OPEN

PRS GLOBAL OPEN

SPECIAL TOPIC

Systematic Review on the Content of Outcome Measurement Instruments on Scar Quality

Michelle E. Carrière, MD*†‡§
Kelly A. A. Kwa, MD*¶
Louise E. M. de Haas, MD*
Anouk Pijpe, PhD, Bsc*
Zephanie Tyack, B Occ Thy, PhD‖
Johannes C. F. Ket**
Paul P. M. van Zuijlen, MD, PhD*†
Henrica C. W. de Vet, PhD§
Lidwine B. Mokkink, PhD§

Background: Measurements of scar quality are essential to evaluate the effectiveness of scar treatments and to monitor scars. A large number of scar scales and measurement devices have been developed, which makes instrument selection challenging. The aim of this study was to provide an overview of the content (ie, included items) of all outcome measurement instruments that measure scar quality in different types of scars (burn, surgical, keloid, and necrotizing fasciitis), and the frequency at which the instruments and included items are used.

Methods: A systematic search was performed in PubMed and Embase.com up to October 31, 2018. All original studies reporting on instruments that measured at least 1 characteristic of scar quality were included and the instrument's content was extracted.

Results: We included 440 studies for data extraction. Included instruments (N = 909) were clinician-reported scales (41%), measurement devices (30%), patient-reported scales (26%), and combined clinician- and patient-reported scales (3%). The Observer scale of the Patient and Observer Scar Assessment Scale, the Cutometer, the Patient Scale of the Patient and Observer Scar Assessment Scale, and the modified Vancouver Scar Scale were the most often used instrument in each of these categories, respectively. The most frequent assessed items were thickness, vascularity, pigmentation, pliability, pain, and itch.

Conclusion: The results of this study lay the foundation for our future research, which includes an international Delphi study among many scar experts, and an international focus group study among scar patients, aiming to elucidate how scar quality must be defined and measured from both professional and patient perspectives.
(Plast Reconstr Surg Glob Open 2019;7:e2424; doi: 10.1097/GOX.0000000000002424; Published online 30 September 2019.)

INTRODUCTION

Measurements of scar characteristics are essential to evaluate the effectiveness of scar treatments and to monitor scars over time. Several visual (eg, color), tactile (eg, pliability), and sensory (eg, itch) characteristics of the scar can contribute to overall scar quality. Measurements of scar quality can be performed by different types of outcome measurement instruments, including measurement devices and scar assessment scales. Measurement devices usually measure 1 characteristic of scar quality. Scar assessment scales allow for the evaluation of multiple characteristics of scar quality, from a clinician's and/or a patient's perspective. Clinician-reported scar scales capture clinicians' opinion on visual and tactile characteristics, whereas patient-reported scales enable the assessment of sensory characteristics in addition to patients' opinion on the visual and tactile characteristics of their scar(s).

The most appropriate outcome measurement instrument to measure scar quality should be chosen based on the following characteristics: (1) the content, (2) the type of instrument, (3) the feasibility, and (4) the measurement properties (eg, re-

*From the *Burn Center and Department of Plastic, Reconstructive and Hand Surgery, Red Cross Hospital, Beverwijk, Noord-Holland, the Netherlands; †Department of Plastic, Reconstructive and Hand Surgery, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam Movement Sciences, Amsterdam, the Netherlands; ‡Association of Dutch Burn Centers, Beverwijk, Noord-Holland, the Netherlands; §Department of Epidemiology and Biostatistics, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam Public Health Research Institute, Amsterdam, Noord-Holland, the Netherlands; ¶Department of Trauma Surgery, Leiden University Medical Center, Leiden, the Netherlands; ‖Centre for Children's Burns and Trauma Research, Child Health Research Centre, the University of Queensland, Brisbane, Australia; and **Medical Library, Vrije Universiteit, Amsterdam, the Netherlands*

Received for publication June 25, 2019; accepted July 9, 2019.

