687.

21. Draaijers LJ, Tempelman FRH, Botman YAM, Kreis RW, Middelkoop E, van Zuijlen PPM. Colour evaluation in scars: tristimulus colorimeter, narrow-band simple reflectance meter or subjective evaluation? Burns 2004; 30: 103–107.

Address:
Petri Välisuo
Department of Electrical Engineering and Automation
University of Vaasa
Wolffintie 34
65320 Vaasa
Finland
Tel: +358 445 804 320
Fax: +35 863 248 677
e-mail: petri.valisuo@uwasa.fi
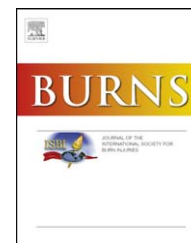
Objective scar assessment—A new method using standardized digital imaging and spectral modelling

Ilkka S. Kaartinen [a,*], Petri O. Välisuo [b], Jarmo T. Alander [b], Hannu O. Kuokkanen [a]

[a] Department of Plastic Surgery, Tampere University Hospital, P.O. Box 2000, Teiskontie 35, FI-33521 Tampere, Finland
[b] Department of Electrical Engineering and Automation, University of Vaasa, P.O. Box 700, FI-65101 Vaasa, Finland

ARTICLE INFO

ABSTRACT

Introduction: Quantitative assessment of scars is needed in clinical practice and in scientific studies. To date, there have been no entirely objective methods available for these purposes. We introduce a new method developed for scar assessment combining standardized digital imaging (SDI) and spectral modelling (SpM). With this method, the estimated concentration changes (ECCs) of haemoglobin and melanin in the scar can be determined quantitatively.
Patient and methods: : In the current study, 22 skin graft donor site (SGDS) wounds were treated with two alternative dressing materials, Suprathel® and Mepilex Transfer®, side by side on the same wound. The SGSD scars were assessed using SDI and SpM. The scars were given subjective ratings by three surgeons using the POSAS and the Vancouver Scar Scale (VSS). The correlations between the ECCs of melanin and haemoglobin and the corresponding subjective ratings were calculated as well as the Intraclass Correlation Coefficient (ICC) of the subjective ratings.
Results: There was a statistically significant correlation between the ECCs of melanin and haemoglobin and the subjective ratings. A single observer could reliably assess pigmentation with the POSAS scale (ICC = 0.75) but not vascularity (ICC = 0.51). The reliability ratings of the VSS were unacceptably low.
Conclusions: The ECC values of haemoglobin and melanin give accurate documentation of the scar status. The results also show that the subjective ratings in this study were unreliable especially when interfering pigmentation and increased vascularity were both present at the same time.

1. Introduction

Even the smallest surgical operation leaves a scar. Some scars heal well, leaving unnoticeable fine lines that can hardly be seen while others develop into hypertrophic or even keloid forms. Large scars produced by traumatic injuries such as burns can cause a significant functional and psychological burden.

There is a lot of scientific interest for scar prevention and scar treatment [1]. While the goal of this research is obvious, the methods for critical analysis of treatments are lacking. Most studies that involve scar assessment are based on subjective evaluation of one or more observers. Different rating scales are used for scar rating, most well-known and widely used scale being the Vancouver Burn Scar Assessment Scale, usually referred to as Vancouver Scar Scale (VSS, Table 1), established in 1995 [2].

More recently, several new scales have been developed to increase the reliability of subjective assessment, including the Patient and Observer Scar Assessment Scale (POSAS, Table 2)

Table 1 – The Vancouver Scar Scale colour subscales.		
Parameter	Descriptor	Points
Vascularity	Normal	0
	Pink	1
	Red	2
	Purple	3
Pigmentation	Normal	0
	Hypopigmentation	1
	Hyperpigmentation	2

[3]. The POSAS adds a second subjective rating done by the patient to increase the reliability of the method.

Subjective evaluation is known to be unreliable when done by a single observer, as usually is the case in clinical practice. In a study setting it is possible to increase the reliability by increasing the number of observers. Still it can only achieve a robust estimate of the scars appearance.

When assessing a scar, the most obvious quantifiable dimension is its colour. The colour changes of a scar are usually divided into two parameters, vascularity and pigmentation [4]. Vascularity is seen as different degrees of redness and is caused by the increase of haemoglobin in the surface of the scar as well as the thinness of the epithelium covering the dermal microvasculature [5]. Pigmentation is a more complex phenomenon that is a consequence of primarily melanin and to a lesser extent, hemosiderin in the surface of the scar. Pigmentation is described as brown discolouration [6–8].

As an objective method, spectrophotometry has been used for over 50 years for skin colour measurement. It has shown good results in the assessment of different skin lesions, both pigmented and vascular [9,10]. The scar colour has been measured with both a tri-stimulus colorimeter (Minolta Chromamameter®) and a narrow-band reflectance meter (Dermaspectrometer®) with fairly good results [11].

Digital images have been used for wound size measurements, blood circulation measurements and to measure the wound healing kinetics [12–15].

We describe a new method combining standardized digital imaging (SDI) with computer controlled lighting and spectral modelling (SpM) for quantitative scar assessment. The method compares the reflectance spectrum of the scar and the adjacent healthy skin and is designed to determine the estimated concentration change (ECC) of haemoglobin and melanin across the whole scar. Acquiring the mean ECC values is important because the scar colour is often irregularly distributed. Moreover, the measurement of the ECC values provides repeatable, quantitative data which is calibrated to indicate melanin and haemoglobin concentration changes instead the mere colour intensity.

We used this method to quantify the colour changes of 22 split-thickness skin graft donor sites (SGDSs) with 3 months

Fig. 1 – Schematic picture of the standardized digital imaging system used in this study.

follow-up. The results of these measurements were then compared to conventional evaluation done by three surgeons using the rating scales previously discussed.

2. Materials and methods

2.1. Standardized digital imaging

A standardized digital imaging (SDI) system, based on the Fuji IS Pro (Fujifilm Corporation, Tokyo, Japan) digital single lens reflex camera (DSLR), was build to accurately measure the reflectance of the skin. The SDI system consists of a rectangular chamber, a lighting system and a camera (Fig. 1). The chamber protects the target form exterior light and keeps the distance between the camera and the skin constant. The camera is attached outside the chamber so that its lens goes through the hole in its roof. The computer controlled lighting system is attached in the roof of the chamber, surrounding the lens hole. The bottom of the chamber is open, and when the chamber is placed on top of the target, the target is illuminated and can be photographed. The digital photographs are transferred into the computer in raw-image format and the images are calibrated to reflectance values using an image of a reference white.

The computer controlled lighting system consists of control electronics and 120 light emitting diodes (LEDs), of which 24 are red, HSMZ-A400-U80M1, 24 green, HSMN-A400-Q8QM2, and 72 blue, HSMM-A400-T8YM2 (Avago Technologies, San Josè, USA). The measured center wavelengths of the LEDs are 470 nm, 520 nm and 640 nm and the -3 dB bandwidths are 19 nm, 25 nm, 15 nm and 50 nm, correspondingly. The imaging software acquires one colour image using red green and blue LEDs for illumination.

Table 2 – The POSAS scale colour subscales.											
Normal skin	1	2	3	4	5	6	7	8	9	10	Worst scar imaginable
Vascularization	○	○	○	○	○	○	○	○	○	○	
Pigmentation	○	○	○	○	○	○	○	○	○	○	

2.2. Patients

In total, 22 SGDS scars were assessed in 14 patients. The donor sites had been treated with two dressing options (Suprathel®, PMI Polymedics Innovations, Germany and Mepilex Transfer®, Mölnlycke Health Care) used side by side on the same wound. This resulted in two scar halves of equal size in each donor site. Both halves were evaluated separately to compare the difference in their appearance. The scars were photographed using SDI at 14, 30 and 90 days after the skin harvest to examine the evolution of scar appearance over time. For statistical analysis, the Suprathel treated scar halves were pooled as Group A and Mepilex Transfer treated halves as Group B.

2.3. Spectral modelling

The colour of the skin is not constant but changes over time due to changes in blood fraction and melanin concentration. Therefore, the characteristics of the skin disorder are here measured as the difference between normal skin colour, Rn, and disorder colour, Rd. The logarithm to base 10 of the inverse of the colour value is believed to be more directly related to skin chromophore concentrations than the colour itself [16,17].

$$A_n = \log_{10}(R_n)$$
$$A_d = \log_{10}(R_d)$$

RGB, CIE $L^*a^*b^*$, XYZ, HSV and other tri-stimulus colour spaces have been used for measuring erythema and pigmentation indices [18]. Here the optimal set of parameters of the mentioned tri-stimulus colour spaces are used for predicting the change of melanosome amount and blood fraction in the skin disorder area. The selection of the optimal predictor variable set is done using Least Angle Reflection method (LARS). The logarithms of X, Y, G and L* colour variables were found to form an optimal set of variables for predicting the melanin and haemoglobin concentration changes creating the colour change in the skin disorder.

$$A = \begin{bmatrix} \log_{10}(X) \\ \log_{10}(Y) \\ \log_{10}(G) \\ \log_{10}(L) \end{bmatrix}$$

The corresponding absorption change due to haemoglobin and melanin concentration changes alone, ΔA_h and ΔA_m, are resolved using a spectral skin model and Monte Carlo Simulation. The melanin and haemoglobin concentration changes, C_h and C_m, in the skin disorder area can be solved from the following equation:

$$\Delta A = C_h \cdot \Delta A_h + C_m \cdot \Delta A_m$$

This is a linearised model around the normal skin colour. Since the skin colour is nonlinearly related to chromophore concentrations, the model is not accurate far away from the normal chromophore concentrations. The root mean square error of prediction (RMSEP) is less than 12% for both melanin and haemoglobin estimation if the concentrations of both chromophores in the disorder area are above 50% of the normal concentrations and below 200% of the normal concentrations. For smaller concentration changes the model is more accurate.

