Perspectives on burn scar evaluation and artificial skin

van Zuijlen, P.P.M.

Citation for published version (APA):
van Zuijlen, P. P. M. (2002). Perspectives on burn scar evaluation and artificial skin.
1

Scar assessment tools: implications for current research

Paul PM van Zuijlen 1,2,3, Adam P Angeles 1, Robert W Kreis 1,2,4,5, Kurt E Bos 3, Esther Middelkoop 1,4.

From the Burn Centre 1 and the Department of Surgery 2 of the Red Cross Hospital, Beverwijk; the Department of Plastic, Reconstructive and Hand Surgery 3, Academic Medical Centre, Amsterdam; the Dutch Burns Foundation 4, Beverwijk; and the Department of Surgery 5, VU University Medical Centre, Amsterdam, the Netherlands.
Abstract  Scarring is considered a major medical problem that leads to cosmetic and functional sequelae. Scar tissue is clinically distinguished from normal skin by an aberrant colour, rough surface texture, increased thickness (hypertrophy), occurrence of contraction, and firmness. Marked histologic differences are the change in dermal architecture and the presence of cells like the myofibroblast. Many assessment tools are available for analysis of pathologic conditions of the skin. However, there is no general agreement as to the most appropriate tools for evaluation of scar tissue in specific. This review critically discusses currently available objective measurement tools and subjective assessment tools as well as potential devices that may be available in the future that make up the scar assessment.
Introduction

The skin acts as a shield to the outer world and has several important functions. It defends the body against invasion of microorganisms, fluid loss, and it must withstand shear and tractive forces. Only decades ago, even small wounds were potentially lethal. It is understandable that, then as well as now, wound healing was aimed at rapid wound closure. The quality of the repaired tissue was therefore of secondary importance. Today however, patients may survive large wounds due to advances in shock therapy and infection prevention. The quality of the product of wound healing is therefore becoming increasingly more important with emphasis on functional and cosmetic outcomes.

The derailment of the wound healing process often results in a raised, firm scar with a poor colour match to its surrounding tissue. It is widely known that scars may lead to an array of cosmetic, functional and psychological consequences for the injured patient whether they result from minor insults or severe injuries such as burns.

Although an increasing number of subjective and objective tools are becoming available, there is no general agreement as to the most appropriate tool(s) for scar evaluation. In this review, we outline problems with scar evaluation and discuss the currently available literature as it pertains to scar assessment tools. Subjective and objective scar assessment techniques are discussed including several types of devices that may be used during clinical evaluation as well as techniques for histologic analysis. The appropriateness of assessment scales and measuring devices are critically reviewed with respect to the concepts of reliability and accuracy. Reliability refers to the reproducibility of measurements or ratings by observers, whereas the accuracy, also termed ‘validity’, refers to the exactness of the measurements or ratings, in other words, do we measure what we want to measure? Clearly, the use of reliable and accurate tools to measure features of scars is of paramount importance if we are to establish empiric evidence, leading to a better understanding of the natural history of scar formation and therapeutic options.
Scar features with relevance for research purposes

The characteristics of a scar depend on its aetiology, depth, size, location, type of treatment, and the individual’s genetic predisposition\(^5\). Frequently scars lead to disturbances in skin function and changes in both macroscopic and microscopic skin architecture. As a consequence, a number of undesired signs generally occur such as colour mismatch, surface texture, increased thickness, firmness, contraction, and contracture.

Colour mismatch and coarse texture of the scar surface pertain to cosmetic outcome. For example, the perception of the mesh pattern after split thickness autografting is a well-known long-term sequel of scar formation and is considered a disturbing feature by both patients and physicians. A clinical example of the ‘patchwork quilt’ in a matured scar is displayed in Figure 1a.