DOI: 10.1097/GOX.0000000000002424

Disclosure: *P.P.M. van Zuijlen is the developer of the Patient and Observer Scar Assessment Scale and Z. Tyack is the developer of the Brisbane Burn Impact Profile. This study was funded by the Dutch Burns Foundation (grant 16.106).*

Related Digital Media are available in the full-text version of the article on www.PRSGlobalOpen.com.

liability, validity).[1] However, this is a challenging task, as a large number of outcome measurement instruments are available, each measuring (characteristics of) scar quality in a slightly different way. To aid users in instrument selection and to promote standardization of measurements, recommendations on the best and most appropriate instrument are necessary, preferably based on a high-quality systematic review of the clinimetric properties of available outcome measurement instruments, thereby updating previous systematic reviews performed for this purpose.[2–7] However, before these recommendations can be provided for the construct scar quality, there must be consensus worldwide on how scar quality should be defined and measured and which scar characteristics are most important from the perspectives of professionals and patients. To reach this agreement, a complete understanding of everything that has been used or is being used to measure scar quality is necessary. Therefore, the aim of our systematic review is to provide an overview of the content (ie, included items) of all the outcome measurement instruments used in studies that measure scar quality in different types of scars and the frequency at which the instruments and included items are reported.

METHODS

Electronic Search Strategy

A systematic literature review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses statement (www.prisma-statement.org). PubMed and Embase.com were searched from inception up to October 31, 2018, by M.E.C. and J.C.F.K. The search strategies are represented in the figure (Supplementary Digital Content 1, which displays systematic search strategy, *http://links.lww.com/PRSGO/B215*). The search query included thesaurus terms and free-text terms for "scars" or "keloid" and "burns" or "skin" or "necrotizing fasciitis" or "surgical" and "assessment" or "tool" and "quality" or "outcome." Animal studies were excluded.

Selection of Studies

Two reviewers (M.E.C. and K.A.A.K.) independently screened all titles and abstracts to select eligible articles. Full-text screening and data extraction were performed by 3 different authors (M.E.C., K.A.A.K., L.E.M.d.H.) and were always checked by one of these authors. Differences between reviewers were discussed and if consensus was not reached, a third reviewer (L.B.M.) was consulted. The screening procedure was conducted with the use of the web-based software platform Covidence (www.covidence.org). To select relevant outcome measurement instruments, we applied the following criteria:

Types of Studies and Language

Original research studies, both quantitative and qualitative, only written in English were included. Conference abstracts, protocols, opinion papers, and systematic reviews were excluded. We did not appraise studies on the basis of their quality, nor did we perform a risk of bias analysis, as the purpose of this review was to determine the content of all of the outcome measurement instruments used in the literature to measure scar quality, including nonvalidated and ad hoc instruments, and not the results of included studies.

Population

Studies concerning humans of all ages with burn, surgical, keloid, and necrotizing fasciitis scars were included. Studies on patients with acne scars, scars resulting from other dermatological diseases, and scarring of tissue other than the skin were excluded.

Outcome Measurement Instruments

We defined scar quality as visual, tactile, and sensory characteristics. Studies were eligible when they used or reported outcome measurement instruments (ie, measurement devices or scar scales) that measured at least 1 specific characteristic of scar quality and defined how these characteristics were quantified. Studies that merely mentioned a scar characteristic without specifying how this characteristic was quantified were excluded. Studies using invasive methods or reporting on biological characteristics of scar quality alone (eg, obtained by histology) were not considered for this study, because these methods are not applicable in clinical practice.

Data Extraction

The following data were extracted from included studies directly into Excel tables: (1) study design; (2) characteristics of the patient population: sample size, age of participants, scar types; and (3) characteristics of the reported outcome measurement instrument(s): the type of instrument, number of items concerning scar quality, and the content of these items. Identified outcome measurement instruments were categorized based on their type (ie, clinician-reported scale, patient-reported scale, combined clinician-reported and patient-reported scale, and measurement device). When a scar assessment scale consisted of different subscales with an own sum score, the subscales were extracted as separate instruments. A scar scale consisting of different subscales with one total sum score was considered one instrument. This was also the case when the scale consisted of both clinician-reported and patient-reported items (ie, combined clinician-reported and patient-reported scales). Scar scales were considered modified versions of original scar scales when items were added, omitted or rephrased, or the response options or answering categories were changed. Each modified version was considered a separate scale. Measurement devices were categorized based on the construct they aim to measure. Of all included outcome measurement instruments, only items referring to scar quality were extracted. Therefore, items concerning physical and psychosocial functioning and items such as "overall opinion" and "satisfaction with treatment" were excluded. The identified items were categorized based on (our perception of) the intended meaning.

