Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/locate/burns

Review

A guide to choosing a burn scar rating scale for clinical or research use

Zephanie Tyack[a,b,*], Jason Wasiak[c], Anneliese Spinks[d,e], Roy Kimble[f,g], Megan Simons[f,h]

[a] Central Queensland Hospital and Health Service, Rockhampton, Queensland, Australia
[b] School of Health and Rehabilitation Sciences, The University of Queensland, St Lucia, Queensland, Australia
[c] Victorian Adult Burns Service and School of Public Health and Preventative Medicine, Monash University, The Alfred Hospital, Commercial Road, Melbourne, Victoria, Australia
[d] Commonwealth Scientific and Industrial Research Organisation, Ecosystem Sciences Division, Queensland, Australia
[e] School of Medicine, Griffith University, Meadowbrook, Queensland, Australia
[f] Centre for Childrens Burns and Trauma Research, Queensland Children's Medical Research Institute, The University of Queensland, Queensland, Australia
[g] Department of Paediatrics and Child Health, Level 3 Foundation Building, Royal Children's Hospital, Herston, Queensland, Australia
[h] Department of Occupational Therapy, Royal Children's Hospital, Brisbane, Queensland, Australia

ARTICLE INFO

Article history:
Accepted 23 April 2013

Keywords:
Burn scar rating measure
Scar assessment
Clinimetric quality
Burn
Outcome measure

ABSTRACT

Introduction: A lack of high quality burn scar rating scales underpins the urgent need to introduce a guide for clinicians and researchers to choose the most appropriate scale for their requirements.

Methods: An updated electronic search of Medline, CINAHL, and EMBASE databases from 2010 to 2011 of a previous published systematic review were used to identify English articles related to burn scar rating scales. The clinimetric properties, content, purpose, characteristics of the subjects tested and feasibility of each scale were critically reviewed.

Results: An additional seven papers were identified by the updated search, bringing the total number of papers reviewed to 36. The majority (88%) covered items pertaining to the physical properties of the skin rated by an observer. All of the scales had been tested for the purpose of discriminating between patient groups; however, only preliminary evidence exists for the ability of the scales to measure change in scar properties over time. The majority of testing of scales occurred using Caucasian subjects, males, upper limb sites and adults.

Conclusions: This paper provides a guide to selecting the most appropriate burn scar rating scale for research and clinical practice by reviewing the content, purpose, test sample characteristics and feasibility of each scale.

Crown Copyright © 2013 Published by Elsevier Ltd and ISBI. All rights reserved.

* *Corresponding author at*: Central Queensland Hospital and Health Service, Rockhampton, Queensland, Australia. Tel.: +61 749 20 7396; fax: +61 749 20 6539.
E-mail address: Zephanie_Tyack@health.qld.gov.au (Z. Tyack).
0305-4179/$36.00. Crown Copyright © 2013 Published by Elsevier Ltd and ISBI. All rights reserved.
http://dx.doi.org/10.1016/j.burns.2013.04.021

Contents

1. Background

Scarring after a burn may lead to reduced body esteem [1], reduced quality of life [2–4], and symptoms such as itch and pain [5,6] which can impact on sleep and return to work for years after the injury has occurred [7]. Thus, the identification of successful interventions to effectively manage scarring after-burn is important for burn patients and the professionals treating them.

Many intervention studies have relied on a burn scar rating scale as the sole method to evaluate scarring [8–17]. However, the ability to evaluate the effectiveness of interventions is hampered by a lack of high quality scar rating scales, as found in a recent systematic review [18] which investigated the clinimetric properties of 18 burn scar rating scales. In that review by Tyack, a high quality rating could be given to only one component of the patient observer scar assessment scale (POSAS).

While systematic reviews of the quality of scales are useful for guiding researchers and clinicians, these reviews should be interpreted in the context of the studies included (i.e., characteristics of the participants, whether there is sufficient support for testing at a single point in time versus testing over time) [19,20]. Recognition of the importance of these contexts can guide further testing and refinement of existing burn scar rating scales or the development and testing of new burn scar rating scales. To date, systematic reviews of burn scar scales have failed to consider these contexts, despite being highlighted in other systematic reviews [21–23].

Thus, this paper provides a guide to choosing the most appropriate burn scar rating scale based on the required content and purpose; and the characteristics of the burn population with whom the scale will be used. The feasibility of administering the scales is also reviewed as clinicians are likely to rely heavily on this feature in practice. In addition, this paper will update the current evidence for the clinimetric quality of available burn scar rating scales.

2. Method

2.1. *Data sources and search strategy*

An electronic search of Medline, CINAHL, and EMBASE databases from 1990 to 2011 were used to identify English articles related to burn scar rating scales in the original

Table 1 – Summary of the clinimetric quality and interpretability of burn scar rating scales.