2.4. Subjective assessment

The digital images captured were used for conventional colour evaluation done independently by three plastic surgeons, the
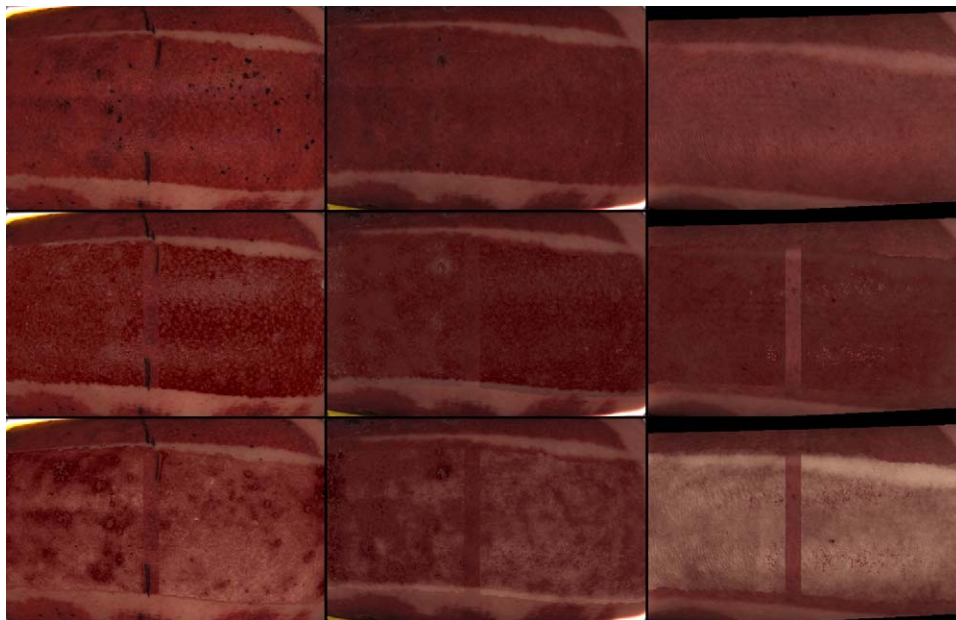
Fig. 2 – Donor site images at 14, 30 and 90 days (left, middle, right, respectively). SpM has been used to separate the interfering pigmentation and vascularity: topmost images are standardized digital images without calibration. The center images are calibrated for visualization of haemoglobin related colour change only. The lowermost images show melanin related colour change. Normally, this kind of scar is very difficult to assess subjectively.

corresponding author and two plastic surgeons of our department serving as independent observers. Scores for vascularity and pigmentation were given according to the VSS and POSAS colour subscales (Fig. 2).

Subjective rating was done twice with 3 weeks interval. At the first assessment, the pictures showing the whole donor site with both scar halves visible were put into a random order and a rating was given for both scar halves using VSS and POSAS scales.

For the second rating, the pictures were cut in half so that only one scar half could be seen at a time and all the pictures were then again randomized. This setting was used to find out if the possibility to compare the adjacent halves would affect the outcome of subjective rating.

2.5. Statistical analysis

Statistical analysis was done using SPSS software version 17.0 (SPSS Inc., Chicago, IL, USA) and R software. The interrater reliability of ratings between the three observers was calculated using Intraclass Correlation Coefficient and measures of >0.75 were considered reliable. The correlation was calculated for both single and average measures. For concurrent validity testing, the correlation between the ECC values of melanin and haemoglobin and the values of the subjective ratings for the corresponding parameters were calculated using Spearman's correlation coefficient. To calculate the statistical differences between the treatment groups A and B and also the difference of pigmentation and vascularity between 2 weeks, 1 month and 3 months, the paired-samples Wilcoxon signed-rank test was used. p value of <0.05 was considered statistically significant.

3. Results

3.1. Outcome measures

The computer scores for the ECCs of melanin and haemoglobin and the ratings of POSAS subscales for the treatment groups A and B at 14, 30 and 90 days are shown in Table 1. The ECC of haemoglobin was significantly higher in group B at 14 and 30 days. The ECC of melanin was significantly higher in group B at all time points (Table 3).

For the first subjective rating, where the scar halves were assessed at the same time, there was a statistically significant difference in all outcome measures of POSAS subscales showing lower vascularity and pigmentation in group A (Table 3). VSS detected lower vascularity in group A at 14 and 30 days and also significantly lower pigmentation in group A at all time points.

In the second rating, when the scar halves had been separated and randomized, there was significantly lower vascularity in group A at 90 days and lower pigmentation in group A at 14 and 30 days on the POSAS subscales. In VSS ratings, there was significantly lower vascularity at 30 days (Table 3). The VSS pigmentation subscale proved very unreliable in the first assessment and was not used in the second rating.

The difference of the mean pigmentation and vascularity ratings at 14, 30 and 90 days and the differences of mean ECCs of haemoglobin and melanin are shown in Figs. 3 and 4. There was a statistically significant difference in the ECCs of haemoglobin between each time point. The same decrease in vascularity was also shown in the POSAS subscale ratings in both assessments (Fig. 3). There was also a statistically significant increase in the ECC of melanin between 14 and
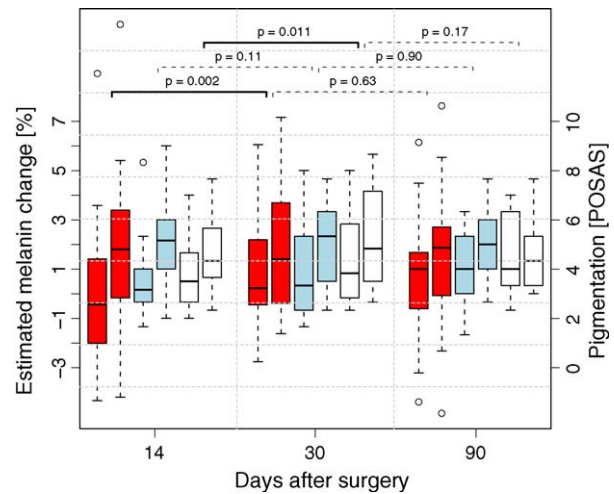
Table 3 – Ratings of the first and the second assessment and the ECC values of melanin and haemoglobin. Mean values ± SD for group A/group B. For the ECC of melanin and haemoglobin compared to normal skin the unit is g/l.

	14 days	30 days	90 days
Results of subjective and objective assessments			
1 Assessment			
POSAS vascularity	$3.56 \pm 0.18/4.42 \pm 0.22$	$3.38 \pm 0.22/3.95 \pm 0.27$	$2.41 \pm 0.19/2.75 \pm 0.20$
	$^*p = 0.000$	$^*p = 0.003$	$^*p = 0.011$
POSAS pigmentation	$3.48 \pm 0.32/5.12 \pm 0.34$	$3.93 \pm 0.43/5.13 \pm 0.38$	$4.08 \pm 0.33/5.08 \pm 0.30$
	$^*p = 0.000$	$^*p = 0.001$	$^*p = 0.002$
VSS vascularity	$1.44 \pm 0.07/1.85 \pm 0.68$	$1.35 \pm 0.11/1.53 \pm 0.11$	$0.65 \pm 0.09/0.68 \pm 0.09$
	$^*p = 0.001$	$^*p = 0.046$	
VSS pigmentation	$1.18 \pm 0.58/1.67 \pm 0.45$	$1.25 \pm 0.73/1.65 \pm 0,59$	$1.65 \pm 0.50/1.87 \pm 0.27$
	$^*p = 0.003$	$^*p = 0.010$	
2 Assessment			
POSAS vascularity	$4.38 \pm 0.23/4.71 \pm 0.19$	$3.85 \pm 0.22/3.93 \pm 0.23$	$2.25 \pm 0.19/2.56 \pm 0.16$
			$^*(p = 0,027)$
POSAS pigmentation	$3.88 \pm 0.33/4.68 \pm 0.30$	$4.40 \pm 0.41/5.20 \pm 0.44$	$4.43 \pm 0.35/4.57 \pm 0.32$
	$^*p = 0.005$	$^*p = 0.001$	
VSS vascularity	$1.95 \pm 0.09/2.18 \pm 0.10$	$1.87 \pm 0.14/2.15 \pm 0.08$	$1.56 \pm 0.16/1.68 \pm 0.12$
		$^*p = 0.010$	
ECC of melanin	$0.067 \pm 0.16/0.44 \pm 0.17$	$0.21 \pm 0.13/0.45 \pm 0.15$	$0.18 \pm 0.14/0.41 \pm 0.15$
	$^*p = 0.000$	$^*p = 0.010$	$^*p = 0.000$
ECC of haemoglobin	$0.27 \pm 0.02/0.31 \pm 0.02$	$0.20 \pm 0.01/0.22 \pm 0.01$	$0.10 \pm 0.01/0.10 \pm 0.01$
	$^*p = 0.028$	$^*p = 0.011$	