RESULTS

Literature Search

The literature search identified 2,506 records: 1,192 through PubMed and 1,314 through EMBASE (Fig. 1).

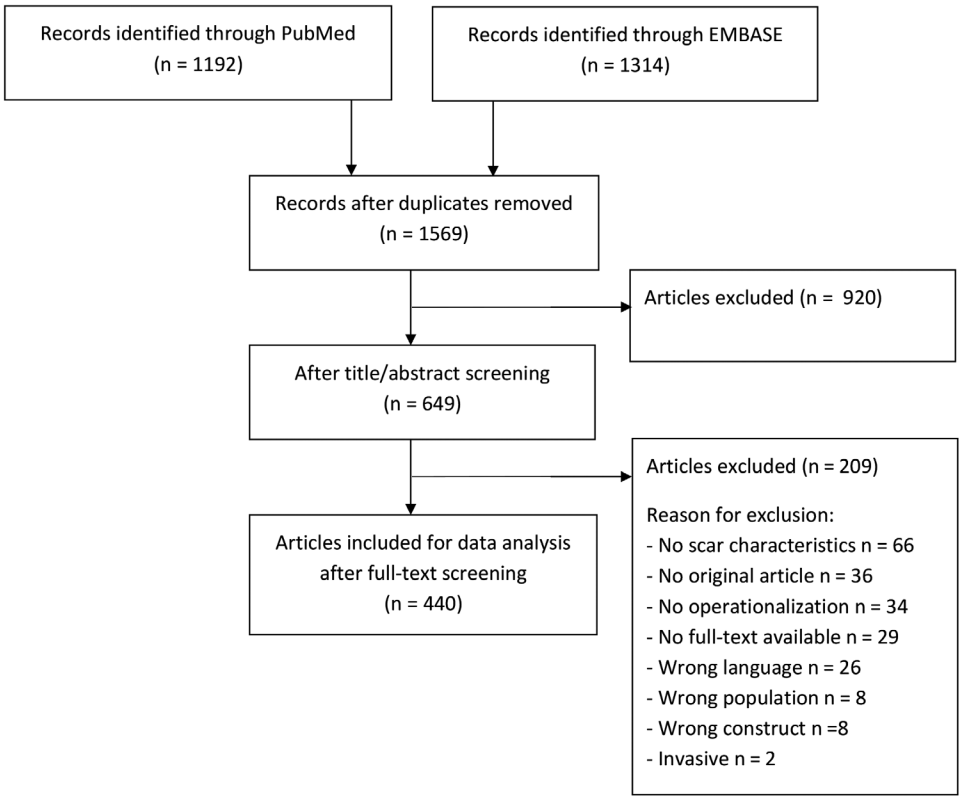

Fig. 1. Flow chart for the selection of studies according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA)-statement. .

After removal of duplicates, 1,569 articles remained, of which the titles and abstracts were screened. This resulted in the selection of 649 articles. After full-text screening, 440 articles were included for data extraction.

Characteristics of the Studies

Table 1 shows the characteristics of the included studies. The majority (88%) were studies in which the outcome measurement instruments were used to clinically evaluate patients. These studies used a median of 2 (range: 1–8) different outcome measurement instruments of scar quality. Other included studies were studies to develop outcome measurement instruments (5%) and clinimetric studies (7%) of individual instruments. Most studies included only patients with surgical scars (40%) or only patients with burn scars (36%). Fewer studies reported only on patients with keloid scars (10%). The remaining articles included 2 or more types of scars in their study (14%).