Measure	Publication year	Content validity	Construct validity	Internal consistency	Reliability[a]	Responsiveness	Interpretability
VSS [34]	1990	0	?	0	?	?	0
MVSS [Baryza, [46]]	1995	0	?	0	0	?	?
MVSS [Draaijers, [39]]	2004	0	?	?	-	0	0
MVSS [Nedelec, [29]]	2000	0	?	0	?	0	0
MVSS [Oliveira, [36]]	2005	0	?	0	0	?	?
POSAS [35]	2004	?	?[b]	?	+[c]	?	?
MAPS [52]	2005	?	?	0	?	?	0
VAS [Masters, [52]]	2005	?	0	0	0	?	0

Rating: + = high quality; ? = indeterminate, – = low quality, 0 = no data available (as per Terwee [19]).
[a] Reliability has been scored for 1 rater only as this is most clinically viable.
[b] Indeterminate evidence in general however 1 study has demonstrated some high quality evidence of validity consistent with the hypothesis.
[c] Vascularity subscale and total scores only.

systematic review by Tyack et al. [18]. Therefore, an updated literature search was performed in MEDLINE, CINAHL, and EMBASE from 2010 to 2012 using keywords such as 'burns' and 'thermal injury' with 'scar', 'cicatrix', 'rating', 'assessment', 'measure', 'tool' or 'Seattle', 'Hamilton', 'pruritis', 'elastic', 'thick', 'volume', 'hypertrophic', 'vascular', 'colour', 'evaluation', 'quantify', 'outcome', 'spectrophometry'. All searches were limited to articles in English and humans only. Only scales that had evidence of testing in two or more clinimetric areas were included (refer to Table 1). These clinimetric areas included validity, reliability, responsiveness (as a component of validity) and interpretability. An additional recent paper [24] was included to review the feasibility of scar rating scales.

2.2. *Clinimetric criteria for reviewing the purpose of each scale*

2.2.1. *Burn scar rating scales for discriminative purposes (distinguishing between patient groups)*

To be recommended as a tool to assess differences between individuals at a single point in time, a scale was required to display both validity and reliability. Validity is ideally demonstrated by showing that the scale is consistent with an accepted external measure at a single point in time [25] while reliability is demonstrated by either test–retest reliability in stable subjects or inter-rater reliability [20]. A reliability coefficient of 0.7 or greater was used as a threshold to judge whether acceptable reliability had been achieved [19].

2.2.2. *Burn scar rating scales for longitudinal evaluative purposes (measuring change over time within patient groups)*

To be recommended as a tool for measuring change over two or more time points, the scale was required to demonstrate validity (ideally longitudinal validity) [19,25]; reliability and responsiveness. Reliability was ideally demonstrated by test–retest reliability in stable subjects [25] for which coefficients exceed 0.5 [20], or a small measurement error determined when absolute measurement error was less than the minimal important change [19,20]. Responsiveness was ideally demonstrated using an external criterion or anchor-based method to determine the clinical importance of change [19]. For some papers responsiveness data were calculated using the means and standard deviations reported in treatment effectiveness studies by the authors of this paper.

2.2.3. *Burn scar rating scales for predictive purposes*

To be recommended for predicting concurrent or future outcomes (predictive) the measure was required to have at least some support for validity (ideally between the measure and an external measure at a single point in time) [25]; in addition to support for reliability (i.e., test–retest reliability in stable subjects or interrater reliability) [20,26].

2.3. *Characteristics of subjects*

For scar assessment, the age of participants, gender, scar site, number of surgical procedures, type of graft, time to wound healing, depth of burn, and time after-burn were important characteristics to consider due to their influence on the physical properties of the skin or scar [17,27–31]. We also chose to include studies that addressed number of surgical procedures, type of graft, and time to wound healing as they were found to be risk factors for hypertrophic scarring [32].

Only those characteristics that varied most across scales (i.e., age, gender, time after-burn) are presented in Table 2 as decisions regarding those characteristics need to be made to guide the choice of scar scale. Where very small numbers of children or lower limb scar sites were tested in studies (i.e., two or fewer), the scale was not considered to have been tested in these sub-groups. Recommendations were made for choosing a scale based on the characteristics of the sample tested (i.e., age, gender), when 50% or more of the studies providing evidence for a clinimetric area reported information on that characteristic. A summary of information across all clinimetric properties is presented in Table 2 when evidence was available for at least two clinimetric properties of the scale (e.g., information regarding gender was available for validity and responsiveness of the POSAS although not for other clinimetric properties so was summarised in Table 2).

3. Results

The 29 papers included in the recent systematic review [18] were included in this paper [15,29,33–59] along with an

Table 2 – Guide to choosing a burn scar rating scale.[a]

Scale	Purpose		Content	Sample characteristics					Feasibility	
	Evaluation at a single time point	Evaluation at two or more time points (longitudinal)	Patient-report itch and pain scale	Children[b]	Adults[c]	Males[d]	Females	Time post-burn	Fast to administer (<10 mins)	Available in a published paper[e]
VSS [34]	×	✓	×	–	✓	✓	✓	0–100 months	–	Yes
MVSS [Baryza [46]]	×	×	×	✓ (16+)	✓	✓	✓	–	–	Yes
MVSS [Draaijers, [35]]	×	×	×	✓ (15+)	✓	–	–	3 months to 360 months	–	Yes
MVSS [Nedelec, [29]]	×	×	✓	–	✓	✓	✓	2–9 months	Yes	Yes
MVSS [Oliveira, [36]]	×	×	✓	✓ 2–17	×	–	–	≤24 months	–	Yes
POSAS [35]	✓	✓	✓	✓ 12+	✓	✓	✓	3 months to >360 months	Yes	Yes
MAPS [52]	×	×	×	×	✓	–	–	–	No	No
VAS [Masters, [52]]	×	×	✓	×	✓	–	–	1–19 months	–	No

MVSS (Oliviera) tested in children only; POSAS 15–85 years; MAPS 18–79 years; VAS Masters no information.