Contraction of the injured area occurs frequently during the wound healing process and can interfere with the range of motion of joints, as illustrated by Figure 1b. The contraction mechanism is hypothesised to result from granulation tissue fibroblasts, the so-called myofibroblast (which exhibits features of smooth muscle cells)\(^6\)–\(^8\). In addition, other non-cellular features have been shown to play a role in scar contraction such as age\(^9\)–\(^10\), anatomical location\(^n\) and the influence of tension and mobility of the skin\(^n\). Despite an extensive number of \textit{in vitro} and animal studies, there is no definite proof provided by clinical studies that the myofibroblast actually is responsible for scar contraction in humans. The absence of such studies is, in part, a consequence of the lack of well-established tools to measure scar surface areas. As shown in Figure 1c, hypertrophy can be a stigmatising feature of scar formation and may reduce the elastic qualities of the skin. Hypertrophy or sometimes atrophy of the scar is quantified by measuring scar thickness or volume. This can be done clinically by assessing its height or by measuring its thickness with devices like ultrasound\(^n\). Scar thickness can also be assessed through histological analysis of biopsied tissue, the current ‘gold standard’, however, less invasive techniques which have a comparable or greater reliability and accuracy are sought.
Collagen structure is one of the many areas of interest in wound healing research as scar tissue and normal tissue display different collagen architecture. Histopathological studies of mature scar tissue show a collagen bundle orientation parallel to the epidermis (Figure 2a) rather than a more randomly organised pattern found in normal dermis (Figure 2b). Typically, histopathological studies are performed by one or two observers using conventional light microscopy in combination with polarised light. The optimal assessment technique of collagen orientation, however, has not been established. Recently, sophisticated computerised image analysis methods and other non-invasive methods have become available for collagen orientation assessment, which will be discussed below. Mechanical qualities of the skin and scar are well appreciated. Firmness and elasticity result from a collagen-elastin network which is impregnated in a ground substance of extracellular matrix molecules (proteoglycans). Mechanical devices and subjective scales are available for the evaluation of elastic qualities of the scar. In addition, clinical tests have been developed to assess joint range of motion or the performance of daily activities in order to establish the consequences of scar formation on functional impairment. Having reviewed the most crucial scar features we will give an outline of the currently available scar assessment tools for the evaluation of scar features.

Clinical (non-invasive) evaluation of scars

Scar colour - colorimetry

A colour mismatch can be a stigmatising feature of a scar. Aberrant skin colour changes are related to disturbances in vascularisation and pigmentation. Colour analysis is part of a burn scar assessment tool also known as the Vancouver Scar Scale, which subjectively measures pigmentation and vascularity (as well as pliability and scar height that are discussed below). Pigmentation initially is scored as normal, hypo- and hyperpigmented. The vascularity score is ranked as follows: normal, pink, red and purple. Because only a moderate inter-rater agreement was documented for each parameter (Cohen's $\kappa$ of approximately 0.5), the authors of the origi-
Figure 1  Features of scar formation

Figure 1a  This figure shows a fully matured 8-year old burn scar with a disturbed surface texture and with a reduced mobility of the shoulder.

Figure 1b  A contracted 6-month old burn scar in the elbow region, leading to impaired range-of-motion of this joint (contracture). Note the visibility of the mesh pattern.

Figure 1c  A hypertrophic 7-month old burn scar on the right forearm.
nal paper noted that the scale needed improvement especially with dark, vascular scars and mixed pigmented scars. Others addressed this criticism by adding the term mixed pigmentation to the list. Their agreement was however only moderate for pigmentation ($\kappa=0.61$).
Beausang et al.\textsuperscript{19} proposed not to rate vascularisation and pigmentation separately but to only use one scale for colour mismatching with the ranking: none, slight, obvious and gross mismatch. In this study no data were given for separate parameters however, a good correlation ($r=0.87$) was obtained for the reliability of ten observers for the complete scale concerning photographic assessment. Although the observer may distinguish thousands of colours, the human brain can not reliably and accurately quantify the colour or its intensity. Moreover, it is difficult to accurately memorise colours, which interferes with the quality of scar colour ratings if they have to be evaluated during different time periods\textsuperscript{20}. Therefore, a colour model based on human perception was developed by the CIE (Commission Internationale de l'Eclairage) and has proven useful for research in the dermatological field\textsuperscript{21}. Different devices have been developed that have integrated the CIE tri-stimulus colour values into their systems. These include the Minolta chromameter\textsuperscript{TM} CR-200 and 300 (Minolta Camera Co., Osaka, Japan), the Labscan 6000\textsuperscript{®} (Hunter Associates Inc., USA), and Dr Lange Micro Color (Dr Bruno Lange GmbH, Dusseldorf, Germany).