Characteristics of the Outcome Measurement Instruments

The majority of instruments used to measure scar quality were clinician-reported scales (41%), measurement devices (30%), and patient-reported scales (26%). Few instruments were combined clinician- and patient-reported scales (3%).

Table 1. Study Characteristics

Included Studies	N = 440 (100%)
Study design	
Clinical studies	386 (88%)
Clinimetric studies	31 (7%)
Development studies	23 (5%)
Population	
Patients with surgical scars	175 (40%)
Patients with burn scars	160 (36%)
Mixed*	62 (14%)
Patients with keloid scars	43 (10%)
Age category population	
Adults	214 (49%)
Mixed	146 (33%)
Children (<18 y)	30 (7%)
Unknown	50 (11%)
Included outcome measurement instrument	909 (100%)
Type of outcome measurement instrument	
Clinician-reported scar scales	375 (41%)
Measurement devices	271 (30%)
Patient-reported scar scales	236 (26%)
Combined clinician- and patient-reported scar scales	27 (3%)

*Mixed population consisting of ≥2 types of scars.

Clinician-reported Scales

Table (Supplemental Digital Content 2 a), *http://links.lww.com/PRSGO/B216*) provides an overview of identified clinician-reported scar scales, including the items used to measure scar quality, and their frequency of occurrence.

Table 2. Five Most Frequently Included Items in Different Types of Outcome Measurement Instruments

	Order of Most Frequently Assessed Items				
Type of Instrument	**First**	**Second**	**Third**	**Fourth**	**Fifth**
Clinician-reported scar scale	Thickness	Pigmentation	Vascularity	Pliability	Surface irregularity
Patient-reported scar scale	Pain	Itch	Color	Thickness	Pliability
Combined	Vascularity	Thickness	Itch	Pliability	Pain
Measurement device	Vascularity	Pliability	Thickness	Pigmentation	Moisture

We identified 31 different (versions of) clinician-reported scales. Most frequently used scales were the second version of the Patient and Observer Scar Assessment Scale (POSAS) (70 times),[8] the Vancouver Scar Scale (VSS) by Sullivan et al.[9] (61 times), and the modified VSS by Baryza and Baryza[10] (58 times). The VSS was modified most often. The 31 clinician-reported scales included a median of 4 (range: 1–7) items to measure scar quality. Besides these 31 scales, we identified 49 different scales without a name, which were categorized as "unlabeled." All unlabeled scales were reported only once. The 3 most frequently used items in all clinician-reported scales were thickness, pigmentation, and vascularity (Table 2).

Patient-reported Scales

Our search identified 39 different patient-reported (sub)scales (see table, Supplemental Digital Content 2b, *http://links.lww.com/PRSGO/B216*). The Patient Scale of the first and second POSAS version[8,11], which are identical, was the most frequently used patient-reported scale (111 times). The POSAS was followed by the Burn Specific Health Scale-brief.[12] This is a quality of life scale including 2 items of scar quality, which was reported 5 times. The following scales were third most reported (4 times): the Appearance scale of the Patient Scar Assessment Questionnaire,[13] the Dermatology Life Quality Index,[14] and the University of North Carolina "4P" Scar Scale.[15] The patient-reported scales used a median of 6 items (range: 1–13) to measure scar quality. Besides these 39 scales, we found 64 scales without a name, which were categorized as "unlabeled." Most frequently assessed items of scar quality within patient-reported scales were pain, itch, and color, as shown in Table 2.

Combined Patient- and Clinician-reported Scales

We found 9 instruments that combined clinician- and patient-reported items (see table, Supplemental Digital Content 3, which displays the overview of combined clinician-reported and patient-reported scar scales, included items and frequency of occurrence, *http://links.lww.com/PRSGO/B217*). The modified VSS by Nedelec et al.[16] was most often used (9 times). The second most reported scale was the modified VSS by Oliveira et al.[17] (6 times), followed by the Kyoto Scar Scale (2 times).[18] In addition to these 9 scales, we identified 4 different unlabeled scales, which were reported only once. Vascularity, thickness, and pliability were the most often assessed clinician-reported items and pain and itch the most often assessed patient-reported items in these scales (Table 2).