[a] ✓ = Some testing; × = no testing; – = insufficient testing or no information provided on testing for that group.

[b] The minimum age of testing in children where information was available: VSS (Sullivan) 3 years+; MVSS (Baryza) 16 years+; MVSS (Draaijers) 15 years+; MVSS (Nedelec) 10 years+; MVSS (Oliviera) 2–17 years; POSAS 12 years+; MAPS and VAS (Masters) tested in adults only.

[c] Age of testing in adults: VSS (Sullivan) 18–75 years; MVSS (Baryza) 18–65 years; MVSS (Draaijers) 18–73 years; MVSS (Nedelec) 18–74 years.

[d] Percentage of males tested across studies that reported information on gender: VSS (Sullivan) 60–100%; MVSS (Baryza) 44–74%; MVSS (Draaijers) no information; MVSS (Nedelec) 66–80%; mVSS (Oliviera) 60%; POSAS 50–57%; MAPS 86%; VAS (Masters) no information.

[e] References where scales are published: VSS (Sullivan) (34); MVSS (Baryza) [46]; MVSS (Draaijers) [35]; MVSS (Nedelec) [29,42,43,48], MVSS (Oliviera) [36]; POSAS [35].

additional seven new papers that were identified in an updated literature search [10,14,60–63].

The updated clinimetric properties of the scar rating scales are summarised in Table 1. The only change to the previously published summary of clinimetric properties was for the POSAS where preliminary evidence of responsiveness from a prospective cohort study [60], showed a moderate effect size between pre-operative and post-operative percutaneous collagen induction therapy scores (Cohen's $d = 0.59$). However, the evidence was rated indeterminate due to the lack of clear hypotheses and with no data indicating whether this change was clinically meaningful.

Additional preliminary evidence for responsiveness was also available for the Vancouver Scar Scale (VSS) (Sullivan) from a study looking at the effectiveness of two silicone gel treatments and topical onion extract, with the smallest effect size being for the topical onion extract group (Cohen's $d = 2.36$; large effect size) [14]. For the modified Vancouver Scar Scale (MVSS) (Baryza), preliminary evidence of responsiveness was available from a single cohort study which examined the effectiveness of combined dermabrasion and thin skin graft with allograft surgical treatment for depigmentation (Cohen's $d = 3.84$; large effect size) [9]. However no changes to the clinimetric properties of scales occurred as a result of additional evidence of validity or responsiveness.

We pose a series of questions below that are designed to assist the reader in choosing the most appropriate burn scar rating scale for their purpose. These questions are supported by further detail in Table 2 which presents information on the characteristics of subjects and the availability of the scale from a published paper. Two scenarios are presented to illustrate the use of information in the text and the tables to guide the selection of a scar rating scale for clinicians and researchers.

3.1. *What is the preferred content of the scale?*

3.1.1. *Patient opinion of scarring and patient reported symptoms*

If the inclusion of the patient's overall opinion of the scar appearance, itch or pain is important then the POSAS is recommended. However, this recommendation is based on minimal testing of reliability [35], no testing of validity in children aged less than 12 years [62], and testing using insufficient numbers for children to provide high quality evidence. As even less testing has occurred using other patient scales (ie., MVSS Oliviera; MVSS Nedelec; VAS Masters) these scales are not recommended.

3.1.2. *Physical properties of scarring*

The majority of scales [the VSS (Sullivan), MVSS (Baryza), MVSS (Draaijers), MVSS (Nedelec), MVSS (Oliviera), and POSAS observer scale] can be considered if vascularity or colour, height or thickness, pliability or pigmentation is important. Only the POSAS observer scale contains a relevant item to assess surface irregularity [35] and the MAPS is the only scale to assess border height of the scar [52].

3.2. *What is the intended purpose of the measure – measurement at a single time point to discriminate between subjects or groups of subjects, or measurement over time to detect change within subjects?*

3.2.1. *Single time point*

For measurement at a single point in time the POSAS is recommended. Evidence from several studies supported inter-rater reliability and validity using an external criterion at a single point in time [35,39,54,62]. The VSS (Sullivan) was not recommended as inter-rater reliability was not demonstrated [29,36] and no evidence of other reliability was available.

3.2.2. *Multiple time points*

For measurement over time to detect change within subjects, either the POSAS or VSS (Sullivan) is recommended. Preliminary evidence was available for responsiveness of the POSAS in the form of a treatment effect over time, using data from Aust [60]. The MVSS (Nedelec) and MAPS had evidence for test–retest reliability however did not have evidence for responsiveness, thus these scales were not recommended.