Spectrophotometric colour measurement is a different method that calculates an erythema and melanin index. The erythema index (EI) is defined as: $EI = 100 \times \log \left( \frac{\text{intensity of reflected red light}}{\text{intensity of reflected green light}} \right)$ and the melanin index (MI) is defined as: $MI = 100 \times \log \left( \frac{1}{\text{intensity of reflected red light}} \right)$. The erythema index of this device is therefore related to the melanin index. The DermaSpectometer\textsuperscript{®} (Cortex Technology, Hadsund, Denmark) is a handheld instrument based on this principle\textsuperscript{22}.

The Minolta Chromameter has been applied for scar evaluation\textsuperscript{23} and was shown to be equal to the Micro Color\textsuperscript{24} for the assessment of quantification of erythema and was found to correlate well with the Dermaspectrometer regarding erythema assessment\textsuperscript{25,26}. The melanin index of the DermaSpectrometer correlated only moderately well with the Chromameter, which measures pigmentation by assessing brightness. The melanin index seems to correlate more specifically with pigmentation itself instead of scoring brightness\textsuperscript{20}. The question that is yet unanswered, is which colour measurement tool has superior accuracy.
If we are interested in scar colour because we want to measure the extent of vascularisation, we might consider the laser Doppler flow meter as this tool is well suited for the evaluation of cutaneous blood flow. The laser Doppler flow meter correlates positively with different colorimeters (Minolta Chromameter and Dr Lange Micro Meter) but is less sensitive than simple visual assessment of erythema. With respect to erythema, it has been concluded that blood volume under a given area, and not the velocity of erythrocytes, determines skin redness. Therefore laser Doppler flow meter should not be the first choice in measuring skin redness.

A major drawback of the reliability studies on colour measurement that are found in the literature is that these studies were performed on pathological conditions of the skin like erythema and not on scar tissue. It remains questionable whether these techniques can also be applied with the same accuracy to the evaluation of scar tissue. For example, it has been stated that the surface texture of the test area influences the measurements by the DermaSpectrometer. This might interfere with measurements on bumpy scar surface areas such as hypertrophic areas or keloids. Further studies are warranted to substantiate the accuracy and reliability of assessment tools for scar colour.

**Surface texture - profilometry**

The surface of scars often becomes irregular. This bumpiness is frequently seen with mesh patterns of burn autograft treatments as the areas between the mesh become raised, forming a 'patchwork quilt'. As shown in Figure 1a such irregularities may remain at least for many years. The degree of irregularity has been evaluated subjectively with good reliability by the Hamilton scale and a scale developed by Smith and colleagues, however data regarding the accuracy of these tools are lacking.

Devices are available that objectively establish the contour of skin surfaces, directly, or indirectly by creating a negative replica of the skin. The contours are analysed with an optical or mechanical profilometer. These devices were designed for applications in the cosmetic industry and have not been studied in the evaluation of scars.
Surface area - planimetry

Measuring surface areas is referred to as planimetry. Planimetry is useful for calculating the extent of contraction over time. Although contraction is considered beneficial to obtain fast wound closure\(^{35}\), the process of contraction frequently continues after the epithelium is restored\(^{35}\), which may lead to a disturbed function of skin and joints. The importance of this complication is well appreciated by numerous animal studies\(^{36-40}\). Several human studies have utilised tracing techniques to evaluate contraction at the acute period after wounding\(^{41-45}\), but there is a lack of studies employing this technique for scar analysis over the long term. This paucity of studies most likely stems from the fact that it is difficult to perform planimetry reliably on scars in clinical situations. For example, scar margins become more difficult to delineate as they mature, or the margins are obscured by blood and the like during operative procedures. Tracing of scar margins on clear plastic film or planimetry by photography\(^{46}\) are the most common methods of planimetry. More sophisticated methods make use of structured light that allow scanning of body areas, which is followed by a three-dimensional computer reconstruction\(^{47}\). Even full body scanners have become available\(^{48}\). Although the latter might become a 'gold standard' in time, we advocate the development of inexpensive and uncomplicated methods that can be clinically validated and which can be applied without limitation. We have found (unpublished data) that planimetry by photography is more accurate and reliable than planimetry by tracing for measurements on flat surfaces like the back and moderately curved surfaces like the thigh. Planimetry by photography is less accurate than planimetry by tracing primarily for extremely curved body parts. This result is probably due to distortion of the highly curved areas. Overall however, we recommend the photographic technique as a standard evaluation method.