Measurement Devices

We found 50 different measurement methods, of which the Cutometer (32 times) was most frequently used, followed by the Dermaspectrometer (25 times) (see table, Supplemental Digital Content 4, which provides an overview of measurement devices used to measure scar quality, http://links.lww.com/PRSGO/B218, and calipers/rulers and ultrasound (both 24 times). The majority of measurement devices measured 1 item, although some of the included devices were able to measure additional items, up to a maximum of 5 items. The most frequently used items were vascularity, pliability, and thickness (Table 2).

DISCUSSION

This is the first study which used a systematic method to provide a complete overview of the content of all validated and nonvalidated outcome measurement instruments of scar quality, including measurement devices and scar assessment scales, and to provide the frequency that instruments and included items are used in the literature. The most frequently assessed items in all included outcome measurement instruments were thickness, color (consisting of vascularity and pigmentation), pliability, pain, and itch, indicating that these scar characteristics are considered most relevant. Heterogeneity in the content of instruments that measure the same construct (ie, scar quality) identified in this study indicates that there is a lack of consensus among patients, clinicians, and researchers on the most important characteristics of scar quality.

Clinician- and Patient-reported Scar Scales

Our search showed that clinician-reported scales were more frequently used than patient-reported scales. A likely explanation could be that clinicians traditionally have been the main assessors of scars. It was not until the last decades that there has been increased awareness of the importance of involving patients in clinical decision-making.[19] This movement toward patient involvement has become apparent in the field of scar assessment by the growing number of scar-specific patient-reported scales.[8,13,20–22] Some of these scales are not yet widely reported because either they are relatively new and unknown or their validation process is still ongoing.

The Observer Scale of the second version of the POSAS was the most frequently reported clinician-reported scar scale. However, if we had not distinguished modified versions of instruments, the VSS, including all its modified versions, would have been the most frequently reported clinician-reported scar scale. The frequent use of

the VSS and POSAS is supported by survey studies which identified these scales as the preferred scales to comprehensively measure scar quality among burn clinicians.[23,24] The Patient Scale of the POSAS was the most used patient-reported scale, with a frequency of 22 times as high as the second most frequently used patient-reported scale, the Burn Specific Health Scale-brief. This difference can be explained by the focus of our study on the construct scar quality and the fact that the Patient Scale of the POSAS is the only patient-reported scale that exclusively consists of items measuring scar quality, as opposed to other included scales which, in addition to items of scar quality, include items on quality of life or satisfaction. The predominance of the Patient Scale of the POSAS has highly influenced our results on the distribution of patient-reported items. To eliminate this effect, we recalculated the items without the inclusion of the Patient Scale of the POSAS. This did not change the order of frequency of items reflecting scar quality in patient-reported scales.

Measurement Devices

Measurement devices were found to be used nearly as often as clinician-reported scar scales. Measurement devices are deemed important in scar evaluation because they are thought of as "objective." However, they are not free of observers' interpretations nor measurement errors. Measurement devices are often expensive, impractical, time-consuming, and only capable of measuring 1 scar quality characteristic. Therefore, clinical studies mostly use a combination of scar scales and measurement devices to measure multiple aspects of scar quality.