* p values of Wilcoxon signed-rank test are shown for statistically significant differences.
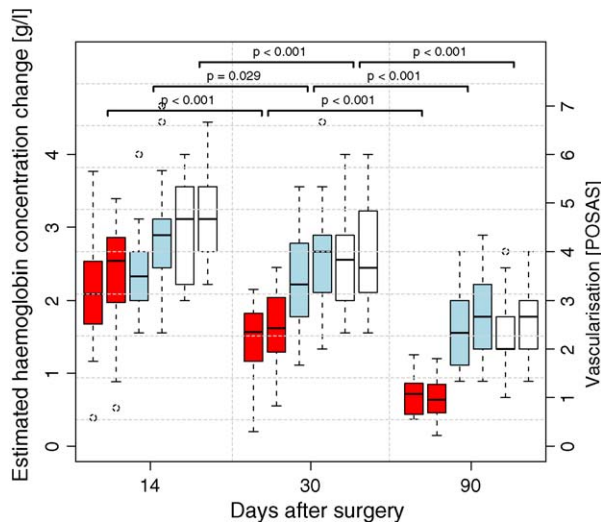
Fig. 3 – Red boxes are mean ECCs of haemoglobin in group A (left) and group B (right) on 14, 30 and 90 days. Blue boxes represent the POSAS vascularity ratings of the first assessment for groups A (left) and B (right). White boxes are POSAS vascularity ratings of the second assessment for groups A (left) and B (right). _p_ values of Wilcoxon signed-rank test are for differences between each time interval for each rating. (For interpretation of the references to colour in this figure caption, the reader is referred to the web version of the article.)

Fig. 4 – See Fig. 3. The red boxes are mean ECCs of melanin, the blue boxes represent the POSAS pigmentation ratings of the first assessment and the white boxes are the POSAS pigmentation ratings of the second assessment for groups A and B, respectively. _p_ values are for differences between each time interval for each rating. (For interpretation of the references to colour in this figure caption, the reader is referred to the web version of the article.)

30 days. The increase was shown in the POSAS rating at the first assessment but not at the second assessment (Fig. 4). As shown in Figs. 3 and 4, especially the POSAS vascularity ratings differed considerably between the two assessments.

3.2. Interrater reliability

The reliability values for the POSAS and the VSS colour subscales are shown in Table 4. The reliability fell below the accepted limits in both VSS subscales, especially the pigmentation subscale, which had no statistically significant correlation between the raters at 30 days. For the POSAS subscales, the reliability for pigmentation ratings was high in both the

first and the second assessment but was lower in the second rating. Also the reliabilities of the vascularity ratings were acceptable in both assessments, when three observers were used. When the ratings were analyzed at different time points, the reliability of the POSAS vascularity rating was the lowest at 30 and 90 days.

3.3. Concurrent validity

The correlations between the ECC values and the subjective ratings are shown in Table 5 for each subscale. There was a statistically significant correlation between the mean values of the POSAS subscales and the haemoglobin and melanin concentration estimates. Also, the concentration values correlated with the VSS vascularity and pigmentation ratings. Also, there was a statistically significant correlation between

Table 4 – Intraclass Correlation Coefficients of subjective ratings at 14, 30 and 90 days for single/average measures.				
	14 days	30 days	90 days	Overall reliability
Reliability ratings				
First assessment				
POSAS vascularity	0.51/0.76	0.43/0.70	0.25/0.50	0.51/0.76
POSAS pigmentation	0.76/0.91	0.73/0.89	0.73/0.90	0.75/0.90
VSS vascularity	0.20/0.42	0.48/0.74	0.45/0.71	0.40/0.67
VSS pigmentation	0.26/0.52	0.09/0.24[*]	0.30/0.56	0.26/0.52
Second assessment				
POSAS vascularity	0.32/0.59	0.43/0.69	0.38/0.64	0.56/0.79
POSAS pigmentation	0.68/0.87	0.82/0.93	0.68/0.86	0.69/0.89
VSS vascularity	0.48/0.73	0.24/0.49	0.26/0.52	0.32/0.58

[*] No evidence of statistically significant correlation.

Table 5 – Concurrent validity. Values of Spearman's estimates and their p values.		
	Spearman estimate	p
Melanin concentration vs POSAS pigmentation subscale	0.6	<0.001
Haemoglobin concentration vs POSAS vascularity subscale	0.63	<0.001
Haemoglobin concentration vs VSS vascularity subscale	0.74	<0.001
Melanin concentration vs VSS pigmentation subscale	0.53	<0.001

the POSAS and the VSS subscales of vascularity and pigmentation ($r = 0.64$ and 0.73, respectively, $p < 0.01$ for both).

4. Discussion

Scar assessment remains a subject of debate. Although quantitative methods have been described, subjective evaluation still remains the primary instrument used by most researchers [19]. The colour is often the most obvious dimension of a scar. Active, hypertrophic scars appear red, reflecting the increased vascularity. Hypo- or hyperpigmentation may significantly affect the cosmetic result. The turn in colour can be measured and gives useful information of the effect of scar management.

The inability to objectively quantify the properties of scars has led to development of a multitude of different observer-based scales. The Vancouver Burn Scar Assessment Scale, usually referred to as Vancouver Scar Scale (VSS), remains the most well-known and mostly used scar rating scale [2]. Originally designed for rating burn scars, it has been used for other purposes such as skin graft donor site studies and to assess linear scars [20]. In this study and also in previous studies, the reliability of the VSS is shown to be unacceptably low. In the current study, the reliability of VSS was low even with three observers. For estimation of pigmentation, the VSS provides only three categories and describes the pigmentation pattern and not its intensity. The overall reliability of VSS pigmentation ratings was below acceptable limits, and at one observation point, there was no evidence of statistically significant correlation between the three observers.

Since the establishing of VSS, newer rating scales have been developed, combining photographs (MAPS) or patient parameters (POSAS) to the assessment [3,21]. In this study, reliability of the pigmentation ratings was high especially in the first assessment whereas the vascularity rating had lower, but still acceptable reliability. Vascularity rating was most unreliable at 30 and 90 days, when the scar had more pigmentation. Interestingly, this is in contradiction with the findings in the study by Draaijers where the pigmentation ratings were less reliable [11]. Using three observers significantly raised the reliability of both vascularity and pigmentation ratings. In the clinical scenario, the scar is usually assessed by a single observer and often by many different individuals over time. The POSAS increases the reliability of the rating by adding a patients rating, thus increasing the number of observers. It can be speculated, how well the patients are able to rate their scars. In fact, in this study, the vascularity rating with POSAS scale was not reliable enough when done by two expert observers.

Eventually, no matter how complicated rating scales or how many observers we use, we cannot overcome the limitations of human as an observer. Changes in both vascularity and pigmentation occur in the scar at the time overlapping each other, a fact that makes colour definition difficult for the observer. Another problem is that the colour changes are usually unevenly distributed across the scar, and estimating a mean value for a certain area is not easily done by a human observer.

The method we described in this study is based on spectrophotometry, which has been used to measure the skin colour for perhaps more than 50 years. The absorption spectra of the skin chromophores, haemoglobin and melanin, are well known. Although it has been proposed, that also hemociderin might be a cause of pigmentation in scars, it plays only a minor role [7]. On the other hand, the absorption spectrum of hemociderin is very close to that of melanin, and we believe that it does not interfere with the measurement significantly.

There are several instruments commercially available for colour estimation at the moment. The DermaSpectrometer® (Cortex Technology, Hadsund, Denmark), the Minolta Chromameter® (Minolta, Osaka, Japan) and the Mexameter® (Courage-Khazaka Electronic, Köln, Germany) have been introduced for scar assessment in studies by Nedelec and Draaijers. These studies have shown the feasibility of spectrophotometry in scar colour assessment. However the measuring head apertures of these instruments are only 5–8 mm in diameter, which makes it impossible to quantitatively measure a larger area, and the user of the instrument has to choose which site of the scar to assess, leaving chance to observational bias. To eliminate this possibility, a mean value of the colour changes or an estimate of the concentration changes of the colour chromophores in the scar is needed [22]. This can be solved by using SDI and SpM. Since the colour changes are proportional to the concentration of the colour pigments, we have chosen to determine the ECC of melanin and haemoglobin in the surface of the scar rather than just the reflectance values at the wavelengths of their absorption peaks, as has been done by Draaijers and Nedelec. In addition, we believe that it is more accurate to measure the ECC values in comparison to the adjacent healthy skin to neutralize the effect of changes in the skin circulation between the measurements [10,23]. The instrumentation used in our study was assembled in the University of Vasa with a total cost of approx. 3000 euros. In comparison, the Mexameter® used by Nedelec et al. costs 4800 euros at the moment.