Thickness

Measuring scar thickness is relevant for the clinical evaluation of hypertrophy, which can be a frequent and cosmetically disfiguring sequel of scarring. The question is whether we are interested in the total thickness of the scar or only in the protruding part of the scar.
The latter is, in general, a subjective tool as observers can only visually assess scars. The parameter ‘scar height’ has been fitted into different scales. The scale of Yeong and the Hamilton scale showed to have a good reliability (weighted \( k > 0.81 \)), however other scales, like the Vancouver Scar Scale, reported only moderate outcome or no specific outcome to this parameter. It should be noted that the reliability of the scoring improved with the number of raters as Smith demonstrated a moderate reliability for one observer \( (r = 0.48) \), and good reliability for four \( (r = 0.79) \) and eight observers \( (r = 0.88) \). It is therefore inappropriate to compare the results of different groups without considering the number of raters.

The application of ultrasound for thickness determination of normal skin has been shown to be reproducible and accurate. Ultrasound has been applied for scar analysis after breast reduction and burn wound treatment that provided quantitative information of scar thickness. Others measured skin thickness with high resolution magnetic resonance imaging, which was found to be reliable and accurate for normal skin. This technique has not been employed for scar analysis. We recommend using ultrasound as a tool to measure scar thickness because it has been tested specifically with scar evaluation and found to be accurate, reproducible, and convenient.

**Tissue Organisation**

Until recently, tissue organisation could only be evaluated by invasive techniques, like analysing biopsies with polarised light microscopy. Such techniques are discussed below. However, a non-invasive clinical approach to tissue organisation has been developed utilising sound waves. The velocity of propagation of ultrasonic vibrations can supply information concerning the mechanical properties of bony tissue and soft tissues. A higher velocity of wave propagation indicates a higher degree of stiffness of the investigated material. Potts et al. were the first to apply this technique to skin measurements and more recently it has become available for clinical purposes. The maximum depth of sound wave propagation through the skin is approximately one millimeter. The method has been applied to hypertrophic scar tissue, which appears to have a markedly higher stiffness compared to normal skin.
This technique enables the detection of anisotropy of tissue. Anisotropy is deduced from the velocity of propagation in several directions. If only one direction is measured it does not tell us whether it is predominantly parallel or perpendicular to the axis of the anisotropy. This implies that points of interest have to be measured always in all directions. Further studies are encouraged to better elucidate the histologic findings which correlate with these sound waves.

**Scar performance: function at skin level**

The mechanical qualities of the skin are well appreciated and studied. Experimental models in *in vivo* animal models and studies in cadaveric skin have been developed to measure a variety of mechanical skin characteristics. By this means, parameters such as the tensile strength, ultimate extension and stress-strain curves are obtained. An extensive overview of these experiments is published by Vogel. Evaluation of mechanical parameters is also part of some scar assessment tools. We outline these methods following four subcategories based on the type of load on the skin: suction; pressure; torsion; and tension.