Content of Outcome Measurement Instruments

When considering all types of outcomes measurement instruments, thickness, color (consisting of vascularity and pigmentation), and pliability were the most assessed visual and tactile items, whereas pain and itch were the most commonly assessed sensory items (Table 2). The items most frequently included in clinician-reported scales correspond to the items noted by clinicians as key elements of scar outcome measurement instruments in qualitative studies.[24,25] If we make the same comparison between items most frequently included in patient-reported scales and the results of qualitative studies among patients reporting on scar characteristics that matter the most to them, we find an agreement on pain, itch,[26–28] thickness, pliability, and color.[27,28] However, in addition to these characteristics, patients have also reported inflammation, sensation, surface irregularities, hardness, hydration, fragility, stretch ability, tightness, burning sensation, and sensitivity to be important aspects of their scars.[26,28–30] These aspects were identified in qualitative studies of which most were content validity studies (ie, interviews or focus groups) conducted to determine the content of patient-reported scales measuring health-related quality of life such as the Patient-Reported Impact of Scars Measure (PRISM)[20] and the Brisbane Burn Scar Impact Profile.[21] This indicates that the patient's perspective is not adequately assessed in the majority of included patient-reported scales. This can be explained by the lack of proper content validity studies as part of patient-reported scale development in many instruments. Qualitative studies are vital in scale development, not only to determine the most relevant items but also to ensure the comprehensiveness and the comprehensibility of the items, the instructions, and the answering categories used in the scale.[31]

Use of Outcome Measurement Instruments

During this systematic review, we have made several noteworthy observations regarding the use of outcome measurement instruments. Based on these observations, we would like to give some recommendations concerning the use and modification of validated scales, the terminology, and definition of items.

Many modifications of the content of established scar scales were found in this study. The most frequently modified scar scale was the VSS. Major modifications, such as adding or omitting items of the original scale, result in a new scale of which the outcomes are incomparable with the outcomes measured by the original scale. An example of a minor modification is the alteration of the answering categories of the item height: "0. flat, 1. <2 mm, 2. 2–5 mm, 3. >5 mm" by changing 3. into 5–6 mm, and by adding an extra fourth category: >6 mm. Although such a modification might seem minor and insignificant, a 5 mm raised scar would be classified differently by each scale, which makes these results less comparable. Furthermore, clinicians and researchers using modified scales might not be aware of the existing differences and may use the instruments interchangeably. Another precarious matter for users is that the specific content of used outcome measurement instruments is often not clearly described in the method section, nor appropriately referred to in the article. This makes it more difficult for readers to correctly assess the content of instruments, especially when several modified versions exist. Therefore, we would advise against modifications of validated scales. However, if done, validation of the new scale is crucial and asking the copyright holders of the original scale for permission.[32] Furthermore, we strongly advise to clearly describe the modifications made to the content of a selected scale in articles, and assigning either a version number or new name to it.

The items included in outcome measurement instruments were categorized according to our understanding of their intent. Items with different terminology but comparable meanings were grouped, resulting in synonyms for aspects of scar quality. Comparable synonyms were also described in recent qualitative studies with patients,[27,29] which confirms that differences exist in the way scar characteristics are defined. During this study, we encountered that the meaning of items was often poorly or not described and that multiple interpretations of a term existed. This complicated categorization of the items, and in some cases compelled us to examine the answering categories to find the meaning of items. For example, the item "contour" was used in different scales to refer to the shape, thickness, or surface irregularities of the scar, and "texture" was used to refer to the surface irregularities or pliability. We would like to recommend authors to provide

a clear definition of each item—especially when it concerns an item that can have different interpretations.

Users must realize that the content (eg, items and answering categories) of instruments are developed for a specific construct, population, and purpose of use.[31] If any of these conditions are changed, the quality of the scale should be assessed in the new context. We would like to give several examples of inappropriate usage of instruments that we came across during this study. First, the "Hollander Wound Evaluation Scale (HWES)," developed for the assessment of wounds, was used to assess scar quality. However, as wound severity is a different construct than scar quality, the items included in the HWES are not appropriate to measure scar quality—an issue which its developers also addressed when they modified the HWES into a scar-specific scale which is known as the Stony Brooks Scar Evaluation Scale.[33] Second is the use of clinician-reported scales as patient-reported or proxy-reported scales. It is debatable if patients or parents of patients value the same scar characteristics as clinicians, and if they are able to understand questions that are designed to be answered by clinicians. Third, multiple studies evaluated scar characteristics such as pliability, thickness, and vascularity from photographs. However, these characteristics are impossible to evaluate on the basis of photographs alone, as palpation of the skin is necessary to provide a sound assessment. With the rise of e-health and telemedicine, we anticipate that photographic scar assessment will become more popular in the future. Therefore, we advise to use instruments that take the pitfalls of 2-dimensional evaluation into account.