3.3. *What are the characteristics of the sample to be assessed?*

The application of clinimetric findings to populations with characteristics other than those tested needs careful consideration. For example, current evidence does not support the extrapolation of findings based on testing of adult males to adult females with burn scarring, as burn scarring in adults has been found to be influenced by gender [32]. However, no such evidence exists in children. Whilst the clinimetric properties of each scale were examined using the characteristics of the groups or subgroups tested (where sufficient information was available) no changes to the clinimetric quality of the scales resulted (Table 1). This was in part due to the small sample size of sub-groups tested in many studies (i.e., less than 50). For example, in the study by Jarrett [55] despite testing sufficient numbers of scar sites in one subgroup (e.g., males) but not another (e.g., females) for a high quality rating; there were other methodological issues that prevented a high quality rating for the subgroup of males. A summary of the characteristics of subjects tested is presented in the proceeding text and Table 2 to guide decisions regarding the characteristics that are important to consider.

3.3.1. *Age*

All scales except for the MVSS (Oliviera) can be considered for assessing burn scars in adults and older children as these scales were tested in these groups (refer Table 2 for details).

Only the MVSS (Oliviera) can be considered for assessing burn scars in children across a wide age-range; with children aged 2–17 tested in a single study. Other scales such as the POSAS can be considered for older children (refer Table 2 for details), however details such as the age and gender of the older children were not available. This lack of information meant the clinimetric properties of scales tested using older children were unable to be examined separately from adults.

3.3.2. *Gender*

The VSS (Sullivan), MVSS (Baryza), MVSS (Nedelec) and POSAS can be considered for testing males or females. However, the majority of testing occurred in males. For example, in studies testing the VSS (Sullivan) 60–100% of the participants were males where information was available.

3.3.3. *Ethnicity*

The VAS (Masters) can be considered for testing Caucasian subjects as it was tested solely using Caucasian subjects. Whilst the MVSS (Nedelec) [42,43] and MVSS (Oliviera) [36] can be considered for testing Fitzpatrick skin types I to VI, minimal testing of Fitzpatrick skin type V and VI occured, thus caution needs to be exercised when generalising results to skin type V and VI.

Most of the scales [VSS (Sullivan), mVSS (Baryza), POSAS, MAPS, VAS] received some testing using Caucasian subjects but information on ethnicity was provided by only one study for most of the scales [MVSS (Baryza), MVSS (Oliviera), MVSS (Nedelec), POSAS, MAPS, VAS]. A lack of information on the breakdown of ethnicity in many studies prevented evaluation of the clinimetric properties using different ethnic groups.

3.3.4. *Burn depth and time after-burn*

The MVSS (Baryza) can be considered for testing deep dermal burns when validity is the area of interest as testing occurred using this depth of injury in a study with evidence for validity [10]. A lack of information on the burn depths of scar sites tested (e.g., the number of scar sites that had full thickness burns versus deep partial burns) in studies pertaining to all of the scales, meant other recommendations were not able to be provided.

With respect to the time after-burn, all of the scales except the MVSS (Baryza), MVSS (Nedelec), and MAPS can be considered for testing in the first months to first few years after-burn (refer Table 2 for details), as testing occurred in the first few months after-burn to years. Wide variation in the information reported in studies testing the MVSS (Baryza) and a lack of information in studies testing the MAPS meant information could not be summarised for these scales. The MVSS (Nedelec) can only be recommended for testing burns two to nine months after-burn as no testing occured outside this period.

3.3.5. *Location of the scarring*

The majority of the scales [VSS (Sullivan), MVSS (Baryza), MVSS (Nedelec), POSAS] were tested predominantly on the upper limb using information from studies in which a breakdown of scar sites was provided. A much smaller number of other scar sites (i.e., lower limb, trunk, face, head and neck) were tested across studies (e.g., 12 other sites for the VSS (Sullivan) compared to 20 upper limb sites and 8 other sites for the MVSS (Baryza) compared to 68 upper limb scar sites. Whilst the MVSS (Oliviera) was tested in thigh and forearm sites only no breakdown of the percentage of each site was provided. No information on scar sites tested was available for the MAPS or VAS (Masters).

3.4. *Is feasibility important (including time to administer and ease of administration)?*

3.4.1. *Administration time*

Based on available information, the MVSS (Nedelec) and POSAS scales are recommended if a short administration time of less than 5 minutes is required. The MAPS may take up to 20 minutes [55]. Administration time was not available for all other reviewed scales.

3.4.2. *Scale availability, readability and comprehensiveness, ease of administration and scoring*

Only the MAPS is not available in a published paper unlike the other scar rating scales, with contact with the original authors required to obtain the scale and the reference colour photographs [18]. The references for published papers that contain the scales are provided below Table 2. No formal testing of the readability and comprehensibility has been completed for any of the scales by intended users [18] thus no recommendation can be provided on these aspects. Ease of completion, scoring and interpretation have been tested for the POSAS and MVSS (Nedelec) with therapists scoring the MVSS (Nedelec) as easier to complete, score and interpret than the POSAS [24].