**a) Suction Methods**

At present, some methods have been published in the effort to objectively evaluate elastic measurements by means of suction principles during clinical trials. The most frequently described technique is the Cutometer® SEM 474 and the latest 575 version (Courage and Khazaka Electronic GmbH, Cologne, Germany). It exerts a controlled negative pressure over a small area of the skin. Data regarding skin deformation are uploaded to a computer. The Cutometer is typically used in the deformation-versus-time mode. This means that a constant negative pressure is used for a short period alternated with periods of normal pressure. A schematic drawing of the Cutometer is given in Figure 3a. In a study on sclerodermal lesions, measurements with this device were found to be highly reliable and have good intra-observer and inter-observer agreement. The Cutometer measurements correlated well with the most widely accepted and utilized scoring scale for skin thickness - the modified skin...
score. When the Cutometer was tested on normal skin reproducible results were obtained. The handheld probes have different aperture diameters ranging from 2 to 8 mm. The 2 mm probe measures qualities of the epidermis and the superficial dermis whereas the 8 mm probes measures the dermis itself together with the slid-
ing mobility of the junction between dermis and the layer of the subcutaneous fat. The 6 mm probe provides the most accurate measurements of dermal elasticity. The Cutometer has been used to study healthy and diseased skin. Its application yielded objective information concerning the influence of aging, sex, and anatomical location on elasticity of healthy skin. Cutometer measurements on the forearm were only minimally influenced by limb position. Of note is the fact that there was a very short-term effect of hydration on the elasticity and retraction parameters already after one minute. It confirms the generally accepted assumption that all elasticity measurements should be taken under comparable circumstances to avert confounding through the influence of humidity and local temperature. More recently, the Cutometer has played a role in the evaluation protocol of burn scars and after reconstructive surgery.

The Dermaflex (Cortex technology, Hadsund, Denmark) has a comparable working mechanism compared to the Cutometer. The main difference is the larger diameter of the suction chamber (10 mm) of the Dermaflex. A good accuracy and reproducibility for this device has been established for elasticity measurements in sclerodermat. We choose to use the Cutometer to evaluate scars but the Dermaflex should also suffice although no trials on scar assessment with this instrument are published.

b) Pressure Methods

Katz and colleagues first introduced tonometry for burn scar rating in 1985. In a small group of patients a specially designed tonometer, the cicatrometer, was tested. The tonometer was originally designed for ophthalmologists to measure intra-ocular pressure. It is composed of an air flow system, a build-in sensor and a membrane that makes contact with the skin surface. Its principle is shown in Figure 3b. When applied to the skin, the amount of pressure is measured that is needed to lock the system. The utility for scar assessment by this technique remains unclear as its reliability and reproducibility have not been sufficiently investigated.

The Durometer (model 1700 Type o Rex Gauge Company Inc., Glenview, Ill., USA) was developed for hardness measurements of metals
and plastics. It provides an indentation load in the vertical direction on the specimen and thereby calculates the hardness of the material, as shown in Figure 3c. It was first introduced for skin elasticity measurements by Falanga et al.\textsuperscript{76} He and his group tested the device for the assessment of skin hardness in scleroderma\textsuperscript{76}, lipodermatosclerosis\textsuperscript{77}, and skin induration in venous diseases\textsuperscript{78}. The primary disadvantage of the device is that it is not applicable to anatomical locations where bony structures are situated directly under the skin (i.e. the hand, fingers or face). In such cases the hardness of the bone influences measurements\textsuperscript{76}. We question whether the type of underlying tissue, other than bony tissue, may also cause confounding of the measurements. Theoretically, the Durometer is applicable for scar assessment. A good correlation with the skin severity score was noted and the Durometer appeared to be a useful device to measure skin hardness but no rigorous testing of its reproducibility, accuracy or inter-observer variability has been conducted. Due to the lack of studies specifically involving scar evaluation and the several possible confounding variables, we do not feel that this device has a place in the scar assessment armamentarium at present.

c) Torsion Methods

The mechanical behavior of the skin can be assessed via torsional force onto the skin\textsuperscript{79-81}. In contrast to the devices that assess elasticity or firmness, torsional force causes a load in the horizontal plane instead of the plane perpendicular to the skin. Although different devices have been developed, they have all been constructed using the same principle that is represented in Figure 3d. Each device has a rotating flat disk which is placed in contact with the skin surface. The disk applies a rotational force along the plain of the skin surface. The rotational load is applied by a motor with controllable voltage, and allows adjustment of the torque. A ring guard was constructed circumferential to the disk leaving a strip of skin in between. Different groups have studied the effect of a constant load\textsuperscript{82} or increasing load followed by relaxation\textsuperscript{83}. The Dermal Torque Meter (Dia-stron Ltd. Andover, UK) is a commercially available device that has been applied for the evaluation of burn scars\textsuperscript{84}. It gives a time-deformation curve that shows a remarkable resem-
blance with Cutometer readings when a constant torsional load is applied. Nevertheless, corresponding parameters of both techniques correlate poorly to one another. Conclusions regarding the suitability of suction techniques are therefore not transferable to measurements by torsional devices. Their appropriateness for scar assessment has yet to be determined.