Digital images have been used in scar assessment previously by some researchers. Despite the high resolution and quality of modern digital imaging, simple digital photography and image analysis cannot be trusted in colour evaluation because of the effect of ambient lighting. In the method described here, the lighting device is integrated to the

camera and external light blocked, thus achieving a stabile instrument, in which lighting, distance and angle to the wound are always nearly the same, leaving little chance of variance between different users.

This study shows that a human observer is good in detecting differences between two sides of a scar. When the scar halves were separated, the sensitivity of subjective rating deteriorated considerably. In the first assessment, the difference between the two groups was actually bigger than the difference in ECC values despite the fact that the pictures were in random order. This observation points out that subjective rating can actually exaggerate differences. The two opposite halves were rarely given the same rating, referring that the observers wanted to point out even very small differences. It can be stated that subjective evaluation is always qualitative rather than quantitative.

We believe that spectrophotometric assessment of scar colour gives essential information of the scars activeness. It is also perhaps the easiest dimension of a scar to assess. However, the method described here does not quantify other properties of scars such as thickness, pliability or texture. These parameters are included in most scar assessment scales, but as is the case with colour assessment, subjective evaluation of the mentioned parameters is unreliable, and better reliability can be achieved by using objective methods, as shown by Nedelec et al. [24].

We have shown the feasibility of this new method in scar assessment. The study includes only skin graft donor site scars, which are an interesting testing site for colour measurement, since they exhibit strong colour changes. New trials are to be conducted to test this method in the assessment of burn scars and hypertrophic linear scars. Also, we intend to investigate the interrater reliability of this method with several users. In the field of scar study, new methods to reduce or even prevent scar formation are constantly under investigation. To critically evaluate these innovations, we need clearly objective methods, like the one described here, to replace the unreliable subjective assessment.

Conflict of interest statement

None of the authors has a financial interest in any of the commercial products mentioned in this article.

REFERENCES

[1] Ferguson MW, Duncan J, Bond J, Bush J, Durani P, So K, et al. Prophylactic administration of avotermin for improvement of skin scarring: three double-blind, placebo-controlled, phase I/II studies. Lancet 2009;373(April (9671)):1264–74.

[2] Baryza MJ, Baryza GA. The Vancouver scar scale. An administration tool and its interrater reliability. J Burn Care Rehabil 1995;16:535–8.

[3] Draaijers LJ, Tempelman FR, Botman YA, Tuinebreijer WE, Middelkoop E, Kreis RW, et al. The patient and observer scar assessment scale: a reliable and feasible tool for scar evaluation. Plast Reconstr Surg 2004;113(June (7)):1960–5 [discussion 6–7].

[4] Sugata K, Kitahara T, Takema Y. Changes of human skin in subepidermal wound healing process. Skin Res Technol 2008;14(November (4)):436–9.

[5] Zijlstra WG, Buursma A, Meeuwsen-van der Roest WP. Absorption spectra of human fetal and adult oxyhemoglobin, de-oxyhemoglobin, carboxyhemoglobin and methemoglobin. Clin Chem 1991;37(9):1633–8.

[6] Tsuji T. Experimental hemosiderosis: relationship between skin pigmentation and hemosiderin. Acta Derm Venereol 1980;60(2):109–14.

[7] Kim W, Kang WH. Role of dermal melanocytes in cutaneous pigmentation of stasis dermatitis: a histopathological study of 20 cases. J Korean Med Sci 2002;17(5):648–54.

[8] Dressler J, Busuttil A, Koch R, Harrison DJ. Sequence of melanocyte migration into human scar tissue. Int J Legal Med 2001;115:61–3.

[9] Takiwaki H, Miyaoka Y, Arase S. Analysis of the absorbance spectra of skin lesions as a helpful tool for detection of major pathophysiological changes. Skin Res Technol 2004;10:130–5.

[10] Takiwaki H. Graphic analysis of the relationship between skin colour change and variations in the amounts of melanin and haemoglobin. Skin Res Technol 2002;8:78–83.

[11] Draaijers LJ. Colour evaluation in scars: tristimulus colorimeter, narrow-band simple reflectance meter or subjective evaluation? Burns 2004;30:103–7.

[12] Molnar JA, Lew WK, Rapp DA, Gordon ES, Voignier D, Rushing S, et al. Use of standardized, quantitative digital photography in a multicenter web-based study. Eplasty 2009;9:e4.

[13] Hansen GL, Sparrow EM, Kommamuri N, Iaizzo PA. Assessing wound severity with color and infrared imaging of reactive hyperemia. Wound Repair Regen 1996;4(July–September (3)):386–92.

[14] Payette JR, Kohlenberg E, Leonardi L, Pabbies A, Kerr P, Liu KZ, et al. Assessment of skin flaps using optically based methods for measuring blood flow and oxygenation. Plast Reconstr Surg 2005;115(February (2)):539–46.

[15] Bon FX, Briand E, Guichard S, Couturaud B, Revol M, Servant JM, et al. Quantitative and kinetic evolution of wound healing through image analysis. IEEE Trans Med Imaging 2000;19(July (7)):767–72.

[16] Välisuo P, Kaartinen I, Kuokkanen H, Alander J, The colour of blood in skin—a comparison of Allen's test and photonics simulations. Skin Res Technol, doi:10.1111/j.1600-0846.2010.00449, in press.

[17] Dawson JB, Barker DJ, Ellis DJ, Grassam E, Cotterill JA, Fisher GW, et al. A theoretical and experimental study of light absorption and scattering by in vivo skin. Phys Med Biol 1980;25(July (4)):695–709.

[18] Clarys P, Alewaeters K, Lambrecht R, Barel AO. Skin color measurements: comparison between three instruments: the Chromameter(R), the DermaSpectrometer(R) and the Mexameter(R). Skin Res Technol 2000;6(November (4)): 230–8.

[19] Idriss N, Maibach HI. Scar assessment scales: a dermatologic overview. Skin Res Technol 2009;15(1):1–5.

[20] Truong PT, Abnousi F, Yong CM, Hayashi A, Runkel JA, Phillips T, et al. Standardized assessment of breast cancer surgical scars integrating the Vancouver Scar Scale, Short-Form McGill Pain Questionnaire, and patients' perspectives. Plast Reconstr Surg 2005;116(October (5)):1291–9.

[21] Masters M, McMahon M, Svens B. Reliability testing of a new scar assessment tool, Matching Assessment of Scars and Photographs (MAPS). J Burn Care Rehabil 2005;26(May–June (3)):273–84.

[22] Takiwaki H, Miyaoka Y, Skrebova N, Kohno H, Arase S. Skin reflectance-spectra and colour-value dependency on

measuring-head aperture area in ordinary reflectance spectrophotometry and tristimulus colorimetry. Skin Res Technol 2002;8(May (2)):94–7.

[23] Latreille J, Gardinier S, Ambroisine L, Mauger E, Tenenhaus M, Guehenneux S, et al. Influence of skin colour on the detection of cutaneous erythema and tanning phenomena using reflectance spectrophotometry. Skin Res Technol 2007;13(August (3)):236–41.

[24] Nedelec B, Correa JA, Rachelska G, Armour A, LaSalle L. Quantitative measurement of hypertrophic scar: interrater reliability and concurrent validity. J Burn Care Res 2008;29(May–June (3)):501–11.
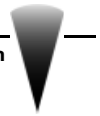
How to assess scar hypertrophy—a comparison of subjective scales and Spectrocutometry: A new objective method

Ilkka S. Kaartinen, MD[1]; Petri O. Välisuo, MSC[2]; Vladimir Bochko[2]; Jarmo T. Alander[2];
Hannu O. Kuokkanen, MD, PhD[1]

1. Department of Plastic Surgery, Tampere University Hospital, Tampere, Finland, and
2. Department of Electrical and Energy Engineering, University of Vaasa, Vaasa, Finland

ABSTRACT

Scar hypertrophy is a significant clinical problem involving both linear scars from elective surgery and scars caused by trauma or burns. The treatment of hypertrophic scars is often time consuming, and patients may need to be followed up for months or even years. The methods for reliable quantification of scar hypertrophy are at present unsatisfying. We have developed a new, objective method, Spectrocutometry, for documentation and quantification of scar hypertrophy. The instrument is based on standardized digital imaging and spectral modeling and calculates the estimated concentration change of hemoglobin and melanin from the entire scar and also provides standardized images for documentation. Three plastic surgeons have assessed 37 scars from melanoma surgery using Spectrocutometry, the Vancouver scar scale, and the patient and observer scar assessment scale. The intraclass correlation coefficient for the Vancouver scar scale and the patient and observer scar assessment scale was lower than required for reliable assessment ($r=0.66$ and 0.60, respectively). The intraclass correlation coefficient for Spectrocutometry was high ($r=0.89$ and 0.88). A Bayesian network analysis revealed a strong dependency between the estimated concentration change of hemoglobin and scar pain. Spectrocutometry is a feasible method for measuring scar hypertrophy. It is shown to be more reliable than subjective rating in assessing linear surgical scars.