4. Application to practice scenarios

4.1. *Scenario (1)*

Researchers are planning to conduct a controlled, longitudinal study in an Australian hospital at multiple time-points; comparing the effectiveness of pressure garments and silicone to a control group who received pressure only. The sample will include Causasian adolescents and adults after-burn aged 13 years or older at the time of skin healing to 18-months. The researchers are particularly interested in assessing itch and pain, as well as the height and pliability of scars. As objective scar assessment tools will also be administered in the study (i.e., ultrasound, cutometer) a short administration time is preferred, although the responsiveness of the measure is the primary consideration.

4.1.1. *Recommended approach to selecting an appropriate scale for scenario 1*

For an adolescent and adults smaple the MVSS (Baryza), MVSS (Draaijers) and POSAS would be suitable. However, as the MVSS (Baryza) and POSAS have been tested predominantly using upper extremity sites the authors decide they will also need to do some clinimetric testing using other sites if they choose one of these measures, as they do not consider it reasonable to assume the estimates of validity and reliability for these scales can be extended to other scar sites in their study (i.e., face/head, torso and lower extremity). In terms of feasibility, the POSAS is likely to take a little longer than the other measures; however, considering the longitudinal nature of the study, the researchers can see from Table 2 that the POSAS is one of the only measures with a patient-report scale including itch and pain that has been tested in adults and adolescents. The researchers also observe that the response

scale of the POSAS has more points than the other scales making it more likely to pick up changes over time. So the POSAS is selected for the study.

4.2. *Scenario 2*

Burn surgeons wish to monitor the severity of scarring routinely in children aged 0 to 14 years with burns attending a paediatric outpatient clinic at 6-months after-burn. They will focus on measuring the physical properties of the worst area of burn scarring. The measure will be administered by occupational and physiotherapists in a busy outpatient burn clinic. Feasibility is a primary consideration, however, it is the intention of the surgeons to publish their findings thus clinimetric properties are also considered important.

4.2.1. *Recommended approach to selecting an appropriate scale for scenario 2*

Using Table 2, the surgeons start by looking at the feasibility column and consider the POSAS and the MVSS (Nedelec) but not the MAPS, which can take up to 20 minutes to administer. However, the surgeons notice from Table 2 that only the POSAS has evidence to support use at a single point in time and that testing has only been conducted in children 12 years or older. Despite this, the surgeons decide to use the MVSS (Nedelec) as they only want to focus on the physical properties of scarring and as they have read that therapists consider the MVSS (Nedelec) easier to administer, score and interpret than the POSAS. However, they decide to test the inter-rater reliability of the MVSS (Nedelec) on their sample of children aged 0 to 14 years with burn scarring before finally committing to use of the MVSS (Nedelec); as multiple therapists will be using the measure, and to strengthen the evidence for use in children younger than 12 years. They decide to conduct reliability testing on upper limb sites and Caucasian children only as they wish to commence their study shortly and believe it will require too many subjects to test reliability of the MVSS (Nedelec) on other sites and with other ethnic sub-groups. Whilst the surgeons are aware of evidence of the influence of gender on scarring in adults they are aware of no such evidence in children so decide not to consider gender when selecting a scar scale for children. The column for time after-burn in Table 2 indicates that most of the scales are suitable for assessing scars at 6-months. From Table 1 the surgeons see that there is indeterminate evidence for the construct validity of the MVSS (Nedelec) which is similar to the rating for the POSAS that further supports their decision to prioritise feasibility.

5. Discussion

The aim of this paper was to provide clinicians and researchers with a guide to choosing a burn scar rating scale with a unique focus on the purpose of the scale and the characteristics of the patient sample with whom the scale will be used. We recommend the POSAS for discriminating between patient groups and both the POSAS and VSS (Sullivan) for measuring change over time.

However, if a scale needs to be used with children, a skin type or ethnicity other than Caucasian, or scar sites other than the upper limb then the choice of scales is more limited due to a lack of testing in these sub-groups. The choice of scale can be made by combining the purpose of the scale (i.e., discriminating between groups or measuring change) with the required content and sample characteristics, and feasibility features (see Table 2); as illustrated by the scenarios presented.

The importance of interpreting the clinimetric properties of a scale in light of

the sample characteristics has been emphasised by experts in the area of

scale development and testing [19,20], as well as in systematic reviews of outcome measures for groups including children experiencing pain [21], bracial plexus injuries [22], and children with congenital hemiplegia [23]. We have presented a synthesis of research in these areas as well as other areas to highlight the importance of considering validity and reliability in the context of the sample characteristics and the purpose for using the scale. The guide presented in Table 2 included the sample characteristics of age, gender and time after burn as there was little variation in other characteristics across studies. However, much of the evidence of the influence of these factors on burn scarring originates from one retrospective study of adult burn patients [32]. Thus, prospective studies are needed to confirm the importance of these characteristics particularly in children.

Where scale testing has not included people with dark brown or black pigmented skin, minimal females, children and scar sites apart from the upper limb, then further testing should occur using sufficient numbers of these groups (refer Terwee [19] for guidelines). Whilst the authors included sample characteristics shown to influence the skin or scarring it may be that there are other factors that are important to consider when selecting a burn scar scale. A greater level of detail regarding sample characteristics in future studies would assist in advancing evidence regarding the scales most suitable for testing groups (or sub-groups).