d) Tension methods

The term skin extension predominantly refers to skin deformation in the horizontal plane. Initially, when tension is applied in a linear direction, collagen bundle convolutions straighten out. Over time, more fibers become aligned in the direction of tension, and finally, the parallelly aligned bundles further extend under increased load. The directional variation depends on the woven pattern of the bundles and the main direction of the fibers, as skin extends maximally in the direction perpendicular to the main direction. Uniaxial tension on the in vivo skin shows the anisotropy of the skin if applied in different directions. Langer demonstrated this effect by making circular wounds in cadavers that became oval due to directional variations in the skin. He developed a map of the human body containing the 'Langer lines'. Since then many surgical incisions are made in the direction of these lines, which are in general, perpendicular to the direction of maximal skin extension, as the wounds gap less in this way and produce a less conspicuous scar.

Devices have been developed for extension evaluation by distracting two loci of skin, as shown in Figure 3e, that concerned ex vivo and in vivo measurements on skin specimen. The 'elastometer' is a hand-held version of the latter device for measurements of skin in vivo skin and has been used to measure hypertrophic burn scars. Unfortunately, the reliability and accuracy of all these devices are sparsely examined.

Larrabee and Tsap propose a different approach to skin and scar assessment: the finite element model. This model attempts to bridge the gap between simplified clinical models and sophisticated but impractical equations that biomechanists use to describe tissue properties. The object of interest is represented as a grid consisting of many small triangular elements. Analysing the deformation of the
framework of each element under tension allows a calculation of deformation. It provides an objective way to calculate elastic properties of burn scars relative to its surrounding areas. As more experience is gained with both these techniques, more information will become available with respect to their suitability for burn scar assessment.

Scar performance: function at joint level

The location and size of scars may have considerable effects on the function of skin and joints, together with the irreparable damage to the joint that is caused by the injury itself. Scars over joints tend to produce mechanical impairment whereas scars on areas such as the back, cause little or no mechanical effects. Impairment is assessed by measuring the range of motion of a joint; it reflects the consequence of a disease or trauma at the organ level. The use of a goniometer has improved the accuracy of measurements compared to observers' estimation. Goniometry is a convenient, accurate and reliable technique. A computer assisted impairment evaluation system to automate and standardise measurements of upper-extremity function has been developed and found to correlate well with, and require less time than, conventional goniometry. Nevertheless, we feel that measurement of disability, which reflects the consequences of the disease on performance of daily functions, may be more valuable for scar evaluation than measuring impairment. If the aim of the study is to improve scar function we should not only measure the range-of-motion but the function itself. The function implies less well-defined parameters such as coordination, strength, skin sensibility and pliability of the scar which may be overlooked by examining only impairment.

Measurement of disability concerning activities in daily life is widely used. These assessment tools might be applicable for evaluating the daily activities in extensively burned patients but are too coarse a tool for testing specific areas. Some tests have been developed to study the disability of hand function like the Jebsen test and SODA (sequential occupational dexterity assessment). The test developed by Jebsen involves seven daily activities of the hand that are performed under standardised conditions: writing, turning over cards, picking up small objects, simulated feeding, stacking checkers, pick-
ing up large light objects and picking up large heavy objects. In the original publication, good test-retest reliability was obtained for all parameters. Bimanual functional abilities were included in the self-care-related tests of SODA\textsuperscript{100}. SODA was shown to have good reliability and accuracy in a sample of patients with rheumatoid arthritis. We feel that these more general examinations would be useful in the global assessment of dysfunction secondary to burn scars.