The process of scar healing is poorly understood. It is known that wounds close under excess tension and, on the other hand, many traumatic wounds, especially burns, are prone to develop hypertrophy.[1,2] Hypertrophic scars (HSs) pose a significant clinical problem.[3] The incidence of postsurgical scar hypertrophy varies between 5 and 44.6% in different studies and among different ethnic populations, with greater risk for developing scar hypertrophy among the more pigmented populations.[4] The rate of postburn scar hypertrophy is found to be even higher because of the larger surface area of the scar.[5]

HSs are managed with a combination of different treatment modalities including both conservative and surgical treatments.[6,7] In the prevention and treatment of HSs, the role of reliable scar assessment in documentation and follow-up is of paramount importance.[8]

One of the essential problems concerning scar treatment and research is the lack of objective instruments to assess and quantify scar hypertrophy. In a clinical setting, this complicates the follow-up of individual patients with HSs; often, there are several professionals who treat the patient, and a precise follow-up of the scars activity would be needed to guide the treatment decisions. In the field of research, where new methods to treat and to prevent scar formation are being developed, objective quantitative instruments are essential.

The number of different observer-based rating scales is vast and new scales are regularly introduced. The Vancouver scar scale (VSS) has been used in studies in both burn scars and linear scars and the patient and observer scar assessment scale (POSAS) has been designed to rate linear scars.[9] Additionally, the Manchester scar scale (MSS) was introduced by Beausung and colleagues, and Duncan proposed the use of the visual analogue scale (VAS) in scar assessment.[10,11]

In the majority of recently published scar studies, subjective evaluation with the VSS has been used as an endpoint.[12] Despite its poor reliability, subjective assessment is considered the "gold standard" in scar assessment.[9,13] This reflects the lack of suitable quantitative methods.

Spectrophotometry has been used for color measurement for over 50 years. In previous studies, it has also been used to measure the color of scars and other skin disorders.[14–16] We have developed an optical instrument for scar assessment, which combines standardized digital imaging (SDI) and spectral modeling (SpM). The color change in scars is caused by variations in hemoglobin and melanin concentrations.[17] The absorbance spectra of these chromophores are well known.[18] From the spectral images, the estimated concentration change (ECC) of melanin and hemoglobin between the scar tissue and adjacent healthy skin can be calculated from the entire scar.[19] This information, especially the quantitative estimation of hemoglobin concentration, is a good indicator of the scar's activity and appearance. At the same time, the method also provides standardized digital images for

documentation and follow-up. We have used this method previously in assessing skin graft donor site scars, and the results have been published in *Burns* journal.[20]

In this study, we have assessed 37 surgical scars after melanoma surgery. The wounds after melanoma excision are often closed with some degree of tension and exhibit more or less hypertrophy, while scars after a sentinel lymph node biopsy and lymphadenectomy usually heal better. This gives us a variation of scars with none to marked hypertrophy. Three plastic surgeons have examined these scars individually using the VSS and the POSAS scar scales and, in addition, the proposed spectrocutometry method described in detail later. This study compares the reliabilities and validity of the different scar assessment methods.

Patients

The patient group in the study included 20 patients (12 males, eight females) who have undergone melanoma surgery in Tampere University Hospital between August 2007 and November 2009. In the study, 37 scars were assessed, including 22 excision scars, 12 sentinel lymph node biopsy scars, and five lymphadenectomy scars. The mean age of the patients was 57 ± 16.9 years and the mean time from melanoma surgery was 17 months (4–30, SD $\pm$ 7.4). All the patients were examined during the same day in the outpatient clinic. Permission for the study was obtained from the Ethical Board of Pirkanmaa Health Care District. A written informed consent was given by all the participants.

MATERIALS AND METHODS

Subjective assessment

The scars were evaluated individually by three plastic surgeons, the corresponding author and two other plastic surgeons from our department serving as independent observers. The assessment was performed under similar lighting in a temperature-controlled room. The VSS (Table 1) and PO-SAS (Tables 2 and 3) scales were used for rating of each scar. The patients also filled the POSAS patient questionnaire of the corresponding scars.

Image acquisition and reflectance measurement

The digital images of the scars were acquired with the same imaging system as used in our earlier research.[20] The imaging system consists of a camera, a protective cover, and a computer-controlled lighting system (Figure 1). The camera is a digital single lens reflex camera, Fuji IS Pro (Fujifilm Corporation, Tokyo, Japan). The cover protects the target from exterior light and preserves the distance and the angle of the camera constant. The lighting system illuminates the target with reproducible, optimized light. The characteristics of the light source and the camera are obtained by imaging a white reference object. The reflectance of the skin is calculated from the acquired color image by scaling the color image with the image of the white reference in the same illumination. The color precision of the original raw images is 12 bits. All arithmetic operations are calculated using 16-bit precision. The resolution of the

Table 1. Vancouver scar scale

Parameter	Descriptor	Points
Vascularity	Normal	0
	Pink	1
	Red	2
	Purple	3
Pigmentation	Normal	0
	Hypopigmentation	1
	Hyperpigmentation	2
Pliability	Normal	0
	Supple	1
	Yielding	2
	Firm	3
	Banding	4
	Contracture	5
Height	Normal (flat)	0
	> 0 and $< 2\,mm$	1
	≥ 2 and $< 5\,mm$	2
	$> 5\,mm$	3
Total score		/13

apparatus used in the study is 225 pixels/mm.[2] In a study by Välisuo et al.,[21] the accuracy of this imaging system in reflectance measurement was found to be similar to that obtained with a spectrophotometer.

Spectral modeling

The concentrations of melanin and hemoglobin were calculated from the reflectance values using similar spectral modeling as in our earlier research.[19] We have calculated the ECC values of hemoglobin and melanin, which means the measured change in the concentration of these chromophores between the scar and the adjacent healthy skin. The main principle of this SpM is the partition of the observed reflectance change into two components, the first corresponding to hemoglobin and the second to melanin concentration change. The rest of the color change, which

Table 2. The POSAS scale observer component

	Normal skin							Worst scar imaginable		
	1	2	3	4	5	6	7	8	9	10
Vascularization	O	O	O	O	O	O	O	O	O	O
Pigmentation	O	O	O	O	O	O	O	O	O	O
Thickness	O	O	O	O	O	O	O	O	0	O
Relief	O	O	O	O	O	O	O	O	O	O
Pliability	O	O	O	O	O	O	O	O	O	O

POSAS, patient and observer scar assessment scale.

Table 3. The POSAS scale patient component

	No, not at all								Yes, very much	
	1	2	3	4	5	6	7	8	9	10
Is the scar painful?	O	O	O	O	O	O	O	O	O	O
Is the scar itching?	O	O	O	O	O	O	O	O	O	O

|	No, just like normal skin								Yes, very different	
	1	2	3	4	5	6	7	8	9	10
Is the scar color different?	O	O	O	O	O	O	O	O	O	O
Is the stiffness of the scar different?	O	O	O	O	O	O	O	O	O	O
Is the thickness of the scar different?	O	O	O	O	O	O	O	O	O	O
Is the scar irregular?	O	O	O	O	O	O	O	O	O	O

POSAS, patient and observer scar assessment scale.

cannot be explained by hemoglobin and melanin change, is the residual.[20]

The images were segmented in three areas: the scar, normal skin, and the rest, including skin disorders, shadows, and foreign objects (hairs, clothes, shadows, etc.).[22] The concentration of melanin and hemoglobin was estimated for each pixel of the scar, and the ECC values were obtained by comparing the mean estimated concentration of the scar and the adjacent healthy skin.