While the patient scale of the POSAS was recommended for evaluating patient-reported symptoms, this recommendation was based on minimal evidence to support the clinimetric properties of the scale. Further testing of patient scales is needed as the importance of evaluating patient reported symptoms of scarring such as itch has been highlighted by studies that have linked itch to wound healing and scarring after-burn [6,64].

The purpose of scales included in this paper was not discussed by the original authors of any of the burn rating scales; however, most have been used for discriminative purposes (determining differences between groups). However, scales are often used for purposes other than what they were originally designed for without issues [20] – a view adopted by the authors of this paper. For example, a scale designed for evaluating scarring at a single point in time in a clinical setting may then be used for predicting quality of life at scar maturation. However, it must be recognised that in order to use the scale for multiple purposes the scale must meet the criteria necessary for each and every measurement purpose for which it will be used (i.e., evaluative, predictive, discriminative) [20]. No scale met all of the ideal criteria for use at a

single point in time (discriminative) or to measure change over time (evaluative). Thus recommendations were made using evidence of other types of validity or reliability.

Whilst feasibility has been reported as a priority for clinicians in their choice of a scar assessment scale [65], it is argued that for researchers a higher priority should be placed upon sound clinimetric properties of a chosen scale. For clinicians a balance needs to be achieved between feasibility and sound clinimetric properties as there is limited benefit in selecting a scale that takes a short time to administer if the information it yields is of poor quality [20]. Surveys of burn health professionals [24,65] have found the MVSS (Nedelec) faster to administer and preferred by therapists over the POSAS due to ease of completion and ease of interpretation of the descriptors. Further testing of the feasibility of scales is required to broaden evidence in this area.

6. Conclusions

The selection of burn scar rating scales needs to be matched to the content and purpose of the scale as well as the characteristics of the sample and the feasibility of the scale. A summary of all of these features has been presented in addition to an update on the clinimetric properties of scales to guide the selection of burn scar rating scales for clinicians and researchers for discriminative or evaluative purposes. No scale met all of the ideal criteria for use at a single point in time (discriminative) or to measure change over time (evaluative).

Conflicts of interest

None.

REFERENCES

[1] Lawrence JW, Fauerbach JA, Thombs BD. A test of the moderating role of importance of appearance in the relationship between perceived scar severity and body-esteem among adult burn survivors. Body Image 2006;3:101–11.

[2] Klein MB, Lezotte DC, Heltshe S, Fauerbach J, Holavanhalli RK, Rivara FP, et al. Functional and psychosocial outcomes of older adults after burn injury: results from a multicentre database of severe burn injury. J Burn Care Res 2011;32(1):66–78.

[3] Moi AL, Wentzel-Larsen T, Salemark L, Wahl AK, Hanestad BR. Impaired generic health status but perception of good quality of life in survivors of burn injury. J Trauma Inj Infect Crit Care 2006;61(4):961–9.

[4] Leblebici B, Adam M, Bagis S, Tarim AM, Noyan T, Akman MN, et al. Quality of life after nurn injury: the impact of joint contracture. J Burn Care Res 2006;27(6):864–8.

[5] Ahee AM, Smith SJ, Pliska-Martyshak G, Cullen ML. When does itching start and stop post-burn. J Burn Care Rehabil 1999;20(1 pt 2):S187.

[6] Van Loey NEE, Bremer M, Faber AW, Middelkoop E, Nieuwenhuis MK, The Research Group. Itching following burns: epidemiology and predictors. Brit J Dermatol 2008;158:95–100.

[7] Oster C, Kildal M, Ekselius L. Return to work after burn injury: burn-injured individuals' perception of barriers and facilitators. J Burn Care Res 2010;31(4):540–50.

[8] Ueda M. Sprayed cultured mucosal epithelial cell for deep dermal burns. J Craniofac Surg 2010;21(6):1729–32.

[9] Oh SJ, Kim Y. Combined AlloDerm and thin skin grafting for the treatment of postburn dyspigmented scar contracture of the upper extremity. J Plast Reconstr Aesthet Surg 2011;64:229–33.

[10] Haslik W, Kamolz L-P, Lumenta DB, Hladik M, Frey BM. The treatment of deep dermal hand burns: how do we achieve better results? Burns 2010;36:329–34.

[11] Haslik W, Kamolz L-P, Manna F, Hladik M, Rath T, Frey BM. Management of full-thickness skin defects in the hand and wrist region: first long-term experiences with the dermal matrix Matriderm. J Plast Reconstr Aesthet Surg 2010;63:360–4.

[12] Gravante G, Sorge R, Merone A, Tamisani AM, Di Lonardo A, Scalise A, et al. Hyalomatric PA in burn care practice: results from a national retroespective survey, 2005 to 2006. Ann Plast Surg 2010;64(1):69–79.

[13] Busche MN, Herold C, Schedler A, Knobloch K, Vogt PM, Rennekampff HO. The Biobrane glove in burn wounds of the hand: evaluation of the functional and aesthetic outcome and comparison of costs with those of conventional wound management. Handchir Mikrochir Plast Chir 2009;41(6):348–54.

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

[15] 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(4):632–42.