**General scar evaluation tools**

Subjective scales have been developed to give an overall impression of the quality of scars. A review of the literature regarding the application of scar assessment tools reveals that there are no generally accepted scales available. The most commonly utilised scale to burn scar evaluation is the Vancouver Scar Scale, which was first described by Sullivan et al. in 1990\textsuperscript{7}. This scale allows assessment of pigmentation, vascularity, pliability, and scar height. During this study and a comparable study with a slightly adapted scale\textsuperscript{8}, a statistically significant inter-observer reliability was demonstrated. The scale has already been applied for burn studies\textsuperscript{101,102}. Although this scale covers many relevant items of scar formation, still some practical problems are encountered. The outcome of the Vancouver Scar Scale is sometimes expressed numerically by adding up the scores of four parameters. However, the parameter ‘pigmentation’ is not an ordinal variable and thus cannot be used in a scale. Pigmentation values thus cannot be added to the three other parameters (vascularity, pliability, and scar height) that are ordinal variables.

The Vancouver Scar Scale measures pliability by testing folding of the scar with a six-step scale: normal, supple, yielding, firm, banding and contracture. The functional mobility of the scar, as relates to scar contraction and pliability, is reflected in this parameter. Since both parameters are considered important for scar evaluation, it would have been useful to separate this item into a parameter on scar contraction and a parameter on elasticity. With the current scale it is possible that significant differences between treatments for elasticity or mobility are lost in the pooled data. The other problem encountered with this scale is that it is sometimes impossible to separate the contribution of the two colour parameters: pigmenta-
tion and vascularity. To overcome this issue Beausang et al. proposed to use only one scale for colour mismatching. They presented a newly devised scale that is appropriate for a wide variety of scars. Their scale included the parameters of scar contour, texture and distortion besides colour assessment. In addition, a visual analogue scale was scored with the overall aspect of the scar together with its matte, or shiny quality. A good correlation was found between the clinical assessment, a panel assessment of the same scar by photographs and histologic findings. Smith et al. performed a study to establish the reliability of burn scar assessment by viewers of colour slide photographs. The ratings included surface regularity, scar thickness and colour as well as the overall cosmetic disfigurement. Again, the assessment of burn scars appeared reliable, however some other interesting conclusions merit highlighting. The inter-rater agreement of one rater was low but increased considerably when several raters were utilised resulting in a good reliability with four or more raters. This finding implies that for clinical studies such scales cannot be performed reliably by less than three raters. Another remarkable finding was that the rater's profession and years of experience with burn treatment did not influence the inter-rater agreement. We feel that the Vancouver Scar Scale and comparative scales are important tools in the assessment of burn scars but the requirement for more than two observers to participate in order to obtain any meaningful data is a drawback in a clinical situation.

**Microscopic evaluation**

What is the role of microscopic research in clinical studies on scar formation? We are convinced that microscopic parameters will gain even more interest and importance. Although the macroscopic evaluation by patients and physicians determines the success rate of the experimental treatment, microscopic evaluations provide answers to the question why a therapy is effective or why it fails. Future therapeutic developments will rely heavily on microscopic evaluations of previous therapies. We confine this section to microscopic evaluation for studies on scar formation. These techniques mainly concern the evaluation of struc-
ture and molecular composition of the extracellular matrix as well as the presence of different cell types. Molecular structures of interest are collagen, proteinases, growth factors, and proteoglycans. The strength and elasticity of dermis is predominantly derived from the dermal collagen network, and to a lesser extent from its elastin and extracellular matrix substances. Histopathological studies of mature scar tissue show collagen bundle orientation parallel to the epidermis rather than a more randomly organised pattern found in normal dermis. Since collagen is the most prevalent protein of the dermis, evaluation of its structure is considered an important parameter of wound healing research. Conventional light microscopy in combination with polarised light is the most prevailing method for the evaluation of collagen structure. One or two observers usually perform analysis of collagen orientation. Observer scores are currently the 'gold standard' for the assessment of collagen orientation. In a yet unpublished study we show that at least three experienced raters are required to obtain good inter-observer reliability, which might be considered as an important drawback of this technique. Besides polarised light evaluation, other techniques are available, such as X-ray diffraction and laser scattering. For these applications, X-rays and laser light were used to obtain a diffraction pattern caused by the orientation of dermal architecture. In a pilot study, we tested a laser scattering technique and Fourier analysis and compared the results with observer assessment. Fourier analysis is a computerised image analysis technique that employs a mathematical calculation of the structure in the image. We established a poor correlation between the laser scattering results with the observer assessment and the Fourier analysis. In contrast to observer ratings and laser scattering, Fourier analysis has a perfect reliability, as the calculation of the same images will continuously reproduce the same outcome. Fourier analysis is superior to multi-observer assessment with respect to accurateness of the assessment (submitted data).