Objective assessment

Each observer used the instrument described above to take one digital image of each scar (Figure 2). After taking the image, the observer checked the calibrated picture from the computer screen for approval. If the observer was not satisfied with the quality of the image, another image was immediately taken to replace the disapproved one until an acceptable image was gained. In most cases, only one picture was needed for approval. As the camera has an incorporated lighting device, the images were shot without any additional lighting. Two parameters were calculated from the spectral images: the ECC of hemoglobin and melanin.
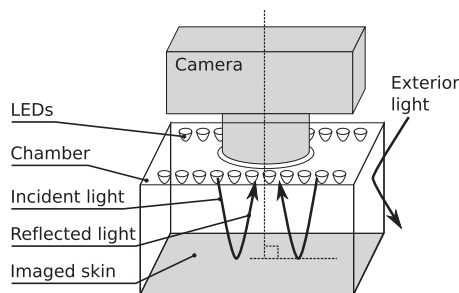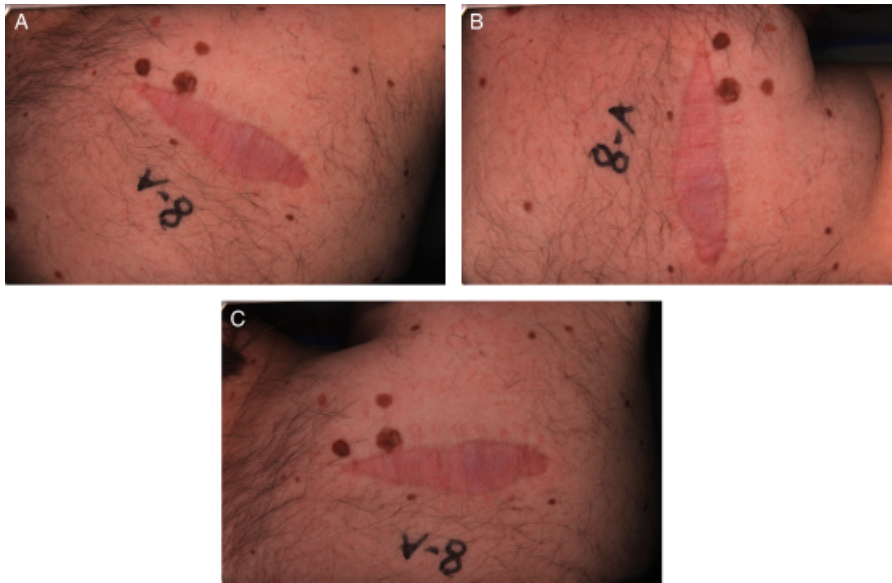

Figure 1. Standardized digital imaging system, picture modified from Zijlstra et al.[18]

Statistical analysis

Statistical analysis was performed using SPSS software version 17.0 (SPSS Inc., Chicago, IL) and R software. The interrater reliability of the measurements between the three observers was calculated using the intraclass correlation coefficient (ICC). The two-way random effect model and consistency type was used in the ICC calculations. The reliability was assessed for both a single measure and average measures including 95% confidence intervals (CI). In this way, the ICC values for single measures indicate the reliability of the ratings of a random set of observers on a single target. The ICC values for average measures, on the other hand, indicate the reliability of the mean ratings of the observers in the target group. The ICC for average measures is always higher than ICC for single measures, but the ICC for single measures is the more relevant measure when testing an instrument designed for use in clinical practice.[23]

When judging the reliability of an instrument, an ICC value of ≥ 0.75 is often considered some kind of a threshold for a reliable instrument, although one must also take into account the CIs. The number of observers needed to achieve a rating with an ICC value of ≥ 0.75 was calculated.

Concurrent validity between the objective and the subjective assessments was analyzed using Pearson's correlation coefficient. To show how the observations and scores depend on each other, the dependency of the overall POSAS and VSS scores of the individual parameters was studied. The Pearson's correlation coefficient, r, can be used for measuring linear and Spearman's correlation coefficient, ρ, for nonlinear monotonic dependencies. Automatic relevance determination (ARD) can also measure nonmonotonic dependencies.[24] Here, the dependencies were first analyzed by constructing a Bayesian network from the data, using the B-Course tool.[25–27] The Bayesian network is a graphical representation of the dependencies between variables, including linear, nonlinear, and also nonmonotonic dependencies. The dependencies shown by the network were checked by calculating r and ρ for all the arcs of the graph. If there was no correlation, the dependency was checked using ARD, too.

Figure 2. (A-C) Three images from the same scar taken individually by the observers (A, B, C) using standardized digital imaging. The distance and angle of the camera and the lighting remain the same; only the rotational angle around the optical axis may vary as can be seen.

RESULTS

Interrater reliability

The ICCs for subjective and objective ratings are shown in Table 4. The objective measurement for the ECC of melanin and hemoglobin showed high reliabilities for both single and average measurements with very narrow CIs. A single observer could perform the measurement for both parameters reliably.

Both the POSAS and the VSS achieved reliability ratings below the acceptable limit ($r=0.60$ and 0.66 $p < 0.001$). The number of observers needed to perform a reliable subjective assessment (ICC > 0.75) with these scales was on average four observers for the VSS and five observers for the POSAS.

Concurrent validity

The correlations between the ECC values and the subjective ratings for the corresponding parameters are shown in

Table 5. There was a statistically significant correlation between the ECC values for melanin and hemoglobin and the subjective ratings for vascularity and pigmentation with both scales. The correlation between ECC of melanin and POSAS pigmentation rating was slightly higher than the correlation between the vascularity measurements ($r=0.68$ and 0.50, respectively). However, it must be noted that the subjective ratings for vascularity do not give measures below zero, whereas the ECC values in some scars achieved a negative measure, indicating that the scar has less hemoglobin than the surrounding normal skin.

Dependency analysis

The dependencies shown in the Bayesian network in Figure 3A are mostly linear and monotonic, as they have a high Pearson's correlation coefficient. The dependency of the overall POSAS score on ECC of hemoglobin is an exception. The r and ρ are low, but the α shows almost as strong a dependency as between VSS height and the

Table 4. ICCs of ratings

	Average measures	Single measures	95% CI	p
Subjective assessments				
POSAS	0.821	0.604	0.69–0.90	< 0.005
VSS	0.855	0.662	0.749–0.92	< 0.005
Spectrocutometry				
ECC of melanin	0.959	0.886	0.928–0.978	< 0.005
ECC of hemoglobin	0.955	0.875	0.920–0.976	< 0.005

The intraclass correlation coefficients (ICC) of subjective ratings and spectrocutometry measurements including the 95% confidence intervals (CI) of average measures and their p-values.
ECC, estimated concentration change; POSAS, patient and observer scar assessment scale; VSS, Vancouver scar scale.

Table 5. Concurrent validity

	r	p
ECC of melanin vs. POSAS pigmentation subscale	0.67	< 0.001
ECC of melanin vs. VSS pigmentation subscale	0.60	0.002
ECC of hemoglobin vs. VSS vascularity subscale	0.40	0.015
ECC of hemoglobin vs. POSAS vascularity subscale	0.50	< 0.001

The Pearson estimates (r) for correlations between the subjective ratings and spectrocutometry measurements.
ECC, estimated concentration change; VSS, Vancouver scar scale; POSAS, patient and observer scar assessment scale.

POSAS score. The α close to zero would have meant no dependency. Therefore, the dependency is true, but it is not monotonic.

The dependencies in the network can be interpreted as follows: the pliabilities of the POSAS and VSS scales are strongly dependent and correlated. The VSS overall score
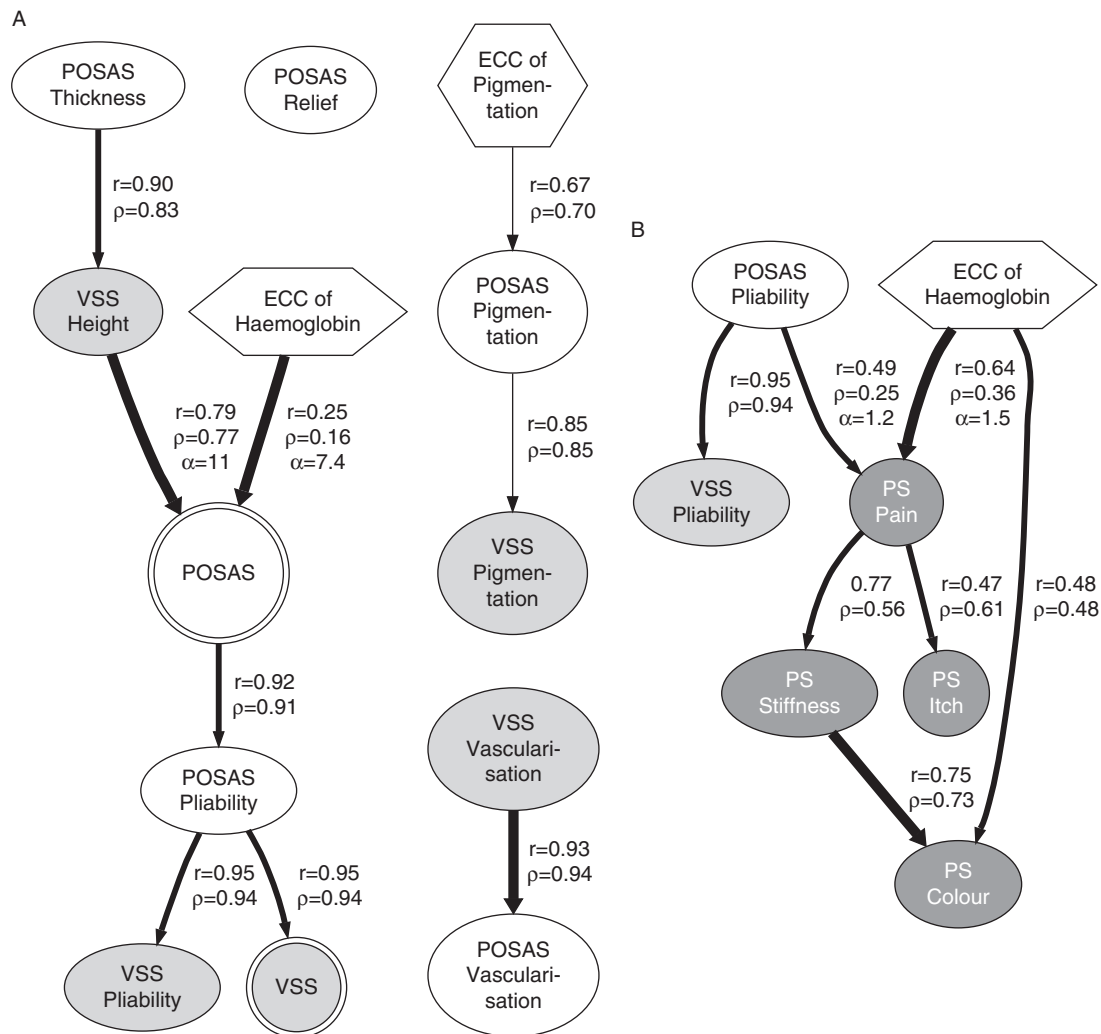
Figure 3. (A) Dependencies of the subjective ratings with the patient and observer scar assessment scale (POSAS) (white ellipses, and double circle for the overall score) and the Vancouver scar scale (VSS) (light gray ellipses, double circle for the overall score) and the spectrocutometry measurements (hexagons) of estimated concentration changes (ECC) of hemoglobin and melanin, (B) the dependencies of the patient score (PS, dark gray ellipses) values to the POSAS, the VSS, and the spectrocutometry values. The widths of the arcs show the strengths of the dependencies. Pearson (r) and Spearman (ρ) estimates and automatic relevancy determination (α) results are shown next to the corresponding arrows.