[16] Maslauskas K, Rimdeika R, Rapoliene J, Saladzinskas Z, Samsanavicius D, Kaikaris V. The comparison of two methods of treatment evaluating complications and deficiency of functions of hands after deep partial skin thickness hand burns. Medicina 2009;45(1):37–45.

[17] Momeni M, Hafezi F, Rahbar H, Karimi H. Effects of silicone gel on burn scars. Burns 2009;35:70–4.

[18] Tyack Z, Simons M, Spinks A, Wasiak J. A systematic review of the quality of burn scar rating scales for clinical and research use. Burns 2012;38:6–18.

[19] Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34–42.

[20] Streiner DL, Norman GR. Health Measurement Scales: A Practical Guide to their Development and Use. New York: Oxford University Press; 2008.

[21] von Bayer CL, Spagrud LJ. Systematic review of observational (behavioral) measures of pain for children and adolescents aged 3 to 18 years. Pain 2007;127:140–50.

[22] Hill BE, Williams G, Bialocerkowski AE. Clinimetric evaluation of questionnaires used to assess activity after traumatic brachial plexus injury in adults: a systematic review. Arch Phys Med Rehabil 2011;92:2082–9.

[23] Gilmore R, Sakzewski L, Boyd R. Upper limb activity measures for 5- to 16-year-old children with congenital hemiplegia: a systematic review. Dev Med Child Neurol 2010;52:14–21.

[24] Simons M, Ziviani J, Thorley M, McNee J, Tyack Z. Exploring reliability of scar rating scales using photographs of burns