The presence of growth factors and extracellular matrix molecules during scar formation is mainly evaluated by in vitro studies or animal studies. Current and past efforts are helping us to shape our understanding of scar formation at the molecular level.
Most studies on cell types concern inflammatory cells, fibroblasts and a specialised type of the latter: the myofibroblast. Many studies are devoted to solving the ‘mystery’ of the myofibroblast and its role in wound contraction. In 1956 it was hypothesised that a connective tissue cell should be able to induce wound contraction. The myofibroblast was discovered in the seventies and recognised as a specialised type of fibroblast that expresses cytoskeletal proteins such as smooth-muscle-actin, sometimes in combination with desmin and smooth muscle myosin heavy chains. The expression of these cytoskeletal proteins does not prove that these cells become highly contractile but at least give the impression that they may be responsible for contraction. Although many papers implied an important role for the myofibroblast in contraction, no definite proof is provided by clinical studies. Also, the role of the fibroblast in relation to scar contraction has been challenged instead of the myofibroblast. In most studies myofibroblasts are evaluated by smooth-muscle-actin and scored qualitatively or by image analysis, that calculates the percentage of positive staining. It appears that the techniques to identify fibroblasts and myofibroblasts are well established but the role of these cells with regard to scar contraction remains less clear.

Conclusions

In the new century, the role of accurate and reproducible scar assessment tools will become more prevalent. This review describes several kinds of measurements that may be considered for scar evaluation. Surprisingly, few scar assessment tools have been studied sufficiently. The best scar measuring devices are yet to be determined for many of the parameters that bear importance for research goals. In general, objective scar measurements are preferred over subjective scar assessment scales. However, subjective scar assessment scales performed by observers have some important advantages. They are convenient, cheap, and give semi-quantitative information along several clinical parameters. Moreover, they can be per-
formed in outpatient clinics, as they require little time to complete. On the other hand, semi-quantitative observer scores require more than one rater to obtain good reliability and to reduce the measuring error of a single observer. Measuring devices provide objective parameters and make errors that are usually consistent and predictable. Whether the objective and subjective measuring tools are accurate, meaning that they measure what we want to measure, needs to be established for many tools that evaluate scar colour, surface texture, and functional qualities. Also we contemplate the development of an improved subjective scar assessment scale that shows a good inter-observer reliability for less than 4 raters. Whether such a scale is suitable for clinical application needs to be determined yet.

In this review, we systematically discussed methods to access scar formation that are considered relevant to research involving scars. Of course there is no need to perform the entire array of clinical and microscopic analyses in a single study. The aim of a study must dictate the methodology and the preferred method of evaluation.

References

Section I: Tools for scar evaluation

32. Viatour M, Henry F, Pierard GE. A comput-
Scar assessment tools: implications for current research

22. Richard S, Querleux B, Bittoun J, Jolivet O, Idy PI, de LO, Leveque JL. Characterization of the
skin in vivo by high resolution magnetic resonance imaging: water behavior and age-related effects. J Invest Dermatol 1993;100(5):705-709.


74. Katz SM, Frank DH, Leopold GR, Wachtel TL.


100. van Lankveld W, van't Pad BP, Bakker J, Terwindt S, Franssen M, van Riel P. Sequential occupational dexterity assessment (SODA):


121. Boyce ST, Supp AP, Harriger MD,
Scar assessment tools: implications for current research