depends mostly on pliability, and to a lesser degree, the height, thickness, and ECC of hemoglobin. When pliability is measured, the other variables have no significance to the overall VSS score. Therefore, the overall VSS value can be calculated from the VSS pliability alone.

Similarly, the overall POSAS score can be calculated if thickness, ECC of hemoglobin, and pliability are known. Thickness can be replaced with VSS height estimate if that is available.

A similar network was constructed from the POSAS patient score (shown in Figure 3B). The patient estimate of scar color depends strongly on stiffness and moderately on ECC of hemoglobin. The patient opinion of pain is strongly dependent on ECC of hemoglobin and moderately on the observer ratings of pliability and patient rating of stiffness. Itching is dependent on pain. The rest of the dependencies were weak. These dependencies are mostly linear, as Pearson's correlation coefficient is also relatively high, but the patient opinion of pain does not strongly correlate with ECC of hemoglobin or observer pliability score. Therefore, the α of these dependencies was determined, too. The α shows that notwithstanding the low correlation, these dependencies are true and that the pain depends on the ECC of hemoglobin slightly stronger than on observer pliability rating.

DISCUSSION

Why is quantitative scar assessment needed? When treating HSs and keloids, quantitative information on the effect of designated treatments is essential. This is emphasized by the fact that the patient is often treated by several physicians and other professionals over time. Such a setting demands an instrument that has a high interobserver reliability.

Secondly, most scar treatments have not been properly tested by well-designed studies. Silicone gel sheeting may be the most widely used conservative method and has been used for over 20 years. Despite its wide use, it was stated in a Cochrane database review from the year 2008 that the evidence of silicone treatment is scarce and studies suggesting its use are susceptible to bias.[28] One reason for the lack of evidence is the fact that the majority of studies investigating scar treatments have been conducted using only subjective rating as an end point.[29,30]

In a recent study by Forbes-Duchart et al.,[31] it was noted that the majority of professionals treating burn scars do not use any instrument, subjective or objective, for scar assessment. In our experience, this makes it difficult to decide whether a treatment has been effective or not. Therefore, we believe there is a demand for a quantitative, objective, easy-to-use instrument for scar assessment both for scientific purposes and for clinical work.

Spectrophotometry has been tested in scar assessment in studies by Nedelec and colleagues and has also been used recently in a clinical study by Bloemen and colleagues.[32–34] The instruments used in these studies, the DermaSpectrometer® (Cortex Technology, Hadsund, Denmark), the Minolta Chromameter® (Minolta, Osaka, Japan), and the Mexameter® (Courage-Khazaka Electronic, Köln, Germany), supply quantitative information of scar color. The problem with these instruments is that they have narrow apertures (5–8 mm) and they can only give a result from a very limited part of the scar. Also, they do not provide an

estimate of the concentration levels of the color chromaphores, hemoglobin and melanin, which cause the color change. A clear benefit of Spectrocutometry is that it can be used to measure the entire scar and its surroundings, yielding an average result of the concentration changes of hemoglobin and melanin compared with the healthy skin. This also minimizes the effect of the change in skin circulation between two measurements.

SIAscopy™ (Biocompatibles, Farnham, UK) is a newer innovation that provides more specific information than the previously mentioned spectrophotometers, as it quantifies the concentrations of hemoglobin, melanin, and collagen. SIAscopy™ gives a score of only a 24×24 mm or smaller area. It may be reliable when used to measure small scars, but is less useful when an average measurement of a larger scar is needed.

Various scar rating scales have been introduced in publications. The VSS is still the most popular and well-known alternative for scar assessment.[31] It is relatively easy to use, but many studies have found it to be too unreliable.[9] Our study makes no exception, and although the VSS showed slightly better reliability than the POSAS, it cannot be considered a reliable tool for clinical practice. Also, to incorporate it into a scientific study would require at least four observers, when calculated from the ICC results in this study, which is supported by findings in previous studies.[35]

The POSAS has been the rating scale of interest in recent years. Truong et al.[36] tested the reliabilities of both the POSAS and VSS in rating scars after breast cancer surgery. The ICC for single measures in their study was far from acceptable ($r=0.54$ and 0.33 for the VSS and the POSAS, respectively). In another study by the same authors, the ICC of the VSS for rating breast cancer surgery scars was 0.64 (0.51–0.74, 95% CI), which is comparable with the levels found in our study, and cannot be considered sufficient.[37] In a study by Draaijers et al.,[38] the POSAS scale showed better interrater reliability, with an ICC of 0.73 for single measurements. It can be stated that standardized training can increase the reliability of subjective scar rating and may produce a better outcome in reliability studies than shown in our study. However, in the study presented here, there was no training involved in the use of the proposed Spectrocutometry method, which still achieved acceptable reliability.

The B-Course tool reveals an interesting aspect of scar assessment, especially with the POSAS scale. The overall score is dependent on only two parameters, the thickness and pliability, while other parameters are independent. Also, the only significant dependency between the patient and the observer scores is between pliability and scar pain. These results indicate that the observer scale is only measuring the thickness and pliability of the scar. Pigmentation, vascularization, and relief are irrelevant and so are all the parameters of the patient scale. The results in this patient group using the POSAS scale question the use of the patient as an additional observer. Durani et al.[39] have presented the Patient Scar Assessment Questionnaire, which showed a better correlation when compared with the clinician assessment carried out with the VAS scale and the MSS.

The properties of a scar change over time, a process known as scar maturation. Scar redness, i.e., increased hemoglobin concentration, is present in the early phase and tends to decrease with time.[11] The maturation process

depends on various factors, such as scar location, etiology, individual response to scarring, wound tightness, etc. The treatment of problematic, HSs, such as burn scars, aims to produce mature scars. It is reasonable to believe that the ECC of hemoglobin value, produced by Spectrocutometry, is a relevant parameter when assessing an immature scar, whereas other measures, such as pliability or volumetric measurements, might be more relevant when rating a mature scar. We believe that Spectrocutometry can be used as a tool in clinical work to gain important information on active, HSs, and on the other hand, in scientific studies, where scar treatment or prevention methods are tested in a prospective, randomized setting.

Logically, it is not possible to gain a full picture of a scar by any single objective method. The VSS and POSAS scales are designed to combine multiple parameters for a thorough scar assessment. Unfortunately, this may result in seemingly different scars in gaining similar scores, as noted by Duncan et al.[11] It would be valuable to obtain an objective measure of such parameters as pliability and thickness, which, based on the finding in our study, seem significant. The Cutometer® (Courage-Khazaka Electronic) has been used by Nedelec et al.[40] to measure the pliability of scars, but the reliability of the method was not sufficient.

The ECC of hemoglobin correlates with the overall POSAS score, and also has a strong dependency on the scar pain, a fact worth noticing. Also, it is known that increased redness, reflecting a high hemoglobin concentration, is strongly associated with scar hypertrophy.[41] The ECC of melanin, on the other hand, as well as the subjective pigmentation ratings, seem to have no correlation with the overall scores, although they are correlated with each other and the patient rating of scar color. It appears that hypo- or hyperpigmentation are factors that may affect the cosmetic outcome of the scars, but do not reflect the scars' activity or degree of hypertrophy.

The proposed Spectrocutometry method showed high reliability in assessing linear scars in this study. A valuable aspect of this method is that it provides standardized image documentation of the scar. Conventional digital photography is dependent on the ambient lighting, and often, the distance and the angle to the target are changed, which makes it difficult to use the digital images for later judgment. As shown in the images in Figure 2, the scar and its surroundings in the images look quite identical even though the images were taken individually by different observers.