Limitations and Future Research

A limitation of this review is that we have only included studies written in English. As a result, we may have failed to capture specific scar aspects that are deemed important in non-English speaking countries. Another important drawback is that this review does not provide recommendations on the best available instrument for measuring scar quality. These recommendations are necessary to encourage clinicians and researchers to start using the same, most suitable instrument, and should be based on critical evaluation of the content, the quality (ie, clinimetric properties), and the feasibility of instruments.[1,34] However, before this can take place, it is of paramount importance that internationally, patients, clinicians, and researchers agree on which scar characteristics are essential for a comprehensive measurement of scar quality. The results of this study have formed the foundation of an international Delphi study among many scar experts (ie, professionals and researchers involved with scar evaluation) and an international focus group study among scar patients, aiming to elucidate how scar quality must be defined and measured from both perspectives. A universally established definition of scar quality will facilitate the evaluation of content validity of outcome measurement instruments, and therefore will assist in the process of providing recommendations on the most appropriate instrument. High-quality, standardized measurements of scar quality will improve the evaluation of treatment strategies, and limit reporting bias in studies. In addition, well-defined constructs help formulate clearly framed research questions, which is highlighted in the Cochrane handbook as the first and most important decision in preparing a systematic review.[35] Furthermore, it will enhance the comparability of study results worldwide, which will not only increase the quality of systematic reviews but also the level of evidence of the outcomes of meta-analyses on scar quality in the future. Moreover, less studies will be needed to come to a conclusion about the efficacy of scar treatment, reducing the amount of research waste in this field.

Michelle E. Carrière, MD
Burn Center and Department of Plastic, Reconstructive and Hand Surgery
Red Cross Hospital
Vondellaan 13, 1942 LE
Beverwijk, the Netherlands
E-mail: m.carriere@amsterdamumc.nl

REFERENCES

1. De Vet HCW, Terwee CB, Mokkink LB, et al. *Measurement in Medicine*. Cambridge, United Kingdom: Cambridge University Press; 2011.
2. Verhaegen PD, van der Wal MB, Middelkoop E, et al. Objective scar assessment tools: a clinimetric appraisal. *Plast Reconstr Surg*. 2011;127:1561–1570.
3. van der Wal MB, Verhaegen PD, Middelkoop E, et al. A clinimetric overview of scar assessment scales. *J Burn Care Res*. 2012;33:e79–e87.
4. Mundy LR, Miller HC, Klassen AF, et al. Patient-reported outcome instruments for surgical and traumatic scars: a systematic review of their development, content, and psychometric validation. *Aesthetic Plast Surg*. 2016;40:792–800.
5. Vercelli S, Ferriero G, Sartorio F, et al. Clinimetric properties and clinical utility in rehabilitation of postsurgical scar rating scales: a systematic review. *Int J Rehabil Res*. 2015;38:279–286.
6. Tyack Z, Simons M, Spinks A, et al. A systematic review of the quality of burn scar rating scales for clinical and research use. *Burns*. 2012;38:6–18.
7. Lee KC, Dretzke J, Grover L, et al. A systematic review of objective burn scar measurements. *Burns Trauma*. 2016;4:14.
8. van de Kar AL, Corion LU, Smeulders MJ, et al. Reliable and feasible evaluation of linear scars by the patient and observer scar assessment scale. *Plast Reconstr Surg*. 2005;116:514–522.
9. Sullivan T, Smith J, Kermode J, et al. Rating the burn scar. *J Burn Care Rehabil*. 1990;11:256–260.
10. Baryza MJ, Baryza GA. The Vancouver scar scale: an administration tool and its interrater reliability. *J Burn Care Rehabil*. 1995;16:535–538.
11. Draaijers LJ, Tempelman FR, Botman YA, et al. The patient and observer scar assessment scale: a reliable and feasible tool for scar evaluation. *Plast Reconstr Surg*. 2004;113:1960–1965; discussion 1966.
12. Kildal M, Andersson G, Fugl-Meyer AR, et al. Development of a brief version of the burn specific health scale (BSHS-B). *J Trauma*. 2001;51:740–746.
13. 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;123:1481–1489.
14. Finlay AY, Khan GK. Dermatology life quality index (DLQI)–a simple practical measure for routine clinical use. *Clin Exp Dermatol*. 1994;19:210–216.