from children aged up to 15 years. J Burn Care Res 2013. http://dx.doi.org/10.1097/BCR.0b013e3182700054.
[25] Guyatt GH, Kirshner B, Jaeschke R. Measuring health status: what are the necessary measurement properties? J Clin Epidemiol 1992;45(12):1341–5.
[26] Kirschner B, Guyatt G. A methodological framework for assessing health indices. J Chron Dis 1985;38(1):27–36.
[27] Bailey SH, Oni G, Brown SA, Kashefi N, Cheriyan S, Maxted M, et al. The use of non-invasive instruments in characterizing human facial and abdominal skin. Lasers Surg Med 2012;44:131–42.
[28] Cheng W, Saing H, Zhou H, Peh HW, Tam PKH. Ultrasound assessment of scald scars in Asian children receiving pressure garment therapy. J Pediatr Surg 2001;36(3):466–9.
[29] Nedelec B, Shankowsky HA, Tredget EE. Rating the resolving hypertrophic scar: comparison of the Vancouver scar scale and scar volume. J Burn Care Rehabil 2000;21(3):205–12.
[30] Nedelec B, Armour AD, Rachelska G, Correa JA. Longitudinal scar quantification. J Burn Care Res 2009;30(2):S67.
[31] Gangemi EN, Gregori D, Berchialla P, Zingarelli E, Cairo M, Bollero D, et al. Epidemiology and rick factors for pathologic scarring after burn wounds. Arch Facial Plast Surg 2008;10(2):93–102.
[32] Gangemi ENG, Berchialla D, Zingarelli P, Cairo E, Bollero M, Ganem D, Capocelli J, Cuccuru R, Cassano F, Risso P, Stella DM. Epidemiology and rick factors for pathologic scarring after burn wounds. Arch Facial Plast Surg 2008;10(2):93–102.
[33] Martin D, Umraw N, Gomez M, Cartotto R. Changes in subjective vs objective burn scar assessment over time: does the patient agree with what we think? J Burn Care Rehabil 2003;24(4):239–44.
[34] Sullivan T, Smith J, Kermode J, McIver E, Courtemanche DJ. Rating the burn scar. J Burn Care Rehabil 1990;11(3):256–60.
[35] Draaijers LJ, Tempelman FRH, Botman YAM, Tuinebreijer WE, Middelkoop E, Kries RW, et al. The patient and observer scar assessment scale: a reliable and feasible tool for scar evaluation. Plast Reconstr Surg 2004;113(7):1960–5.
[36] Oliviera GV, Chinkes D, Mitchell C, Oliveras G, Hawkins HK, Herndon DN. Objective assessment of burn scar vascularity, erythema, pliability, thickness, and planimetry. Dermatol Surg 2005;31(1):48–58.
[37] Bray R, Forrester K, Leonard C, McArthur R, Tulip J, Lindsay R. Laser doppler imaging of burn scars: a comparison of wavelength and scanning methods. Burns 2003;29:199–206.
[38] Davey RB, Sprod RT, Neild TO. Computerised colour: a technique for the assessment of burn scar hypertrophy. A preliminary report. Burns 1999;25:207–13.
[39] Draaijers LJ, Tempelman FRH, Botman YAM, Kries RW, Middelkoop E, van Zuijlen PPM. Colour evaluation in scars: tristimulus colorimeter, narrow-band simple reflectance meter or subjective evaluation? Burns 2004;30:103–7.
[40] Lye I, Edgar DW, Wood F, Carroll S. Tissue tonometry is a simple, objective measure for pliability of burn scar: is it reliable? J Burn Care Res 2006;27(1):82–5.
[41] Spann K, Mileski WJ, Atiles L, Purdue G, Hunt J. The 1996 clinical research award use of a pneumatonomter in burn scar assessment. J Burn Care Rehabil 1996;17(6 (part1)):515–7.
[42] 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(3):489–500.
[43] 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.
[44] Schwanholt CA, Ridgway CL, Greenhalgh DG, Staley MJ, Gaboury TJ, Morress CS, et al. A prospective study of burn scar maturation in pediatrics: does age matter? J Burn Care Rehabil 1994;15(5):416–20.
[45] Branski LK, Herndon DN, Peteira C, Mlcak RP, Celis MM, Lee JO, et al. Longitudinal assessment of integra in primary burn management: a randomized pediatric clinical trial. Crit Care Med 2007;35(11):2615–23.
[46] Baryza MJ, Baryza GA. The Vancouver scar scale: an administration tool and its interrater reliability. J Burn Care Rehabil 1995;16(5):535–8.
[47] Forbes-Duchart L, Marshall S, Stock A, Cooper JE. Determination of inter-rater reliability in pediatric burn scar assessment using a modified version of the Vancouver scar scale. J Burn Care Res 2007;28(3):460–7.
[48] Fraulin FOG, Illmayer SJ, Tredget EE. Assessment of cosmetic and functional results of conservative versus surgical management of facial burns. J Burn Care Rehabil 1996;17(1):19–29.
[49] Tredget EE, Shankowsky H, Rajeet P, Nedelec B, Takashi LW, Ghahary A, et al. Transforming growth factor-[beta] in thermally injured patients with hyper scars: effects of interferon [alpha]-2b. Plast Reconstr Surg 1998;102(5):1317–28.
[50] Yeong EK, Mann R, Engrav LH, Goldberg M, Cain V, Costa B, et al. Improved burn scar assessment with use of a new scar-rating scale. J Burn Care Rehabil 1997;18(4):353–5.
[51] Fong SSL, Hung LK, Cheng JCY. The cutometer and ultrasonography in the assessment of postburn hypertrophic scar – a preliminary study. Burns 1997;23(Suppl. 1):S8–12.
[52] 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(3):273–84.
[53] Moieman NS, Vlachou E, Staiano JJ, Thawy Y, Frame JD. Reconstructive surgery with integra dermal regeneration template: histologic study, clinical evaluation, and current practice. Plast Reconstr Surg 2006;117(7S):74S–160S.
[54] Draaijers LJ, Botman YAM, Tempelman FRH, Kreis RW, Middelkoop E, Zuijlen PPM. Skin elasticity meter or subjective evaluation in scars: a reliability assessment. Burns 2004;30:109–14.
[55] Jarrett M, McMahon M, Stiller K. Physical outcomes of patients with burn injuries – a 12 month follow-up. J Burn Care Res 2008;29(6):975–84.
[56] Musgrave MA, Umraw N, Fish JS, Gomer M, Cartotto RC. The effect of silicone gel sheets on perfusion of hypertrophic burn scars. J Burn Care Rehabil 2002;23(3):208–14.
[57] Chou T-D, Chen S-L, Lee T-W, Chen S-G, Cheng T-Y, Lee C-H, et al. Reconstruction of burn scar of the upper extremities with artificial skin. Plast Reconstr Surg 2001;108(2):378–84.
[58] Lau JCM, Li-Tsang CWP, Zheng YP. Application of tissue ultrasound palpation system (TUPS) in objective scar evaluation. Burns 2005;31:445–52.
[59] Stewart CJ, Frank R, Forrester KR, Tulip J, Lindsay R, Bray RC. A comparison of two laser-based methods for determination of burn scar perfusion: laser doppler versus laser speckle imaging. Burns 2005;31:744–52.
[60] Aust MC, Knobloch K, Reimers K, Redeker J, Ipaktchi R, Altintas MA, et al. Percutaneous collagen induction therapy: an alternative treatment for burn scars. Burns 2010;36:836–43.
[61] van der Wal MBA, van Zuijlen PP, van de Van P, Middelkoop E. Topical silicone gel versus placebo in promoting the maturation of burn scars: a randomized controlled trial. Plast Reconstr Surg 2010;126(2):524–31.

[62] Bloemen MCT, Van Gerven MS, van der Wal MBA, Verhaegen PDHM, Middelkoop E. An objective device for measuring surface roughness of skin and scars. J Am Acad Dermatol 2011;64(4):706–15.
[63] Bloemen MCT, van Leeuwen MCE, van Vucht NE, van Zuijlen PPM, Middelkoop E. Dermal substitution in acute burns and reconstructive surgery: a 12-year follow-up. Plast Reconstr Surg 2010;125:1450–9.
[64] van der Kar A, Corion LUM, Smeulders MJCLJD, van der Horst CMAM, van Zuijlen PPM. Reliable and feasible evaluation of linear scar by the patient and observer scar assessment scale. Plast Reconstr Surg 2005;116(2): 514–22.
[65] Simons M, Tyack Z. Health professionals' and consumers' opinion: what is considered important when rating burn scars from photographs? J Burn Care Res 2011;32(2):275–85.