The mean age of the scars in this study was 17 months, which means that most of the scars were mature. It would be beneficial to study this method in a group of more immature scars. In the future, we aim to study the reliability and feasibility of Spectrocutometry on burn scars. We believe that this is an easy and reliable way to quantitatively measure and document changes in scar activity, and recommend it for both clinical and scientific work.

REFERENCES

1. Junker JP, Kratz C, Tollback A, Kratz G. Mechanical tension stimulates the transdifferentiation of fibroblasts into myofibroblasts in human burn scars. *Burns* 2008; 34: 942–6.
2. Shin D, Minn KW. The effect of myofibroblast on contracture of hypertrophic scar. *Plast Reconstr Surg* 2004; 113: 633–40.
3. Sheridan RL, Tompkins RG. What's new in burns and metabolism. *J Am Coll Surg* 2004; 198: 243–63.
4. Li-Tsang CW, Lau JC, Chan CC. Prevalence of hypertrophic scar formation and its characteristics among the Chinese population. *Burns* 2005; 31: 610–6.
5. Bombaro KM, Engrav LH, Carrougher GJ, Wiechman SA, Faucher L, Costa BA, Heimbach DM, Rivara FP, Honari S. What is the prevalence of hypertrophic scarring following burns? *Burns* 2003; 29: 299–302.
6. Ogawa R. The most current algorithms for the treatment and prevention of hypertrophic scars and keloids. *Plast Reconstr Surg* 2010; 125: 557–68.
7. Mustoe TA, Cooter RD, Gold MH, Hobbs FD, Ramelet AA, Shakespeare PG, lla M, Téot L, Wood FM, Ziegler UE, International Advisory Panel on Scar Management. International clinical recommendations on scar management. *Plast Reconstr Surg* 2002; 110: 560–71.
8. Bloemen MC, van der Veer WM, Ulrich MM, van Zuijlen PP, Niessen FB, Middelkoop E. Prevention and curative management of hypertrophic scar formation. *Burns* 2009; 35: 463–75.
9. Durani P, McGrouther DA, Ferguson MW. Current scales for assessing human scarring: a review. *J Plast Reconstr Aesthet Surg* 2009; 62: 713–20.
10. Beausang E, Floyd H, Dunn KW, Orton CI, Ferguson MW. A new quantitative scale for clinical scar assessment. *Plast Reconstr Surg* 1998; 102: 1954–61.
11. Duncan JA, Bond JS, Mason T, Ludlow A, Cridland P, O'Kane S, Ferguson MW. Visual analogue scale scoring and ranking: a suitable and sensitive method for assessing scar quality? *Plast Reconstr Surg* 2006; 118: 909–18.
12. Chan KY, Lau CL, Adeeb SM, Somasundaram S, Nasir-Zahari M. A randomized, placebo-controlled, double-blind, prospective clinical trial of silicone gel in prevention of hypertrophic scar development in median sternotomy wound. *Plast Reconstr Surg* 2005; 116: 1013–20; discussion 21–2.
13. Nguyen DQ, Potokar T, Price P. A review of current objective and subjective scar assessment tools. *J Wound Care* 2008; 17: 101–2, 4–6.
14. Draaijers LJ, Tempelman FR, Botman YA, Kreis RW, Middelkoop E, van Zuijlen PP. Colour evaluation in scars: tristimulus colorimeter, narrow-band simple reflectance meter or subjective evaluation? *Burns* 2004; 30: 103–7.
15. Jung B, Choi B, Durkin AJ, Kelly KM, Nelson JS. Characterization of port wine stain skin erythema and melanin content using cross-polarized diffuse reflectance imaging. *Lasers Surg Med* 2004; 34: 174–81.
16. Latreille J, Gardinier S, Ambroisine L, Mauger E, Tenenhaus M, Guehenneux S, Morizot F, Tschachler E, Guinot C. Influence of skin colour on the detection of cutaneous erythema and tanning phenomena using reflectance spectrophotometry. *Skin Res Technol* 2007; 13: 236–41.
17. Takiwaki H, Miyaoka Y, Kohno H, Arase S. Graphic analysis of the relationship between skin colour change and

variations in the amounts of melanin and haemoglobin. *Skin Res Technol* 2002; 8: 78–83.

18. Zijlstra WG, Buursma A, Meeuwsen-van der Roest WP. Absorbtion spectra of human fetal and adult oxyhemoglobin, de-oxyhemoglobin, carboxyhemoglobin and methemoglobin. *Clin Chem* 1991; 37: 1633–8.

19. Välisuo P, Kaartinen I, Kuokkanen H, Alander J. The colour of blood in skin—a comparison of Allen's test and photonics simulations. *Skin Res Technol* 2010; 16: 390–6.

20. Kaartinen IVP, Kuokkanen H., Alander J. Objective scar assessment—a new method using standardized digital imaging and spectral modelling. *Burns* 2011; 37: 74–81.

21. Välisuo P, Harju T, Alander J. Reflectance measurement using digital camera and a protecting dome with built in light source. *J Biophotonics* 2010; DOI: 10.1002/jbio.201000113

22. Bochko VVP, Harju T., Alander J. Lower extremity ulcer image segmentation of visual and near-infrared imagery. *Skin Res Technol* 2010; 16: 190–7.

23. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; 86: 420–8.

24. Nabney I. *Netlab: algorithms for pattern recognition.* Berlin: Springer Verlag, 2002.

25. Cooper GF. The computational complexity of probabilistic inference using Bayesian belief networks. *Artif Intell* 1990; 42: 393–405.

26. Pearl J. *Probabilistic reasoning in intelligent systems: Publisher: Networks of Plausible Inference.* Los Altos, CA: Morgan Kaufmann, 1988.

27. Myllymäki P, Silander T, Tirri H, Uronen P. B-course: a web-based tool for Bayesian and causal data analysis. *J Artificial Int Tools* 2002; 11: 369–88.

28. O'Brien L, Pandit A. Silicon gel sheeting for preventing and treating hypertrophic and keloid scars. *Cochrane Database Syst Rev* 2006: CD003826.

29. Harte D, Gordon J, Shaw M, Stinson M, Porter-Armstrong A. The use of pressure and silicone in hypertrophic scar management in burns patients: a pilot randomized controlled trial. *J Burn Care Res* 2009; 30: 632–42.

30. Karagoz H, Yuksel F, Ulkur E, Evinc R. Comparison of efficacy of silicone gel, silicone gel sheeting, and topical onion extract including heparin and allantoin for the treatment of postburn hypertrophic scars. *Burns* 2009; 35: 1097–103.

31. Forbes-Duchart L, Cooper J, Nedelec B, Ross L, Quanbury A. Burn therapists' opinion on the application and essential characteristics of a burn scar outcome measure. *J Burn Care Res* 2009; 30: 792–800.

32. Bloemen MC, van Leeuwen MC, van Vucht NE, van Zuijlen PP, Middelkoop E. Dermal substitution in acute burns and reconstructive surgery: a 12-year follow-up. *Plast Reconstr Surg* 2010; 125: 1450–9.

33. Draaijers LJ. Colour evaluation in scars: tristimulus colorimeter, narrow-band simple reflectance meter or subjective evaluation? *Burns* 2004; 30: 103–7.

34. Nedelec B, Correa JA, Rachelska G, Armour A, LaSalle L. Quantitative measurement of hypertrophic scar: interrater reliability and concurrent validity. *J Burn Care Res* 2008; 29: 501–11.

35. van Zuijlen PP, Angeles AP, Kreis RW, Bos KE, Middelkoop E. Scar assessment tools: implications for current research. *Plast Reconstr Surg* 2002; 109: 1108–22.

36. Truong PT, Lee JC, Soer B, Gaul CA, Olivotto IA. Reliability and validity testing of the Patient and Observer Scar Assessment Scale in evaluating linear scars after breast cancer surgery. *Plast Reconstr Surg* 2007; 119: 487–94.

37. Truong PT, Abnousi F, Yong CM, Hayashi A, Runkel JA, Phillips T, Olivotto IA. Standardized assessment of breast cancer surgical scars integrating the Vancouver Scar Scale, Short-Form McGill Pain Questionnaire, and patients' perspectives. *Plast Reconstr Surg* 2005; 116: 1291–9.

38. Draaijers LJ, Tempelman FR, Botman YA, Tuinebreijer WE, Middelkoop E, Kreis RW, van Zuijlen PP. The patient and observer scar assessment scale: a reliable and feasible tool for scar evaluation. *Plast Reconstr Surg* 2004; 113: 1960–5; discussion 6–7.

39. Durani P, McGrouther DA, Ferguson MW. The Patient Scar Assessment Questionnaire: a reliable and valid patient-reported outcomes measure for linear scars. *Plast Reconstr Surg* 2009; 123: 1481–9.

40. Nedelec B, Correa JA, Rachelska G, Armour A, LaSalle L. Quantitative measurement of hypertrophic scar: intrarater reliability, sensitivity, and specificity. *J Burn Care Res* 2008; 29: 489–500.

41. Li-Tsang CW, Lau JC, Liu SK. Validation of an objective scar pigmentation measurement by using a spectrocolorimeter. *Burns* 2003; 29: 779–84.