15. Hultman CS, Edkins RE, Wu C, et al. Prospective, before-after cohort study to assess the efficacy of laser therapy on hypertrophic burn scars. *Ann Plast Surg*. 2013;70:521–526.
16. 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;21:205–212.
17. Oliveira GV, Chinkes D, Mitchell C, et al. Objective assessment of burn scar vascularity, erythema, pliability, thickness, and planimetry. *Dermatol Surg*. 2005;31:48–58.
18. Yamawaki S, Naitoh M, Ishiko T, et al. Keloids can be forced into remission with surgical excision and radiation, followed by adjuvant therapy. *Ann Plast Surg*. 2011;67:402–406.
19. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango). *Soc Sci Med*. 1997;44:681–692.
20. Brown BC, McKenna SP, Solomon M, et al. The patient-reported impact of scars measure: development and validation. *Plast Reconstr Surg*. 2010;125:1439–1449.
21. Tyack Z, Ziviani J, Kimble R, et al. Measuring the impact of burn scarring on health-related quality of life: development and preliminary content validation of the Brisbane burn scar impact profile (BBSIP) for children and adults. *Burns*. 2015;41:1405–1419.
22. Klassen AF, Ziolkowski N, Mundy LR, et al. Development of a new patient-reported outcome instrument to evaluate treatments for scars: the SCAR-Q. *Plast Reconstr Surg Glob Open*. 2018;6:e1672.
23. Lumenta DB, Siepmann E, Kamolz LP. Internet-based survey on current practice for evaluation, prevention, and treatment of scars, hypertrophic scars, and keloids. *Wound Repair Regen*. 2014;22:483–491.
24. Simons M, Tyack Z. Health professionals' and consumers' opinion: what is considered important when rating burn scars from photographs? *J Burn Care Res*. 2011;32:275–285.
25. Forbes-Duchart L, Cooper J, Nedelec B, et al. Burn therapists' opinion on the application and essential characteristics of a burn scar outcome measure. *J Burn Care Res*. 2009;30:792–800.
26. Brewin MP, Homer SJ. The lived experience and quality of life with burn scarring-the results from a large-scale online survey. *Burns*. 2018;44:1801–1810.
27. Jones LL, Calvert M, Moiemen N, et al; PEGASUS team. Outcomes important to burns patients during scar management and how they compare to the concepts captured in burn-specific patient reported outcome measures. *Burns*. 2017;43:1682–1692.
28. Brown BC, McKenna SP, Siddhi K, et al. The hidden cost of skin scars: quality of life after skin scarring. *J Plast Reconstr Aesthet Surg*. 2008;61:1049–1058.
29. Simons M, Lim PCC, Kimble RM, Tyack Z. Towards a clinical and empirical definition of burn scarring: a template analysis using qualitative data. *Burns*. 2018;44:1811–1819.
30. Simons M, Price N, Kimble R, et al. Patient experiences of burn scars in adults and children and development of a health-related quality of life conceptual model: a qualitative study. *Burns*. 2016;42:620–632.
31. Terwee CB, Prinsen CAC, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res*. 2018;27:1159–1170.
32. Anfray C, Arnold B, Martin M, et al; ISOQOL Translation and Cultural Special Interest Group (TCA-SIG). Reflection paper on copyright, patient-reported outcome instruments and their translations. *Health Qual Life Outcomes*. 2018;16:224.
33. Singer AJ, Arora B, Dagum A, et al. Development and validation of a novel scar evaluation scale. *Plast Reconstr Surg*. 2007;120:1892–1897.
34. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27:1147–1157.
35. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration; 2011. www.handbook.cochrane.org.